ASSESSMENT OF FREQUENCY, DEGRADATION, NORMALIZATION, INHIBITION, AND POTENTIAL SURROGATES OF THE SARS-COV-2 GENE TARGETS

by

EMILY HUNAWILL

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN BIOMEDICAL SCIENCES: HEALTH AND ENVIRONMENTAL CHEMISTRY

2023

Oakland University
Rochester, Michigan

Doctoral Advisory Committee:

David Szlag, Ph.D., Chair
Colin Wu, Ph.D.
Adam Avery, Ph.D.
Doug Wendell, Ph.D.
Randal Westrick, Ph.D.

To my loving parents and grandparents,

who gave me strength and support throughout this process

ABSTRACT

ASSESSMENT OF FREQUENCY, DEGRADATION, NORMALIZATION, INHIBITION, AND POTENTIAL SURROGATES OF THE SARS-COV-2 GENE TARGETS

by

EMILY HUNAWILL

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the virus responsible for the coronavirus disease 2019 (COVID-19) pandemic. Our study focuses on identifying and quantifying the presence of SARS-CoV-2 in the sewage based on identifying three gene targets which target the nucleocapsid gene (N1 and N2) and the envelope gene (E). The first set of experiments looked at the degradation of these targets at 4°C, 25°C, and 35°C. As temperature increased so did the rate constants. Based on the half-life data, N1 degraded faster than N2 at 4°C, all targets degraded at the same rate at 25°C, and E degraded faster than N1 and N2 at 35°C. Frequency of the gene targets was then assessed. E occurred less frequently than N1 and N2 adding 2.3% more SARS-CoV-2 detections leading to its removal from testing. The N1 and N2 gene targets were both necessary with removal of one resulting in minimally 9.6% loss of SARS-CoV-2 detections. We investigated normalization of the N1 and N2 gene targets with four different fecal indicators (pepper mottled mild virus, HF183, lachno3, and crAssphage) to improve the correlation between the sewage signal and clinical cases. Normalizing the

data did not result in an increased correlation between the clinical and sewage data. Bovine coronavirus (BCoV) and Phi6 viruses were evaluated as surrogates to estimate SARS-CoV-2 inhibition in a single duplex reaction. Duplexing these targets was successful without significant signal loss, and these targets were used to estimate SARS-CoV-2 inhibition. Compared to BCoV, Phi6 suggested over 10% more fully inhibited samples. Several workflow modifications including bovine serum albumin (BSA), dilution, polyvinylpyrrolidone, and pasteurization were applied to the sewage samples in an attempt to reduce sample inhibition. BSA was able to reduce inhibition for both N1 and N2. Lastly, BCoV, Phi6, and murine hepatitis virus (MHV) retrieval were compared to the retrieval of the SARS-CoV-2 gene targets in sewage to determine if they accurately depicted the amount of inhibition the SARS-CoV-2 targets exhibited. BCoV and MHV were better surrogates to assess SARS-CoV-2 inhibition. SARS-CoV-2 RNA concentration in sewage were shown to correlate with COVID-19 cases.

TABLE OF CONTENTS

TABLE OF CONTENTS—Continue

TABLE OF CONTENTS—Continue

TABLE OF CONTENTS—Continue

TABLE OF CONTENTS—Continue

LIST OF TABLES

LIST OF TABLES—Continued

LIST OF FIGURES

LIST OF ABBREVIATIONS

SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
COVID-19	Coronavirus Disease 2019
CoVs	Coronaviruses
WHO	World Health Organization
ddPCR	Droplet Digital Polymerase Chain Reaction
CP/L	Gene Copies per Liter of Sewage
C	Concentration at Any Time
Co	Concentration at Time Zero
Std. Dev	Standard Deviation
CV	Correlation of Variation
T90	Shelf Life
t1/2	Half-Life
A	Pre-Exponential Factor
Ea	Activation Energy
R	Universal Gas Constant
CDC	Centers for Disease Control and Prevention
MDL	Method Detection Limit
RdRp	RNA-dependent RNA Polymerase
PMMoV	Pepper Mild Mottle Virus
WWTPs	Waste Water Treatment Plants
BCoV	Bovine Coronavirus
ZEN/IABKFQ	ZEN™/Iowa Black® Fluorescent Quencher
HCoVs	Human Coronaviruses
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
P/P	Primer/Probe
MERS	Middle East Respiratory Syndrome

SARS	Severe Acute Respiratory Syndrome
BSA	Bovine Serum Albumin
PVP	Polyvinylpyrrolidone
MHV	Murine Hepatitis Virus
PEG	Polyethylene Glycol 8,000
NaCl	Sodium Chloride

CHAPTER ONE

TEMPERATURE BASED DEGRADATION OF THE E, N1, AND N2 SARS-COV-2 GENE TARGETS IN WASTEWATER OVER TIME

1.1 Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a virus that primarily causes a respiratory disease in humans known as coronavirus disease 2019 (COVID-19). In addition to the respiratory system, the virus often infects the gut and is shed in the feces. This virus is detected in a clinical setting with a nasopharyngeal swab, but it can also be detected in sewage. We determined the presence of SARS-CoV-2 in a community by measuring and detecting three gene targets (E, N1, and N2) in wastewater. Both clinical and wastewater sampling have their advantages. Clinical testing is advantageous because each individual is tested, and a negative or positive result is obtained. However, clinical testing is limited to the individuals who choose to get tested which could greatly underestimate the number of detected cases. In contrast, testing sewage monitors whole communities regardless of community participation and can detect asymptomatic cases. Sewage surveillance may be problematic given that the gene targets may be highly diluted, and there can be many inhibitors present in the sewage matrix. Our study looked at the degradation of the SARS-CoV-2 gene targets (E, N1, and N2) over time at 4°C, 25°C, and 35°C. As expected, the rate constants were the lowest at 4°C and the highest at 35°C. For the unpasteurized sewage, the first-order rate constants were 0.0032hrs^{-1} (4°C), 0.0138hrs^{-1} (25°C), and 0.0256hrs^{-1} (35°C) for E, 0.00323hrs^{-1} (4°C), 0.0142hrs^{-1} (25°C), and 0.0108hrs^{-1} (35°C) for N1, and 0.00297hrs^{-1} (4°C), 0.0135hrs^{-1} (25°C), and

$0.0225hrs^{-1}$ (35°C) for N2. At 4°C N1 degraded faster than N2. Other combinations were not statistically different. At 25°C all gene targets degraded at a similar rate. At 35°C the E gene target appeared to degrade faster than N2. Other rates were not statistically different from one another. Pasteurizing the sewage caused the RNA targets to degrade more slowly.

1.2 Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an RNA virus that is enveloped and single-stranded (Brant et al., 2021). In general, coronaviruses (CoVs) including SARS-CoV-2 have large genomes with a length generally near 30kb (Bartas et al., 2022). The SARS-CoV-2 virus is a pathogenic and very transmissible leading to coronavirus disease 2019 (COVID-19) that arose in China towards the end of 2019 (Hu et al., 2021). Initial diagnosis of what is now known as COVID-19 ranged from December 18^{th} through the 29^{th} in 2019 (Worobey et al., 2022). The origin of this virus is purported to be a market in Wuhan, China known as the Hunan seafood market which sells various creatures including snakes, palm civets, bats, and raccoon dogs (Shereen et al., 2020). The likely origin of COVID-19 is from bats. Three species of horseshoe bats harbor viruses (BANAL-236, BANAL-103, and BANAL-52) that were over 95% identical to SARS-CoV-2 (Mallapaty, 2021).

COVID-19 is disease that mainly causes vascular and respiratory illness (Cascella et al., 2022a). This disease has many symptoms in common with the cold such as: sore throat, cough, and runny nose, but COVID-19 may also cause nausea, diarrhea, and vomiting which are not associated with the common cold (*COVID-19, Cold, Allergies and the Flu*, n.d.). In general, COVID-19 produces mild symptoms in 81% of the symptomatic

infections (Hernandez Acosta et al., 2022). Mild disease was associated with sore throat, malaise, fever, and myalgia (Hernandez Acosta et al., 2022). Despite occurring in late 2019, COVID-19 spread across the world and continues to infect a large fraction of the population. The World Health Organization (WHO) has a COVID-19 dashboard which shows the collective total of COVID-19 cases across the world. As of March 23, 2023, the WHO has reported over 760 million cases of COVID-19 and over 6.8 million deaths caused by COVID-19 globally (*WHO Coronavirus (COVID-19) Dashboard*, n.d.). This corresponds to a death rate of about 1% across the globe.

To detect SARS-CoV-2 in a clinical setting, nasopharyngeal swab testing for SARS-CoV-2 is considered the gold standard due to it being used in the past to detect other respiratory viruses (Tan et al., 2021). Although nasopharyngeal swab testing is the gold standard, some studies suggest saliva as an alternative (Beyene et al., 2021). A study by Tsujimoto et al. found that nasopharyngeal swabs, nasal swabs, and saliva samples had the following amount of sensitivity for SARS-CoV-2 detection: 100.0% sensitivity (nasopharyngeal swabs), 86.4% sensitivity (nasal swabs), and 63.6% sensitivity (saliva samples) (Tsujimoto et al., n.d.). Other human bodily fluids can contain SARS-CoV-2 and sewage contains viruses shed in urine, sputum, feces, vomit, and blood (Wolfe, n.d.)., Urine and feces will likely be the most common human bodily fluids found in the sewage. Urine samples usually do not test positive for SARS-CoV-2 with only 4.5% detection of SARS-CoV-2 in urine when testing 533 individuals compiled from 39 studies (Mumm et al., 2021). One study noted that 27/62 (43.5%) of stool samples analyzed had SARS-CoV-2 signal (Cerrada-Romero et al., 2022). A pooled study incorporating 95 articles was in agreement with Cerrada-Romero et al. suggesting

934/2,149 (43.5%) of anal swabs or stool samples were positive for SARS-CoV-2 (van Doorn et al., 2020). Therefore, it is likely that feces will be the main source of SARS-CoV-2 in sewage.

Sewage surveillance is able to monitor the wastewater collected from select buildings and wastewater treatment plants allowing a group sample to be tested for SARS-CoV-2 or other viruses (Haque et al., 2022). This type of SARS-CoV-2 detection can assess if specific areas are at risk of COVID-19 outbreaks to guide public health decisions in institutional settings such as prison, college, and nursing homes and in community settings (Harris-Lovett et al., 2021). The ability of SARS-CoV-2 to be detected in the sewage was first noted in April 2020 in Spain, the Netherlands, Australia, and the United States (F. Wu et al., 2022). However, sewage surveillance had been used earlier to detect polio and other enteroviruses (F. Wu et al., 2022). Other viruses such as hepatitis A and E, enterovirus D, and mastadenovirus have been monitored in sewage to determine which strains were present in certain areas of Australia and France (Farkas et al., 2020). They suggested that it may be advantageous to monitor the sewage to act as an early warning system of disease outbreaks (Farkas et al., 2020). Sewage surveillance is cost-effective and able to preemptively indicate rising COVID-19 cases allowing health interventions to be implemented earlier (Karthikeyan et al., n.d.). Universities can benefit tremendously from sewage surveillance because campuses have the potential to be COVID-19 hot spots and early detection allows targeted diagnostic testing in buildings and implementation of measures to slow the spread. (Karthikeyan et al., n.d.). Testing wastewater also has the purported advantage over clinical testing based on its ability to detect rising cases of COVID-19 earlier than clinical testing (Daughton, 2020). One study detected SARS-

CoV-2 variants of concern two weeks earlier in sewage than in the clinical data (Karthikeyan et al., 2022). They also identified many occasions of SARS-CoV-2 spread that was not identified in the clinical data (Karthikeyan et al., 2022). Because so many cases of COVID-19 are asymptomatic or mild (80%), sewage surveillance can detect more of these cases that would otherwise go unreported (Melvin et al., 2021). A meta study (N=28) found that the percentage of asymptomatic infections was between 1.4-78.3% of patients with an average of 25% (Alene et al., 2021). Testing for SARS-CoV-2 in wastewater overcomes other biases present in clinical testing (Larsen & Wigginton, 2020). For instance, deaths and hospitalizations lag behind COVID-19 transmission by weeks (Larsen & Wigginton, 2020). Also, clinical testing generally relies on willing participation that may be limited by fear of quarantine and isolation, and the ability for willing individuals to pay (Larsen & Wigginton, 2020). Nevertheless, wastewater surveillance also has issues. Clinical testing provides an exact number of individuals that are and are not infected with COVID-19. Wastewater data cannot provide an exact number of infected people. However, Gerrity et al. provide a formula that can estimate the number of COVID-19 infected people based on the SARS-CoV-2 gene target concentration (Gerrity et al., 2020). Another challenge to detecting SARS-CoV-2 in wastewater is that the SARS-CoV-2 signal may be highly diluted in sewage and the presence of chemicals that interfere with the SARS-CoV-2 PCR chemistry (Sims & Kasprzyk-Hordern, 2020). Consequently, clinical and wastewater detection of SARS-CoV-2 should be considered complimentary methods, and both should be applied judiciously.

In order to better understand the relationship between SARS-CoV-2 gene targets and clinical cases it is essential to understand how rapidly the SARS-CoV-2 gene targets (N1, N2, and E) degrade in sewage. We used droplet digital polymerase chain reaction (ddPCR) that is purported to be more sensitive than quantitative PCR (qPCR) and less sensitive to inhibition (Sedlak et al., 2014). We measured degradation rates for each gene target, N1, N2, and E, at 4°C, 25°C, and 35°C.

1.3 Materials and Methods

1.3.1 Sample Delivery, Sample Selection, and Sample Concentration

Our workflow used the methods described by Flood et al (Flood et al., 2021) as a starting point. Sewage samples were delivered to the lab in 500mL portions. Samples were transported in a cooler with icepacks while in transit to our lab. Samples were processed immediately upon arrival. The 500mL bottle of sewage was mixed by inverting the bottle between 40 and 60 times to minimize foaming. A 45 mL homogenous sewage sample was decanted into a 50mL conical bottom centrifuge tube. Approximately 4.0g of polyethylene glycol 8000 (catalog number 97061-100 from VWR) and 0.59g sodium chloride (catalog number 470302-520 from VWR) were added to the sample yielding a concentration of 8% polyethylene glycol and 0.2M sodium chloride. After the solids and solutions were combined, the centrifuge tubes were transferred to a rotator where the tubes were rotated at 70rpm for 2hrs. The rotator was held at 4°C. After mixing, the samples were centrifuged at 3,650 rpm for 1 hr. Once centrifugation was completed, the supernatant was discarded, and a 1-2 mL pellet remained.

The degradation experiments utilized 20mL portions instead of 45mL portions of sewage to conserve the stock sewage. Consequently only 1.6g of polyethylene glycol and 0.234g

of sodium chloride were added to the sewage samples for the degradation experiments. Stock sewage was held in an incubator at 25°C or 35°C or stored in a refrigerator for the 4°C experiments. Some samples were pasteurized before taking the original sample at time zero. The samples were pasteurized for 2hrs at 60°C and inverted every 20-30mins. After pasteurization, the sample at time zero was collected and the rest of the sewage was incubated at the prescribed temperature for the duration of the experiment. For samples that were to be pasteurized, the delivered sewage was first split into two bottles (250mL of sewage in each). After this, half of the sample was pasteurized, and the other half was not pasteurized. Sewage was stirred at 60rpm while samples were collected. The remaining parts of the procedure were the same for the samples run for the degradation experiments, and the regular SARS-CoV-2 detection-based experiments.

1.3.2 RNA Elution and Purification

The QIAamp Viral RNA Mini Kit (catalog number 52904) was purchased from QIAGEN to purify and elute the RNA from the sewage samples. To begin the process, buffer AVL and carrier RNA were mixed in portions of 800µL and 8µL per sample, respectively. A 200µL portion of pellet was then pipetted into a 2mL cryovial containing the buffer and RNA mix. The cryovial was vortexed for 10s before incubation for 10mins at room temperature. After incubation, 200proof ethanol was added to the cryovial in an 800µL portion, and the cryovial was vortexed for 10s. At this point in the procedure there was 1.8mL solution in the cyrovial. This solution was run through the QIAamp mini spin column in order to obtain the extracted RNA. The 1.8mL solution was added in three 630µL quantities to the spin column. Three additions were done because the spin column could not hold enough solution for a single addition. Each 630µL portion of solution was

added to the spin column and centrifuged sequentially. The flow through was not retained. The centrifugation settings were 1min at 8,000rpm. These centrifuge settings were the same for every addition except when the buffer AW2 was added. The waste collected in the collection tube under the spin column was discarded after every addition until the elution step (buffer AVE). After the whole 1.8mL solution was added to the spin column and centrifuged, the wash buffers were added. A 500µL portion of buffer AW1 was added to the spin column and centrifuged. Next, a 500µL portion of buffer AW2 was added to the spin column. At this point in the process (and only during this step) the centrifuge speed and duration was increased to 14,000rpm for 4min. After centrifugation, centrifugation settings were returned to normal, and the spin column was transferred to a 1.5mL microcentrifuge tube. The microcentrifuge tube was used to collect the RNA eluted from the spin column. Buffer AVE was added to the spin column in a 40µL portion and incubated at room temperature for 1min. After incubation, the spin column was centrifuged. This step was repeated which resulted in a final volume of 80µL of RNA eluate.

1.3.3 Amplification of DNA and Detection of Target Sequences

The assay mix was made using the solutions obtained from the One-Step RT-ddPCR Advanced Kit for Probes (catalog number 1864021) produced by Bio-Rad. This kit provided everything needed for the assay mix except the water and the primer/probe mix. Table 1.1 shows the primer/probe sequences used to detect the SARS-CoV-2 gene targets.

Table 1.1: SARS-CoV-2 Gene Target Sequences

Target	Sequences (5'-3')	References
E	ACAGGTACGTTAATAGTTAATAGCGT (forward primer) ATATTGCAGCAGTACGCACACA (reverse primer) FAM-ACACTAGCCATCCTTACTGCGCTTCG-BHQ1 (probe)	(Corman et al., 2020a)
N1	GACCCCAAAATCAGCGAAAT (forward primer) TCTGGTTACTGCCAGTTGAATCTG (reverse primer) FAM-ACCCCGCATTACGTTTGGTGGACC-BHQ1 (probe)	(CDC, 2020c)
N2	TTACAAACATTGGCCGCAAA (forward primer) GCGCGACATTCCGAAGAA (reverse primer) HEX-ACAATTTGCCCCCAGCGCTTCAG-BHQ1 (probe)	(CDC, 2020c)

Table 1.1: Sequences used to detect the SARS-CoV-2 gene targets N1, N2, and E.

The assay mix composition varied if the reactions were run for an E singleplex compared to an N1-N2 duplex. Table 1.2 shows the components that were used for a singleplex reaction and a duplex reaction. The values in Table 1.2 were based on a single reaction.

Table 1.2: Assay Mix Calculations for a Single Well

Reagent	Singleplex	Duplex
Supermix	5.5µL	5.5µL
Reverse transcriptase	2.2µL	2.2µL
Dithiothreitol	1.1µL	1.1µL
Primer/probe mix 1	3.3µL	3.3µL
Primer/probe mix 2	N/A	3.3µL
Water	4.4µL	1.1µL
Total	16.5µL	16.5µL

Table 1.2: The assay mix calculations utilized for a singleplex and a duplex reaction. These calculations were based on the amount of assay mix required for one reaction (one well).

Singleplex and duplex reactions only varied based on the amount of water and the additional primer/probe mix which would be added for a duplex reaction. Samples and controls (negative, no template, and positive controls) were pipetted into a 96-well plate with the correct assay mix. The N1 and N2 gene targets were ran as a duplex reaction, and the E gene target was run as a singleplex reaction. Each well had 5.5µL of sample/control and 16.5µL of assay mix. Each sample and control were run in triplicate per assay mix. Once plating was finished, the plate was sealed with foil. This plate was then vortexed to mix the sample and the assay mix. Lastly, the plate was spun in a plate spinner to collect the liquid at the bottom of the 96-well plate. This plate was then placed into the Automated Droplet Generator (catalog number 1864101 Bio-Rad) to generate droplets. The Automated Droplet Generator generated droplets into another 96-well plate. This new 96-well plate was sealed with foil and placed into the thermocycler. The thermocycler settings are listed in Table 1.3.

Table 1.3: Thermocycling Settings for SARS-CoV-2 Gene Targets

Step	**Temperature**	**Time**	**Cycles**
1	25°C	3mins	1
2	50°C	1hr	1
3	95°C	10mins	1
4	95°C	30s	40*
5	55°C	1min	40*
6	98°C	10min	1
7	4°C	30min	1
8	4°C	Infinity	1

Table 1.3: The thermocycling settings used for the N1, N2, and E gene targets. *Step 4 and 5 alternate back and forth 40 times. Steps are completed in chronological order.

The thermocycling settings were based on the instructions for the 1-Step RT-ddPCR Advanced Kit for Probes for general settings, and the annealing temperatures from Flood et al. (*One-Step RT-DdPCR Advanced Kit for Probes*, n.d.; Pearson et al., 2021). Once thermocycling was complete, the plate was held at room temperature for 15mins before reading the plate with the QX200 Droplet Reader (catalog number 1864003) from Bio-Rad. Once the plate was read, thresholds for positive and negative droplets were set and the plate was quantified.

1.3.4 Analyzing the Data

The droplet reader was paired with QuantaSoft™ Software to score the data. The data obtained from this software provided information on the number of droplets (accepted, positive, and negative), concentration, and copies per 20µL well. The positive droplets were assessed to determine if the sample was positive for SARS-CoV-2. Three or more positive droplets were considered a positive hit. Accepted droplets were assessed to determine the validity of the values obtained from the QuantaSoft™ Software. Wells that had greater than 10,000 accepted droplets were considered acceptable. The copies per 20µL well was the only value used from the data produced by the QuantaSoft™ Software to calculate the concentration of the various SARS-CoV-2 gene targets. This value was utilized along with the volume of the pellet which was determined experimentally and other experimental values that were constant such as the eluted RNA volume, amount of pellet used, and initial volume. Equation 1.1 shows the formula utilized to calculate the concentration of the SARS-CoV-2 gene targets. The concentration was in units of gene copies per liter of sewage (CP/L).

(Equation 1.1)

$$\frac{CP}{L} = copies\ per\ \mu L \times \frac{pellet\ volume\ (mL)}{extraction\ volume\ (0.2mL)} \times \frac{eluate\ volume\ (80\mu L)}{initial\ volume\ (45mL)} \times \frac{1000mL}{1L}$$

Equation 1.1: Concentration Calculation for the SARS-CoV-2 Gene Targets
This formula shows how the concentrations for the N1, N2, and E gene targets were calculated.

To obtain the degradation rate constants, the CP/L at any given time (C) was divided by the CP/L at time 0 (Co). Next, the natural logarithm of that value was taken (ln(C/Co)). These natural logarithm values were then plotted against time (hrs). The absolute value of the slope was the first-order degradation rate constant (hrs^{-1}). Rate constants were compared for the targets at three temperatures to determine if they all degraded equally. Arrhenius plots were made by plotting the natural logarithm of the degradation rate constant against the inverse of temperature in Kelvin. The Arrhenius plots allow the degradation rate constants for the SARS-CoV-2 gene targets to be calculated at any temperature.

1.4 Results

Zero-order, first-order, and second-order reactions were assessed to determine which order had the best linear fit to our data. Figure 1.1 shows an example of each graph used to assess the reaction order. Trend lines were not shown in the graph to allow the curves in graphs to be viewed without interference from the trendline.

Figure 1.1: Reaction Order Example Based on One Sewage Sample

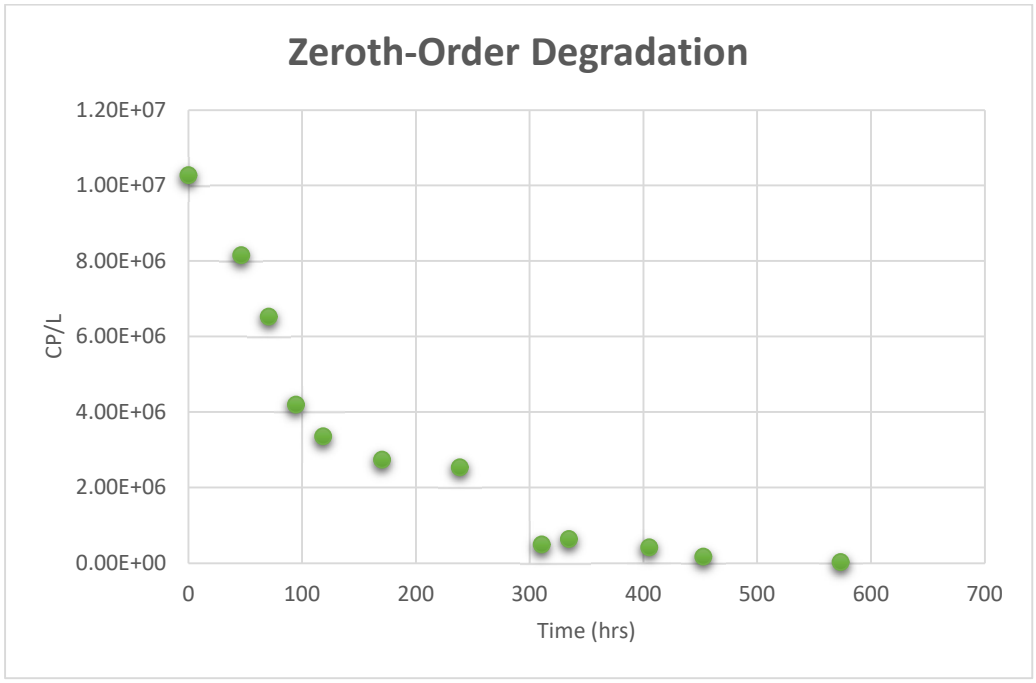


Figure 1.1: Reaction Order Example Based on One Sewage Sample (Continued)

Figure 1.1: This figure shows an example of the graphs obtained when the data was plotted for one of our samples. This sample was plotted to observe zero-order, first-order, and second-order reaction rates.

As shown, the graphs for zero-order and second-order reactions have curvature. The data plotted for a first-order reaction were linear. The following R^2 values were obtained for each reaction order: 0.76 (zero-order), 0.96 (first-order), 0.45 (second-order). The zero-order reaction had a reasonable R^2 value, but in general, the first-order reaction was more linear. Based on the curvatures and R^2 values we concluded degradation was best represented by a pseudo-first-order reaction.

Stirred and non-stirred sewage samples were compared at 25°C to determine if this slow speed changed the rate constant. Samples were stirred at 60rpm to simulate the movement of the sewage. This speed was also used to mix sewage in another study while samples were taken, but their study focused on anaerobic digestion of sewage (Barati Rashvanlou et al., n.d.). On average, the stirred samples degraded 1.2 times faster than the non-stirred samples. Their average values are listed in Table 1.4. The stirred sewage degraded statistically faster for both the N1 ($p=0.022$) and N2 gene targets ($p=0.031$), but the E gene target did not degrade statistically ($p=0.090$) faster when stirred. There was a high correlation between the stirred and non-stirred sewage with Pearson's correlation coefficients effectively 1.0.

Table 1.4: Stirred and Non-Stirred Sewage Comparison

Gene Target	**Average k value (hrs^{-1}) (stirred)**	**Average k value (hrs^{-1}) (non-stirred)**	**stirred/non-stirred**
E	0.0138	0.0112	1.18
N1	0.0142	0.0114	1.25
N2	0.0135	0.0109	1.23

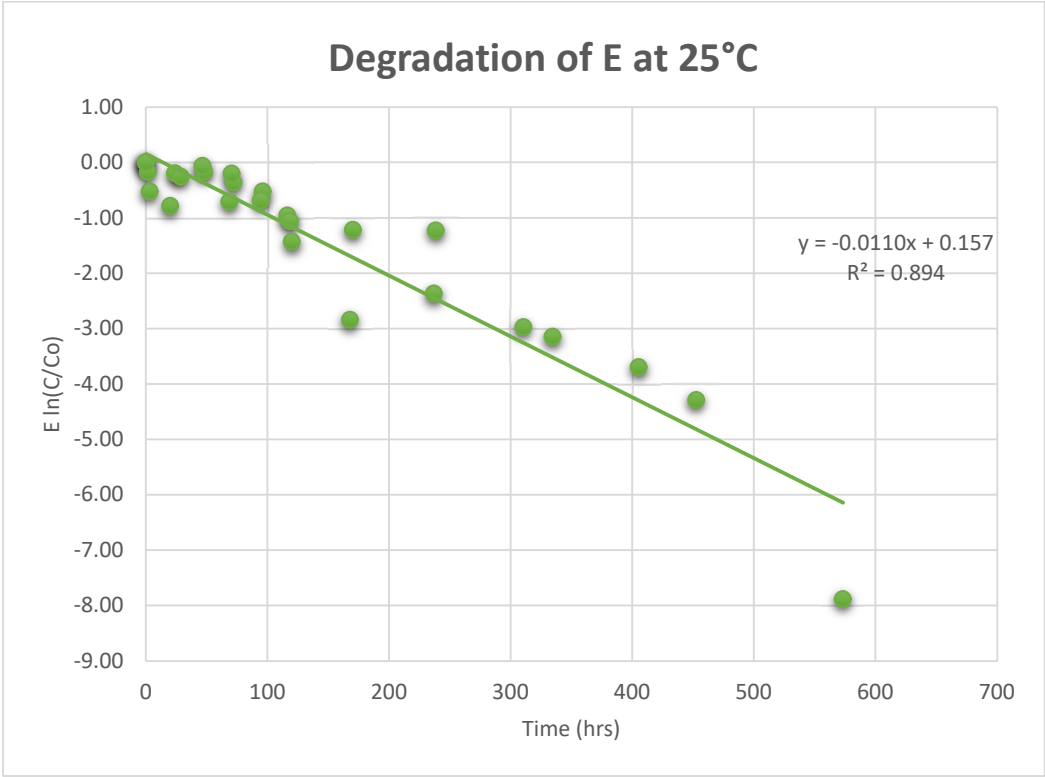
Table 1.4: Comparison of the sewage that had its stock stirred at 60rpm or was not stirred. These reactions were carried out in triplicate, and both stirred and non-stirred sewage were held at 25°C.

As expected, sewage that is quiescent will degrade more slowly than sewage that is stirred. Moving forward all experiments were stirred at 60 rpm to simulate the movement in sewers.

The benefit of pasteurization was also assessed. The 500mL sewage stock was split in half. Half of the sewage was pasteurized, and half of the sewage was not pasteurized. Pasteurized samples had their degradation rates measured at 25°C. These experiments generally produced noisy data. Thus, R^2 values generally greater than 0.7 were accepted with the lowest R^2 value accepted being 0.66. Figure 1.2 graphically shows how the three gene targets degrade over time at 25°C when the sewage is unpasteurized. These graphs utilize the data points from all three sewage samples to create a single graph for each gene target. All three gene targets degraded similarly at 25°C. The degradation rate constants were 0.0138hrs^{-1} (E), 0.0142hrs^{-1} (N1), and 0.0135hrs^{-1} (N2). Statistically, these rate constants were effectively the same. There was a p-value of 0.17 or higher when the targets were compared to each other. Samples were highly correlated with Pearson's correlation coefficients as follows: 0.93 (E vs. N1), 0.93 (E vs. N2), 1.0 (N1 vs. N2). The first-order rate constants at 25°C are listed in Table 1.5.

The R^2 values from the three trials ranged from 0.66-0.90 for the E gene target, 0.74-0.91 for the N1 gene target, and 0.70-0.84 for the N2 gene target. Combining the data produced much higher R^2 values which were near or above 0.9. However, combining the data suggested lower rate constant values. The data based on the three trials suggested a degradation rate that was 1.3 times faster (E), 1.4 times faster (N1), and 1.2 times faster (N2) than the rate constant from the graph with the combined data.

Figure 1.2: Gene Target Degradation at 25°C with Unpasteurized Sewage



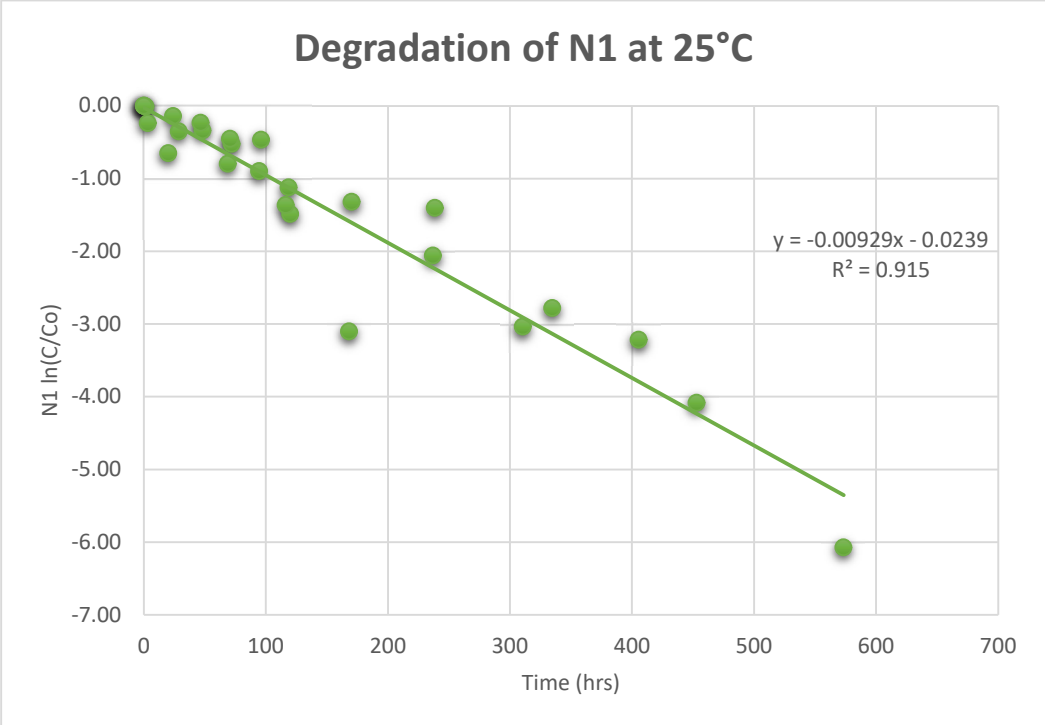


Figure 1.2: Gene Target Degradation at 25°C with Unpasteurized Sewage (Continued)

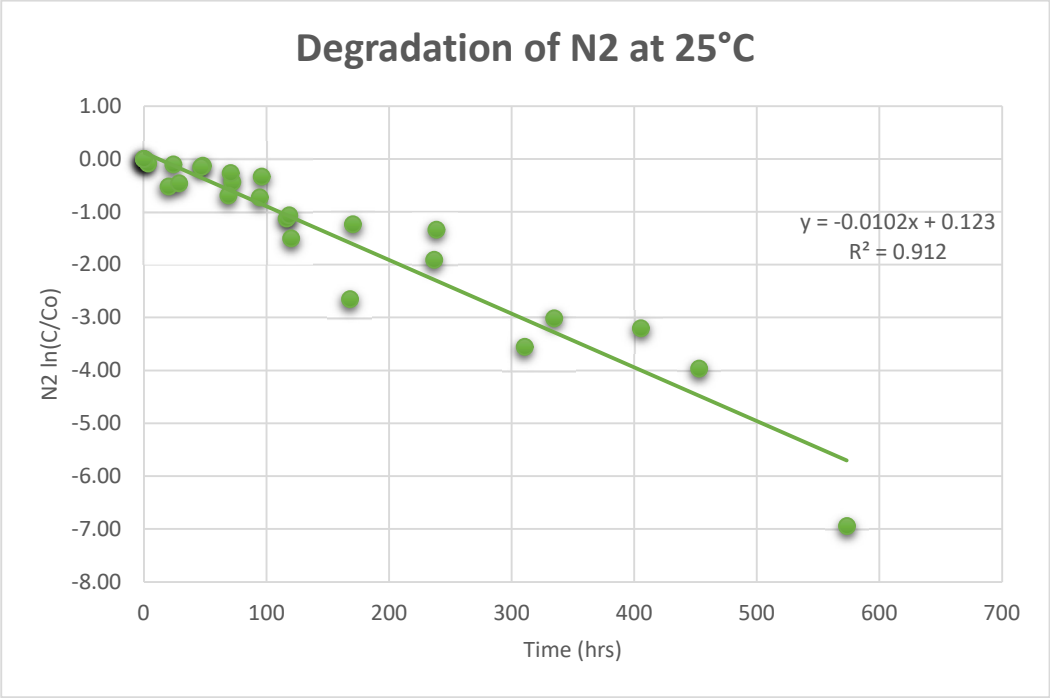


Figure 1.2: This figure shows three graphs that utilize the compiled data points from three different sewage samples to compose a single graph. Sewage samples were stirred throughout the collection of samples while being held at 25°C. These samples were not pasteurized.

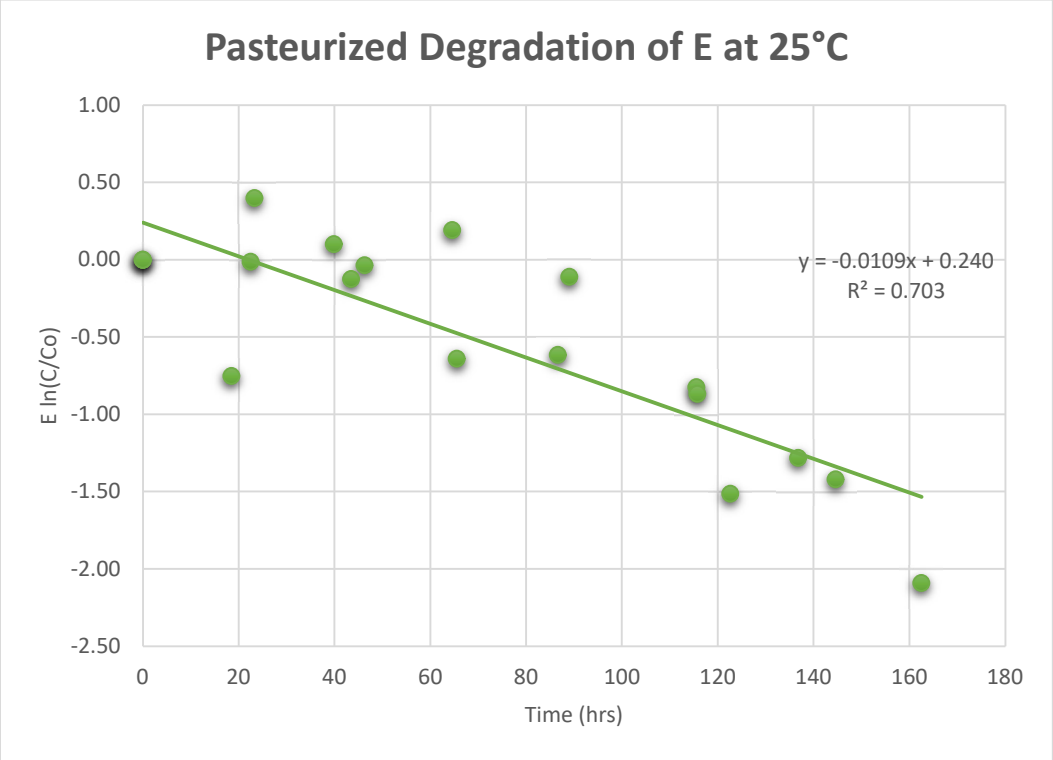
Table 1.5: SARS-CoV-2 Gene Target Degradation Rate Constants at 25°C (Unpasteurized)

Target	k (hrs^{-1}) Trial 1	k (hrs^{-1}) Trial 2	k (hrs^{-1}) Trial 3	Average k (3 Trials) (hrs^{-1})	Std. Dev. (hrs^{-1})	CV (%)	k (Graph) (hrs^{-1})
E	0.0183	0.00860	0.0146	0.0138	0.00490	35.4	0.0110
N1	0.0199	0.00990	0.0128	0.0142	0.00515	36.2	0.00900
N2	0.0181	0.0100	0.0124	0.0135	0.00416	30.8	0.0102

Table 1.5: List of degradation rate constants obtained at 25°C for the N1, N2 and E gene targets. Each temperature was run in triplicate for each gene target. This table also includes the average rate constant value, standard deviation (Std. Dev), correlation of variation (CV), and the rate constant obtained from graphing all data points (CV and Std. Dev not based on this).

The pasteurized sewage degradation was analyzed at 25°C. Figure 1.3 graphically represents each gene target by combining the data points from all three sewage samples. At 25°C the E, N1, and N2 gene targets in the pasteurized samples degraded at the same rate. There was no gene target that degraded significantly faster. The p-values were as follows: 0.061 (E vs. N1), 0.063 (E vs. N2), and 0.21 (N1 vs. N2). Again, the gene targets had a high correlation with each other with Pearson's correlation coefficients as follows: 1.0 (E vs. N1), 0.99 (E vs. N2), and 0.97 (N1 vs. N2). Table 1.6 shows the first-order degradation rate constants obtained for the pasteurized sewage collected at 25°C. The rate constant obtained from the compiled plot and the rate constant obtained from averaging the three slopes from the three independent trials were nearly identical/identical. The R^2 values from the three trials ranged between 0.73-0.96 (E), 0.80-0.83 (N1), 0.78-0.80 (N2). The compiled degradation graphs had lower R^2 values than the individual data sets with R^2 values between 0.68-0.76. Compared to unpasteurized sewage, the pasteurized sewage had a slower rate of degradation and a lower correlation of variation.

Figure 1.3: Gene Target Degradation at 25°C with Pasteurized Sewage



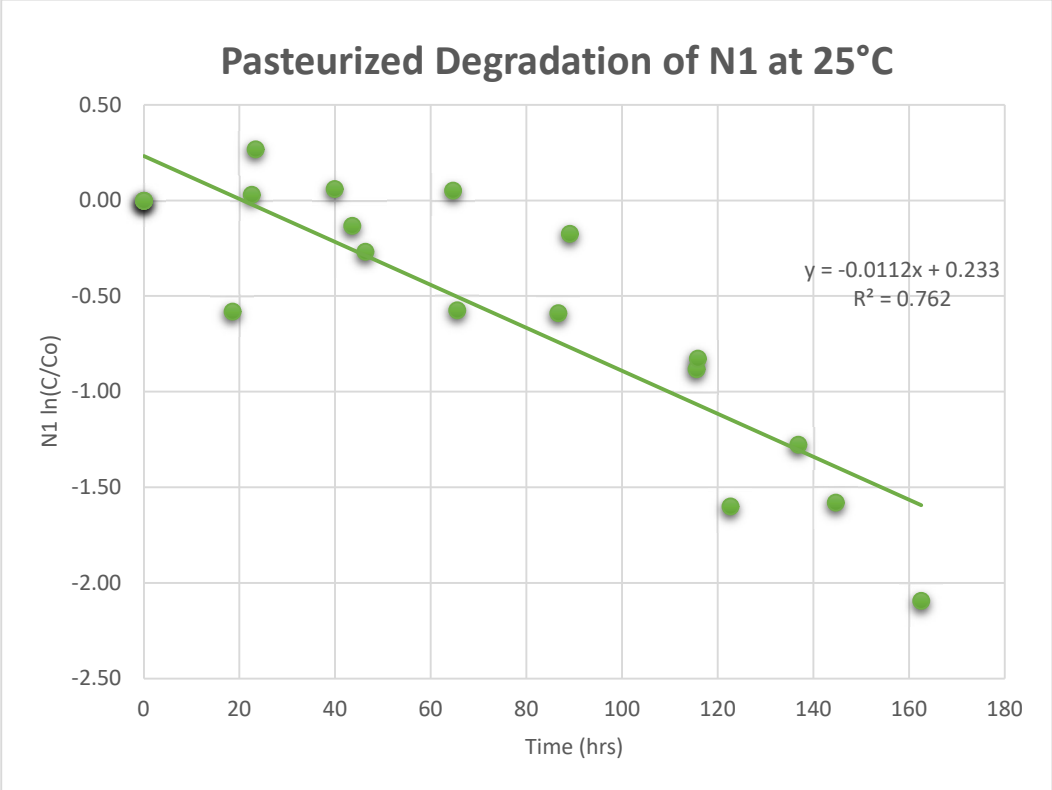


Figure 1.3: Gene Target Degradation at 25°C with Pasteurized Sewage (Continued)

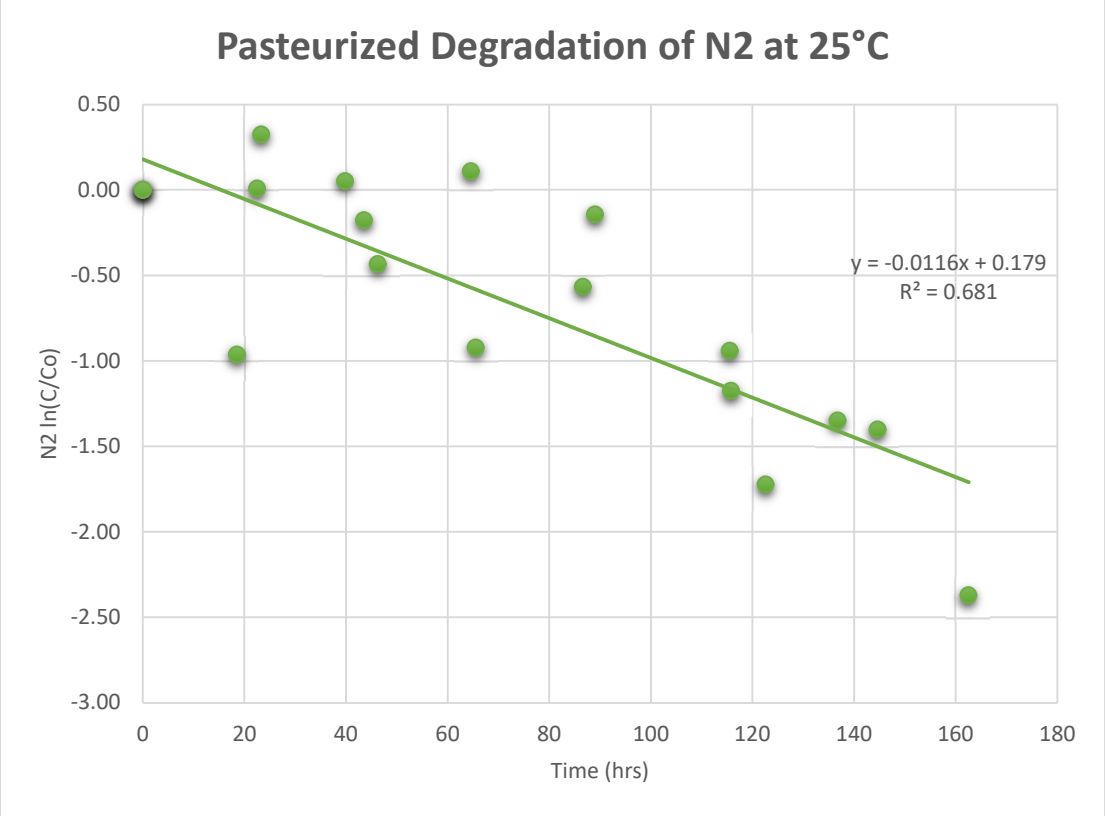


Figure 1.3: This figure shows three graphs that utilize the compiled data points from three different sewage samples to compose a single graph. Sewage samples were stirred throughout the collection of samples while being held at 25°C. These samples were pasteurized.

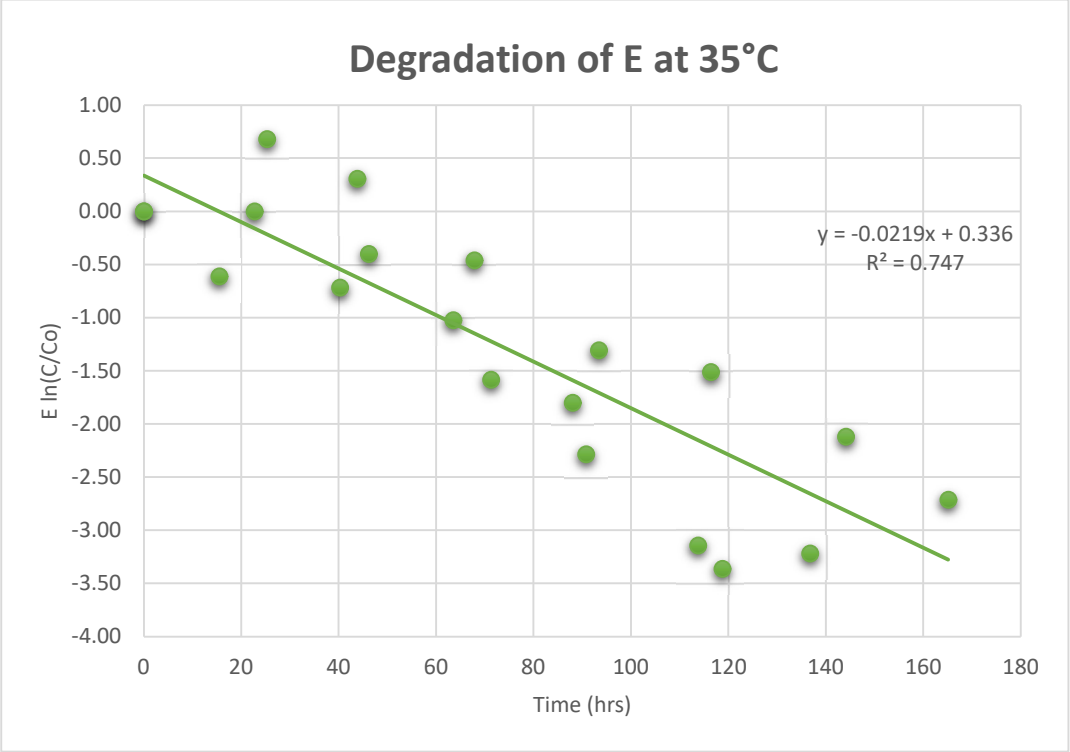
Table 1.6: SARS-CoV-2 Gene Target Degradation Rate Constants at 25°C (Pasteurized)

Target	k (hrs^{-1}) Trial 1	k (hrs^{-1}) Trial 2	k (hrs^{-1}) Trial 3	Average k (3 Trials) (hrs^{-1})	Std. Dev. (hrs^{-1})	CV (%)	k (Graph) (hrs^{-1})
E	0.0134	0.00860	0.0106	0.0109	0.00241	22.2	0.0109
N1	0.0139	0.00890	0.0107	0.0112	0.00253	22.7	0.0112
N2	0.0136	0.00930	0.0117	0.0115	0.00216	18.7	0.0116

Table 1.6: List of degradation rate constants obtained at 25°C for the N1, N2, and E gene targets when the sewage was pasteurized. Each temperature was run in triplicate for each gene target. This table also includes the average rate constant value, standard deviation (Std. Dev), correlation of variation (CV), and the rate constant obtained from graphing all data points (Std. Dev and CV not based on this).

The average correlation of variation for all unpasteurized gene targets was 34.1% (range: 30.8%-36.2%), and the average correlation of variation for the pasteurized gene targets was 21.2% (range: 18.9%-22.6%). The unpasteurized sewage had rate constants that were 1.3 times greater for E, 1.3 times greater for N1, and 1.2 times greater for N2 when compared to the rate constants from the pasteurized sewage. However, despite having an apparent faster rate of degradation, the scatter in the data led to no statistical difference between the unpasteurized sewage and pasteurized sewage. The p-values obtained when the unpasteurized and pasteurized sewage at 25°C were compared were as follows: 0.09 (E), 0.09 (N1), and 0.13 (N2). Pasteurized sewage and unpasteurized sewage were highly correlated with Pearson's correlation coefficients as follows: r=0.97 (E), r=1.0 (N1), and r=0.96 (N2). The main focus of this study was to determine how the gene targets degrade under natural conditions. However, if samples need to be stored, pasteurizing the sample can stabilize the sample by killing bacteria that degrade the targets leading to a smaller rate constant.

Figure 1.4: Gene Target Degradation at 35°C with Unpasteurized Sewage



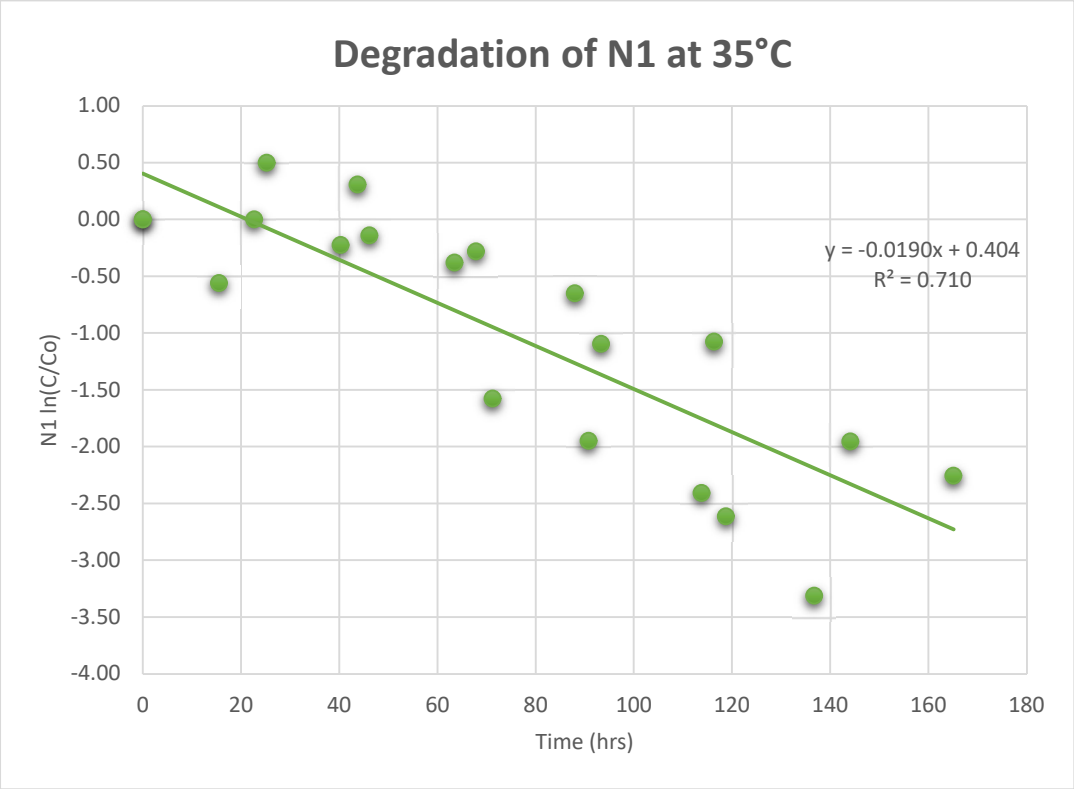


Figure 1.4: Gene Target Degradation at 35°C with Unpasteurized Sewage (Continued)

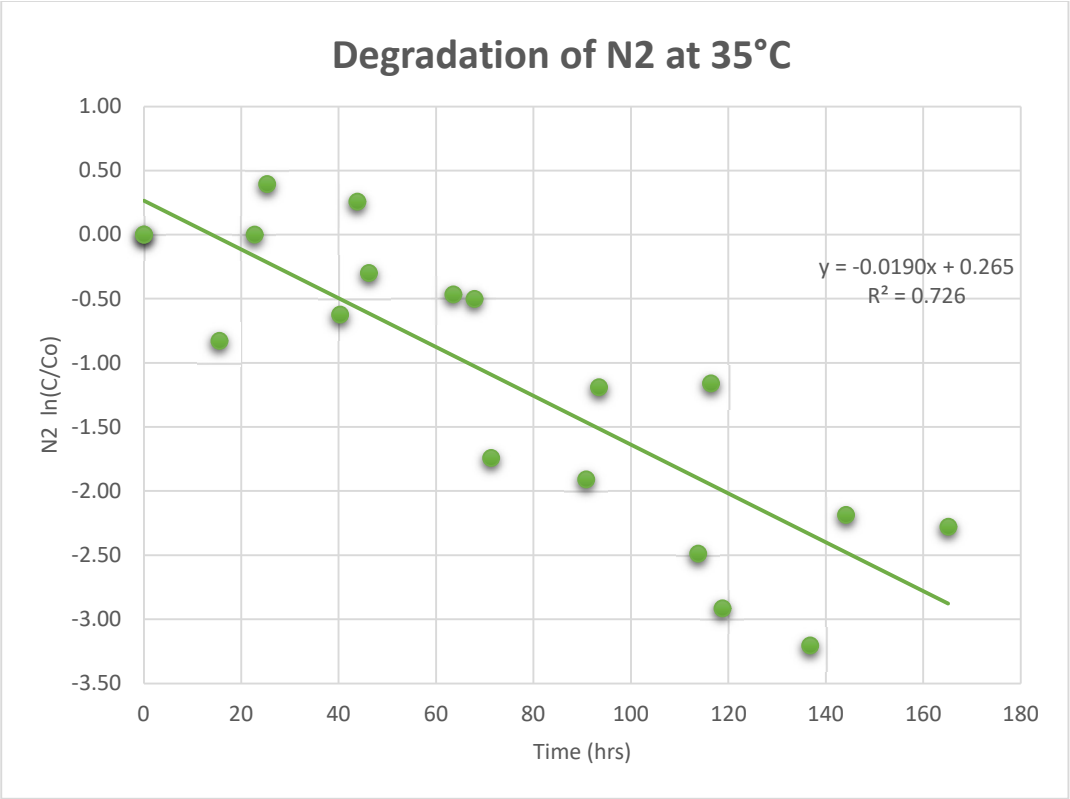


Figure 1.4: This figure shows three graphs that utilize the compiled data points from three different sewage samples to compose a single graph. Sewage samples were stirred throughout the collection of samples while being held at 35°C. These samples were not pasteurized.

Because the degradation was expected to be very slow at 4°C, pasteurized degradation experiments were carried out at 35°C to see if this preservation effect was present. Degradation experiments of the SARS-CoV-2 gene targets were run at 35°C for both pasteurized and unpasteurized sewage. Starting with the unpasteurized sewage, as expected, as the temperature increased, the rate constant increased indicating the samples degraded faster at a higher temperature. On average the rate constants increased by 1.9-fold (E), 1.5-fold (N1), and 1.7-fold (N2) at 35°C for the unpasteurized sewage. Figure 1.4 lists the rates of degradation at 35°C when the sewage was unpasteurized.

All gene targets degraded similarly at 35°C, but the E gene target appeared to degrade slightly faster. The E gene target degraded 1.2 times faster than the N1 gene target and 1.1 times faster than the N2 gene target. Statistically, the E gene degraded faster than the N2 gene target ($p=0.04$), but it did not degrade faster than the N2 gene target ($p=0.06$). The N1 and N2 gene targets degraded at the same rate ($p=0.19$). The Pearson's correlation coefficients were 1.0 for all gene target pairings.

Table 1.7: SARS-CoV-2 Gene Target Degradation Rate Constants at 35°C (Unpasteurized)

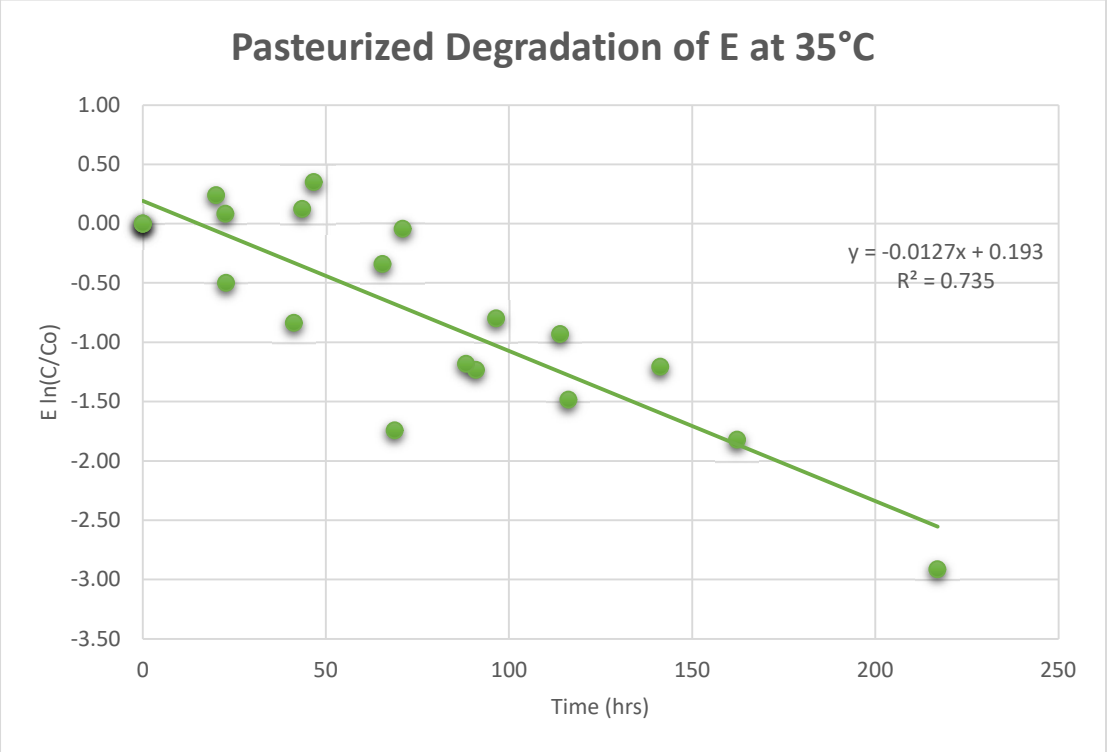
Target	k (hrs^{-1}) Trial 1	k (hrs^{-1}) Trial 2	k (hrs^{-1}) Trial 3	Average k (3 Trials) (hrs^{-1})	Std. Dev. (hrs^{-1})	CV (%)	k (Graph) (hrs^{-1})
E	0.0242	0.0188	0.0339	0.0256	0.00765	29.8	0.0219
N1	0.0217	0.0168	0.0274	0.0220	0.00531	24.1	0.0190
N2	0.0217	0.0169	0.0290	0.0225	0.00609	27.0	0.0190

Table 1.7: List of degradation rate constants obtained at 35°C for the N1, N2, and E gene targets. Each temperature was run in triplicate for each gene target. This table also includes the average rate constant value, standard deviation (Std. Dev, correlation of variation (CV), and the rate constant obtained from graphing all data points (Std. Dev and CV not based on this).

Table 1.7 shows the various first-order rate constants obtained from degrading sewage at 35°C. The R^2 values for the three trials were between 0.84-0.94 for E, 0.75-0.83 for N1, and 0.84-0.96 for N2. The combined data had lower R^2 values that were near 0.70. There was a slight difference (1.2 times for all three gene targets) between the combined data's rate constant, and rate constant obtained from averaging the independent trials rate constants. Interestingly, despite the 35°C sewage degrading over 1.5 times faster than the 25°C sewage, there was only a statistical significance for the E gene target when the temperature was increased 10°C. The following p-values were obtained when the 25°C and 35°C rate constants were compared: 0.048 (E), 0.086 (N1), and 0.073 (N2). This p-values may not be significant because the data set was too small with only three data points.

Next, the unpasteurized sewage was compared to the pasteurized sewage at 35°C. Pasteurizing the sewage samples lowered the degradation rate to considerably. When the samples were pasteurized before incubating them at 35°C, the SARS-CoV-2 gene targets degraded 1.9 times slower (E), 2.0 times slower (N1), and 2.0 time slower (N2) than the unpasteurized samples held at 35°C. The E gene target was not shown to have a statistical difference between the pasteurized and unpasteurized sewage, but it was borderline significant ($p=0.05$). The N1 ($p=0.01$) and N2 gene targets ($p=0.03$) did degrade significantly slower when pasteurized. Interestingly, there was only a high amount of correlation ($r=0.99$) between the pasteurized and unpasteurized sewage for the N2 gene target. The E and N1 gene targets had a correlation that was near 0. Figure 1.5 graphically shows the degradation of the SARS-CoV-2 targets against time at 35°C when the sewage was pasteurized.

Figure 1.5: Gene Target Degradation at 35°C with Pasteurized Sewage



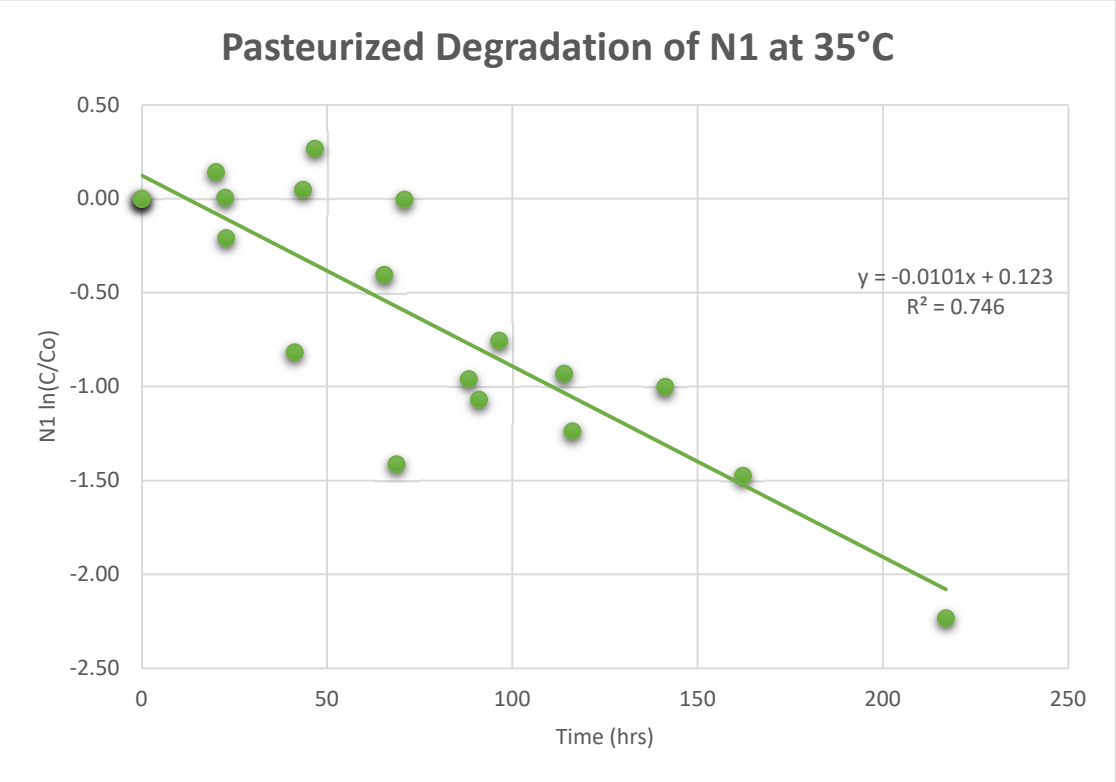


Figure 1.5: Gene Target Degradation at 35°C with Pasteurized Sewage (Continued)

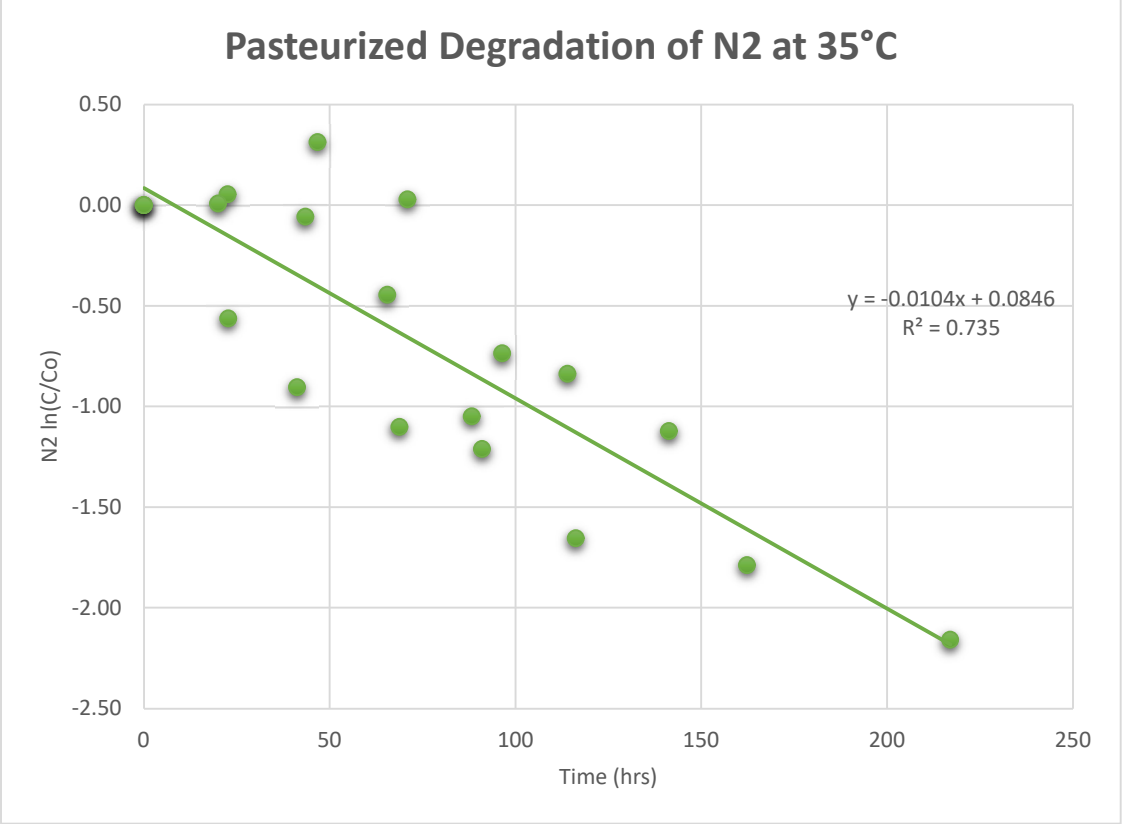


Figure 1.5: Three graphs that utilize the compiled data points from three different sewage samples to compose a single graph. Sewage samples were stirred throughout the collection of samples while being held at 35°C. These samples were pasteurized.

The R^2 values for each data set ranged between 0.74-0.89 (E), 0.73-0.88 (N1), and 0.87-0.92 (N2). This set of data had varying correlations between targets. The N1 and E gene targets were highly correlated (r=0.82). However, there was not high correlation with the N2 pairings yielding Pearson's correlation coefficients of -0.01 (N2 and E) and 0.56 (N1 and N2). Table 1.8 includes all the first-order rate constants for the pasteurized sewage at 35°C.

The graphically obtained rate constant and the average rate constant were similar and at most the averaged data was 1.1 times larger than the rate constant obtained from combined data points. The pasteurized sewage's rate constant at 35°C was 1.2 times greater for E, and 1.0 times greater for N1 and N2 compared to the pasteurized sewage at 25°C. The E gene target's degradation rate at 35°C was statistically higher (p=0.02) than the E gene target's degradation rate at 25°C according to the pasteurized sewage. No statistical difference was observed for the N1 and N2 gene targets.

Table 1.8: SARS-CoV-2 Gene Target Degradation Rate Constants at 35°C (Pasteurized)

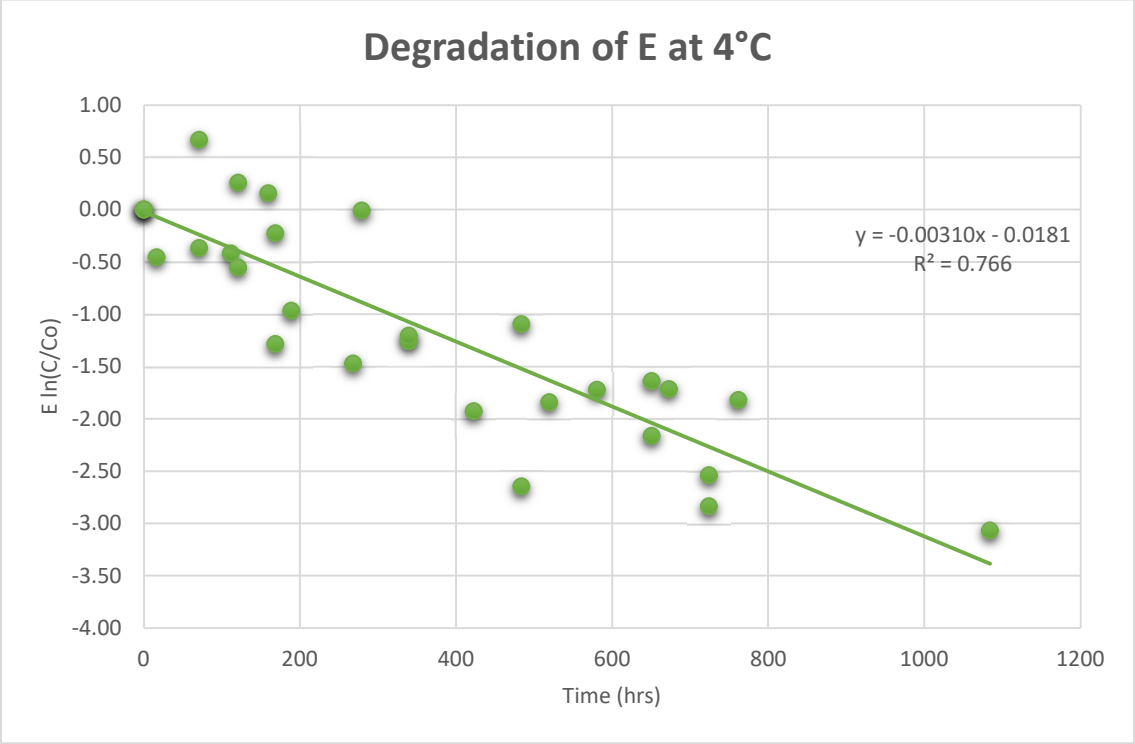
Target	k (hrs^{-1}) Trial 1	k (hrs^{-1}) Trial 2	k (hrs^{-1}) Trial 3	Average k (3 Trials) (hrs^{-1})	Std. Dev. (hrs^{-1})	CV (%)	k (Graph) (hrs^{-1})
E	0.0149	0.0122	0.0129	0.0133	0.00140	10.5	0.0127
N1	0.0114	0.0990	0.0111	0.0108	0.000794	7.35	0.0101
N2	0.0112	0.0108	0.0123	0.0114	0.000777	6.79	0.0104

Table 1.8: List of degradation rate constants obtained at 35°C for the N1, N2, and E gene targets when the sewage was pasteurized. Each temperature was run in triplicate for each gene target. This table also includes the average rate constant value, standard deviation (Std. Dev), correlation of variation (CV), and the rate constant obtained from graphing all data points (Std. Dev and CV not based on this).

Three samples were run to determine the first-order degradation of the gene targets at 4°C. No pasteurization experiments were run because the degradation rate was very slow, and it was impractical to run these samples for weeks. Reducing the temperature to 4°C reduced the rate by 4.3-fold (E), 4.4-fold (N1), and 4.6-fold (N2) when compared to the unpasteurized sewage at 25°C. The compiled data points for the degradation of the SARS-CoV-2 gene targets at 4°C are shown graphically in Figure 1.6.

The R^2 values for these graphs ranged from 0.68-0.92 (E), 0.79-0.99 (N1), and 0.71-0.96 (N2). Interestingly, there was not a high correlation between the E gene target with either the N1 (r=0.26) or N2 gene target (r=0.41). However, there was a high correlation between the N1 and N2 gene targets (r=0.99). The N1 gene target degraded the fastest, but this gene target only had a statistically higher degradation rate constant (p=0.047) than N2. Overall, a 1.1-fold difference was the largest difference between the rate constant values for all three gene targets. Table 1.9 lists the first-order rate constant at 4°C. The average rate constant values were similar amongst gene targets (0.93-1.0-fold difference between them). The temperatures and rate constants were utilized with the Arrhenius equation to make Arrhenius plots. By plotting the natural logarithm of the rate constants against the inverse of temperature reported in Kelvin, an apparent activation energy (E_a) and preexponential factor (A) can be obtained from the graph to allow the degradation rate constant at any temperature to be calculated for each gene target. Figure 1.7 shows the Arrhenius plots for each SARS-CoV-2 gene target based on the three temperatures used in this study.

Figure 1.6: Gene Target Degradation at 4°C with Unpasteurized Sewage



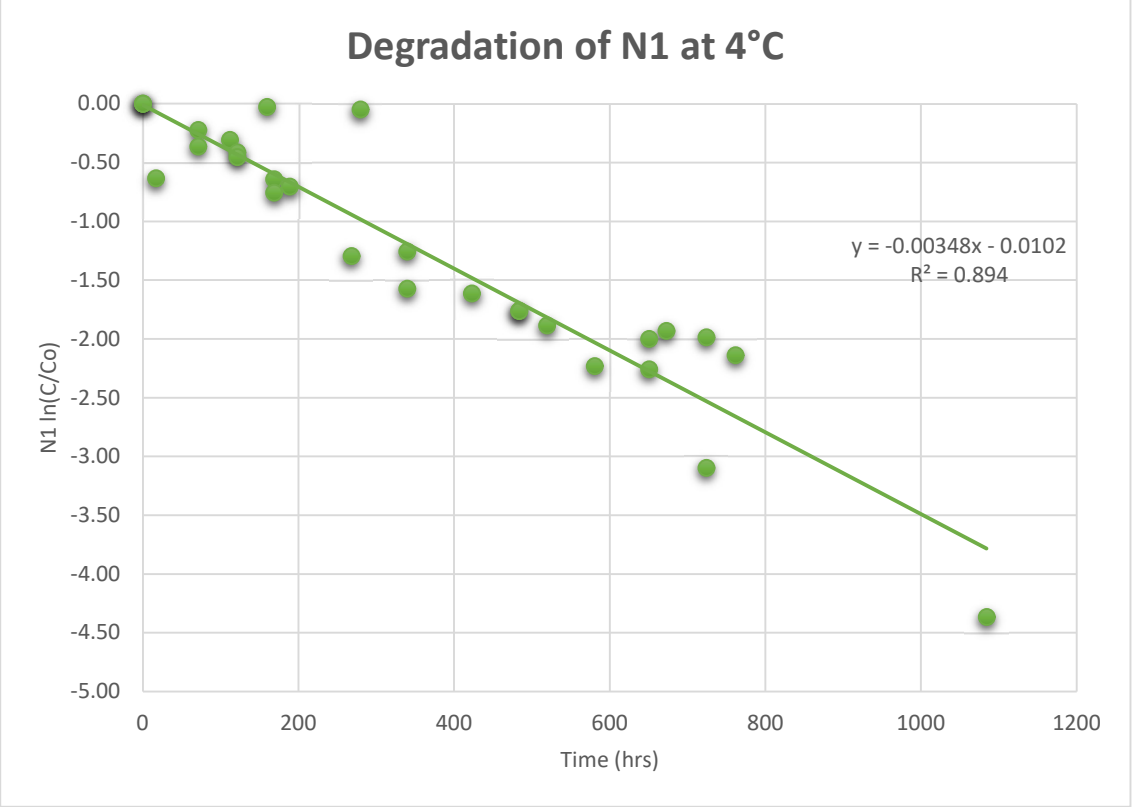


Figure 1.6: Gene Target Degradation at 4°C with Unpasteurized Sewage (Continued)

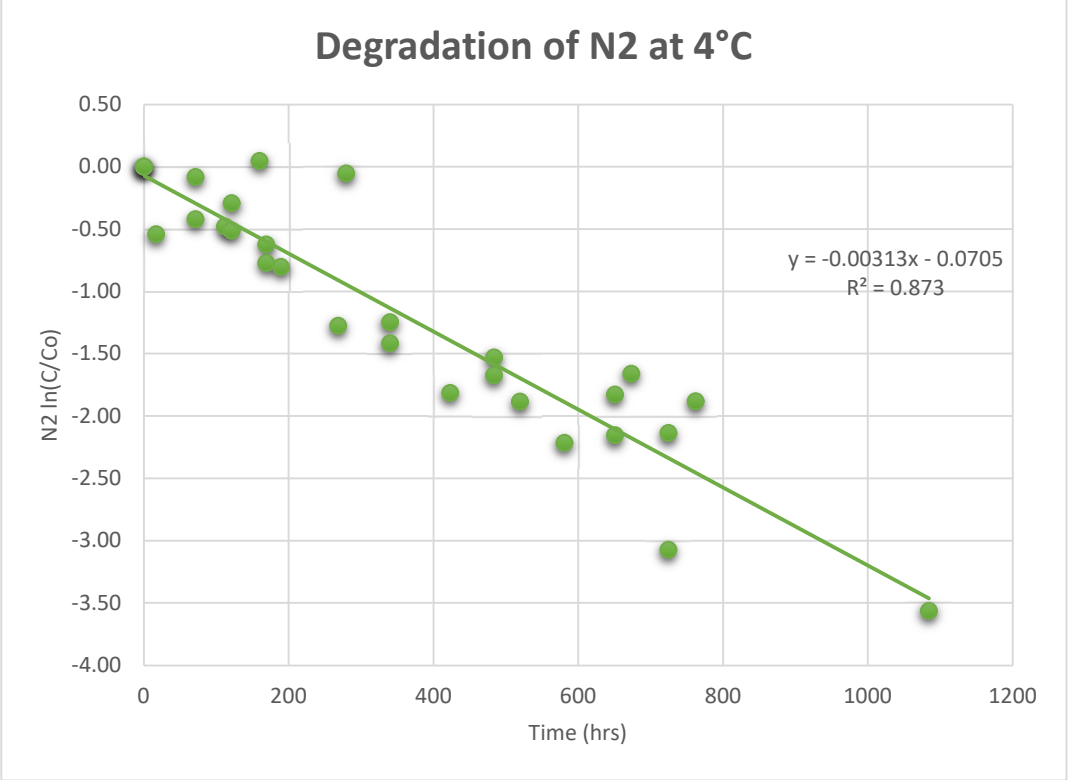
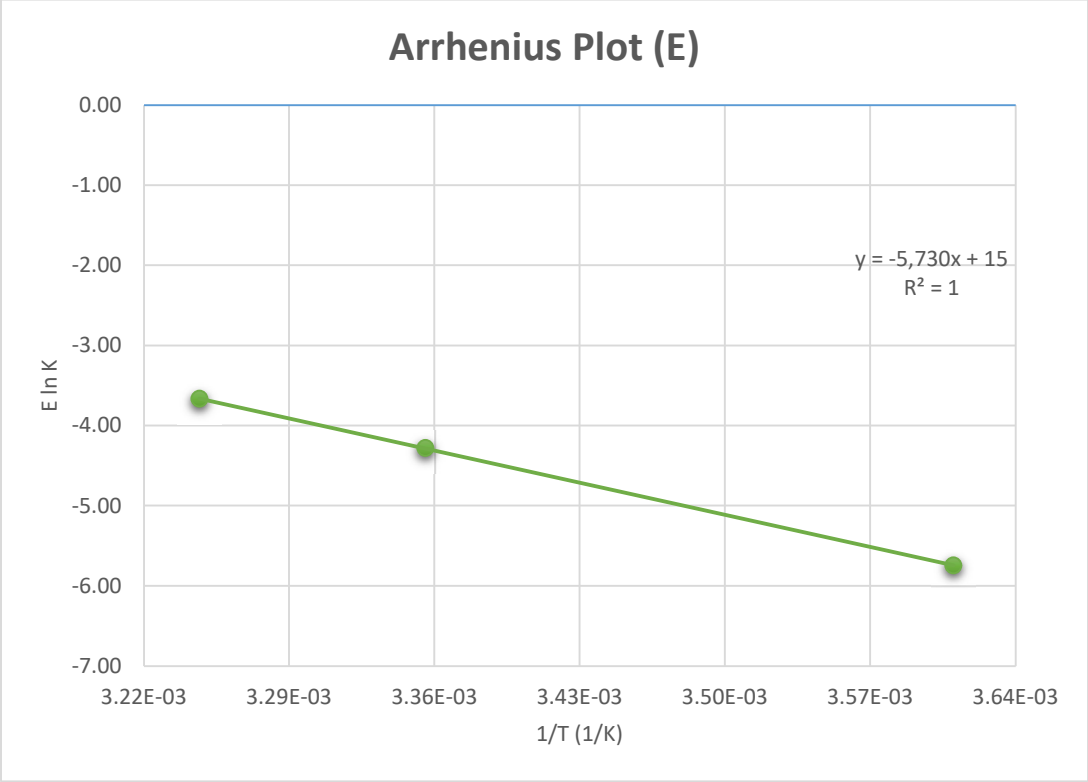


Figure 1.6: This figure shows three graphs that utilize the compiled data points from three different sewage samples to compose a single graph. Sewage samples were stirred throughout the collection of samples while being held at 4°C. These samples were not pasteurized.

Figure 1.7: Arrhenius Plots for Each SARS-CoV-2 Gene Target



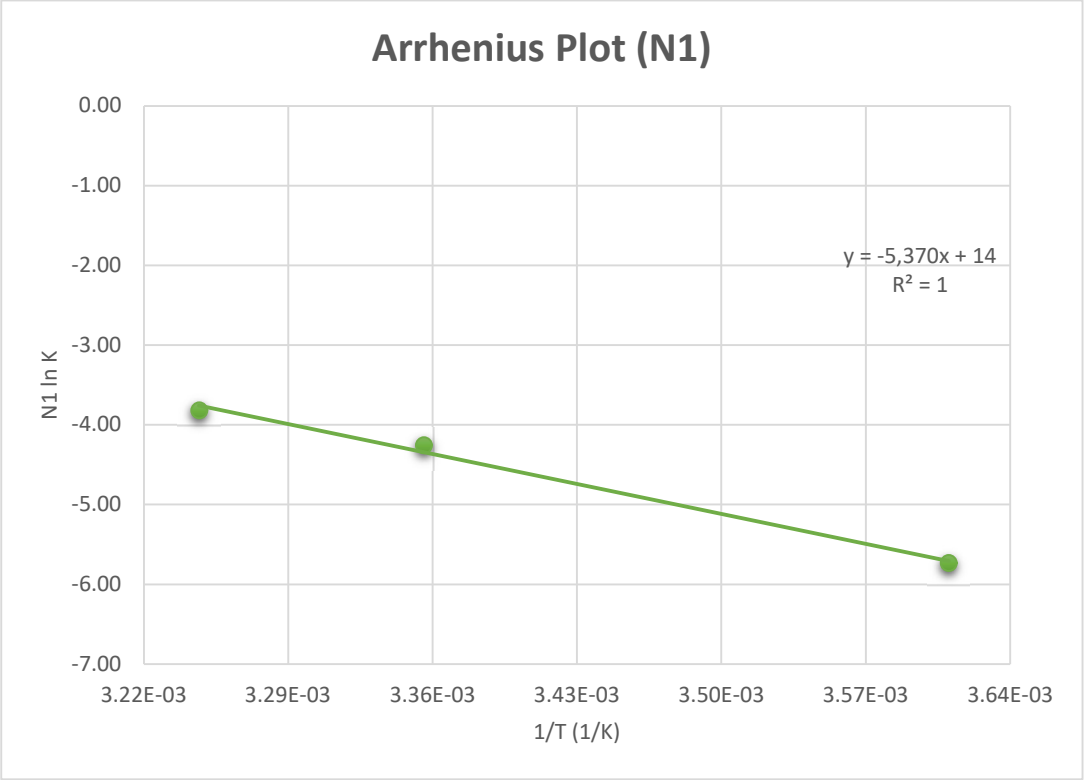


Figure 1.7: Arrhenius Plots for Each SARS-CoV-2 Gene Target (Continued)

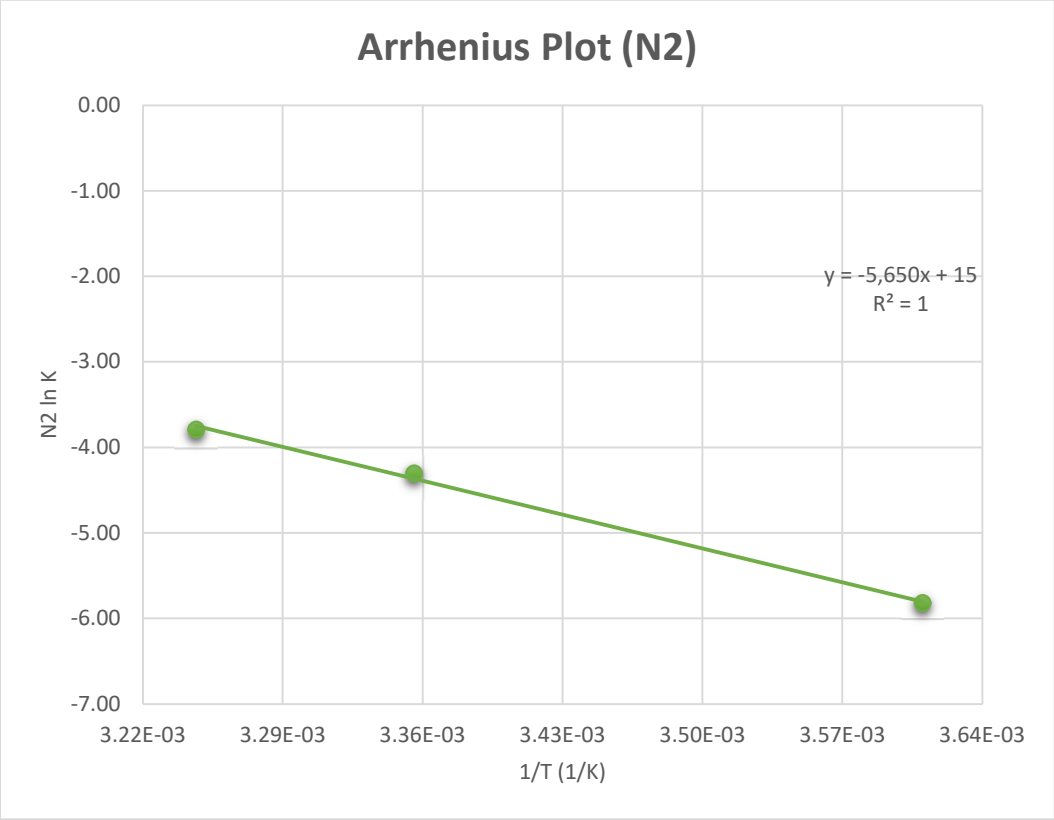


Figure 1.7: This figure shows three Arrhenius plots for the E, N1, and N2 gene targets.

Table 1.9: SARS-CoV-2 Gene Target Degradation Rate Constants at 4°C (Unpasteurized)

Target	k (hrs^{-1}) Trial 1	k (hrs^{-1}) Trial 2	k (hrs^{-1}) Trial 3	Average k (3 Trials) (hrs^{-1})	Std. Dev. (hrs^{-1})	CV (%)	k (Graph) (hrs^{-1})
E	0.00260	0.00360	0.00340	0.00320	0.000529	16.5	0.00310
N1	0.00290	0.00280	0.00400	0.00323	0.000667	20.6	0.00348
N2	0.00260	0.00270	0.00360	0.00300	0.000551	18.6	0.00313

Table 1.9: List of degradation rate constants obtained at 4°C for the N1, N2, and E gene targets. Each temperature was run in triplicate for the three gene targets. This table also includes the average rate constant value, standard deviation (Std. Dev), correlation of variation (CV), and the rate constant obtained from graphing all data points (Std. Dev and CV not based on this).

The y-intercept is equivalent to the natural logarithm of the pre-exponential factor. The slope is a negative value that is equivalent to activation energy divided by the universal gas constant (R). Table 1.10 shows the values of the slopes and the pre-exponential factors for the three gene targets. We calculated the shelf life (T90) which represents the time it would take for the gene targets to decrease in concentration to 90% of their original concentration and the half-life (t1/2) of our gene targets. Table 1.11 shows the half-lives, shelf lives, and rate constants for all temperatures.

Table 1.10: Pre-Exponential Factor and Slope Values for the SARS-CoV-2 Gene Targets

Gene Target	A (hrs^{-1})	-Ea/R (K)
E	3.10×10^{6}	-5730
N1	8.70×10^{5}	-5370
N2	2.15×10^{6}	-5650

Table 1.10: A table of pre-exponential factors and slope values for the E, N1, and N2 gene targets. These values can be used to calculate the first-order degradation rate constants at any temperature.

Table 1.11: Half-Lives, Shelf Lives, and Rate Constants for the SARS-CoV-2 Gene Targets

Gene Target	K avg (hrs^{-1})	Temperature (°C)	Pasteurized or Unpasteurized	T90 =0.105/k (hrs)	T1/2=0.693/k (hrs)
E	0.0138	25	Unpasteurized	7.59	50.1
N1	0.0142	25	Unpasteurized	7.39	48.8
N2	0.0135	25	Unpasteurized	7.78	51.3
E	0.0109	25	Pasteurized	9.66	63.8
N1	0.0112	25	Pasteurized	9.40	62.1
N2	0.0115	25	Pasteurized	9.10	60.1
E	0.0256	25	Unpasteurized	4.10	27.0
N1	0.0220	35	Unpasteurized	4.78	31.5
N2	0.0225	35	Unpasteurized	4.66	30.8
E	0.0133	35	Pasteurized	7.88	52.0
N1	0.0108	35	Pasteurized	9.72	64.2
N2	0.0114	35	Pasteurized	9.18	60.6
E	0.00320	4	Unpasteurized	32.8	217
N1	0.00323	4	Unpasteurized	32.5	214
N2	0.00297	4	Unpasteurized	35.4	233

Table 1.11: This table shows the half-lives, shelf lives, and average rate constants at 4°C, 25°C, and 35°C for E, N1, and N2. All data had sewage samples that were stirred.

1.5 Discussion

RNA degradation may be caused by combinations of three mechanisms which include: chemical (paraffin, aldehydes, etc.), physical (high temperature, UV light, etc), and enzymatic pathways (Sidova et al., 2015). One property that can influence RNA degradation is pH. RNA is in its most stable form when the pH is between 4 and 5, but as the pH becomes alkaline, the RNA becomes unstable (Bernhardt & Tate, 2012). RNA is susceptible to both acid hydrolysis (pH less than 2) and alkaline hydrolysis (pH higher than 6) (Bernhardt & Tate, 2012). Secondary wastewater effluents have a pH range between 6 and 8 (Wei et al., 2008). Many personal care and cleaning products that end up in sewage have a pH greater than 5. Shampoos often have a pH between 6 and 7 and

many soaps have a pH between 9 and 10 (Tarun et al., 2014) or higher (Boonchai & Iamtharachai, 2010). Common cleaners which also exhibit high alkalinity include Clorox bleach (pH of 11.4) and powdered laundry detergents (pH range of 10.8-11.7) (Ross & Spiller, 1999). Therefore, these products can contribute to increasing the pH of the sewage. Due to the general pH of the sewage leaning towards a neutral to basic pH, alkaline hydrolysis is probably a significant factor leading to RNA degradation. Pesticides, trash, automotive chemicals, fertilizers, rain, and many other things also go down the sewers which could interfere with the RNA signal (*How Sewage Pollution Ends Up In Rivers*, n.d.). Next, the 2'-hydroxyl group located on RNA is much more susceptible to hydrolysis which will cause RNA to be degraded faster than DNA (Fordyce et al., 2013). Hydrolysis of RNA is further enhanced under alkali conditions and when cations including transitional metals and calcium are present (Fordyce et al., 2013). A physical factor which can contribute to RNA degradation is an increase in temperature. In general RNA samples are frozen at either -20 °C or -80 °C, or liquid nitrogen is used to prevent RNA degradation (Fabre et al., 2014). However, ribonucleases (RNases) and ribozymes can still be active at cold temperatures so there may always be some RNA degradation, but lowering temperature is generally used to reduce RNA degradation (Fabre et al., 2014). RNA degradation was found to increase linearly as the temperature was raised from 55°C to 92°C (Brisco & Morley, 2012). Their study measured RNA degradation as induced lesions/1000 bases (Brisco & Morley, 2012). When the temperature was 55°C the induced lesions/1000 bases was ~0.6, but this value elevated to ~9 when the temperature increased to 92°C (Brisco & Morley, 2012). Wastewater in the

sewers has an average temperature of ~15.6°C (*Waste Wattage*, 2012). Thus, RNA will degrade more slowly with the low temperature.

RNases are the enzymes that are mainly responsible for the degradation of RNA (Ohyama et al., 2014). These degradative enzymes are nonspecific to substrates (Bechhofer & Deutscher, 2019). For example, humans have around 25,000 genes in total, but there are less than 100 RNase genes to degrade those genes meaning one RNase degrades hundreds or even thousands of RNA (Ohyama et al., 2014). An RNase known as RNase 7 is the most plentiful RNase in skin (Koczera et al., 2016). Sewage contains many human components such as dead skin cells, fingernails, hair, and bodily fluids (urine and feces) (*Sludge, Farmer's Friend or Toxic Slime?*, 2009). The human body secretes RNases into bodily fluids such as mucus, tears, perspiration, and salvia; however, RNases are also found in hair which has been shed (both human and pet hair) and flaked skin (*The Basics*, n.d.). Thus, sewage can have many RNases just from human waste. Therefore, there will be many types of RNases in the sewer system leading to enhanced RNA degradation. For instance, plant RNases in the T2 family are very broadly distributed and can be found in a wide-range of things such as bacteria, viruses, and mammals (Pauline A. Bariola et al., 1999).

Based on the Arrhenius plot from our study, the rate-constant at any temperature can be calculated for the N1, N2, and E gene targets. Sewage has a temperature ranging from 10 to 21°C (Jamal, n.d.) with an average temperature of approximately 15.6°C (*Waste Wattage*, 2012). The degradation rates at the range of sewer temperatures are given in Table 1.12.

Table 1.12: Gene Target Degradation Rates at Typical Sewer Temperatures

Temp (°C)	E (hrs^{-1})	N1 (hrs^{-1})	N2 (hrs^{-1})
10	0.00497	0.00500	0.00465
15.6	0.00736	0.00723	0.00685
21	0.0106	0.0102	0.00982

Table 1.12: Degradation rates that the gene targets would exhibit at common sewage temperatures based on the Arrhenius plots from our data.

In sewage it is more likely that the gene targets will degrade similarly and the small differences we observed will not be significant. Table 1.13 shows the shelf-life and half-life values at the typical sewage temperatures. Besides our study, there are a limited number of studies that investigated SARS-CoV-2 degradation. Mahlknecht et al. discussed that infectious SARS-CoV-2 is not as stable or abundant as inactivated SARS-CoV-2 RNA in various types of liquids (Mahlknecht, 2022). Infectious SARS-CoV-2 is quickly inactivated in wastewater, sea water, and river water in comparison to tap water (Mahlknecht, 2022).

Table 1.13: Half-life and Shelf-Life Values at Typical Sewage Temperatures

Target	Temperature (°C)	Shelf-life (hrs)	Half-life (hrs)
E	10.0	21.1	139
E	15.6	14.3	94.1
E	21.0	9.90	65.4
N1	10.0	21.0	138
N1	15.6	14.5	95.8
N1	21.0	10.3	68.1
N2	10.0	22.6	149
N2	15.6	15.3	101
N2	21.0	10.7	70.6

Table 1.13: Half-life and shelf-life values of the SARS-CoV-2 gene targets N1, N2, and E typically seen under normal sewage temperatures.

They observed that SARS-CoV-2 is very stable in sewage when temperatures are between 4°C-37°C which means that temperature is unlikely to play a role if a sewage sample is negative for SARS-CoV-2 (Ahmed, Tscharke, et al., 2021). Another study mentioned that the N1 and N2 gene targets detected in wastewater influent did not have significant degradation at 4°C when samples were held for 19 days (Beattie et al., 2022). They also found that pasteurization did not have much of an impact on the SARS-CoV-2 gene targets concentrations in 2/3 wastewater treatment plants, but pasteurization did cause concentration loss in one of the wastewater treatment plants serving the smallest population (Beattie et al., 2022). They pasteurized their samples for 1hr at 65°C (Beattie et al., 2022). Our lab pasteurized the samples for 2hrs at 60°C. In our study pasteurization was never found to decrease the SARS-CoV-2 signal. Pasteurizing the sample reduced the rate of degradation when compared to the unpasteurized sewage. Our study also contradicted their study when it came to the stability of the sewage sample at 4°C because our study suggested that there was signal loss at 4°C while their study did not. Averaging the T90 values for all three gene targets, we found that the sewage sample lost 10% of the gene target concentration after 33hrs (1.4 days) at 4°C. Another study analyzing the N1 gene target agreed with the Beattie et al. study suggesting that storing wastewater at 4°C for as long as nine days would not cause a significant reduction in the SARS-CoV-2 concentration (Markt et al., 2021). However, they mentioned that freezing samples at -20°C significantly decreased the SARS-CoV-2 N1 gene target (Markt et al., 2021). Therefore, they suggested storing the sewage at 4°C instead of freezing the samples (Markt et al., 2021). A study by Takeuchi et al. looked at SARS-CoV-2 degradation at 4°C, 25°C, 37°C, and 45°C but they used a sequence targeting the

nonstructural protein 15 (nsp15) of SARS-CoV-2 (Takeuchi et al., 2022). Also, their experiments utilized nucleic acid amplification tests rather than sewage surveillance (Takeuchi et al., 2022). The nucleic acid amplification tests use specimens collected from the lower or upper respiratory tract (CDC, 2020b). The first-order rate constants in their study were $0.017hrs^{-1}$ (4°C), $0.137hrs^{-1}$ (25°C), $0.206hrs^{-1}$ (37°C), and $0.255hrs^{-1}$ (45°C) (Takeuchi et al., 2022). The half-lives were 40.1hrs (4°C), 5.08 (25°C), 3.37 (37°C), 2.72 (45°C) (Takeuchi et al., 2022). Their degradation rate was 5.4-fold higher than the average rate constant value at 4°C. Their degradation rate was 9.9-fold higher than the average rate constant value at 25°C. As the samples analyzed were different (respiratory sample vs. sewage) from our study and the target was different, it is not unexpected that it would degrade differently. Weidhaas et al. suggested a high first-order rate constant indicating that the N1 and N2 gene targets would degrade at a rate of $0.09\text{-}0.12hrs^{-1}$ when the temperature was between 4°C and 35°C (Weidhaas et al., 2021). They reported that the SARS-CoV-2 signal could no longer be detected after 6hrs at a temperature of 35°C, while SARS-CoV-2 could be detected at 4°C and 10°C after 22hrs (Weidhaas et al., 2021). Hokajärvi et al. looked at the first-order degradation of the E gene target and the N2 gene target of SARS-CoV-2 at 4°C, -20°C, and -75°C (Hokajärvi et al., 2021). They reported degradation rate constants at 4°C as $0.04days^{-1}$ (E) and $0.06days^{-1}$ (N2) (Hokajärvi et al., 2021). They also reported T90 values of 52days (E) and 36days (N2) indicating N2 degraded much faster than the E gene target at 4°C (Hokajärvi et al., 2021). We report comparable values of degradation between the E and N2 gene targets at 4°C that are not statistically different. Our shelf-life values suggested a much faster rate of degradation than suggested by Hokajärvi et al. with T90 values of 32.8hrs (E) and

35.4hrs (N2) at 4°C. Another study looked at SARS-CoV-2 degradation in tap water and wastewater with high and low amounts of SARS-CoV-2 (Bivins et al., 2020). They found first-order degradation rate constants as follows: $1.1days^{-1}$ (20°C, low concentration, wastewater), $1.4days^{-1}$ (20°C, high concentration, wastewater), $0.15min^{-1}$ (50°C, high concentration, wastewater), $1.0min^{-1}$ (70°C, high concentration, wastewater), and $1.2days^{-1}$ (20°C, high concentration, tap water) (Bivins et al., 2020). They reported T90 values between 2.1days and 2.2mins based on the conditions SARS-CoV-2 faced (Bivins et al., 2020). This study did not mention what part of SARS-CoV-2 they targeted. However, on average our gene targets had a T90 value of 2.1days at 25°C which was the same as the T90 value they reported for their sewage that was inoculated with a low concentration of SARS-CoV-2 and held at 20°C (Bivins et al., 2020). Another study looked at the degradation of the N1 gene target of SARS-CoV-2 at 4°C and 20°C in seawater and river water (Sala-Comorera et al., 2021). They reported first-order rate constants of $0.61days^{-1}$ (4°C) and $1.01days^{-1}$ (20°C) in river water and $1.07days^{-1}$ (4°C) and $2.02days^{-1}$ (20°C) in seawater (Sala-Comorera et al., 2021). The T90 values they reported were as follows: 3.8days (4°C) and 2.3days (20°C) in river water and 2.2days (4°C) and 1.1days (20°C) in seawater (Sala-Comorera et al., 2021). These values were similar to the prior study despite being sea and river water and even suggested seawater degraded the SARS-CoV-2 signal faster than the sewage at 20°C. Roldan-Hernandez et al. looked at the degradation of the N1 and N2 gene targets and pepper mild mottle virus at two wastewater treatment plants at three temperatures (Roldan-Hernandez et al., 2022). They found that as temperature increased so did the first order rate constants (Roldan-Hernandez et al., 2022). They listed their rate constants as follows: $0.024days^{-1}$ (4°C),

0.027days^{-1} (22°C), and 0.063days^{-1} (37°C) for N1, 0.011days^{-1} (4°C), 0.021days^{-1} (22°C), and 0.047days^{-1} (37°C) for N2, and 0.010days^{-1} (4°C), 0.040days^{-1} (22°C), and 0.045days^{-1} (37°C) for pepper mild mottle virus for wastewater treatment plants A (Roldan-Hernandez et al., 2022). For wastewater treatment plant B, they listed their rate constants as follows: 0.036days^{-1} (4°C), 0.063days^{-1} (22°C), and 0.091days^{-1} (37°C) for N1, 0.031days^{-1} (4°C), 0.089days^{-1} (22°C), and 0.098days^{-1} (37°C) for N2, and 0.059days^{-1} (4°C), 0.077days^{-1} (22°C), and 0.091days^{-1} (37°C) for pepper mild mottle virus (Roldan-Hernandez et al., 2022). They looked at pepper mild mottle virus because it can be used to assess fecal strength, and therefore, may be used to normalize the SARS-CoV-2 signal in other studies (Roldan-Hernandez et al., 2022). Lastly, Ahmed et al. looked at the degradation rate of untreated sewage, autoclaved sewage, and tap water for N1 and murine hepatitis virus at 4°C, 15°C, 25°C, and 37°C (Ahmed, Bertsch, Bibby, et al., 2020). For N1 they found first-order rate constants of 0.084-0.286days^{-1} (4°C-37°C) for untreated sewage, 0.054-0.405days^{-1} (4°C-37°C) for autoclaved sewage, and 0.039-0.245days^{-1} (4°C-37°C) for tap water (Ahmed, Bertsch, Bibby, et al., 2020). Overall, our study was bracketed by the other studies that found lower or higher rates than our study. Our study suggests T90 values near 33.6hrs (4°C), 7.6hrs (25°C) and 4.5hrs (35°C) for all gene targets. In contrast, other studies reported T90 values in units of days, and suggested it would take days or weeks such as Ahmed et al. who suggested T90 values of 27.8days-8.04days (4°C-37°C) for N1 in sewage (Ahmed, Bertsch, Bibby, et al., 2020). Our study looked at pasteurization and found pasteurization significantly reduced the rate of degradation. Therefore, pasteurization may prove beneficial in stabilizing samples and increasing safety at the cost of time.

1.6 Conclusion

This study looked at degradation rates at 4°C, 25°C, and 35°C over time. We found that at the lowest temperature the RNA degraded slowly, and as the temperature increased so did the degradation rate. For the unpasteurized sewage at 4°C, gene targets degraded at $0.0032hrs^{-1}$ (E), $0.00323hrs^{-1}$ (N1), and $0.00297hrs^{-1}$ (N2). As the temperature increased to 25°C, the rate constants increased over four-fold for each of the gene targets: 0.0138 hrs^{-1} (E), 0.0142 hrs^{-1} (N1), and 0.0135 hrs^{-1} (N2). At 35°C the rate constants were as follows: $0.0256hrs^{-1}$ (E), $0.0220hrs^{-1}$ (N1), and $0.0225hrs^{-1}$ (N2). There was a statistical difference in rate as the temperature increased. Also, based on the rates, the gene targets in pasteurized sewage had a slower rate of degradation than the unpasteurized sewage. Comparing the half-lives of the unpasteurized and pasteurized sewage, the average half-lives for all gene targets were 50hrs (unpasteurized) and 62hrs (pasteurized) at 25°C and 30hrs (unpasteurized) and 59hrs (pasteurized) at 35°C. Lastly, there was a set of non-stirred experiments carried out at 25°C. The half-lives were as follows: 55hrs (stirred) and 62hrs (non-stirred) for E, 53hrs (stirred) and 66hrs (non-stirred) for N1, and 54hrs (stirred) and 68hrs (non-stirred) for N2.

The rates of degradation are low at 4°C indicating that there will be negligible signal loss for the first day or two of storage at 4°C Pasteurization will preserve the signal if analysis cannot be completed right away. Pasteurization may also enhance the SARS-CoV-2 gene target's signal by reducing bacteria or enzymatic action on the samples. However, if pasteurization is implemented into the workflow, the analysis this will take. Arrhenius parameters were determined so that rate constants may be calculated at any temperature for use in modelling and case study correlations.

CHAPTER TWO

ASSESSING SARS-COV-2 GENE TARGETS AND THE NECESSITY OF MULTIPLE TARGETS FOR RELIABLE SARS-COV-2 DETECTION IN WASTEWATER SAMPLES SPANNING SOUTHEASTERN MICHIGAN

2.1 Abstract

Coronavirus disease 2019 (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which is an RNA virus. Clinical testing focuses on testing nasopharyngeal swabs while our work focuses on testing for SARS-CoV-2 within sewage. Wastewater surveillance is a non-invasive way to track COVID-19 trends in whole communities. It can provide data on individuals that may not be captured with clinical testing such as individuals unwilling to be tested, asymptomatic individuals, or individuals that do not report their at-home test results. Our work used three SARS-CoV-2 gene targets derived from the envelope and nucleocapsid genes of SARS-CoV-2. Our research worked to reduce the number of targets tested to save cost and time with minimal loss of SARS-CoV-2 detections. The first set of experiments compared the N1, N2, and E gene targets in 1,430 samples over nine months for 29 different sites. It was found that E was detected in over 200 less samples (655) than N1 (861) and N2 (873). All three gene targets were effective and highly correlated with each other. However, when evaluating each individual targets contribution, the E gene target was only able to increase the number of detections by 2.3% as compared to the N1 and N2 gene targets which increased the detections by 8.3% and 10.7% respectively. After the first set of experiments, E was removed from analysis, and N1 and N2 were analyzed for an

additional year. A total of 4,527 samples were assessed for 44 sites. Of these samples, 2,559 samples were positive for N1, and 2,453 samples were positive for N2. This extra data showed that N1 was detected more frequently as opposed to the first nine months of data that favored N2. Both N1 (13.3%) and N2 (9.6%) added a substantial amount of SARS-CoV-2 detections. Overall, these data suggest that removing E, and retaining N1 and N2 will achieve a high level of SARS-CoV-2 detection while substantially lowering the cost of analysis.

2.2 Introduction

Coronavirus disease 2019 (COVID-19) is an acute respiratory disease which is caused by a highly transmissible coronavirus (CoV) called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Hu et al., 2021). The virus SARS-CoV-2 has similarities to the other CoVs including being a single-stranded and enveloped RNA virus; the genome size approaches 30,000 nucleotides (Alexandersen et al., 2020). These viruses were first discovered in 1937 when they were isolated from chickens (Song et al., 2022). In addition to chickens, other animals have been infected with CoVs such as: cats, cows, dogs, pigs, and camels (Alluwaimi et al., 2020). Overall, there are many CoVs (hundreds) that infect various creatures, but there are only seven CoVs that are known to infect humans (Healthline, 2020). When humans are infected with a CoV, symptoms often resemble the common cold, but some individuals develop a severe case with severe inflammation, internal organ dysfunction, high fever, and even death (Dhama et al., 2020, p. 19). Based on some data released in China, the majority of individuals (81%) were found to have mild to moderate symptoms of COVID-19, while severe disease and critical disease found in 14% and 5% of individuals, respectively (Gandhi et al., 2020).

There are also some cases of COVID-19 that produce no symptoms. Numerous studies have noted the large fluctuation in the estimated number of asymptomatic cases, but overall, the data suggests a high fraction of asymptomatic cases. One article examined 95 studies with nearly 30 million individuals. When they analyzed data with confirmed cases of COVID-19, the number of individuals that were asymptomatic was 40.50% (Ma et al., 2021). When they further analyzed the COVID-19 positive group they found that 54.11% of pregnant woman were asymptomatic, 47.53% of nursing home staff or residents were asymptomatic, and 52.91% of cruise or air travelers were asymptomatic (Ma et al., 2021). The disease COVID-19 is generally diagnosed using PCR on a nasopharyngeal swab specimen (Gopaul et al., 2020). Nasopharyngeal swabs exhibit high specificity (99.5%) and had a high amount of sensitivity (84.6%) (Gopaul et al., 2020). Although nasopharyngeal swabs are the main way to test for COVID-19 in a clinical setting, SARS-CoV-2 can also be detected by analyzing sewage. Feces are the main contributor of the SARS-CoV-2 signal in sewage. Urine is not a major source of SARS-CoV-2. A review that combined 33 studies with 430 patients found that 3.7% of the patients tested positive for SARS-CoV-2 in their urine (Kashi et al., 2020). SARS-CoV-2 is detected in feces more frequently than urine, but fecal samples are not comparable to nasopharyngeal samples. Only 32-67% of patients that have a clinically confirmed case of COVID-19 have positive SARS-CoV-2 detection in their fecal samples (Szymczak et al., 2020). Sewage surveillance of SARS-CoV-2 is an advantageous way to assess COVID-19 levels in communities to check for outbreaks of COVID-19 and may indicate increases in COVID-19 cases one to two weeks ahead of the clinical data. Wastewater surveillance can be used to identify facilities that have high COVID-19 infection rates and allow

public health interventions directly in those buildings to reduce the spread of COVID-19 (Deng et al., 2022). Sewage surveillance has been previously used to monitor salmonellosis and polio, and this technique has been shown to provide SARS-CoV-2 data that had a consistently high association with the clinical cases of COVID-19 (Wolfe, n.d.). Consequently, sewage surveillance provides a quick, cost-effective way to screen whole communities for trends in COVID-19 cases. Due to the ongoing nature of the COVID-19 pandemic, participation in clinical testing is not as strong as it was in the beginning. Therefore, it is likely that the clinical cases provide a less accurate depiction of COVID-19 infection due to under participation, high percentages of asymptomatic cases, and underreported at-home tests. Due to this, sewage surveillance is important and will be increasingly important as participation in clinical testing dwindles.

Currently, the Centers for Disease Control and Prevention (CDC) recommends two different SARS-CoV-2 gene targets named N1 and N2 (CDC, 2020c). There are three nucleocapsid gene-based targets used to target CoVs. However, the nucleocapsid based targets N1 and N2 were specifically designed with the intent of detecting SARS-CoV-2 while the N3 gene target was designed to be a universal detector of viruses that belong to the subgenus *Sarbecovirus* (Lu et al., 2020). To diversify the study, another SARS-CoV-2 gene target that targeted the envelope gene of SARS-CoV-2 was utilized. This E gene target was used by Corman et al (Corman et al., 2020a). Therefore, three different gene targets of SARS-CoV-2 were utilized to determine if SAS-CoV-2 was present in the sewage. As there are two different genes being detected, there could be an advantage found when targeting one gene over the over.

The focus of this study was to determine which gene target was able to detect SARS-CoV-2 the most frequently. If there is a superior gene target, one or two gene targets could potentially be removed from analysis. Due to the fast-spreading nature of this disease, it is vital to analyze targets that are sensitive and reliable enough to test for SARS-CoV-2 in the sewage. This study could find that all three gene targets are necessary for the detection of SARS-CoV-2, but it also has the potential to save time and money if all three gene targets are not necessary for analysis. Unfortunately, COVID-19 continues to produce outbreaks even though it originated almost two years ago. As of its origin that occurred near the end of December 2019, this virus has spread and made its way to the United States by January and made its way to minimally 25 countries as of February (Y.-C. Wu et al., 2020). According to the World Health Organization (WHO), 3.8million cases of COVID-19 were reported across the globe in the last seven days as of September 7th, 2022 (*WHO Coronavirus (COVID-19) Dashboard*, n.d.). As of September 6th, 2022 over 603 million cases of COVID-19 across the globe have been reported (*WHO Coronavirus (COVID-19) Dashboard*, n.d.). Therefore, COVID-19 still has a heavy presence within the community and monitoring this virus is still vital for the public's health. Our research works to find a more cost effective and time efficient way to monitor COVID-19 in the sewage without loss or with minimal loss of SARS-CoV-2 detection.

2.3 Materials and Methods

2.3.1 Sample Delivery, Sample Selection, and Sample Concentration

Our workflow used the methods described by Flood et al (Flood et al., 2021) as a starting point. Sewage samples were delivered to the lab in 500mL portions. Samples were

transported in a cooler with icepacks while in transit to our lab. Samples were processed immediately upon arrival. The 500mL bottle of sewage was mixed by inverting the bottle between 40 and 60 times to minimize foaming. A 45 mL homogenous sewage sample was decanted into a 50mL conical bottom centrifuge tube. Approximately 4.0g of polyethylene glycol 8000 (catalog number 97061-100 from VWR) and 0.59g sodium chloride (catalog number 470302-520 from VWR) were added to the sample yielding a concentration of 8% polyethylene glycol and 0.2M sodium chloride. After the solids and solutions were combined, the centrifuge tubes were transferred to a rotator where the tubes were rotated at 70rpm for 2hrs. The rotator was held at 4°C. After mixing, the samples were centrifuged at 3,650 rpm for 1 hr. Once centrifugation was completed, the supernatant was discarded, and a 1-2 mL pellet remained.

2.3.2 RNA Elution and Purification

The QIAamp Viral RNA Mini Kit (catalog number 52904) was purchased from QIAGEN to purify and elute the RNA from the sewage samples. To begin the process, buffer AVL and carrier RNA were mixed in portions of 800µL and 8µL per sample, respectively. A 200µL portion of pellet was then pipetted into a 2mL cryovial containing the buffer and RNA mix. The cryovial was vortexed for 10s before incubation for 10mins at room temperature. After incubation, 200proof ethanol was added to the cryovial in an 800µL portion, and the cryovial was vortexed for 10s. At this point in the procedure there was 1.8mL solution in the cyrovial. This solution was run through the QIAamp mini spin column in order to obtain the extracted RNA. The 1.8mL solution was added in three 630µL quantities to the spin column. Three additions were done because the spin column could not hold enough solution for a single addition. Each 630µL portion of solution was

added to the spin column and centrifuged sequentially. The flow through was not retained. The centrifugation settings were 1min at 8,000rpm. These centrifuge settings were the same for every addition except when the buffer AW2 was added. The waste collected in the collection tube under the spin column was discarded after every addition until the elution step (buffer AVE). After the whole 1.8mL solution was added to the spin column and centrifuged, the wash buffers were added. A 500µL portion of buffer AW1 was added to the spin column and centrifuged. Next, a 500µL portion of buffer AW2 was added to the spin column. At this point in the process (and only during this step) the centrifuge speed and duration was increased to 14,000rpm for 4min. After centrifugation, centrifugation settings were returned to normal, and the spin column was transferred to a 1.5mL microcentrifuge tube. The microcentrifuge tube was used to collect the RNA eluted from the spin column. Buffer AVE was added to the spin column in a 40µL portion and incubated at room temperature for 1min. After incubation, the spin column was centrifuged. This step was repeated which resulted in a final volume of 80µL of RNA eluate.

2.3.3 Amplification of RNA and Detection of Target Sequences

The assay mix was made using the solutions from the One-Step RT-ddPCR Advanced Kit for Probes (catalog number 1864021) produced by Bio-Rad. Table 2.1 shows the primer/probe sequences used to detect the SARS-CoV-2 gene targets. The assay mix composition varied if the reactions were run for an E singleplex versus an N1-N2 duplex. Table 2.2 shows the components that were used for a singleplex reaction and a duplex reaction.

Table 2.1: SARS-CoV-2 Gene Target Sequences

Target	Sequences (5'-3')	References
E	ACAGGTACGTTAATAGTTAATAGCGT (forward primer) ATATTGCAGCAGTACGCACACA (reverse primer) FAM-ACACTAGCCATCCTTACTGCGCTTCG-BHQ1 (probe)	(Corman et al., 2020a)
N1	GACCCCAAAATCAGCGAAAT (forward primer) TCTGGTTACTGCCAGTTGAATCTG (reverse primer) FAM-ACCCCGCATTACGTTTGGTGGACC-BHQ1 (probe)	(CDC, 2020c)
N2	TTACAAACATTGGCCGCAAA (forward primer) GCGCGACATTCCGAAGAA (reverse primer) HEX-ACAATTTGCCCCCAGCGCTTCAG-BHQ1 (probe)	(CDC, 2020c)

Table 2.1: Sequences used to detect the SARS-CoV-2 gene targets N1, N2, and E.

The values in Table 2.2 were based on a single reaction. Singleplex and duplex reactions only varied based on the amount of water and the additional primer/probe mix which would be added for a duplex reaction. Samples and controls (negative, no template, extraction, and positive controls) were run on each 96-well plate.

Table 2.2: Assay Mix Calculations for a Single Well

Reagent	Singleplex	Duplex
Supermix	5.5µL	5.5µL
Reverse transcriptase	2.2µL	2.2µL
Dithiothreitol	1.1µL	1.1µL
Primer/probe mix 1	3.3µL	3.3µL
Primer/probe mix 2	N/A	3.3µL
Water	4.4µL	1.1µL
Total	16.5µL	16.5µL

Table 2.2: This table shows the assay mix calculations utilized for a singleplex and a duplex reaction. These calculations were based on the amount of assay mix required for one reaction (one well).

The extraction control utilized buffer AVE instead of pellet, and the negative control utilized water instead of pellet otherwise the procedure was the same. The N1 and N2 gene targets were ran as a duplex reaction, and the E gene target was run as a singleplex reaction. Each well had 5.5µL of sample/control and 16.5µL of assay mix. Each sample and control were run in triplicate per assay mix. After pipetting, the plate was sealed with foil. This plate was then vortexed to mix the sample and the assay mix and spun in a plate spinner to collect the liquid at the bottom of the 96-well plate. This plate was then inserted into the Automated Droplet Generator (catalog number 1864101 Bio-Rad) to generate droplets. The Automated Droplet Generator generated droplets into another 96-well plate. This new 96-well plate was sealed with foil and placed into the thermocycler. The thermocycler settings are listed in Table 2.3. The thermocycling settings were based on the instructions for the 1-Step RT-ddPCR Advanced Kit for Probes for general settings, and the annealing temperatures from Flood et al. (*One-Step RT-DdPCR Advanced Kit for Probes*, n.d.; Pearson et al., 2021).

Table 2.3: Thermocycling Settings for SARS-CoV-2 Gene Targets

Step	Temperature	Time	Cycles
1	25°C	3mins	1
2	50°C	1hr	1
3	95°C	10mins	1
4	95°C	30s	40*
5	55°C	1min	40*
6	98°C	10min	1
7	4°C	30min	1
8	4°C	Infinity	1

Table 2.3: This table includes the thermocycling settings used for the N1, N2, and E gene targets. *Step 4 and 5 alternate back and forth 40 times. Steps are completed in chronological order.

Once thermocycling was complete, the plate was held at room temperature for 15mins before reading the plate with the QX200 Droplet Reader (catalog number 1864003 Bio-Rad). After reading, thresholds for positive and negative droplets were set and the plate was quantified.

2.3.4 Analyzing the Data

The droplet reader was paired with QuantaSoft™ Software to identify positive droplets. The data obtained from this software provided information on the number of droplets (accepted, positive, and negative), concentration, and copies per 20µL well. The positive droplets were assessed to determine if the sample was positive for SARS-CoV-2. Three or more positive droplets were considered a positive hit. Accepted droplets were assessed to determine the validity of the values obtained from the QuantaSoft™ Software. Wells that had greater than 10,000 accepted droplets were considered acceptable and droplet data was converted to gene copies per 20µL well. This value was utilized along with the volume of the pellet, the eluted RNA volume, volume of pellet used, and initial volume.

(Equation 2.1)

$$\frac{CP}{L} = copies\ per\ \mu L \times \frac{pellet\ volume\ (mL)}{extraction\ volume\ (0.2mL)} \times \frac{eluate\ volume\ (80\mu L)}{initial\ volume\ (45mL)} \times \frac{1000mL}{1L}$$

Equation 2.1: Concentration Calculation for the SARS-CoV-2 Gene Targets
This formula shows how the concentrations for the N1, N2, and E gene targets were calculated.

Equation 2.1 shows the formula utilized to calculate the concentration of the SARS-CoV-2 gene targets. The concentration was in units of gene copies per liter of sewage (CP/L). Values above the method detection limit (MDL) were positive for SARS-CoV-2 while concentrations equivalent to the MDL were considered negative for SARS-CoV-2. The goal of this study was to analyze the frequency of the gene targets leading to a positive detection and to determine if one or two gene targets could be removed from analysis.

2.4 Results

For the first analysis, the frequency of the N1, N2, and E gene targets were analyzed from September 16th, 2020 to May 26th, 2021 for 29 different sites from four areas in South Eastern Michigan. There was a total of 1,430 samples analyzed for the N1, N2, and E gene targets. The N1 gene target had the highest average concentration $10.4x10^4$ CP/L or $16.7x10^4$ CP/L without the negative samples i.e averaged without MDL values. The N2 target had similar concentrations ($9.1x10^4$ CP/L or $14.4x10^4$ CP/L without MDL values) and the E target had a significantly lower concentration ($6.4x10^4$ CP/L or $12.9x10^4$ CP/L without MDL values). Out of the 1,430 samples, there were 1,019 detections. The sample was considered a positive detection for SARS-CoV-2 if there was at least one gene target that was positive. Overall, there were 861 samples positive for N1, 873 samples positive for N2, and 655 samples positive for E. These samples were also analyzed to determine what pair of targets were able to detect the most SARS-CoV-2 positive samples. It was found that using both N1 and N2 resulted in the largest number of SARS-CoV-2 positive detections. The number of SARS-CoV-2 positive detections per combination of gene targets were as follows: 996 (N1 and/or N2), 908 (N1 and/or E), and 931 (N2 and/or E).

Table 2.4: N1, N2, and E Detections per Site

Site	N1 detections	N2 detections	E detections	Total Detections/All Samples	City, Township, or University
CT1	64	60	54	64/70	Township
CT2	56	53	49	59/66	Township
CT3	57	55	51	64/70	Township
CT4	48	43	30	56/67	Township
CT5	57	54	49	64/71	Township
CT6	52	49	36	57/70	Township
CT7	49	42	39	56/71	Township
EMU1	8	11	4	16/47	University
EMU2	35	38	31	38/48	University
EMU3	8	15	4	17/48	University
EMU4	25	24	21	30/48	University
EMU5	12	13	7	17/46	University
EMU6	35	35	23	39/49	University
EMU7	25	26	18	30/47	University
EMU8	36	36	29	41/48	University
EMU9	9	17	10	18/47	University
EMU10	12	11	7	13/49	University
EMU11	29	30	18	35/52	University
OU1	16	17	9	21/38	University
OU2	11	17	8	19/38	University
OU3	21	25	12	29/39	University
OU4	27	30	18	36/39	University
OU5	23	23	18	27/38	University
OU6	21	18	11	26/38	University
WN1	30	26	22	32/39	City
WN2	25	28	19	32/38	City
WN3	9	12	3	15/38	City
WN4	31	34	27	36/38	City
WN5	30	31	28	32/33	City

Table 2.4: This table shows the number of detections that each gene target had based on the 1,430 samples collected between September 2020 and May 2021. There were 29 different sites located in four different areas. The samples that were derived from the same area will be represented with a number. The samples were taken from cities, universities, or townships.

Using the N1 and N2 gene targets accounted for 97.7% of the SARS-CoV-2 positive samples while the other combinations were only able to detect 89.1% (N1 and/or E) and

91.4% (N2 and/or E) of the SARS-CoV-2 positive samples. Table 2.4 shows the breakdown on a site-by-site basis on the quantity of SARS-CoV-2 positive samples that were detected per target. Samples came from four areas. These areas included Clinton Township (CT), Oakland University (OU), Eastern Michigan University (EMU) and Warren (WN). The seven Clinton Township samples came from lift stations and divided the township into sewer sheds. The university samples and Warren 1-4 were generally at the building level. Warren 5 was a wastewater treatment plant. Of the 29 sites analyzed, the E gene target had the fewest detections (SARS-CoV-2 positive samples) across all sites. However, there was one site where the E gene target had more detections than the N1 gene target, but the E gene target was never superior to the N2 gene target. For 11/29 sites, the N1 gene target had the most detections while the N2 gene target had the most detections for 15/29 sites. The remaining three sites had an equal number of detections for N1 and N2. On a site-by-site basis, site CT1 had the highest number (100%) of detections for N. Site EMU2 was the site with the most N2 detections with 100% of the detections in this site being attributed to N2. Site WN5 had the most E detections with 87.5% of the detections in this site being attributed to E. Site WN5 also was the site that had the highest number of detections to samples ratio with 32/33 (97.0%) of its samples positive for SARS-CoV-2. In contrast, site EMU10 was the site with the fewest number of detections with only 13/49 (26.5%) of its samples positive for SARS-CoV-2. The lowest percentage of gene target detections were found at the following sites: 47.1% of EMU3 detections (N1), 68.8% of EMU1 detections (N2), and 20.0% of WN3 detections (E).

The most frequent gene target was assessed for city, university, and township samples. From these data we found that the N1 gene target was most frequent in the township locations while the N2 gene target was most frequent in the city and university locations. The gene target detections per SARS-CoV-2 positive samples were as follows for the township locations: 91.1% (N1), 84.5% (N2), and 72.9% (E). The gene target detections per SARS-CoV-2 positive samples were as follows for the university locations: 75.1% (N1), 84.5% (N2), and 51.6% (E). The gene target detections per SARS-CoV-2 positive samples were as follows for the city locations: 82.3% (N1), 88.0% (N2), and 62.1% (E). Lastly, the township sites had the highest percentage of detections per samples analyzed (86.6%) followed by the city (79.5%) and university (60.0%) samples. Overall, the township sites favored N1 in all 7 sites. The university sites favored N2 in 11/17 sites, N1 in 3/17 sites, and both N1 and N2 in 3/17 sites. The city sites favored N2 in 4/5 sites and N1 in 1/5 site. Thus, the N1 gene target had the most detections in 11 sites, the N2 gene target had the most detections in 15 sites, and N1 and N2 had equal frequencies of detections for three sites.

Next, the gene targets were compared with a paired t-test to determine if they had a statistically similar concentration. Comparing N1 to E, a p-value of 1.4×10^{-7} was obtained suggesting that the N1 gene target was found at statistically higher concentrations than the E gene target. When N2 was compared to E, N2 was also found to have a statistically higher concentration than E ($p=4.2 \times 10^{-6}$). The N1 and N2 gene targets were compared and the N1 gene target was found in higher concentrations than the N2 gene target ($p=2.3 \times 10^{-6}$). For all three gene target combinations, the targets were found to have extremely high positive correlations with each other. The Pearson's correlation

coefficients were as follows: 1.00 (N1 and N2), 0.99 (N1 and E), and 0.99 (N2 and E). Overall, the N1 gene target had the highest concentration for most of the samples with 621/1,430 samples having the highest concentration for N1. The N2 gene target had the highest concentration in 312/1,430 samples. The E gene target only had the highest concentration in 63/1,430 samples. There were 4 samples that had equal concentration for N1, N2, and E, and 400 at the MDL.

Next, the number of samples that were only positive for one gene target were assessed. Each gene target increased SARS-CoV-2 detection by the following percentage: 10.7% (109/1,019) more SARS-CoV-2 detection when N2 is analyzed, 8.3% (85/1,019) more SARS-CoV-2 detection when N1 is analyzed, and 2.3% (23/1,019) more SARS-CoV-2 detection when E is analyzed. Based on these data, the E gene target was dropped from the workflow.

After the first phase of testing, the lab increased the surveillance in Southeastern Michigan. Therefore, instead of 29 sites, 44 sites were analyzed from 14 distinct areas. Some additional sites were added to EMU and OU as well as 10 other sites. These 10 other sites included Delhi, Grosse Ile (GILE), South Lyon (Lyon), Mount Clemons (Mtclms), New Baltimore (Baltimore), Richmond, Romeo, St. Clair, Wixom, and the Michigan Health Department (MCHD). Data for the N1 and N2 gene targets were collected from September 16th, 2020 to July 19th, 2022. A total of 4,527 samples were analyzed for the N1 and N2 gene targets.

There were 4,527 samples tested in the next phase, and there were2,833 detections for SARS-CoV-2. Of these samples, 2,559 samples were positive for N1, and 2,453 samples were positive for N2.

Table 2.5: N1 and N2 Detections per Site

Site	N1 detections	N2 detections	Total Detections/All Samples	City, Township, or University
CT1	152	136	155/183	Township
CT2	142	129	147/181	Township
CT3	141	127	146/178	Township
CT4	123	110	136/179	Township
CT5	141	131	151/185	Township
CT6	119	97	124/180	Township
CT7	57	50	60/104	Township
EMU1	20	25	32/98	University
EMU2	35	38	38/49	University
EMU3	31	37	42/96	University
EMU4	52	49	58/100	University
EMU5	33	33	41/96	University
EMU6	50	48	58/100	University
EMU7	56	55	65/95	University
EMU8	65	62	71/97	University
EMU9	21	32	36/97	University
EMU10	20	21	23/93	University
EMU11	49	50	59/103	University
EMU12	5	5	8/48	University
OU1	39	38	46/96	University
OU2	27	34	37/93	University
OU3	40	42	50/97	University
OU4	44	46	53/95	University
OU5	53	50	59/96	University
OU6	29	28	39/94	University
OU7	19	19	24/60	University
OU8	13	13	17/50	University
WN1	47	43	50/76	City
WN2	49	49	57/78	City
WN3	14	18	23/79	City
WN4	47	53	55/74	City
WN5	120	111	125/140	City
MCHD1	16	19	23/98	City
MCHD2	41	40	43/95	City
MCHD3	15	18	22/97	City
Delhi	80	74	85/100	Township
GILE	24	24	26/37	Township
Lyon	65	63	67/81	City
Mtclms	80	79	84/106	City

Table 2.5: N1 and N2 Detections per Site (Continued)

Site	N1 detections	N2 detections	Total Detections/All Samples	City, Township, or University
Baltimore	83	74	85/100	City
Richmond	68	69	72/100	City
Romeo	69	65	71/92	City
St. Clair	87	79	89/104	City
Wixom	76	67	78/121	City

Table 2.5: This table shows the number of detections that the N1 and N2 gene targets had based on the 4,527 samples collected between September 2020 and July 2022. There were 44 different sites located in 14 different areas. The samples that were derived from the same area but in different locations of that area will be represented with a number. The samples were taken from cities, universities, or townships.

In this phase, N1 was detected more frequently than N2. The N1 gene target also had a higher average concentration ($8.1x10^4$ CP/L or $13.6x10^4$ CP/L without MDL values) than the N2 gene target ($6.6x10^4$ CP/L or $11.5x10^4$ CP/L without MDL values). Table 2.5 shows the number of samples that were positive for the N1 and N2 gene targets in the second phase of testing. For the 44 sites analyzed, the N1 gene target had a greater number of detections than the N2 gene targets for 24 sites. The N2 gene target had a greater number of detections than the N1 gene target for 14 sites. The remaining six sites had the same number of detections for both the N1 and N2 gene targets. Interestingly, the sites with the highest number of N1 detections, N2 detections, and total detections were the same as the sites listed for the first set of experiments despite over an additional year's worth of data and an extra 15 sites. However, the sites with the fewest number of detections were different for N1, N2, and the overall detections. Site CT1 had the highest number of N1 detections with 92.1% of its detections positive for N1 while the site EMU9 had the fewest number of detections that were positive for N1 (58.3%). Site

EMU2 attributed 100.0% of its detections to the N2 gene target while site EMU12 had the fewest number of detections for the N2 gene target with only 62.3% of its detections being attributed to the N2 gene target. Lastly, the site WN5 had the highest percentage of samples with detections with 89.3% of its samples positive for SARS-CoV-2. The site EMU12 had the lowest number of samples positive for SARS-CoV-2 with only 16.7% of its samples positive for SARS-CoV-2.

City, township, and university data were combined to determine if there was a target that was more frequent in each area. There were 15 city samples, nine township samples, and 20 university samples. For the city samples, the N1 and N2 gene targets had a very similar percentage of detections. The N1 gene target was responsible for 88.8% of the detections in the city sites, and the N2 gene target was 88.7% of the detections in the city sites. Next, the township sites favored the N1 gene target with 94.7% of the detections attributed to the N1 gene target while 85.7% of the detections were attributed to the N2 gene target. Lastly, the university sites favored the N2 gene target with 83.6% of detections attributed to the N2 gene target and 79.7% of detections attributed to the N1 gene target. Overall, the nine city sites favored the N1 gene target, five city sites favored the N2 gene target, and one site had the same number of N1 and N2 detections. The township sites heavily favored the N1 gene target with the N1 gene target having the most detections in eight sites with the remaining site having the same number of N1 and N2 detections. The university sites had seven sites that favored N1, nine sites that favored N2, and four sites with equal N1 and N2 detections. The township sites had the highest percentage of detections with 76.4% of the samples positive for SARS-CoV-2. The city

sites had 64.6% of samples test positive for SARS-CoV-2 while the university sites only had 48.1% of the samples test positive for SARS-CoV-2.

Paired t-tests were carried out to determine if one gene target was found at a statistically higher concentration than the other gene target. The N1 gene target had a statistically higher concentration than the N2 gene target with a p-value of 8.2×10^{-17}. These two gene targets also had a very strong positive correlation with a Pearson's correlation coefficient value of 0.99. Examining all samples, it was found that the N1 gene target had a higher concentration than the N2 gene target in 1,999/4,527 samples. The N2 gene target had 779/4,527 with a higher concentration than N. Sixty-four samples had the same concentration for the N1 and N2 gene targets. About 1,685) had both N1 and N2 gene target concentrations equivalent to the MDL.

The frequency of the targets was inspected to determine if one gene target could be removed from analysis without losing a high number of SARS-CoV-2 detection. There was a total of 377 samples that were only positive for the N1 gene target out of the 2,833 SARS-CoV-2 positive samples. There was a total of 271 samples that were only positive for the N2 gene target out of the 2,833 SARS-CoV-2 positive samples. The N1 gene target was found to be superior to the N2 gene target both in terms of concentration and frequency of detection from the total 4,527 sample. Although N1 was superior to N2, the N2 gene target added detections that would otherwise go unreported. Removing the N2 gene target would have resulted in 9.6% SARS-CoV-2 fewer detections while removing the N1 gene target would have resulted in 13.3% fewer detections. Therefore, it is not suggested that either target be removed from analysis. Overall, it is suggested that the N1

and N2 gene targets should be duplexed while the E gene target should be removed from analysis which will result in minimal loss of SARS-CoV-2 detection in sewage.

2.5 Discussion

The prevalence of COVID-19 within communities can be monitored with sewage surveillance in addition to clinical testing. Sampling wastewater is a routine occurrence at wastewater treatment plants in order for these plants to meet their permit compliance requirements (McClary-Gutierrez et al., 2021). Adding sampling for SARS-CoV-2 testing is generally easy to implement into wastewater treatment plants (McClary-Gutierrez et al., 2021). Surveilling sewage in communities allows monitoring many people with a single sample in a quick and cost-effective way (Geraghty, 2022). Wastewater surveillance was first implemented in 1939 for detecting poliovirus and it proved useful today as it was found in the sewage of three cities in North America which were experiencing an outbreak (Thompson et al., 2020). Sewage surveillance is valuable because it is not based on behavioral changes so bias is minimized in comparison to clinical testing (Paterson & Durrheim, 2022). For example, everyone showers and uses the toilet, but individuals often choose not to be tested in a clinical setting (Paterson & Durrheim, 2022). In general, nasopharyngeal swabs are much more sensitive than fecal testing. The sewage signal will be hampered by the high amount of dilution in the sewage and potential inhibitors in the sewage. However, sewage surveillance is advantageous because it can pick up on cases that are mildly symptomatic or asymptomatic which are still contagious and will be underreported (Thompson et al., 2020). Overall, testing sewage detects asymptomatic cases which account for a significant number of COVID-19 cases, and provides data where there is reluctance to be tested clinically or insufficient

funding to test (O'Keeffe, 2021). Due to these factors, it is thought that clinical testing may underestimate the community burden of COVID-19 making sewage surveillance an even more useful tool to guide community strategies to mitigate COVID-19 spikes (O'Keeffe, 2021). Wastewater surveillance data may be used to allow students to return safely to campus and to provide data to public health officials who can follow the COVID-19 trends to determine if and when safety measures need to be taken in specific communities (Safford et al., 2022).

Our research utilized three gene targets recommended by the CDC and Corman et al. to detect SARS-CoV-2 in sewage. The E and N gene targets appear to be the most suggested targets for SARS-CoV-2 detection. The E and N gene targets were said to be more sensitive when compared to the ORF1ab gene target even though their genes were less conserved than the ORF1ab gene (Zhou et al., 2020). The ORF1ab gene is the most conserved gene of SARS-CoV-2, but when detecting SARS-CoV-2 using a target derived from this gene, there was low sensitivity (Zhou et al., 2020). Another study mentioned that the RNA-dependent RNA polymerase (RdRp) assay was not as sensitive for detection of SARS-CoV-2 as the E and N gene target based assays (Mardian et al., 2021). They suggested that the decrease in detection may be attributed to the reverse primer used in the RdRp assay mismatching with the SARS-CoV-2 sequence (Mardian et al., 2021). There have been few studies that compare these three targets in sewage. A study by Colton et al. analyzed 12,015 clinical samples (Colton et al., 2021). Of these samples, 2,593 were SARS-CoV-2 positive with 2,273 (87.7%) of the samples positive for both E and RdRp targets, 319 (12.3%) positive for only the E target, and 1 (<0.1%) positive for only the RdRP target (Colton et al., 2021). They emphasized the need for dual testing to

improve the sensitivity allowing earlier decisions to combat the spread of disease (Colton et al., 2021). Ahmed et al. looked at the amplification efficiencies of the E, N1, and N2 gene targets. Based on the Minimum Information for Publication of Quantitative Real-Time PCR Experiments guidelines which suggest a range of 90-110%, the N1 (98.6-106%) and N2 (94.6-103%) gene targets were within range while the E gene target was slightly outside of this range (89.6-96.5%) (Ahmed, Bertsch, Angel, et al., 2020). This suggested that the N1 and N2 gene targets would outperform the E gene target (Ahmed, Bertsch, Angel, et al., 2020). A study by Pabbaraju et al. suggested that the E and N2 gene targets were able to detect SARS-CoV-2 at lower concentrations than the N1 and RdRp targets (Pabbaraju et al., 2021). The E and N2 gene targets were able to detect SARS-CoV-2 at as low as 31.5 viral copies while N1 was not able to identify SARS-CoV-2 until there was 63 viral copies, and RdRp needed 10-fold more copies (630 viral copies) in order to detect SARS-CoV-2 (Pabbaraju et al., 2021). Lastly, Acosta et al. had the best data to compare to our study because it analyzed all three gene targets our study analyzed. Overall, they analyzed wastewater samples and found that the N1 gene target was detected in 64.1% of samples, the N2 gene target was detected in 49.7% of samples, and the E gene target was detected in 10% of samples (Nicole Acosta et al., 2021). In their study, they found that 9/96 samples were positive for E, N1, and N2, 28/96 samples were positive for N1 and N2, and 39/96 samples were positive for N1 (Nicole Acosta et al., 2021). From this information, they dropped the E gene target from their future testing (Nicole Acosta et al., 2021). Our study was similar but carried out for over nine months with three SARS-CoV-2 gene targets. We found that 861/1,430 samples were positive for N1, 873/1,430 samples were positive for N2, and 655/1,430 samples were positive for E.

The Acosta et al. study found that between August and October the N1 gene target could identify some SARS-CoV-2 positive samples without the N2 or E gene target (Nicole Acosta et al., 2021). Our study also showed that a gene target could identify a SARS-CoV-2 positive sample alone but would miss some positive samples. There were 1,019 detections out of the 1,430 samples analyzed. From this group of samples, E was detected in 23/1,019 (2.3%) of samples by itself, N1 was detected in 85/1,019 (8.3%) samples by itself, and N2 was detected in 109/1,019 (10.7%) samples by itself. Although the N2 gene target was detected the more than the other gene targets, our data showed that the N1 gene target ($10.4x10^4$ CP/L) had a statistically higher concentration than both N2 ($9.1x10^4$ CP/L) and E ($6.4x10^4$ CP/L) with p-values of $2.3x10^{-6}$ (N1 vs. N2) and $1.4x10^{-7}$ (N1 vs. E). Overall, the E gene target was detected in over 200 less samples than the N1 and N2 gene targets and only added 2.3% more SARS-CoV-2 detection. Therefore, the E gene target may be removed from workflows with little risk (~2.3%) of making a false negative determination. Removing the E gene target will reduce the cost of SARS-CoV-2 detection significantly with only a small amount of SARS-CoV-2 signal loss. As three targets were run, one target had to be run in a singleplex to obtain separation between the positive and negative droplets. Although all three targets could be detected from the same RNA eluate obtained from the spin column, the cost of the assay mix was doubled to include this additional target. The assay mix is the most expensive part of analysis. In addition to the extra assay mix, additional oils were needed to generate the droplets and analyze the wells when the E gene target was analyzed. Therefore, dropping the E gene target significantly reduced cost and saved time associated with the Automated Droplet Generator, pipetting the samples, and the QX200™ Droplet Reader reading the data.

Also, dropping the E gene target allowed room on the plate for the inhibition controls to be be run simultaneously with the SARS-CoV-2 gene targets allowing results on concentration and inhibition to be processed together. Running inhibition and SARS-CoV-2 targets together allowed samples to be rerun immediately if they were fully inhibited.

After the E gene target was removed, the N1 and the N2 gene targets were compared side-by-side to determine if one gene target was sufficient to detect SARS-CoV-2 without reducing SARS-CoV-2 detection and increasing the number of false negative determinations. The study by Acosta et al. also compared solely the N1 and N2 gene targets after removing the E gene target from analysis. They analyzed 69 wastewater samples and found that both gene targets were found in 51 samples while the N1 gene target was found in 63 wastewater samples (Nicole Acosta et al., 2021). Again, they conclude that the N1 gene target can detect SARS-CoV-2 alone. Our larger N1 and N2 comparison obtained data for nearly two years. A total of 4,527 samples were collected with 2,833 samples positive for SARS-CoV-2. Our data agreed with the Acosta et al. study suggesting that the N1 gene target was superior to the N2 gene target. We found that 2,559 wastewater samples were positive for N1, and 2,453 samples were positive for N2. Both the N1 and N2 gene targets were able to detect SARS-CoV-2 by themselves. There were 377/2,833 (13.3%) SARS-CoV-2 positive samples that were only positive for N1, and 271/2,833 (9.6%) SARS-CoV-2 positive samples that were only positive for N2. Therefore, we suggest that both gene targets are necessary for a high level of SARS-CoV-2 detection in wastewater. The N1 and N2 gene targets can be duplexed making the only additional cost the extra primers and probe. In addition, little additional time would

be required. Because it is effectively the same price to run a duplex as it is to run a singleplex, and running both targets would add around 10% more SARS-CoV-2 detections, both gene targets should be duplexed for the best SARS-CoV-2 detection. Other studies have also looked at both SARS-CoV-2 gene targets. One of these studies analyzed 1,029 samples from patients and determined that both N1 and N2 had no differences between their genome copies or cycle threshold values (Klein et al., 2020). In contrast, our study suggested that N1 ($8.1x10^4$ CP/L) had a statistically higher concentration than N2 ($6.6x10^4$ CP/L) with a p-value of $8.2x10^{-17}$.

We initially hypothesized that the location will influence the contribution of the SARS-CoV-2 gene targets. In this study we had 29 different sites analyzed for the N1, N2, and E gene targets in the first set of experiments, and the N1 and N2 gene targets were analyzed for 44 different sites when all the data were included. The samples analyzed were broken into university, township, and city samples. For the first set of data (1,430 data points), the townships favored N1, while the city and university sites favored the N2 gene target. The E gene target was not favored at any site. Overall, the N1 gene target was favored in 11 sites, the N2 gene target was favored in 15 sites, and they tied for best target in three sites. When analyzing the full 4,527 samples, the N1 gene target was most frequent in the township and city data, and the N2 gene target was most frequent in the university samples. Overall, there were 24 sites that favored N1, 14 sites that favored N2, and six sites that the gene targets tied for best target. The gene target favored switched in 11/29 sites including sites that showed that N1 and N2 were equivalent then changed to favoring one target and vice versa. Of these sites, nine were university sites and two were city sites. The township data were consistent. Overall, this fluctuation suggests that it

may be difficult to try to pinpoint a gene target that will be the most effective based on location. These gene targets will be broken down based on the composition of the sewage. Due to this, it is feasible that sites could have a superior gene target. However, this data showed that over a third of the sites (11/29) fluctuated in the superior gene based on comparing the data that were taken for around nine months and the data that were collected for nearly two years. This could suggest that there were changes in the sewage matrix due to human behaviors which caused the N1 gene target to be most frequent in some cases while the N2 gene target was most frequent in other cases. The fluctuation in which gene target was superior at a site could be due to which RNases are present and the main RNase in the sewage matrix. RNases are very common and can be found in almost every organism such as fungi, plants, humans, viruses, bacteria, and animals (Hillwig et al., 2009). Although RNases can degrade many RNA molecules making them nonspecific, RNases can degrade specific base sequences such as RNase E that cleaves AU-rich locations in the RNA molecule (Baek et al., 2019). Therefore, the RNase composition could influence which gene target is dominant. Also, if a mutation arises which prevents the primers or probes from binding, the sequence of the gene target would not amplify. This could also cause the most frequent gene target to change at a site. Therefore, we suggest that it is important to use both N1 and N2 gene targets to detect SARS-CoV-2. One gene target may work fine for a period of time, but it is feasible that the dominant gene target could change over the course of time due to fluctuations in the sewage matrix or mutations.

The N1 and N2 gene targets were closer in the number of samples they detected compared to the E gene target. It reasonable that the N1 and N2 gene targets would be

detected at a similar rate because they are derived from the same gene. The E gene target could be detected less frequently if it degraded faster than the N1 and N2 gene targets. However, the temperature of sewage is generally between 10-21°C with an average temperature of approximately 15.6°C (Jamal, n.d.; *Waste Wattage*, 2012). This temperature is colder than room temperature. Therefore, the rate of degradation would be expected to be slow and likely similar amongst the targets. Thus, it is not expected that degradation played a big role in the reduced frequency of the E gene target. Other possibilities for the reduced frequency and concentration of the E gene target could be due to a lower production of the E gene compared to the N gene or the human body favoring the breakdown of the E gene. One study mentioned that the N gene had a higher transcription efficiency than the E gene (Y. Zhao et al., 2020). The methodology used to detect SARS-CoV-2 in sewage requires that SARS-CoV-2 be converted into DNA by reverse transcription. This DNA is then amplified and detected by ddPCR. Thus, the E genes reduced efficiency for transcription could cause a reduction in the E genes targets replication which would account for the lower amount of E gene target detected in our study. Next, the E gene target could degrade quicker in the human body. However, there are no articles that support this idea. Next, the E gene target may be less prone to amplification than the N gene targets because the N gene targets require fewer base pairs to replicate than the E gene target. The E gene is only 227 nucleotides long which is almost six times smaller than the N gene which is 1260 nucleotides long (Bland et al., 2021). Therefore, if the E gene is cut by an enzyme there is a high chance that the segment cut is needed for amplification. The primers and probes need 125 of these nucleotides to amplify the E gene target which is over half of the E gene

(*Whoinhouseassays.Pdf*, n.d.). Thus, enzymes cutting the E gene could cause this portion of target to not be amplifiable. Because the N gene is so much larger than the E gene, an enzyme cutting the N gene is less likely to render it non-amplifiable. The N1 gene target amplifies 73 nucleotides of the N gene, and the N2 gene target amplifies 67 nucleotides of the N gene (Lu et al., 2020). The N1 gene target only makes up 5.8% of the N gene, and the N2 gene target is only 5.3% of the N gene. This suggests that the N gene targets have a higher probability of staying intact than the E gene target which requires 55.1% of the E gene to be intact. However, there are studies that show that the E gene target was better at detecting SARS-CoV-2 than the N gene. Corman et al. found the N gene to be the least sensitive and dropped it from further analysis (Corman et al., 2020a). However, the N gene target in his study is not the N1 or N2 gene suggested by the CDC. The selected primers and probes to amplify the N gene target in that study may not be as sensitive as the N1 and N2 gene targets used for our study. Several clinical comparisons suggest that the E gene target is superior to N1 and N2. Unfortunately, the sequences of the primers and probes are different or could not be found so it is uncertain if the same primers and probes that our study uses were comparable (Chu et al., 2021)(Liotti et al., 2021).

Another reason that the E gene target could be detected less frequently could be a mutation occurred in the E gene target which would not allow amplification with our primers and probe. There were two different studies that mentioned mutations in the E gene which caused issues with amplification leading to false negatives. One study found a mutation in the E gene which was caused by a single-nucleotide polymorphism which converted the base cytosine into a uracil at nucleotide 26340 (Artesi et al., 2020). This

mutation was associated with false negatives when the Roche cobas SARS-CoV-2 assay was used (Artesi et al., 2020). However, this study mentioned that they did not find evidence that suggested this mutation had an adverse effect on identification of the E gene target with the primers and probe suggested by Corman et al. which is used in our study (Artesi et al., 2020). Although this study said that the primers and probe from Corman's study were not impacted for them, there is no certainty that it could never cause issues with detection. Nevertheless, another study mentioned a different mutation at nucleotide 26372 in which the base guanine was changed into a uracil (Tahan et al., n.d.). This study states that the ability of the primers and probes used in our research could be impacted resulting in false negative results (Tahan et al., n.d.). There could also be mutations in the N gene targets too. There is one study that found a mutation which caused the base guanine to change to a thymine at nucleotide 29104 (Vanaerschot et al., 2020). This portion of the N gene had a mutation where the forward primer binds to amplify the DNA (Vanaerschot et al., 2020). Thus, this can cause issues with amplification and lead to false negatives. However, this N gene target did not match either of the N gene targets used in our research. In general, mutations can occur in the genes which could prevent primer/probe binding leading to false negatives. Thus, it is important to analyze at least two targets. One of these studies suggested that the N gene is found in larger amounts than the E gene (blog, n.d.). Additionally, this article suggested that due to the lower viral load of the E gene it is the first gene to become undetectable (blog, n.d.). Another study stated that the N gene is the gene found in the highest amounts during infection, and that the E gene is present in the lowest amounts during infection (Mollaei et al., 2020). Their research showed that the N gene target was the most specific

and sensitive gene target (Mollaei et al., 2020). Specificity was defined as the ability of the gene target to test negative in an individual without COVID-19 (Mollaei et al., 2020). The N gene target had 100% specificity, and the E gene target only had 66.7% specificity (Mollaei et al., 2020). Sensitivity was defined as the ability of the gene target to test positive when an individual had COVID-19 (Mollaei et al., 2020). The N gene target had 96.6% sensitivity, and the E gene target only had 66.7% sensitivity (Mollaei et al., 2020). The primers and probes in this study do not match ours. However, this shows that the N gene targets could be better gene targets than the E gene target. Overall, the reduced capability of the E gene to undergo transcription, mutations in the E gene, degradation, and inhibition could be the main reasons that the E gene target in our study is not detected as frequently as the N gene targets.

Due to mutations in this virus, it is possible that either one or both N1 and N2 gene targets could become compromised due to a mutation which prevents the primers/probe from binding to the RNA. These mutations could become the dominant strain of SARS-CoV-2. Thus, new gene targets may be used in the future. The spike gene has been reported to have around 50% of the mutations even though it only makes up 12.7% of the genome (*Technical-Bulletin_sars-Cov-2-Variants-Mutations.Pdf*, n.d.). Thus, the spike gene would not be the best candidate for SARS-CoV-2 detection. The ORF1ab gene had a mutation rate around 3/10,000 which was much lower than the spike gene's mutation rate of around 2.3/1,000. This article was not clear about the units for the mutation rate (*Technical-Bulletin_sars-Cov-2-Variants-Mutations.Pdf*, n.d.). However, the mutation rates per gene can be rated from highest to lowest mutation rate as follows: ORF8 > S > N = E > ORF1ab > ORF3a (*Technical-Bulletin_sars-Cov-2-Variants-Mutations.Pdf*,

n.d.). Thus, if the N gene becomes undetectable due to mutations, one of the genes with a lower mutation rate could be used. On the other hand, if a mutation becomes dominant new primer/probe sequences may be created and used for routine detection.

2.6 Conclusion

This study aimed to identify a combination of gene targets to detect SARS-CoV-2 that minimized false negatives and provided an efficient workflow. For the first phase, 1,430 samples were tested with 1,019 samples of these samples positive for at least one gene target. The focus of this study was to determine if one or more gene targets could be eliminated. Our data showed that the E gene target was by far the most inferior gene target detecting over 200 less samples than either the N1 or N2 gene target. The E gene target only detected 655/1,019 SARS-CoV-2 positive samples while the N1 and N2 gene targets detected 861/1,019 and 873/1,019 of the SARS-CoV-2 positive samples, respectively. We determined what percentage of the SARS-CoV-2 signal was detected by a single gene target. It was found that the N1 gene target by itself detected 8.3% of the SARS-CoV-2 positive samples, the N2 gene target by itself detected 10.7% of the SARS-CoV2 positive samples, and the E gene target detected 2.3% SARS-CoV-2 positive samples by itself. Thus, the E gene target had a statistically lower concentration, a lower number of positive detections than the N1 and N2 gene targets, and only detected 2.3% more SARS-CoV-2 positive samples. The reduced detection from the E gene target could be due to the E gene degrading faster (unlikely based on Chapter 1 results) than the N1 and N2 genes, the E gene being produced less, mutations in the E gene, the E genes reduced ability to undergo transcription, or the E gene being more prone to inhibition. Based off this, the E gene target was removed from our protocol to reduce time and cost

with minimal reduction of SARS-CoV-2 detection. The N1 and N2 gene targets were analyzed further to assess if another gene target could be eliminated, or to determine if the N2 gene target would remain the superior gene target. A total of 2,833/4,527 samples were positive for SARS-CoV-2 with 2,559 samples positive for N1 and 2,453 samples positive for N2. The N1 gene target added 13.3% more SARS-CoV-2 detection, and the N2 gene target added 9.6% more SARS-CoV-2 detection. Therefore, these data suggest that both gene targets should be used otherwise there would be around 10% fewer SARS-CoV-2 detections. Using our data, we also found that even when sewage was taken from the same site, the dominate gene target could fluctuate over time. It was seen that 11/29 sites switched which target was most frequent. This also suggests that two gene targets should be analyzed. Our work suggests that the E gene target can be removed with minimal loss of signal, but the N1 and N2 gene targets should both be retained. The N1 and N2 gene targets can be duplexed essentially requiring little additional cost or time to process.

CHAPTER THREE

FLOW AND FECAL INDICATOR BASED NORMALIZATION OF THE N1 AND N2 SARS-COV-2 GENE TARGETS AND ESTIMATION OF COVID-19 CASES

3.1 Abstract

Coronavirus disease 2019 is a disease which is generally mild, but it could be deadly in some cases. This disease is caused by an RNA virus called Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Generally, COVID-19 is monitored clinically with nasopharyngeal swabs, but it can also be monitored in sewage. Clinical testing provides more certainty because each individual is tested while sewage surveillance only provides a trend in COVID-19 cases. Sewage surveillance is advantageous because it may be able to detect trends before clinical testing, and it can test large groups of people quickly without consent. This work used fecal based and flow-based normalization to determine if either could provide a better correlation between the clinical and sewage data. For all three sites, it was found that the best correlation was between the unaltered and clinical data while pepper mild mottle virus (PMMoV) normalized data provided the worst correlation. The fecal indicators HF183 and Lachno3 provided the best fecal based normalization to the clinical data. However, normalizing with flow was superior to normalizing with fecal indicators in 2/3 sites. Statistically, crAssphage always had thehighest concentration, and Lachno3 always had the lowest concentration. The concentrations of the SARS-CoV-2 gene targets were utilized to estimate the number of COVID-19 cases. There was a moderate to high correlation between the clinical and sewage estimated cases, but there was a statistical difference between them. Therefore,

the cases can be estimated, but the estimated cases were statistically higher than the clinically reported cases. Lastly, peaks in the graph were compared to determine if sewage surveillance provided an early warning of rising cases. At best, the sewage data were 11 days ahead of the clinical data. On average, sewage data preceded clinical data by 2.2 days.

3.2 Introduction

Coronavirus disease 2019 (COVID-19) is a respiratory disease that is highly transmissible which arose in December of 2019 in Wuhan, China (Mohapatra et al., 2020). Quickly after its discovery, the World Health Organization (WHO) labeled COVID-19 a Public Health Emergency of International Concern and a global pandemic as of January 30th, 2020 and March 11th, 2020, respectively (Hiscott et al., 2020). Mild symptoms associated with COVID-19 are cough, sore throat, diarrhea, headache, runny nose, fatigue, fever, and body aches (CDC, 2020a). This disease is also associated with neurological symptoms which can be mild including anosmia and dizziness or more severe including seizures, varied levels of consciousness, ataxia, and acute cerebrovascular events (Lee et al., 2020). As of September 2nd, 2022 the WHO has reported over 601 million cases of COVID-19 with a 1.1% death rate across the globe (*WHO Coronavirus (COVID-19) Dashboard*, n.d.). As of August 30th, 2022, the Michigan government website reported over 2.7 million confirmed and probable cases of COVID-19 with a 1.4% death rate in Michigan (Michigan.gov, n.d.-a). The disease COVID-19 is attributed to a single-stranded and enveloped virus known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (V'kovski et al., 2021).

Coronaviruses (CoVs) not only infect humans, but they can also infect birds and mammals (V'kovski et al., 2021).

In a clinical setting, the WHO suggests that SARS-CoV-2 be detected via throat or nasopharyngeal swabs (Uddin et al., n.d.). However, SARS-CoV-2 can be detected in other fluids. Uddin et al. suggested COVID-19 testing via salvia which is a noninvasive test with a sensitivity of 80.4% that falls in line with the criteria set by the WHO (Uddin et al., n.d.). Wang et al. investigated SARS-CoV-2 in human bodily fluids with the following prevalence: 63% of nasal swabs, 32% of pharyngeal swabs, 0% of urine, 93% of bronchoalveolar lavage fluid, 46% of fibrobronchoscope brush biopsies, 29% of feces, and 72% of sputum (W. Wang et al., 2020). We were interested in fecal samples because feces will be the main component carrying SARS-CoV-2 in the sewage. Jones et al. found SARS-CoV-2 in 30-60% of fecal samples while urine rarely had SARS-CoV-2 (Jones et al., 2020). Respiratory tract samples had more detections than fecal samples with 70-100% of samples testing positive for SARS-CoV-2 (Jones et al., 2020). The focus of our study was to detect the N1 and N2 SARS-CoV-2 gene targets and various fecal indicators in the sewage using Droplet Digital polymerase chain reaction (ddPCR). The primary goal was to compare the data obtained from the N1 and N2 gene targets to the clinical cases of COVID-19 to determine if there was a quantitative relationship between the sewage and the clinical data. Unfortunately, we were also aware that inflow and infiltration can dilute sewage with surface or groundwater and convoluting the relationships. The data for the clinical cases was chosen on a county by county basis from Michigan's government website (Michigan.gov, n.d.-b). Three different counties in Southeastern Michigan were assessed for the correlation between the SARS-CoV-2 gene

target data and the reported clinical cases of COVID-19 on a county-by-county basis. All three locations were wastewater treatment plants (WWTPs). Because the WWTPs do not receive sewage from the whole county, the sites are expected to have some inconsistencies, but they should follow a similar trend because they are derived from the same area. It is estimated that Delhi's WWTP accounts for 22,500/284,034 (7.9%) of the total population in Ingham County, St. Clair accounts for 20,000/160,053 (12.5%) of the total population in St. Clair County, and Warren accounts for 135,000/876,792 (15.4%) of the total population in Macomb County (*U.S. Census Bureau QuickFacts*, n.d.-a; *U.S. Census Bureau QuickFacts*, n.d.-b; *U.S. Census Bureau QuickFacts*, n.d.-c).

Cases of COVID-19 can be crudely estimated from the gene target data, flow data, feces excreted daily, and viral load. The formula used was provided in a paper by Gerrity et al. (Gerrity et al., 2020). Their article used the values of 126g feces/(person*day) and a fecal shedding rate of $10^{8.9}$ gene copies (CP)/g of feces (Gerrity et al., 2020). Our study used the values 349g feces/(person*day) and a viral load of 10^{7} CP/g of feces (S.m.k & S.d, 2000; Wölfel et al., 2020). Using this formula, the actual COVID-19 cases reported in the various Michigan counties were compared to the calculated cases of COVID-19.

In addition to comparing the clinical cases of COVID-19 from the government data to the unaltered SARS-CoV-2 gene target concentrations, these gene targets were also normalized and compared to the clinical cases of COVID-19. The concentration of the SARS-CoV-2 gene targets were normalized by dividing the SARS-CoV-2 concentration by the concentration of the various human fecal indicators or by the flow rate. The fraction of feces in the sewage is likely to change due to temporary workers, weekday commuters, and tourism (CDC, 2022d). Graham et al., further elaborated on this topic

adding the concentration of feces can vary over time due to precipitation, human behaviors, sewershed populations, and changing waste streams (Graham et al., 2020). If the sample contains a low amount of feces, the concentration of SARS-CoV-2 will also be lower. Thus, normalization will take the amount of feces into account allowing a better representation of the SARS-CoV-2 gene target concentration to be made. The CDC mentions that *Lachnospiraceae* Lachno3, crAssphage, *Bacteroides* HF183, and pepper mild mottle virus (PMMoV) can be used as fecal indicators for SARS-CoV-2 (CDC, 2022b). One study looked at human fecal pollution markers in environmental water samples seeded with human wastewater (Ahmed et al., 2016). They determined that *Bacteroides* HF183 was the most sensitive marker for determining human fecal pollution compared to polyomaviruses and human adenoviruses (Ahmed et al., 2016). However, because HF183 can sometimes be present in animal feces, they recommend that HF183 and another fecal marker be ran for the best results (Ahmed et al., 2016).

PMMoV appears to be the most commonly used fecal marker to normalize the SARS-CoV-2 signal. PMMoV was found to be the RNA virus of highest abundance in human feces in one study (Wu et al., 2020). PMMoV is very stable in sewage showing minimal seasonal variation in its concentration (Wu et al., 2020). PMMoV cannot generally be found in animal feces, but PMMoV has been detected in some non-human feces in animals which include geese, chickens, cows, and gulls but in much lower concentrations compared to concentrations seen in human feces (Kitajima et al., 2018).

Lachno3 and HF183 can also be used as human fecal indicators. Lachno3 in sewage had a concentration that was 3.0 ± 1.3-fold higher than Lachno12 and is much more human specific than Lachno12 which is considered a human preferred fecal marker because it

was present in 25% of cow fecal samples (Feng et al., n.d.). In another study Lancho3 and HF183 were found in all tertiary treated effluent except for one sample suggesting they are useful fecal indicators in sewage (Codello et al., 2021). HF183 was found in higher concentrations than Lachno3 (Codello et al., 2021).

CrAssphage is also thought to be a suitable human fecal indicator. It was found that 71.4% of human fecal samples, 60% of the collected hand rinse samples, and 56.2% of environmental swabs had crAssphage (G. W. Park et al., 2020). CrAssphage was not detected in animal stool specimens (G. W. Park et al., 2020). In another study HF183 and crAssphage CPQ_064 had concentrations of 8.07 and 8.35 log10 copy/100 ml, respectively in sewage and were not detected in dog or bird feces (Jennings et al., 2020). One study ran a duplex PCR assay using crAssphage CPQ_056 and HF183 for sewage samples and generated data that was highly accurate, sensitive, and reproducible (Ahmed, Payyappat, et al., 2019). They suggest that the use of two fecal indicators is vital to reduce the chance of false negatives when fecal concentration is low (Ahmed, Payyappat, et al., 2019). False positives are rare, but crAssphage (CPQ_056 and CPQ_064) were found in gull and dog fecal samples (Stachler et al., 2017). Lastly HF183, crAssphage CPQ_056, Lachno3, human polyomavirus, human adenovirus, and BacV6-21 were compared (Ahmed, Gyawali, et al., 2019). HF183, crAssphage CPQ_056, and Lachno3 provided the best results for fecal tracking in sewage (Ahmed, Gyawali, et al., 2019). CrAssphage CPQ_056 was the most specific (98%) compared to HF183 (96%) and Lachno3 (95%) (Ahmed, Gyawali, et al., 2019). However, when 29 human stool samples were tested, HF183 had the highest detection, and crAssphage CPQ_056 had the lowest detection (Ahmed, Gyawali, et al., 2019). Percentages of fecal indicators in the 29

samples were as follows: 72.4% (HF183), 68.9% (Lachno3) and 37.9% (crAssphage CPQ_056) (Ahmed, Gyawali, et al., 2019).

From a workflow perspective, PMMoV appears to be the best option for sample normalization because it is an RNA virus, it is present at high concentrations in sewage, and it has a high stability in sewage. In addition, HF183, Lachno3, and crAssphage were tested as fecal indicators. All these fecal indicators have been shown to be found in high concentrations in human feces, but their concentration in sewage will depend on the sewage matrix. Two different gene targets have been suggested for crAssphage including crAssphage CPQ_056 and CPQ_064. A study by Ahmed et al., showed that CPQ_056 had a slightly higher concentration in sewage than CPQ_064 (9.43 ± 0.14 log10 CP/L compared to 8.91 ± 0.17 log10 CP/L) (Ahmed et al., 2018). Also, CPQ_056 was more specific to human feces than CPQ_064 with a 95% specificity verses 93% (Ahmed et al., 2018). Thus, crAssphage CPQ_056 was used in our study. Overall, this study compares the concentration of the SARS-CoV-2 gene targets to the clinical cases of COVID-19 from Michigan's government website. The concentration of the SARS-CoV-2 gene targets will be compared to the clinical cases from the county that housed the sewage delivered to our lab. The N1 and N2 gene copies per liter of sewage were compared to the clinical cases as unaltered, normalized by fecal indicators, normalized by flow rate, or converted into estimated cases of COVID-19. Finally, we determined which fecal indicators had the highest concentration to determine which indicator might provide the highest sensitivity.

3.3 Materials and Methods

3.3.1 Sample Preparation and Concentration

Our workflow used the methods described by Flood et al (Flood et al., 2021) as a starting point. Samples were generally delivered twice a week. However, holidays and vacations led to some weeks with one or no samples. Wastewater samples were delivered in either 500mL portions (Warren) or a 250mL portions (Delhi and St. Clair). Samples were transported in a cooler with icepacks while in transit to our lab. Samples were processed immediately upon arrival. The bottle of sewage was mixed by inverting it between 40 and 60 times to minimize foaming. A 45 mL homogenous sewage sample was decanted into a 50mL conical bottom centrifuge tube. Approximately 4.0g of polyethylene glycol 8000 (catalog number 97061-100 from VWR) and 0.59g sodium chloride (catalog number 470302-520 from VWR) were added to the sample yielding a concentration of 8% polyethylene glycol and 0.2M sodium chloride. After the solids and solutions were combined, the centrifuge tubes were transferred to a rotator where the tubes were rotated at 70rpm for 2hrs. The rotator was held at 4°C. After mixing, the samples were centrifuged at 3,650 rpm for 1 hr. Once centrifugation was completed, the supernatant was discarded, and a 1-2 mL pellet remained. This allowed some pellet to be saved (frozen at -80°C) while a 200µL portion was immediately used to determine the concentration of the N1 and N2 gene targets. A 1mL portion of pellet was pipetted into a 1.5mL microcentrifuge tube and stored at -80°C. This frozen pellet was thawed to test for the four fecal indicators.

3.3.2 Purification and Elution of RNA

The QIAamp Viral RNA Mini Kit (catalog number 52904) was purchased from QIAGEN to purify the RNA from the concentrated sewage pellets. To begin the process, buffer AVL and carrier RNA were mixed in portions of 800µL and 8µL per sample, respectively. A 200µL portion of pellet was then pipetted into a 2mL cryovial containing the buffer and RNA mix. The cryovial was vortexed for 10s before incubation for 10mins at room temperature. After incubation, 800µL of 200proof ethanol was added to the cryovial and vortexed for 10s. At this point in the procedure there was 1.8mL solution in the cyrovial. This solution was run through the QIAamp mini spin column in order to obtain the extracted RNA. The 1.8mL solution was added to the spin columns as three 630µL aliquots. Three additions were done because the spin column could not hold enough solution for a single addition. Each 630µL portion of solution was added to the spin column and centrifuged sequentially. The flow through was not retained. The centrifugation settings were 1min at 8,000rpm. After the whole 1.8mL solution was added to the spin column and centrifuged, the wash buffers were added. A 500µL portion of buffer AW1 was added to the spin column and centrifuged. Next, a 500µL portion of buffer AW2 was added to the spin column. At this point in the process (and only during this step) the centrifuge speed and duration was increased to 14,000rpm for 4min. After centrifugation, centrifugation settings were returned to normal, and the spin column was transferred to a 1.5mL microcentrifuge tube. The microcentrifuge tube was used to collect the RNA eluted from the spin column. Buffer AVE was added to the spin column in a 40µL portion and incubated at room temperature for 1min. After incubation, the spin column was centrifuged. This step was repeated which resulted in a final volume of 80µL

of RNA eluate. This eluate was the portion tested to determine if N1, N2, and/or PMMoV were present in the original sewage sample.

3.3.3 Purification and Elution of DNA

To purify and eluate DNA, the QIAamp Fast DNA Stool Mini Kit from QIAGEN was used. This process was similar to the elution of RNA, but there were different buffers, different centrifugation times and speed, and incubation at one step was carried out at a higher temperature. A 1mL portion of InhibitEX Buffer was added to a 1.5mL microcentrifuge tube. Next, 200µL of pellet was added to the InhibitEX buffer and vortexed for 10-15s. Once thoroughly mixed into a homogenous solution, the microcentrifuge tube was placed into the centrifuge and spun at 14,000rpm for 1min. Next, Proteinase K (25µL) was pipetted into a clean 2mL cryovial. The supernatant from the InhibitEX buffer/pellet mix (600µL) was later pipetted on top of the Proteinase K. The cryovial was vortexed for 10-15s. Buffer AL (600µL) was added to the cryovial, and the cryovial was vortexed for 10-15s. This solution was then incubated at 70°C for 10min. Lastly, 200proof ethanol (600µL) was added to the cryovial and vortexed for 10-15s. This solution was then run through the spin column. Again, it took three runs through the spin column to use up the entirety of the solution. Each addition was in the amount of 630µL. All centrifugation steps for the DNA set of spin columns were carried out at a speed of 14,000rpm. The solution was centrifuged for 1min. All the flow-through from the spin column was discarded. Next, the wash buffers were added to the spin column in 500µL portions. Buffer AW1 was added and ran through the spin column for 1min. Buffer AW2 was added after buffer AW1 and ran through the spin column for 6min. The flow-through was discarded from both wash buffer steps. Next, the DNA was

eluted from the spin column. Thus, the spin column was transferred to a 1.5mL microcentrifuge tube to keep the eluate for analysis. Lastly, buffer ATE (200µL) was added to the spin column and incubated at room temperature for 1min. After incubation, the spin column was centrifuged for 1min. The solution obtained was analyzed for Lachno3, HF183, and crAssphage.

3.3.4 Plating and Amplification of RNA and DNA

Samples and controls (positive, negative, and no template) were pipetted in triplicate. The 1-Step RT-ddPCR Advanced Kit for Probes from Bio-Rad was used to create an assay mix for the RNA samples, and the ddPCR Supermix for Probes (No dUTP) from Bio-Rad was used to create an assay mix for the DNA samples. The ddPCR Supermix for Probes (No dUTP) contains only Supermix while the 1-Step RT-ddPCR Advanced Kit for Probes contains Supermix, reverse transcriptase, and dithiothreitol. Besides the kits, primers, probes and water were needed to make up each assay mix. Singleplex and duplex reactions varied in the amount of primer and probe added and the amount of water added per reaction well. For a duplex reaction, the assay mix for the RNA samples consisted of 5.5µL of Supermix, 1.1µL of water, 2.2µL of reverse transcriptase, 1.1µL of dithiothreitol, 3.3µL of primer/probe mix 1, and 3.3µL of primer/probe mix 2. For a singleplex reaction, the assay mix for the RNA samples consisted of 5.5µL of Supermix, 4.4µL of water, 2.2µL of reverse transcriptase, 1.1µL of dithiothreitol, and 3.3µL of primer/probe mix. The final volume of assay mix was always 16.5µL. The assay mix for the DNA samples was simpler. For the duplex and singleplex reactions the assay mix for the DNA samples received 11µL Supermix, 3.3µL primer/probe mix, and 2.2µL of water.

Table 3.1: Primer and Probe Sequences Used for SARS-CoV-2 Detection, SARS-CoV-2 Normalization, and Positive Controls for Normalization

Target	Sequence	Reference
CrAssphage Probe	5'-HEX-AATAACGAT-ZEN-TTAC GTGATGTAAC-IABkFQ-3'	(Z. Wu et al., 2020)
CrAssphage Forward Primer	CAGAAGTACAAACTCCTAAAA AACGTAGAG	(Z. Wu et al., 2020)
CrAssphage Reverse Primer	GATGACCAATAAACAAGCCAT TAGC	(Z. Wu et al., 2020)
HF183 Probe	5'-6-FAM-CTAATGGAA-ZEN-C CGCATCCC-IABkFQ-3'	(Cao et al., 2017)
HF183 Forward Primer	ATCATGAGTTCACATGTCCG	(Cao et al., 2017)
HF183 Reverse Primer	CTTCCTCTCAGAACCCCTATCC	(Cao et al., 2017)
Lachno3 Probe	5'-HEX-CTCTGACCGGTCTTTAA TCGGA-BHQ1-3'	(Feng et al., n.d.)
Lachno3 Forward Primer	CAACGCGAAGAACCTTACCAAA	(Feng et al., n.d.)
Lachno3 Reverse Primer	CCCAGAGTGCCCACCTTAAAT	(Feng et al., n.d.)
PMMoV Probe	5'-6-FAM-CCTACCGAA-ZEN-G CAAATG-IABkFQ-3'	(Haramoto et al., 2013)
PMMoV Forward Primer	GAGTGGTTTGACCTTAACGTT TGA	(Haramoto et al., 2013)
PMMoV Reverse Primer	TTGTCGGTTGCAATGCAAGT	(Haramoto et al., 2013)
N1 Probe	5'-FAM-ACCCCCGCATTACGTTTGGTGGACC-BHQ1-3'	(CDC, 2020c)
N1 Forward Primer	5'-GACCCCAAAATCAGCGAAAT-3'	(CDC, 2020c)
N1 Reverse Primer	5'-TCTGGTTACTGCCAGTTGAATCTG-3'	(CDC, 2020c)
N2 Probe	5'-HEX-ACAATTTGCCCCCAGCGCTTCAG-BHQ1-3'	(CDC, 2020c)
N2 Forward Primer	5'-TTACAAACATTGGCCGCAAA-3'	(CDC, 2020c)
N2 Reverse Primer	5'-GCGCGACATTCCGAAGAA-3'	(CDC, 2020c)

Table 3.1: Primer and Probe Sequences used for SARS-CoV-2 Detection, SARS-CoV-2 Normalization, and Positive Controls for Normalization (Continued)

Control	Sequence
PMMoV Positive Control	GGTGGCAGCAAAGGTAATGGTAGCTGTGGTTTCAAATGAGAGTG GTTTGACCTTAACGTTTGAGAGGCCTACCGA AGCAAATGTCGCACTTGCATTGCAACCGACAATTACATCAAAGG AGGAAGGTTCG
DNA Positive Control	ATCATGAGTTCACATGTCCGCATGATTAAAGGTATTTTCCGGTAG ACGATGGGGATGCGTTCCATTAGCTCGAGATAGTAGGCGGGGTA ACGGCCCACCTAGTCAACGATGGATAGGGGTTCTGAGAGGAAGG GAAGCAACGCGAAGAACCTTACCAAATCTTGACATCCCTCTGAC CGGTCTTTAATCGGACCTTCTCTTCGGAGCAGAGGTGACAGGTG GTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGT CCCGCAACGAGCGCAACCCCTATCCTCAGTAGCCAGCATTTAAG GTGGGCACTCTGGGGAGCTAATGCAGAAGTACAAACTCCTAAAA AACGTAGAGGTAGAGGTATTAATAACGATTTACGTGATGTAACT CGTAAAAAGTTTGATGAACATACTGATTGTAATAAAGCTAATGG CTTGTTTATTGGTCATCTTGAAGATGTTAAAGTTGATTGGGCTAC A

Table 3.1: Primers and probes used for normalization and standard detection of SARS-CoV-2 in sewage. This table also shows the positive controls used for the normalization targets. These positive controls were designed by Integrated DNA Technologies. The DNA positive control was used as a positive control for Lachno3, crAssphage, and HF183.

For the duplex reaction the primer/probe mix was made up of both sets of primers and probes. Again, the final volume was 16.5µL. The sequences used for the primers, probes, and positive controls are listed in Table 3.1. The assay mixes were pipetted into a 96 well plate (plated) along with 5.5µL of sample or control. In total there were four assay mixes. One assay mix for the N1 and N2 gene targets, one assay mix for HF183 and CrAssphage, one assay mix for Lachno3, and one assay mix for PMMoV were made. Once the samples were plated, the plate was sealed with foil, vortexed, and spun down to collect liquid in the bottom of the 96-well plate.

Table 3.2: Thermocycling Settings for DNA and RNA

Step	RNA	DNA	Cycles
1	25°C for 3min	25°C for 3min	1
2	50°C for 60min	N/A	1
3	95°C for 10min	95°C for 10min	1
4	95°C for 30s	94°C for 30s	40
5	55°C for 1min	52.7°C for 1min	40
6	98°C for 10min	98°C for 10min	1
7	Hold at 4°C for ∞	Hold at 4°C for ∞	1

Table 2: This table shows the thermocycling steps for both DNA and RNA experiments. Steps 4 and 5 have 40 back-to-back cycles. This means that the temperature cycles between step 4 and step 5 a total of 40 times. For all DNA steps there was a ramp speed of 2°C/s.

The plate was placed into Bio-Rad's Automated Droplet Generator, and droplets were generated and transferred to another 96-well plate. This 96-well plate was sealed and placed into the thermocycler. The thermocycler settings for the DNA and RNA plate are listed in Table 3.2. Thermocycling settings for the general settings were based on the documents listed on Bio-Rad's website for the corresponding assay mix, and annealing temperatures were found in a research article (*DdPCR Supermix for Probes (No DUTP)*, n.d.; *One-Step RT-DdPCR Advanced Kit for Probes*, n.d.; Pearson et al., 2021). For the DNA based targets, temperature gradients were analyzed, and the annealing temperature with the best separation was chosen.

3.3.5 Data Analysis

After thermocycling was completed, the plate was placed onto the QX200 Droplet reader. After the plate was read, the data was analyzed with QuantaSoft™ software. After setting thresholds the software would calculate gene copies per 20µL well. Wells with less than 10,000 acceptable droplets were rejected. The copies of gene target per liter of sewage

were calculated for RNA as shown in Equation 3.1. For PMMoV the concentration was multiplied by 100 because the sample was diluted 100-fold.

(Equation 3.1)

$$\frac{CP}{L} = copies\ per\ uL \times eluate\ volume\ (80uL) \times \frac{\frac{final\ concentrate\ volume}{extraction\ volume\ (0.2mL)}}{initial\ volume\ (45mL)} \times \frac{1000mL}{1L}$$

Equation 3.1: Gene Copies per Liter of Sewage Calculation for RNA
This equation was utilized for all of the RNA calculations. There was an additional multiplication of 100 for PMMoV only.

For DNA the equation is similar, but slightly different from the calculation for RNA. These equations are slightly different due to the additional step of the InhibitEX buffer. Equation 3.2 shows the concentration calculation for the DNA samples. The dilution factor was added to Equation 3.2.

(Equation 3.2)

$$\frac{CP}{L} = copies\ per\ uL \times eluate\ volume\ (80uL) \times \frac{1.2mL\ (inhibitEX + pellet)}{0.2mL\ (pellet)} \times \frac{\frac{final\ concentrate\ volume\ (1-2mL)}{extraction\ volume\ (0.2mL)}}{initial\ volume\ (45mL)} \times \frac{1000mL}{1L} \times \frac{100uL}{50uL}$$

Equation 3.2: Gene Copies per Liter of Sewage Calculation for DNA
This equation was utilized for all DNA calculations. There was a 2x dilution. Therefore, the formula was multiplied by 100/50 which represents the final volume of 100µL (50µL sample and 50µL water) divided by the number of samples used.

The RNA equation was used to calculate both PMMoV and the SARS-CoV-2 gene targets N1 and N2. The extra dilution factor was left off in Equation 3.1 to avoid confusion for the calculations of the N1 and N2 gene targets which do not need to be multiplied by 100 because they were not diluted. Cases of COVID-19 in each county were estimated with Equation 3.3. This formula was mentioned in a study by Gerrity et al (Gerrity et al., 2020).

(Equation 3.3)

$$COVID-19\ infected\ people = \frac{SARS-CoV-2\ target\ concentration\ \left(\frac{CP}{L}\right) \times avg\ flow\ rate\ (\frac{L}{day})}{amount\ of\ feces\ (\frac{g\ of\ feces}{day \times person})\ \times viral\ load\ (\frac{CP}{g\ of\ feces})}$$

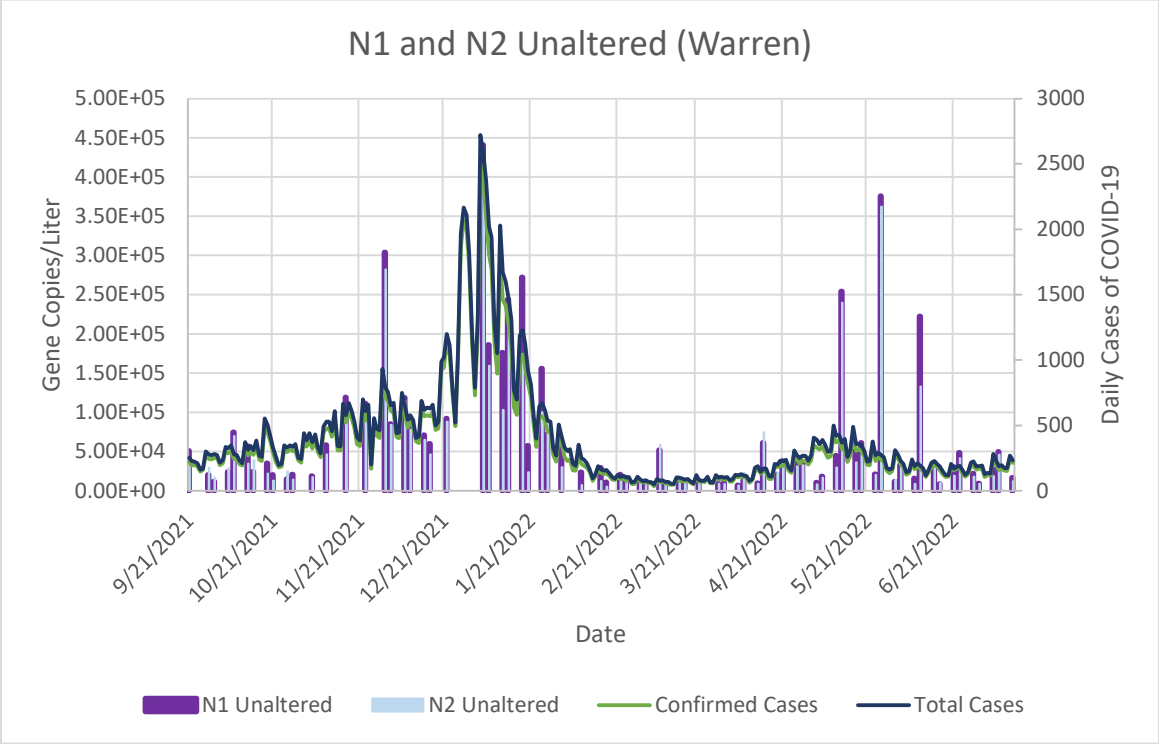
Equation 3.3: Estimated Cases of COVID-19
This equation provides an estimate of the number of COVID-19 cases that could be present in the areas analyzed.

Gene target data were compared to the clinical cases of COVID-19 obtained from the Michigan government website. Variation is expected because the WWTPs do not represent the whole county as the government data does. There should be a correlation, but there will be variation because the government data will cover a larger area. Therefore, the clinical data should have a higher number of COVID-19 cases because the area covered is larger. A double y-axis graph was used to compare the SARS-CoV-2 gene targets to the clinical cases of COVID-19. A Pearson's correlation coefficients were calculated to determine if there were similar trends between the concentrations of the gene targets and the COVID-19 cases.

3.4 Results

Clinical cases of COVID-19 were compared to the estimated cases of COVID-19 and to the N1 and N2 gene target concentrations that were normalized and unaltered. Michigan's government data provided the daily clinical cases of COVID-19. The county data was directly compared to the concentrations of the N1 and N2 gene targets, and the values obtained when these gene targets were normalized. This study used PMMoV, Lachno3, CrAssphage, and HF183 to normalize the SARS-CoV-2 gene targets. To compare the gene targets of SARS-CoV-2 to the clinical cases of COVID-19 a double y-axis graph was utilized. The date was the x-axis, and the y-axis was both the gene target concentrations and the daily clinical cases of COVID-19. After plotting the data, a trend could be compared between the target's concentrations and cases of COVID-19 to determine if the sewage and clinical data had a correlation. For all three sites in Southeastern Michigan, the unaltered N1 and N2 gene targets and the normalized targets were assessed to determine if normalizing the N1 and N2 signal led to an increase in correlation between the sewage and clinical data. Figure 3.1 graphically represents the comparison of the clinical cases of COVID-19 and the concentrations of the N1 and N2 gene targets that were normalized and unaltered. The clinical data were more detailed than the sewage data because every day had a reported number of cases. At best, the sewage data had two data points a week. Data was plotted from September 21st, 2021 – July 12th, 2022 for all three sites. Warren had 74 unique sewage samples, Delhi had 75 unique sewage samples, and St. Clair had 78 unique sewage samples.

Figure 3.1: Gene Target Concentration vs. Clinical Cases of COVID-19



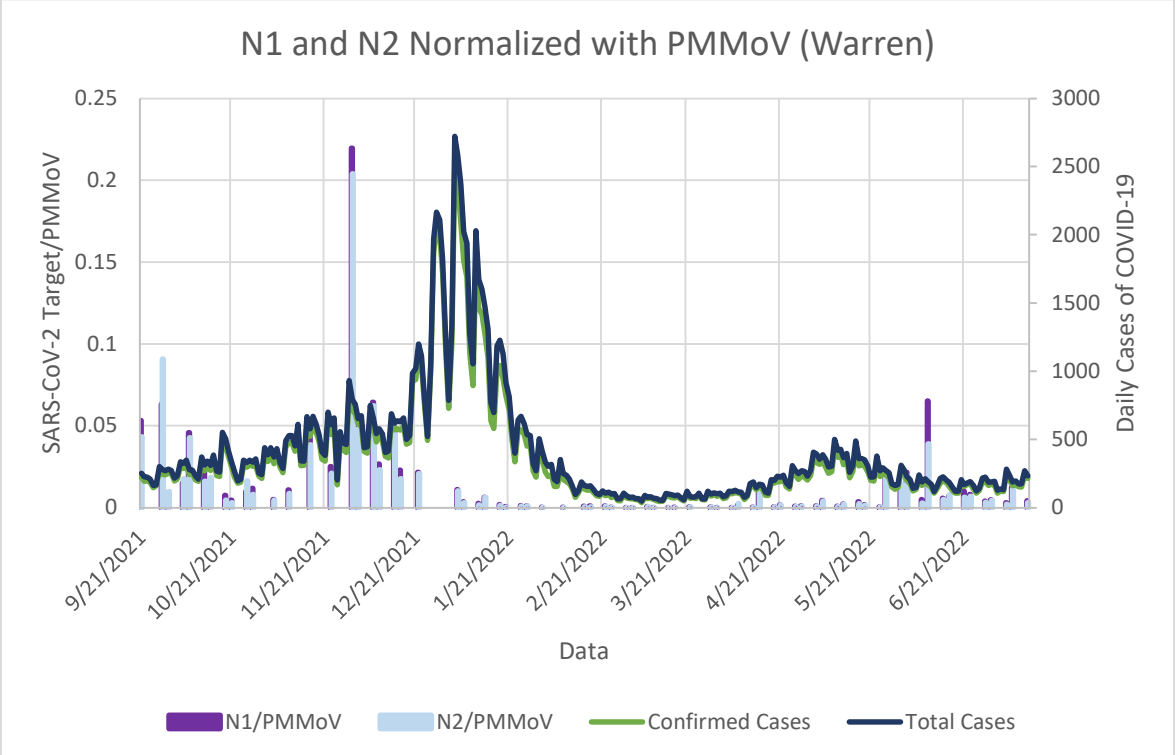
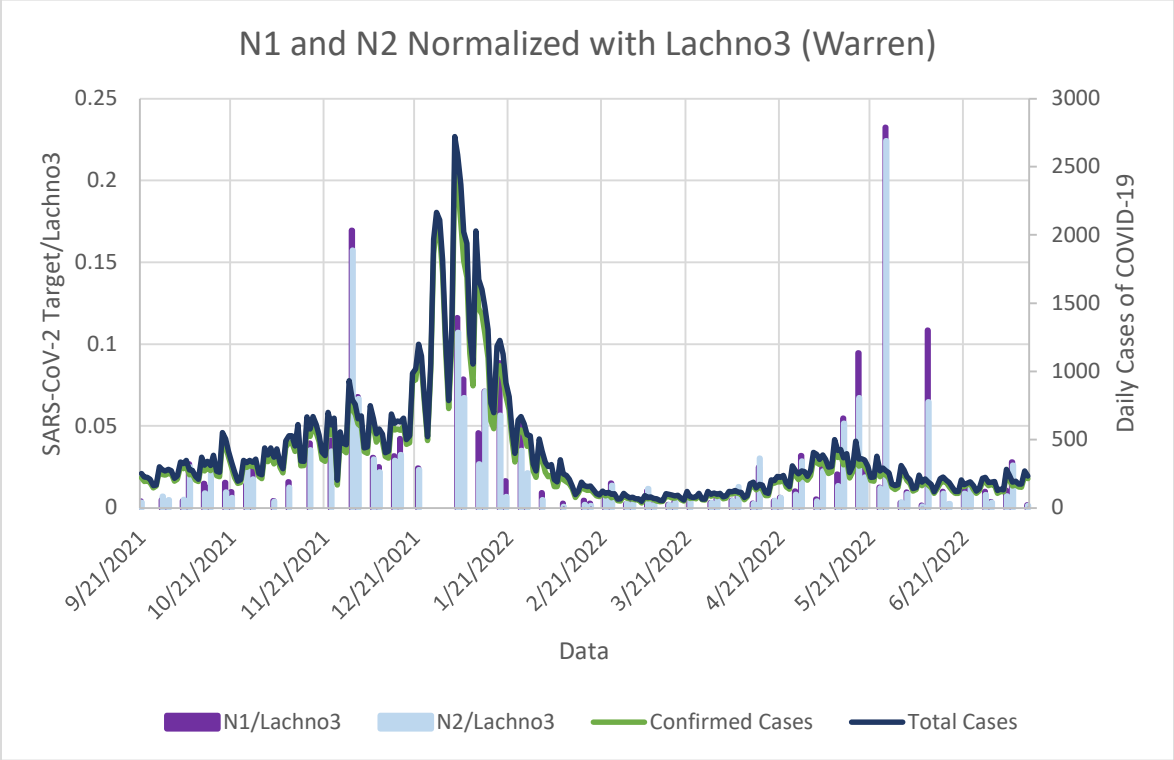


Figure 3.1: Gene Target Concentration vs. Clinical Cases of COVID-19 (Continued)



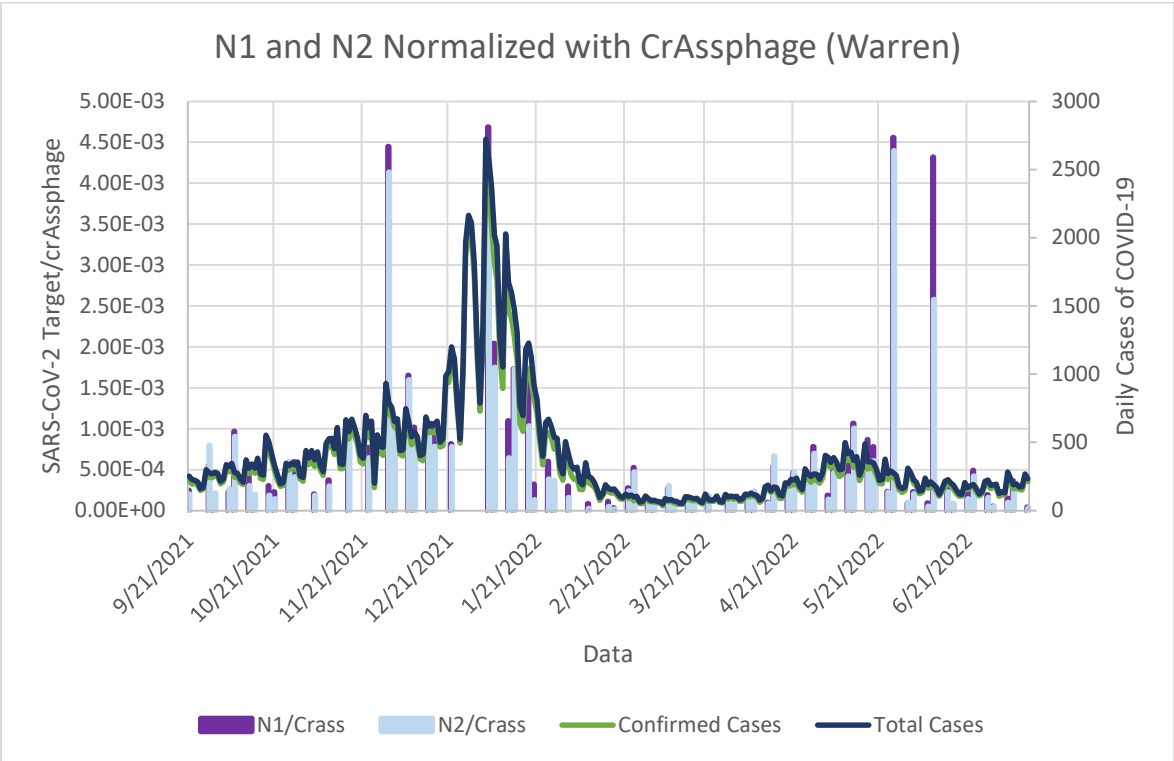
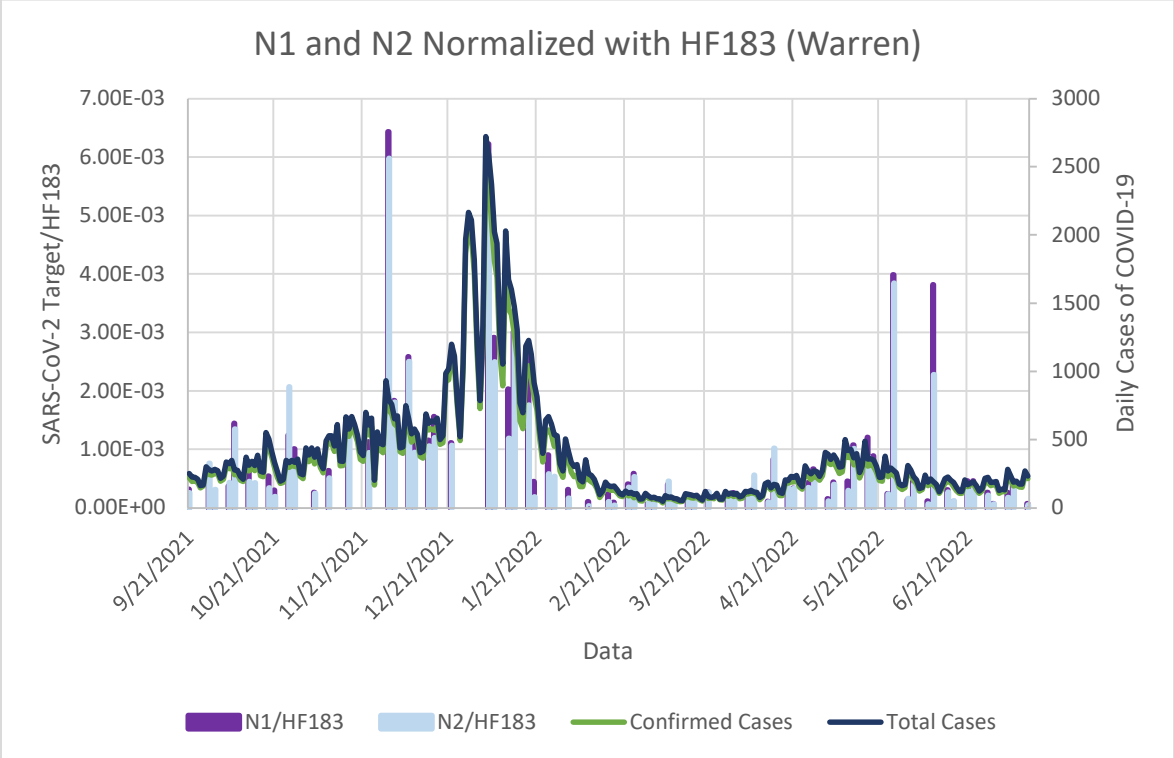


Figure 3.1: Gene Target Concentration vs. Clinical Cases of COVID-19 (Continued)



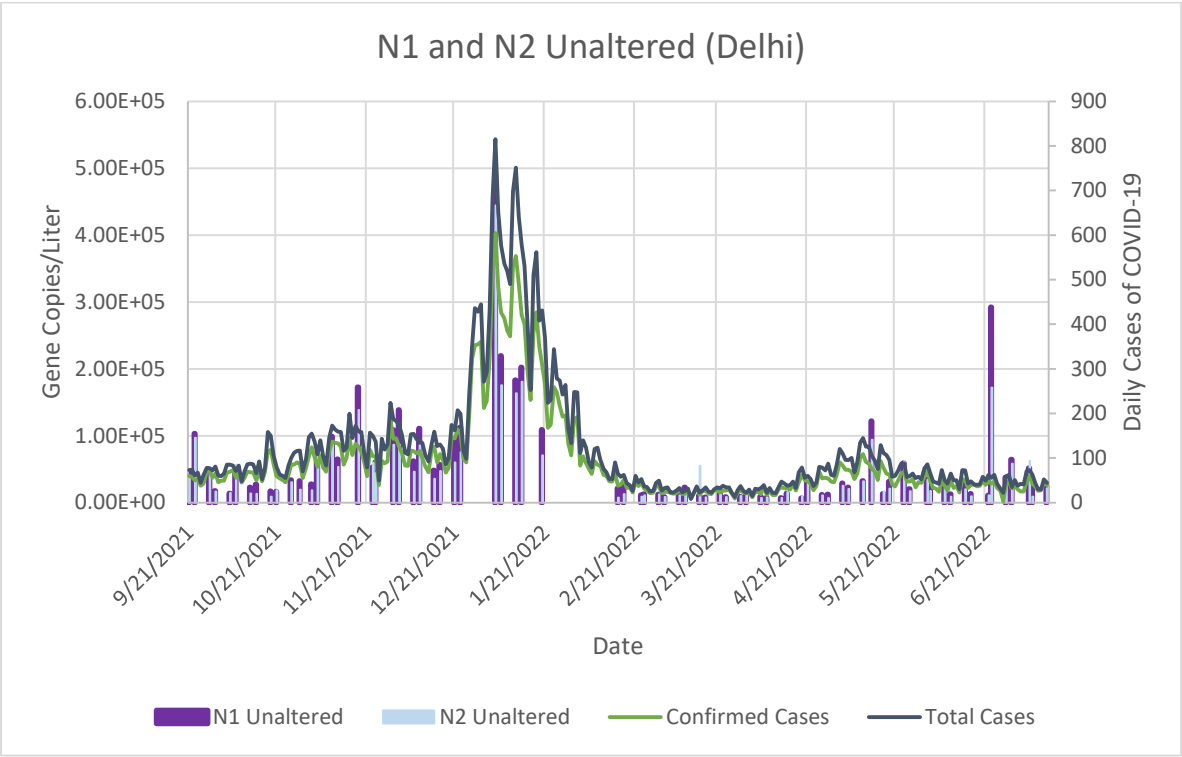
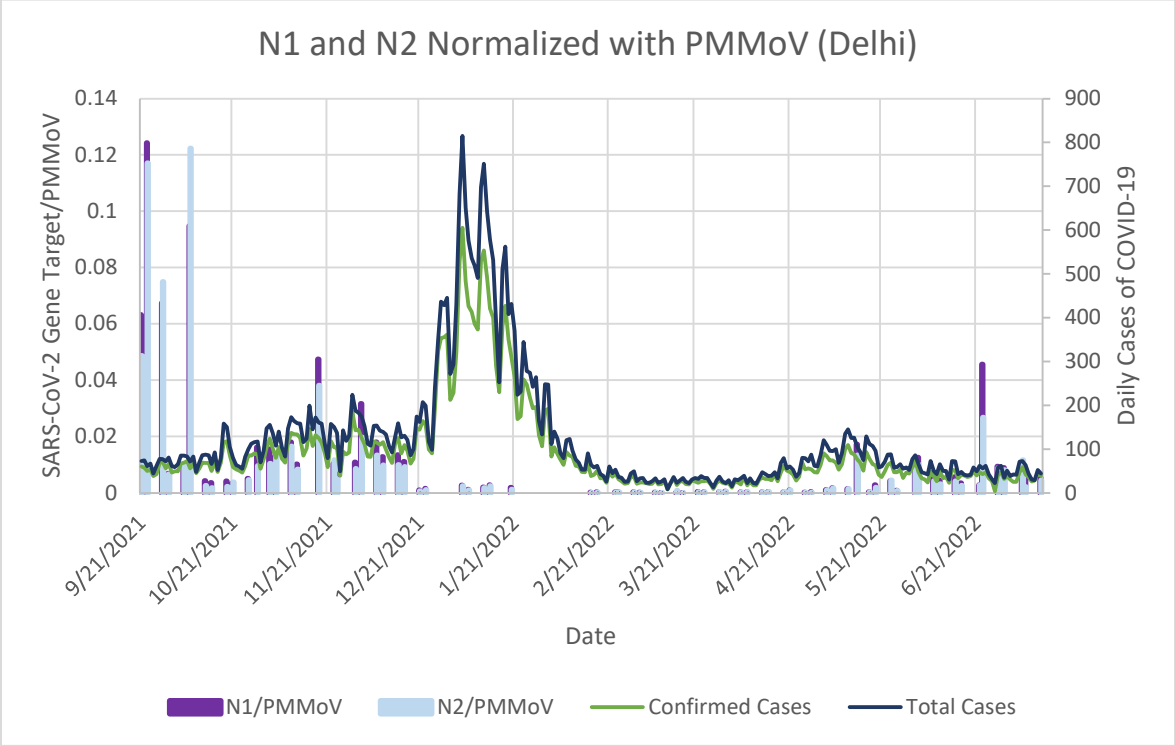


Figure 3.1: Gene Target Concentration vs. Clinical Cases of COVID-19 (Continued)



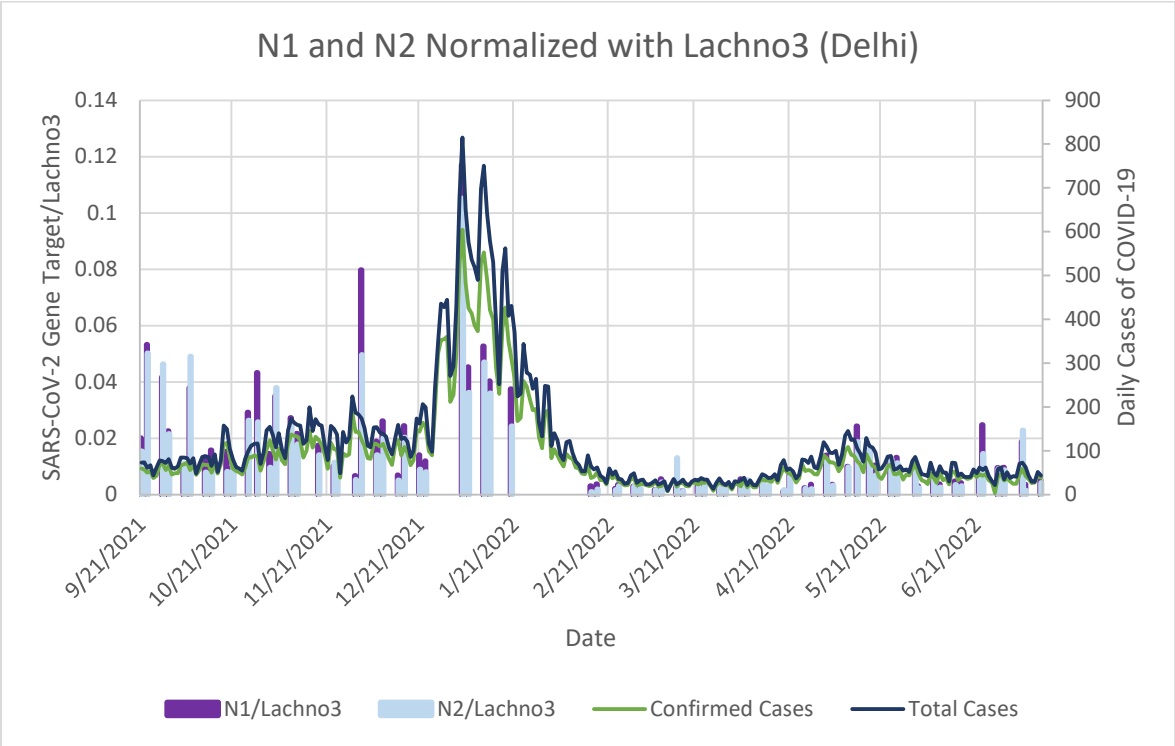
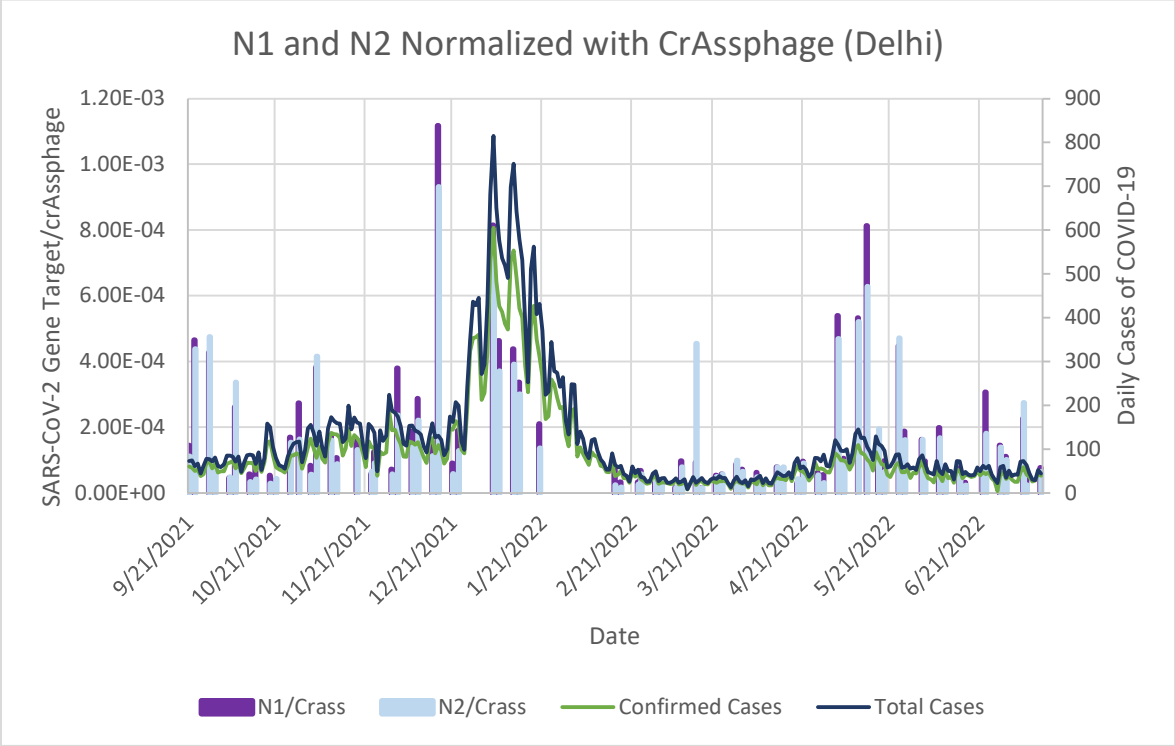


Figure 3.1: Gene Target Concentration vs. Clinical Cases of COVID-19 (Continued)



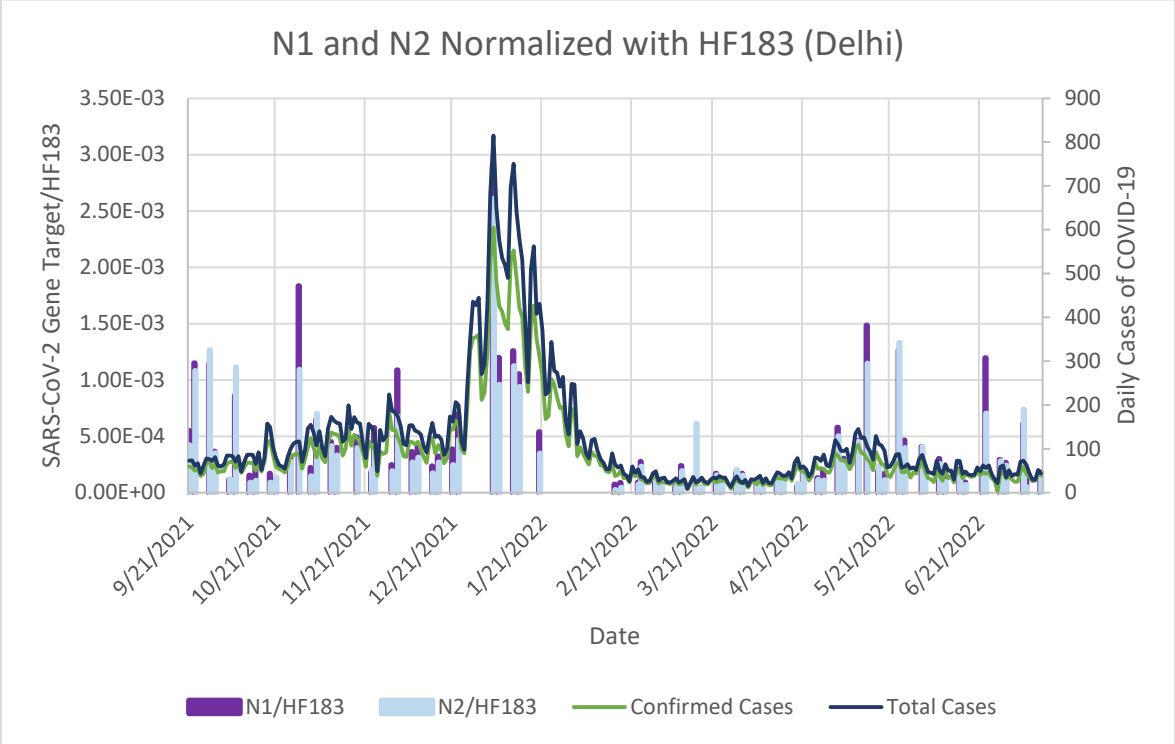
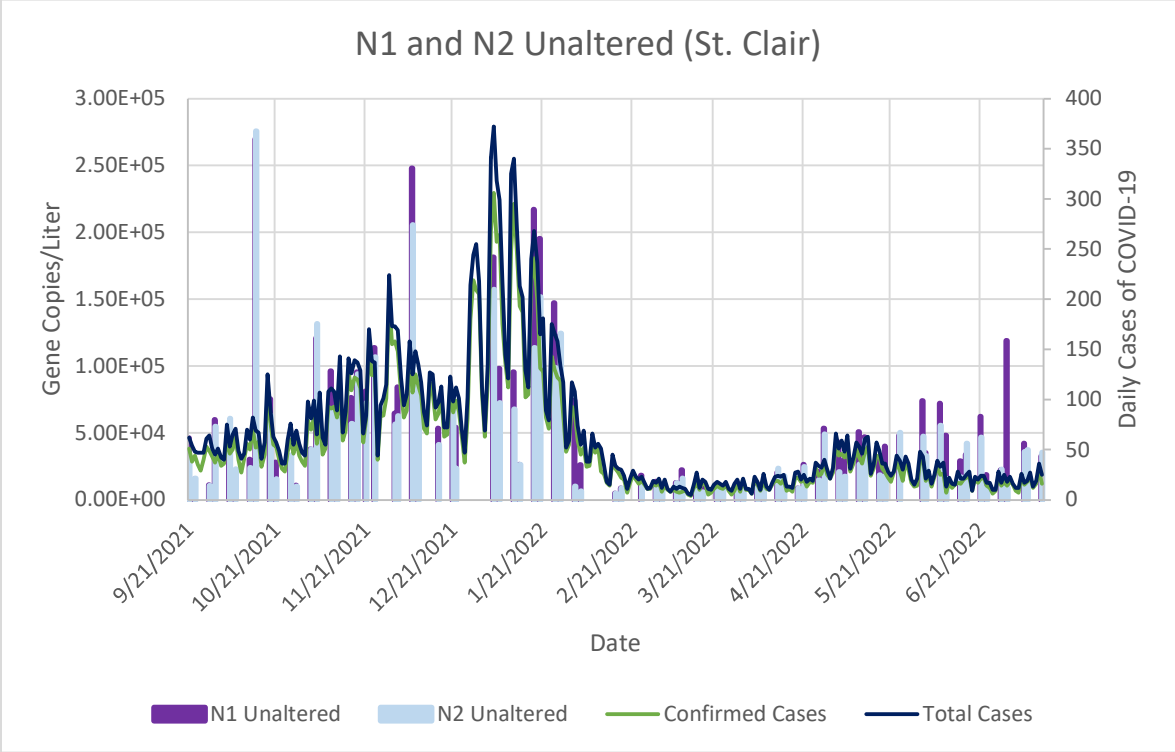


Figure 3.1: Gene Target Concentration vs. Clinical Cases of COVID-19 (Continued)



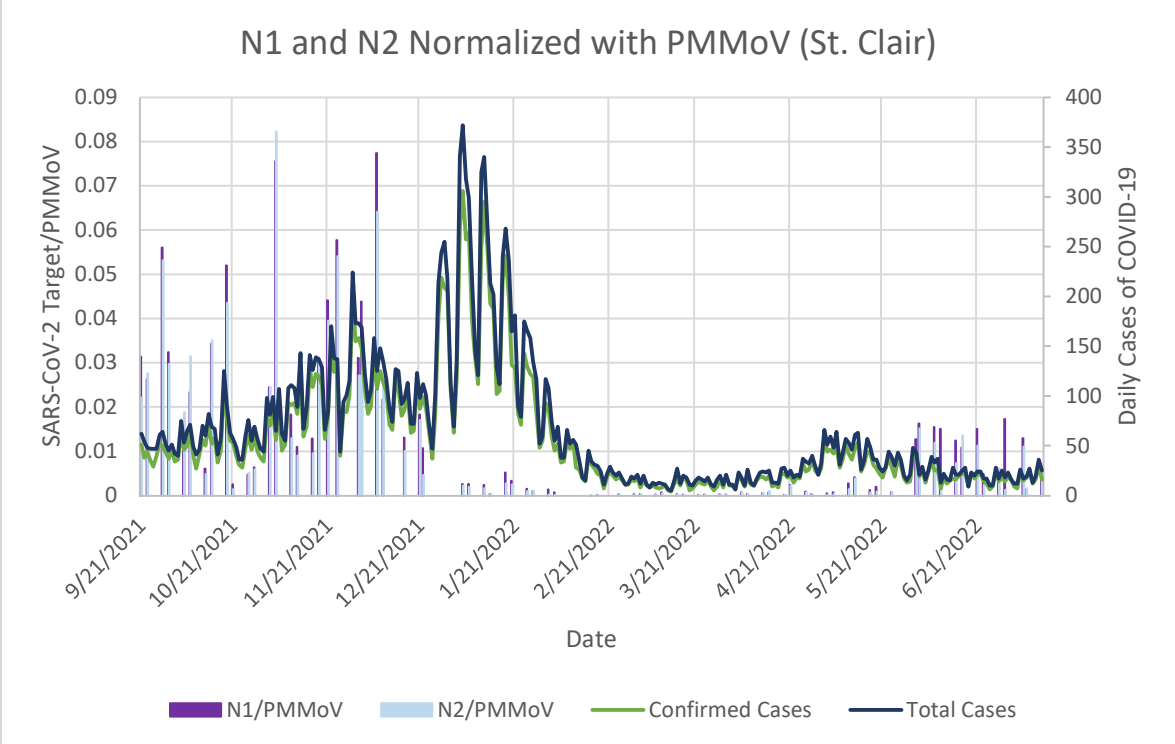
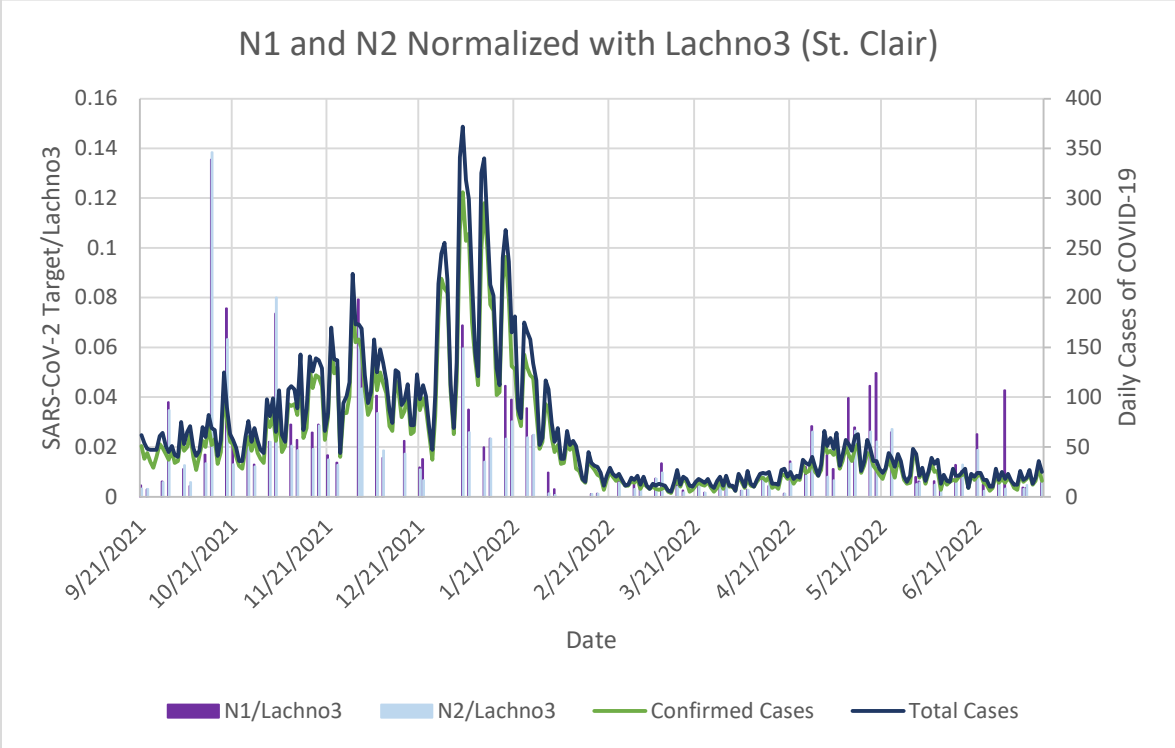


Figure 3.1: Gene Target Concentration vs. Clinical Cases of COVID-19 (Continued)



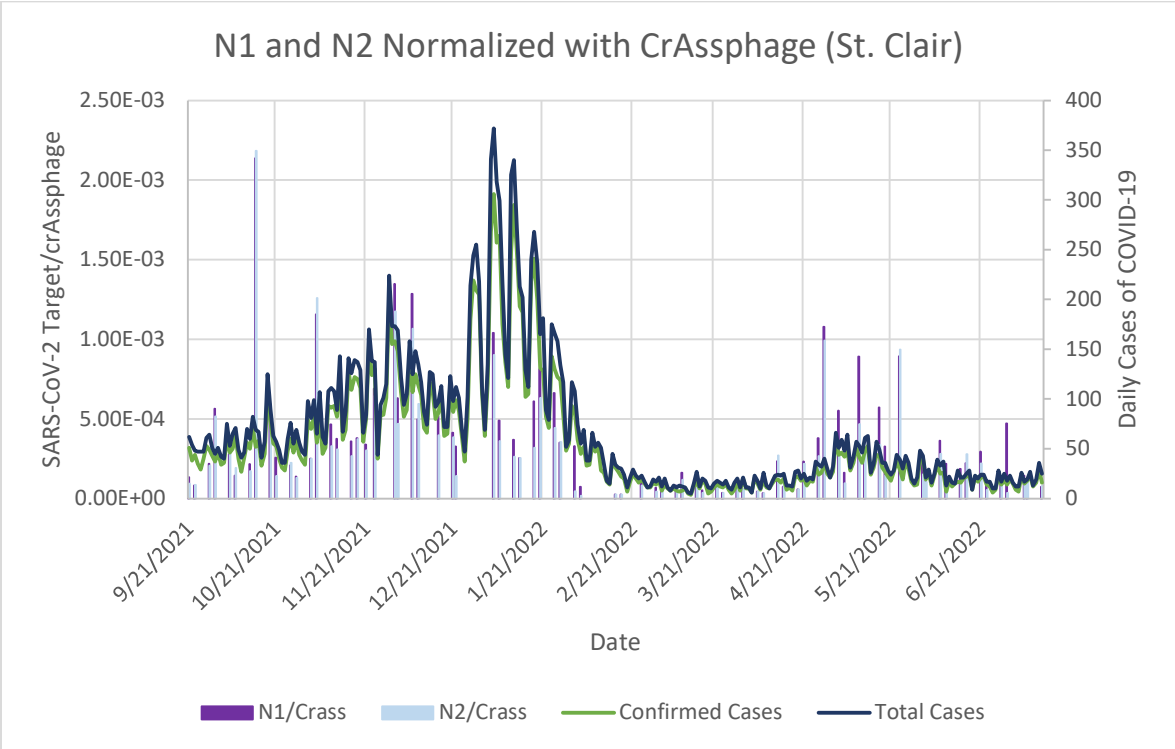


Figure 3.1: Gene Target Concentration vs. Clinical Cases of COVID-19 (Continued)

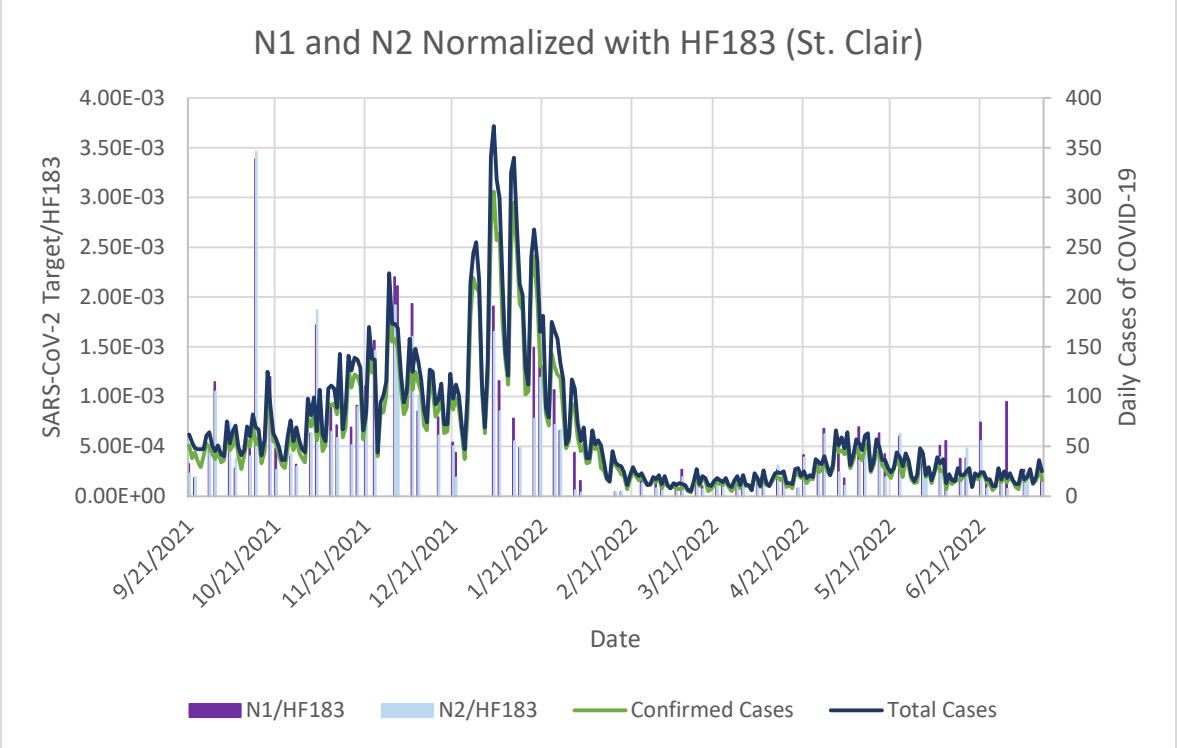
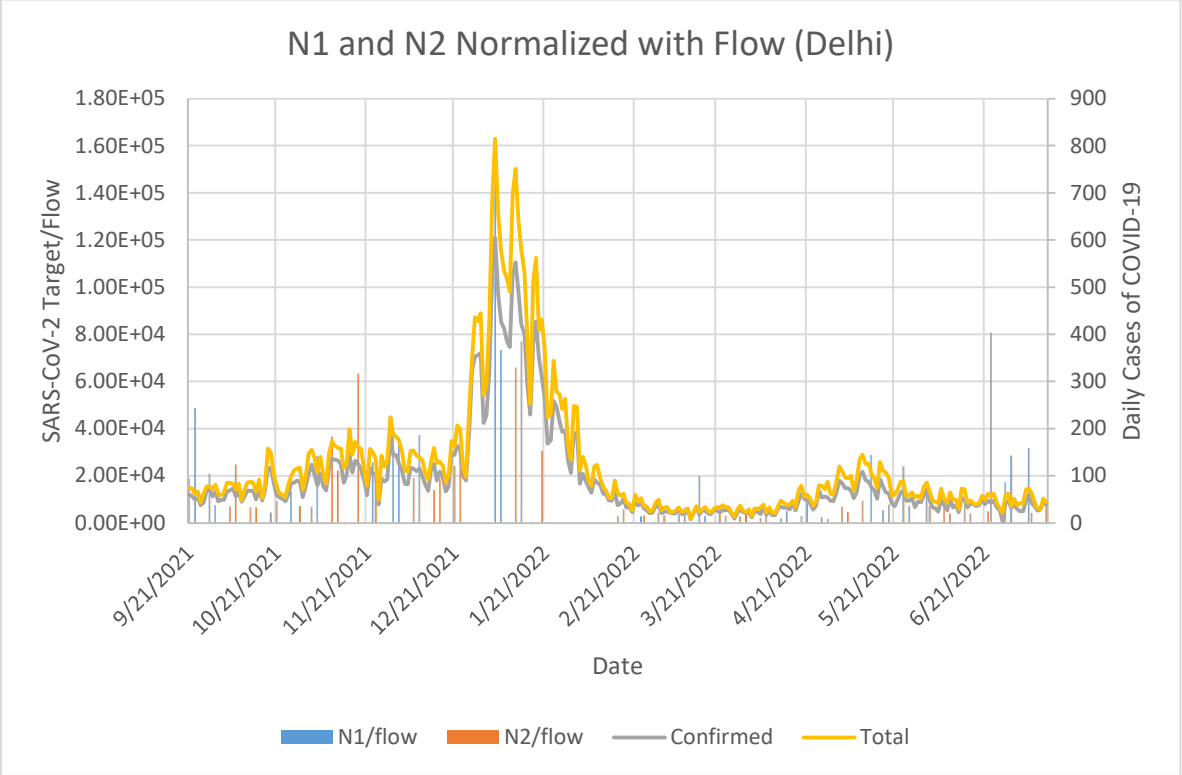


Figure 3.1: The concentrations of the N1 and N2 gene targets compared to the daily cases of COVID-19 found in the counties that provided the sewage. The N1 and N2 gene targets concentrations were compared to the data found on Michigan's government website. The concentrations of the N1 and N2 gene targets were left unaltered or normalized with PMMoV, Lachno3, crAssphage, or HF183 by dividing by the concentration of the gene targets by the concentration of the fecal indicators. These comparisons were made for all three sites. Total cases were defined as confirmed and probable cases of COVID-19.

To determine which data set best fit the clinical data, Pearson's correlation coefficients were calculated. The Pearson's correlation coefficients calculated were only based on the confirmed cases. Both normalized and unaltered data had a positive correlation with the confirmed cases of COVID-19 excluding the PMMoV normalized data for Delhi. The only RNA target used to normalize the signal was PMMoV while the remaining targets were DNA based. From a visual inspection of the graphs, PMMoV normalization did not improve correlation between the clinical data and N1/N2. DNA normalized data and the unaltered N1 and N2 gene targets had a similar look. The clinical data showed the largest peak in COVID-19 cases around January 21st, 2022. However, the PMMoV normalized data looked linear without any increase for the duration of the peak for all three sites. Therefore, the PMMoV normalized data did not mirror the increase seen in the clinical cases and suggested that COVID-19 was not prevalent in the community throughout the whole duration of the peak. When the gene targets were normalized with the DNA fecal indicators or unaltered, visual inspection confirmed a general correlation. However, it was not a perfect match as the sewage data had some high data points that were not in the clinical data. Overall, there was no correlation between the confirmed cases of COVID-19 and the PMMoV normalized gene targets. The Pearson's correlation coefficients for the N1 and N2 gene targets normalized by PMMoV for Delhi were as follows: r=–0.06 (N1/PMMoV) and r=–0.06 (N2/PMMoV). The Pearson's correlation coefficients between the clinical and sewage data for the remainder of the sites at Delhi were as follows: 0.78 (N1), 0.81 (N2), 0.72 (Lachno3/N1), 0.69 (Lachno3/N2), 0.47 (crAssphage/N1), 0.43 (crAssphage/N2), 0.64 (N1/HF183), and 0.62 (N2/HF183). The gene target concentrations with the best correlation to least correlation with the clinical

cases of COVID-19 for Delhi were as follows: unaltered>Lachno3>HF183>crAssphage>PMMoV. Warren and St. Clair had positive correlations for all normalized and unaltered data. For Warren the unaltered data had the highest correlation to the reported clinical cases of COVID-19. Pearson's correlation coefficients were as follows: 0.69 (N1) and 0.66 (N2). This shows a moderate positive correlation between the clinical data and the sewage data. Next, the lowest Pearson's correlation coefficients for Warren were again attributed to the PMMoV normalized data. The Pearson's correlation coefficients were 0.14 (N1/PMMoV) and 0.13 (N2/PMMoV) which indicates almost no correlation between the sewage and clinical data. The DNA fecal indicators used to normalize the N1 and N2 signal at Warren had a low to moderate correlation with the clinical data. Their Pearson's correlation coefficients were as follows: 0.47 (N1/Lachno3), 0.44 (N2/Lachno3), 0.55 (N1/crAssphage), 0.56 (N2/crAssphage), 0.68 (N1/HF183), and 0.65 (N2/HF183). The ranking of most to least correlation for Warren were as follows: unaltered>HF183>crAssphage>Lachno3>PMMoV. Lastly, St. Clair also had the highest correlation with the unaltered data (0.59 for N1 and 0.52 for N2) and the lowest correlation with the PMMoV normalized data (0.15 N1/PMMoV and 0.14 N2/PMMoV) which was consistent with all three sites in this study. The remaining Pearson's correlation coefficients at St. Clair were in between the PMMoV normalized and unaltered data. This site had the lowest correlation between the sewage data and the clinical data. The correlations between the gene target concentrations and the clinical cases were positive with weak to moderate correlations for the DNA fecal indicators.

Figure 3.2: Gene Targets Normalized with Flow



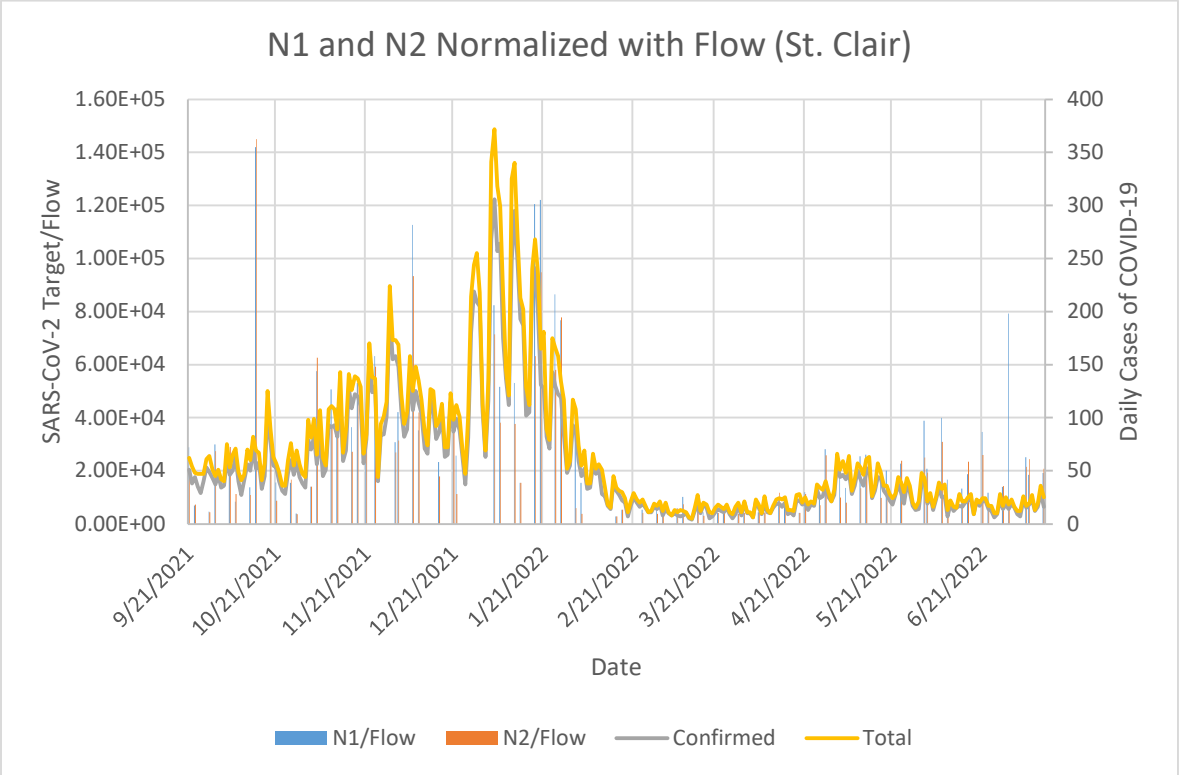


Figure 3.2: Gene Targets Normalized with Flow (Continued)

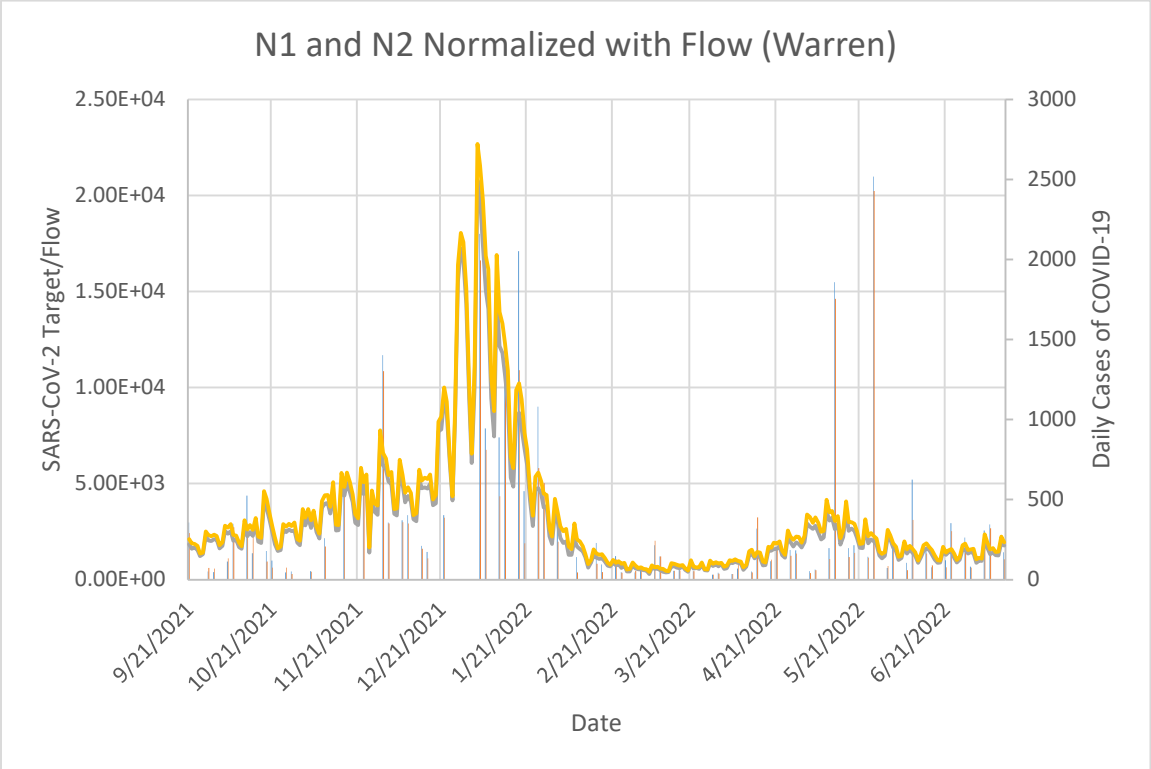
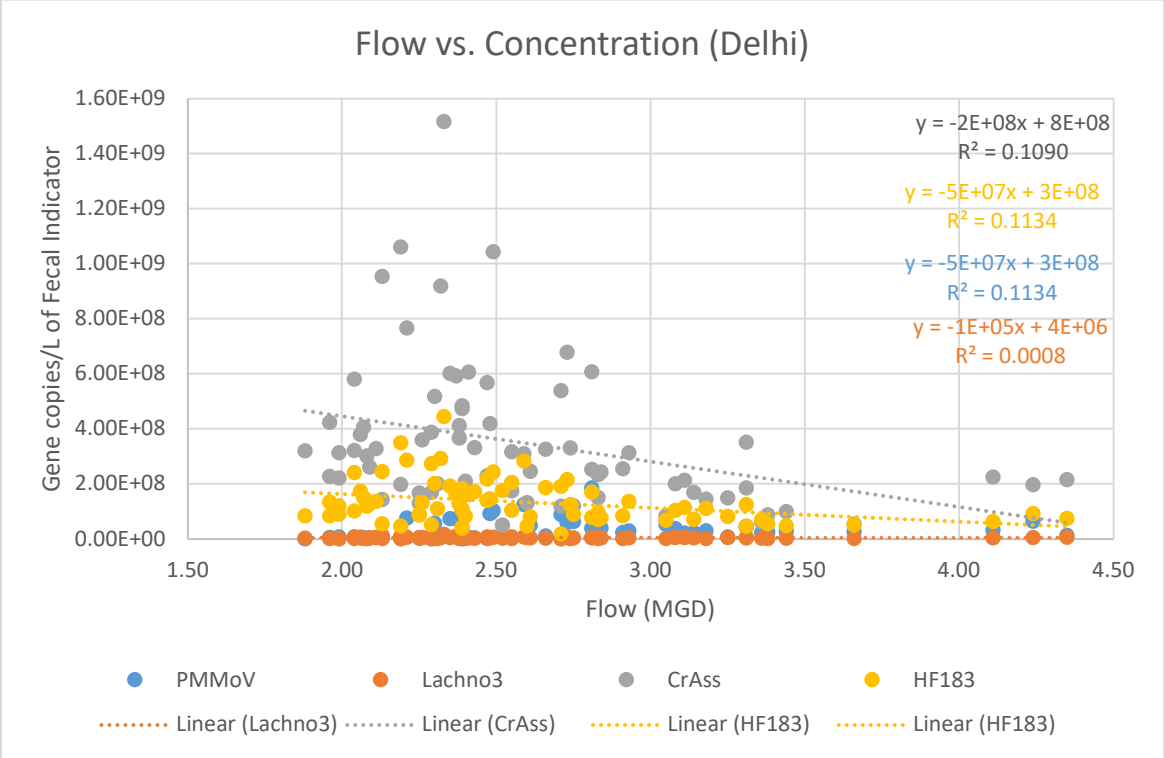


Figure 3.2: The N1 and N2 gene targets normalized by flow. These values are visually compared to the clinically reported cases of COVID-19.

The remaining Pearson's correlation coefficients were as follows: 0.42 (N1/Lachno3), 0.37 (N2/Lachno3), 0.40 (N1/crAssphage), 0.34 (N2/crAssphage), 0.56 (HF183/N1), and 0.47 (HF183/N2). The order from strongest to weakest correlation for St. Clair was as follows: unaltered>HF183>Lachno3, crAssphage>PMMoV.

Next, the SARS-CoV-2 targets were normalized by the flow rate to determine if there was a greater correlation between the clinical cases of COVID-19, and the sewage data when the gene targets were normalized by the flow rate. Figure 3.2 graphically shows the flow normalized SARS-CoV-2 gene targets when compared to the clinical cases of COVID-19. Visually, there is a similar trend between the flow normalized gene target data and the clinical cases of COVID-19 at all three sites. At Warren there were Pearson's correlation coefficients of 0.61 (N1/flow), 0.57 (N2/flow), 0.69 (N1), and 0.66 (N2) between the clinical and sewage data. For Delhi and St. Clair, the gene targets normalized by the flow rate had the best fit to the clinical data when comparing the correlation coefficients obtained from normalizing the signal. For Delhi, the Pearson's correlation coefficients were as follows: 0.75 (N1/flow), 0.79 (N2/flow), 0.78 (N1), and 0.81 (N2). For St. Clair, the Pearson's correlation coefficients were as follows: 0.57 (N1/flow), 0.51 (N2/flow), 0.59 (N1), and 0.52 (N2). Overall, the data show that the normalized data have a lower correlation to the clinical data than the unaltered data regardless of being normalized by flow or by fecal indicator. Therefore, these data suggest normalizing the data will not increase the gene target's correlation to the clinical cases of COVID-19.

Figure 3.3: Flow Rate Compared to the Concentration of the Fecal Indicators



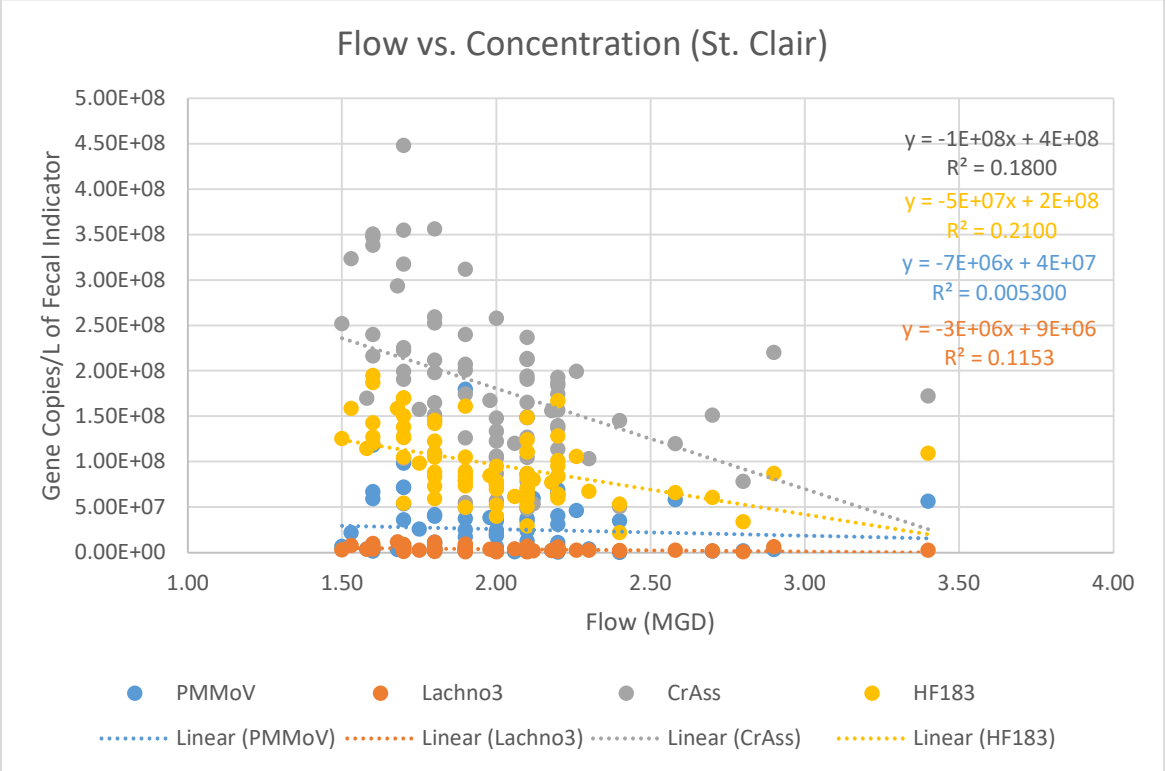


Figure 3.3: Flow Rate Compared to the Concentration of the Fecal Indicators (Continued)

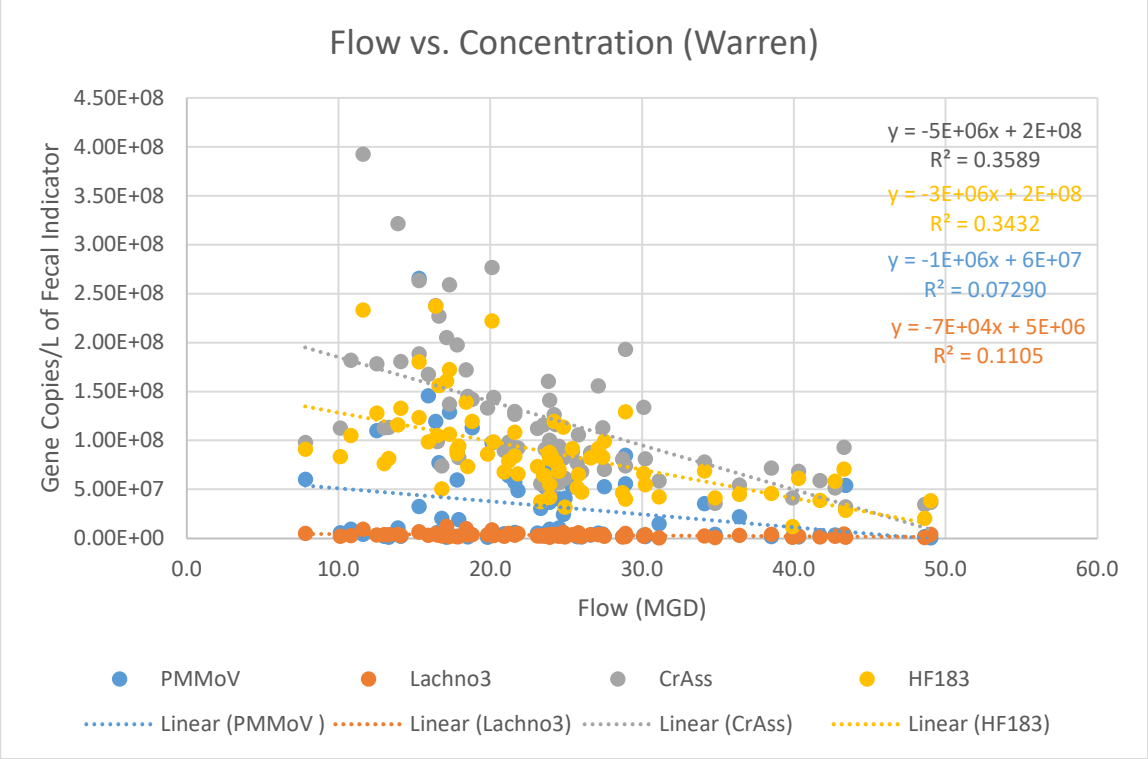


Figure 3.3: On a site-by-site basis the flow rate and fecal indicator concentrations were compared. Some of the x-axes are not set to zero for a better view of the data points.

Fecal indicators have been suggested to normalize the SARS-CoV-2 signal, but flow rate is typically used to normalize the virus's signal in sewage. To establish that fecal indicators can normalize the SARS-CoV-2 gene targets, the flow rate and the concentration of the fecal indicators were compared. It is expected that the trendline should have a downward trend if the fecal indicators are reliable for sample normalization. Figure 3.3 shows the comparison of the flow rate and the concentration of the fecal indicators.

Based on a visual inspection of the plots, HF183 and CrAssphage fecal indicators may have utility for SARS-CoV-2 normalization.

The fecal indicator concentration data was assessed to determine if there was variation on a site-by-site basis. For all three sites the order of highest to lowest concentration for the fecal indicators had the same pattern. The pattern of highest to lowest concentration were as follows for all three sites: crAssphage>HF183>PMMoV>Lachno3. The concentration of crAssphage was statistically higher than the other fecal indicators with p-values in the range of $p=4.6x10^{-24} – 5.3x10^{-11}$ (Warren), $p=4.8x10^{-18} – 3.3x10^{-14}$ (Delhi), and $p=5.1x10^{-31} – 2.3x10^{-12}$ (St. Clair). For all sites, HF183 had the largest p-value (shown), Lachno3 had the smallest p-value (shown), and PMMoV had a p-value between HF183 and Lachno3. Combing all the data points from all three sites, crAssphage compared against the other fecal indicators had the following p-values which shows that crAssphage had a statistically higher concentration than the others: $4.4x10^{-25}$ (HF183), $1.7x10^{-35}$ (PMMoV), $1.3x10^{-42}$ (Lachno3). In contrast, Lachno3 was the fecal indicator with the least presence in all three sites. The concentration of Lachno3 was statistically

lower than the other fecal indicators with p-values in the range of $p=3.6x10^{-26} – 2.3x10^{-7}$ (Warren), $p=1.4x10^{-23} – 7.2x10^{-7}$ (Delhi), $p=1.2x10^{-35} – 2.3x10^{-8}$ (St. Clair). HF183 had a statistically higher concentration than PMMoV for every site and when all sites were averaged. Overall, the average concentrations obtained per fecal indicator were within a half order of magnitude of each other. However, Delhi had the highest concentrations for the DNA fecal indicators, and Warren had the highest concentration of PMMoV. All indicators had less than a 3-fold difference in their averages based on site. The average concentrations for all samples (for all three sites) were as follows: $29.6x10^6$ CP/L (PMMoV), $3.7x10^6$ CP/L (Lachno3), $215.9x10^6$ CP/L (crAssphage), and $105.6x10^6$ CP/L (HF183).

The DNA fecal indicators had a closer resemblance to each other than PMMoV. The amount of correlation varied by site. Pearson's correlation values are listed in Table 3.3. Table 3.3 shows many combinations of fecal indicators and the corresponding correlation between them based on site.

Table 3.3: Correlation between Fecal Indicators

Site	CrAssphage vs. PMMoV	CrAssphage vs. Lachno3	CrAssphage vs. HF183	PMMoV vs. Lachno3	PMMoV vs. HF183	HF183 vs. Lachno3
1	0.30	0.59	0.85	0.18	0.48	0.66
2	0.14	0.72	0.86	0.16	0.13	0.70
3	-0.026	0.68	0.77	-0.080	0.22	0.69
Combined	0.11	0.58	0.83	0.090	0.24	0.66

Table 3.3: This table shows the Pearson's correlation coefficients obtained when the fecal indicators were compared to each other. It shows the correlation between fecal indicators on a site-by-site basis.

Every fecal indicator paired with PMMoV had the worst correlation compared to the other DNA pairings. Overall, PMMoV had no to low correlation with the DNA fecal indicators. In contrast, all of the DNA fecal indicators had a moderate to high amount of correlation with each other. CrAssphage and HF183 always had the highest correlation with each other compared to the other fecal indicator pairings. than any of the other fecal indicators.

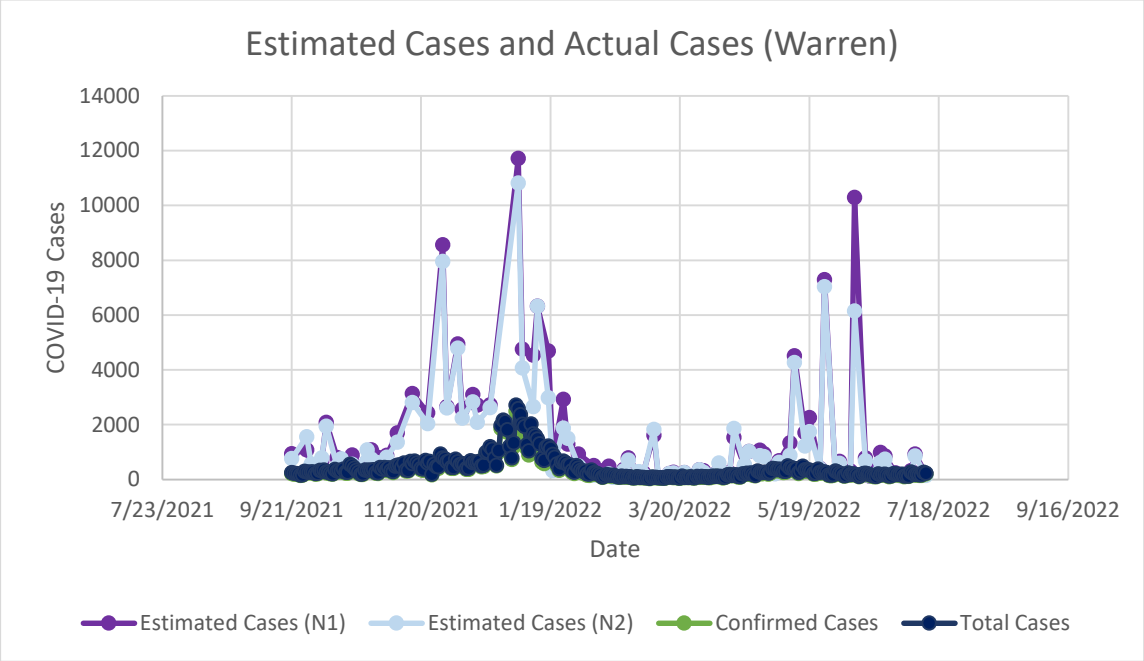
The graphs were visually viewed for peaks with each having three distinct peaks. These peaks were identified and compared between the clinical and sewage data to determine if either data set preceded the other. The date of the peak values was observed and compared to determine if the highest cases were found on the same day for both clinical and sewage samples. If one peak was identified first, that means that set of data suggested the maximum COVID-19 cases first and gave a warning of COVID-19 increase before the other. The data used for this analysis was the unaltered N1 and N2 gene target data. Three peaks were assessed. There were two small peaks and one large peak. One small peak between October 28th, 2021 and December 16th, 2021 (peak 1) and another small peak between April 14th, 2022 and May 24th, 2022 (peak 3). There was one large peak between December 10th, 2021 and March 2nd, 2022 (peak 2). For Warren the date of the peaks were as follows: November 29th, 2021 (clinical) and November 30th, 2021 (sewage) for peak 1, January 4th, 2022 (clinical and sewage) for peak 2, and May 12th, 2022 (sewage) and May 16th, 2022 (clinical) for peak 3. For Delhi the dates of the peaks were as follows: November 18th, 2021 (sewage) and November 29th, 2021 (clinical) for peak 1, January 4th, 2021 (clinical and sewage) for peak 2, and May 10th, 2021 (clinical) and May 13th, 2021 (sewage). For St. Clair the dates of the peaks were as follows: November 24th,

2021 (sewage) and November 29th, 2021 (clinical) for peak 1, January 4th, 2022 (clinical and sewage) for peak 2, and April 28th, 2022 (sewage) and May 2nd, 2022 (clinical) for peak 3. In general, the sewage data slightly preceded the clinical data. Lastly, the number of COVID-19 cases were estimated from the N1 and N2 gene target concentrations and compared to the clinical cases. Figure 3.4 shows the plots for each site with the estimated cases based of the N1 and N2 gene target concentrations and the confirmed and total (confirmed and probable) clinical cases of COVID-19.

Averaging the cases of COVID-19 for Warren, there were 1,748 estimated cases of COVID-19 based on the N1 gene target, 1,507 estimated cases of COVID-19 based on the N2 gene target, and 370 confirmed clinical cases of COVID-19. There was a moderate correlation between the cases of COVID-19 determined from the clinical and sewage data. The Pearson's correlation coefficients were 0.67 (N1 and confirmed cases) and 0.68 (N2 and confirmed cases). The values of estimated and clinical cases of COVID-19 were statistically different. For Warren there were p-values of 1.3×10^{-7} (N1) and 2.1×10^{-7} (N2) which determined that the gene target data estimated significantly more COVID-19 cases than the clinically confirmed cases. When the cases of COVID-19 were averaged, it was found that there were 4.7 times (4.2 times) more cases of COVID-19 cases estimated via N1 and 4.1 times (3.6 times) more cases of COVID-19 estimated via N2 than the confirmed (total – probable and confirmed cases) cases of COVID-19.

For Delhi the Pearson's correlation coefficients based on the confirmed cases were 0.79 for N1 ($p=4.4 \times 10^{-4}$) and 0.81 for N2 ($p=6.9 \times 10^{-4}$). The average number of people infected with COVID-19 was 149 (N1), 125 (N2), 92 (confirmed cases), and 122 (total cases).

Figure 3.4: Estimated Cases vs. Clinical Cases of COVID-19



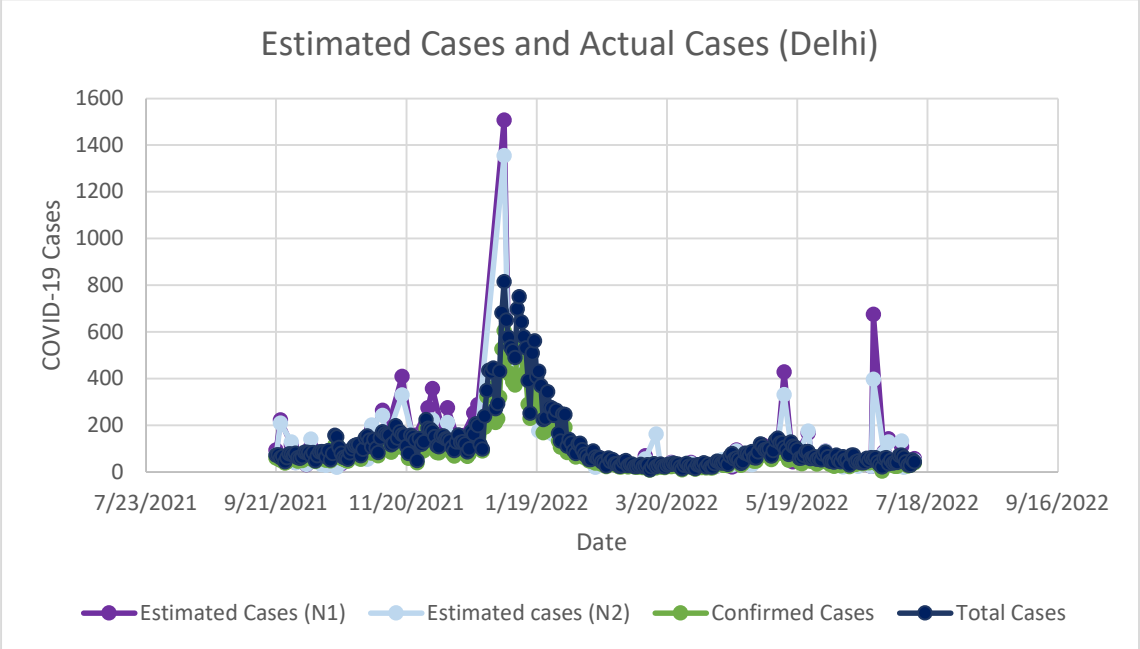


Figure 3.4: Estimated Cases vs. Clinical Cases of COVID-19 (Continued)

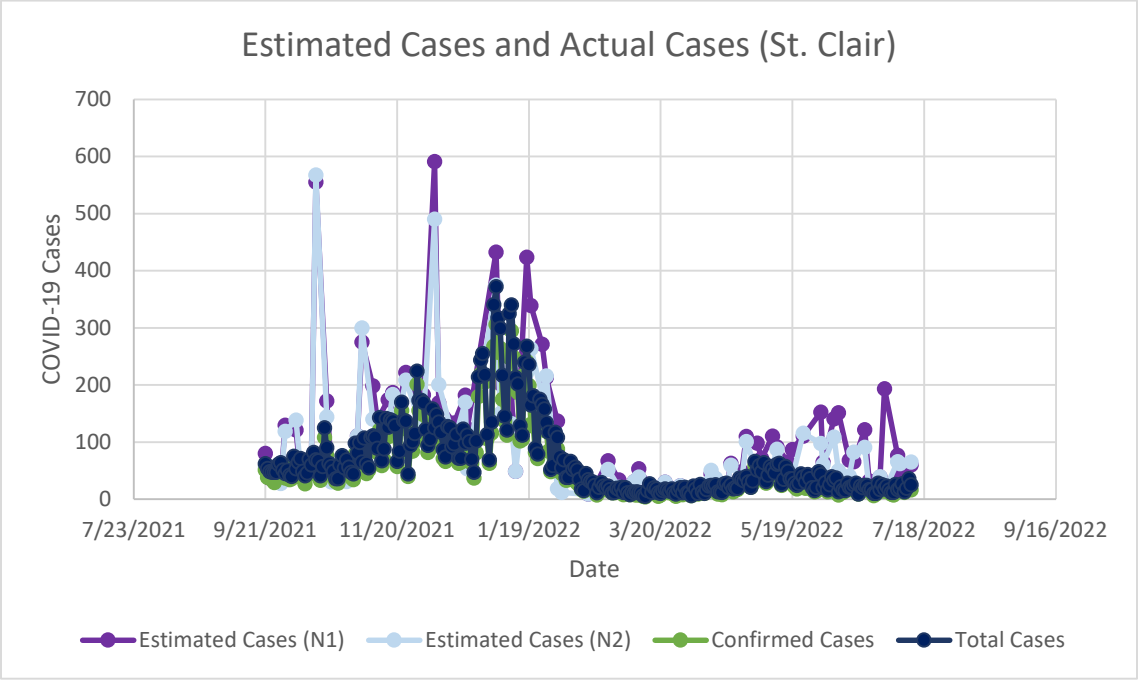


Figure 3.4: This figure shows the number of cases of COVID-19 estimated by the N1 and N2 gene targets tested in the sewage. It also shows the clinical data for the confirmed and total (confirmed and probable) cases of COVID-19. This data was plotted over a span of nearly ten months for all three sites.

For Delhi, the total COVID-19 cases compared to the N2 data showed no statistical difference. When the cases of COVID-19 were averaged, it was found that there were 1.6 times (1.2 times) more cases of COVID-19 cases estimated via N1 and 1.4 times (1.0 times) more cases of COVID-19 estimated via N2 than the confirmed (total) cases of COVID-19. For St. Clair the Pearson's correlation coefficients between the confirmed cases and the gene targets were 0.58 for N1 ($p=5.9x10^{-7}$) and 0.51 for N2 ($p=8.1x10^{-4}$). The average number of cases of COVID-19 were 119 (N1), 95 (N2), 63 (confirmed cases), and 73 (total cases). When the cases of COVID-19 were averaged, it was found that there were 1.9 times (1.6 times) more cases of COVID-19 cases estimated via N1 and 1.5 times (1.3 times) more cases of COVID-19 estimated via N2 than the confirmed (total) cases of COVID-19. Overall, the estimated COVID-19 cases had a moderate to high positive correlation with the clinical cases of COVID-19. The estimated cases using N1 were always statistically higher than the estimated cases of COVID-19 using the N2 gene target. The p-values for all three sites indicating that N1 had a higher number of estimated cases of COVID-19 than N2 were as follows: $4.4x10^{-4}$ (Warren), $4.0x10^{-5}$ (Delhi), and $1.3x10^{-6}$ (St. Clair).

3.5 Discussion

Within hours, the majority (almost 80%) of households in the United States transport fecal waste from houses to various WWTPs (Kirby, 2021). Wastewater analysis of SARS-CoV-2 can assess large communities for COVID-19 by analyzing groups of people from less than 2,000 or up to nearly three million people per sewage sample (Kirby, 2021). Clinical testing is useful because it directly identifies individuals with COVID-19. However, because feces from both asymptomatic and presymptomatic

individuals exhibit a SARS-CoV-2 signal, sewage surveillance can pick up on cases of COVID-19 which may not be reported by clinical testing (Hata & Honda, 2020). Individuals can take 5.1days (median value) to show signs of COVID-19, and asymptomatic cases make up a significant portion of the population with an estimated 18-31% of COVID-19 cases having no symptoms (Hata & Honda, 2020). This means that many COVID-19 cases may go unreported in a clinical setting giving sewage surveillance an edge over clinical testing. Sewage surveillance is not reliant on participation of individuals and is a non-invasive way to monitor large groups of people with a single sample. It is thought that testing wastewater can warn of an increase in rising COVID-19 cases in a specific population before clinical testing. One review found that the sewage data preceded clinical data by 0-16 days (Xiao et al., 2022). For example, in the city of Amersfoort in the Netherlands the sewage signal preceeded COVID-19 being clinically detected by six days (Medema et al., 2020). Another study found that the sewage data was five days ahead of clinical cases of COVID-19 (Galani et al., 2022). There was also a study that said wastewater data were eight days behind the onset of COVID-19 symptoms but two days ahead of clinical testing for COVID-19 (Nemudryi et al., 2020). These data were closer to our data which suggested on average that the sewage data were 2.2 days ahead of clinical data. However, these data had a range that showed that the sewage data peaks ranged by -3 to 11 days relative to the clinical peaks. On average, sewage data was ahead of clinical data by five days for peak 1, zero days for peak 2, and 1.7 days for peak 3. There are far less sewage data points than clinical points which contributes to the variation. The sewage data may have earlier peaks if more samples had been taken. At most, the sewage data had two points a week, but the clinical

data had daily results. Also, the sewage data did not fully represent the clinical data as the sewage collected did not represent the entire county. Therefore, the data for the county took into consideration a larger population than the sewage data. Trends should be similar, but there could be slight peak variation based on the samples not being completely from the same pool of people. Overall, many papers report differences in how early the sewage samples can precede the clinical cases with some studies suggesting a small advance such as two days while other papers suggest that it could precede clinical data by more than two weeks (Nemudryi et al., 2020; Randazzo et al., 2020).

Besides looking at peak data, our research looked at the general trend between the reported cases of COVID-19 and the sewage data. Both unaltered and altered gene target concentrations were used to compare the trend of the sewage data and the clinical cases of COVID-19 reported on a county-by-county basis. These data were altered by normalizing the SARS-CoV-2 signal with four different fecal indicators or by normalizing with flow. Normalization was used to counteract variations of fecal content within the sewage sample. Feces concentration in the sewage matrix fluctuates so various human fecal indicators were used to identify how the fecal content in the sewage changed over time. The concentrations of the SARS-CoV-2 gene targets were then divided by the concentration of the fecal indicators to account for fluctuations in SARS-CoV-2 gene target concentration due to amount of feces in the sewage. Hypothetically this would eliminate the possibility of underestimating the SARS-CoV-2 gene target concentration due to low fecal content. This study used four different fecal indicators. The human fecal indicators that were used in this study were all suggested by the centers for disease control and prevention (CDC) to normalize the SARS-CoV-2 gene target's concentration

(CDC, 2022b). The CDC lists PMMoV which is an RNA based fecal indicator and Lachno3, crAssphage, and HF183 which are DNA based fecal indicators (CDC, 2022b). The way normalized and clinical cases were compared was based on the paper by D'Aoust et al (D'Aoust et al., 2021). Our study found that normalization did not enhance the correlation at any site between the clinical cases of COVID-19 and the SARS-CoV-2 gene target's concentrations. The unaltered data had the best correlation to the clinical data. The D'Aoust et al. study only used PMMoV to normalize their data. Their study provided the correlation between the PMMoV normalized SARS-CoV-2 gene targets and the clinical cases of COVID-19. When compared to clinically detected cases of COVID-19, their PMMoV normalized data had Pearson's correlation coefficients of 0.67 (N1/PMMoV) and 0.65 (N2/PMMoV) (D'Aoust et al., 2021). For the three sites in this study, we saw Pearson's correlation coefficients as follows for PMMoV: 0.14 (N1/PMMoV) and 0.13 (N2/PMMoV) for Warren, -0.06 (N1/PMMoV) and -0.06 (N2/PMMoV) for Delhi, and 0.15 (N1/PMMoV) and 0.14 (N2/PMMoV) for St. Clair. Our PMMoV normalized data essentially showed there was no correlation between the clinical cases of COVID-19 and the sewage data. However, the data by D'Aoust et al. shows a moderate correlation to the clinical data with PMMoV normalized gene targets. A study by Feng et al. was somewhat in agreement with our study suggesting that normalization may not be very beneficial. Their study utilized both HF183 and PMMoV to normalize the SARS-CoV-2 signal. They found that 5/12 WWTPs (HF183) and 8/12 (PMMoV) WWTPs had reduced correlation to the clinical data in comparison to the unaltered data (Feng et al., 2021). They stated that normalization decreased the correlation in many WWTPs and that this could suggest that the variability in

measurements of the SARS-CoV-2 gene targets and the fecal indicators may produce a higher amount of variability than fecal variation in the sample (Feng et al., 2021). They concluded that WWTPs may not need to be normalized with fecal indicators (Feng et al., 2021). Instead they suggested that fecal indicators may be more suitable in cases where there is known drastic changes in the fecal content (Feng et al., 2021). Also, PMMoV may not be the best fecal indicator to normalize the sample in a small population because PMMoV concentration in the sewage is based on a diet of eating foods such as tomatoes and peppers (D'Aoust et al., 2021). There may not be enough individuals eating these foods leading to less PMMoV signal in the sewage (D'Aoust et al., 2021). Therefore, they suggest PMMoV may be able to normalize the signal better for WWTPs pool sewage over large area (D'Aoust et al., 2021). Another study used HF183 to normalize the N1 gene target's signal along with bovine coronavirus (BCoV). They used Spearman rank correlation coefficients to compare correlation between the average cases of COVID-19 on a weekly basis to the sewage data. Between the gene target data and the clinical cases of COVID-19 they reported Spearman rank correlation coefficients of 0.75 (unaltered), 0.72 (BCoV normalized), and 0.39 (HF183 normalized) for Salitrillo WWTP and 0.68 (unaltered), 0.61 (BCoV normalized), and 0.16 (HF183 normalized) for Martinez II WWTP (Vadde et al., 2022). The unaltered data had the best correlation with the clinical cases of COVID-19, while BCoV normalization had a slightly inferior correlation to the clinical data than the unaltered data (Vadde et al., 2022). However, both the unaltered and BCoV normalized data had a moderate to strong correlation with the clinical data while the HF183 normalized data had a very weak to weak correlation with the clinical data (Vadde et al., 2022). This contrasts with our data which suggested that

HF183 had a higher degree of correlation to the reported cases of COVID-19. The Pearson's correlation coefficients for HF183 and the unaltered data compared to the clinical cases of COVID-19 were as follows in our study: 0.68 (N1/HF183), 0.65 (N2/HF183), 0.69 (unaltered N1), 0.66 (unaltered N2) for Warren, 0.64 (N1/HF183), 0.62 (N2/HF183), 0.78 (N1 unaltered) and 0.81 (N2 unaltered) for Delhi, and 0.56 (N1/HF183), 0.47 (N2/HF183), 0.59 (N1 unaltered), and 0.52 (N2 unaltered) for St. Clair. Therefore, our study indicates that HF183 generally has a moderate correlation to the clinical data and is inferior but somewhat similar to the correlation seen between the unaltered SARS-CoV-2 data and the clinical data. A study by Greenwald et al. used crAssphage, human 18S ribosomal RNA, PMMoV, and *Bacteroides* rRNA to normalize the SARS-CoV-2 signal. The 18S rRNA was later excluded as a fecal indicator for sample normalization due to the variability of this RNA in sample concentrations, it is a common lab contaminant, and it degrades very fast (Greenwald et al., 2021). The Greenwald et al. study found that the unaltered SARS-CoV-2 gene target data were superior to the normalized data. To test for correlation of clinical COVID-19 cases and the SARS-CoV-2 gene target N1's concentration, they used Kendall's Tau-b values and obtained the following Kendall's Tau-b values: 0.43 (unaltered), 0.38 (crAssphage), 0.35 (*Bacteroides* rRNA), and 0.18 (PMMoV) (Greenwald et al., 2021). A strong correlation was found between the clinical and sewage data for the unaltered data, crAssphage, and *Bacteroides* rRNA, but a weak correlation was found between the clinical and sewage data when PMMoV was used to normalize the signal (Greenwald et al., 2021). The Greenwald study again agreed with our study by showing that normalization with various fecal markers does not produce a correlation to the clinical data that is superior to the

unaltered data. However, their work did show that crAssphage was able to provide a strong correlation to the clinical data, but it was not able to provide as high as a correlation as the unaltered data. Our data suggested that crAssphage was never the best DNA fecal indicator to normalize the data. Overall, crAssphage had a weak to moderate correlation to the clinical cases of COVID-19 with Pearson's correlation coefficients as follows: 0.55 (N1/crAssphage) and 0.56 (N2/crAssphage) for Warren, 0.47 (N1/crAssphage) and 0.43 (N2/crAssphage) for Delhi, and 0.40 (N1/crAssphage) and 0.34 (N2/crAssphage) for St. Clair. No studies using Lachno3 to normalize their data could be found. Our data suggests a highly variable correlation between clinical cases of COVID-19 and Lachno3 normalized SARS-CoV-2 data. Our data showed that Lachno3 normalized data when compared to the clinical data had Pearson's correlation coefficients as follows: 0.47 (N1/Lachno3) and 0.44 (N2/Lachno3) for Warren, 0.72 (N1/Lachno3) and 0.69 (N2/Lachno3) for Delhi, and 0.42 (N1/Lachno3) and 0.37 (N2/Lachno3) for St. Clair. The unaltered data had the best correlation to the reported cases of COVID-19, and PMMoV normalization had extremely poor correlation. There was no consistency as to which which fecal indicator provided the best normalized data when compared to the clinical data. The fecal indicators that provided the best normalized signal to the clinical data were as follows: HF183 (Warren), Lachno3 (Delhi), and HF183 (St. Clair).

One study found that normalizing the SARS-CoV-2 gene targets by flow did enhance the correlation to the clinical cases of COVID-19 but only slightly. Their Spearman's rank correlation coefficient increased from 0.73 to 0.74 when normalized by flow (Feng et al., 2021). Normalization by PMMoV, HF183, and BCoV recovery reduced the correlation coefficients to 0.56, 0.66, and 0.59, respectively (Feng et al., 2021). Another study found

that only 3/12 WWTPs improved when normalized with flow (Maal-Bared et al., 2023). Our data did not improve upon normalization with flow. For the normalized data, HF183 provided the best correlation between the sewage and clinical data for Warren, and flow provided the best correlation between the sewage and clinical data for Delhi and St. Clair. However, the unaltered data always provided the best correlation to the clinically reported cases of COVID-19.

Figure 3.3 suggests that the fecal indicators could be used in place of flow which is less reliant on human behavior. The downward trend in the graph indicated that as flow increased the concentration of feces apparently decreased as expected; there are a relatively fixed number of individuals in a sewer shed or building. It is likely that the increased flow indicates that more water was introduced into the system via inflow and infiltration or some commercial process. Because fecal mass of the sewage remains the same while inflow and infiltration can cause an abnormally high flow rate, the downward trend in fecal concentration was expected. Because all the fecal indicators have a negative slope, this suggests that the fecal indicators are accounting for inflow and infiltration, albeit with a high variability. Although this implies that they could be used for normalizing the SARS-CoV-2 gene targets, our study with a fairly narrow flow range did not benefit from fecal indicator normalization. Our data showed that neither fecal normalization nor flow normalization were able to correlate with the clinically reported cases as well as the unaltered gene targets. Generally, normalizing the SARS-CoV-2 signal with the flow rate is thought to be the best way to normalize the SARS-CoV-2 signal. A higher flow indicates that the sample is more diluted, and a lower flow indicates that the sample is less diluted. Using feces is dictated by human behaviors more than the

flow rate. When there is a low amount of feces present in the sewage, it is uncertain if there are less feces due to less excretion of feces or due to individuals using bathrooms in other locations. Indicators such as PMMoV are diet based which means that this indicator may not be present if the individual does not eat particular foods. Also, sewage from restaurants that use peppers may contribute to more PMMoV being in the sewage which would skew the amount of PMMoV detected. Lastly, although these fecal indicators are used for detection of human feces, it is feasible that some of the signal may be due to animal feces. Flow rate is not subject to any of these biases. However, flow rate is not easy to obtain. Specific instrumentation is required to collect flow data which is expensive and difficult to maintain leading to many sites not having flow data. Because flow data is not always available, fecal indicators were assessed to determine if any fecal indicator could be used in place of flow rate. Overall, HF183 and flow normalized data had the best correlation to the clinical data, but neither were superior to the unaltered data.

The fecal indicators were compared to determine if one fecal indicator was more sensitive than the others, or if one fecal indicator was always detected in much lower quantities than the others. Our study suggested that crAssphage always had a statistically higher concentration than the other fecal indictors, and Lachno3 always had a statistically lower concentration than the other fecal indicators regardless of site. The trend for every site from highest concentration to lowest concentration was as follows: crAssphage>HF183>PMMoV>Lachno3. One study showed that the concentrations of both crAssphage and HF183 were very similar in sewage samples when run as both singleplex and duplex reactions (Ahmed, Payyappat, et al., 2019). The concentrations

they reported were 9.39 ± 0.10 log10 CP/L (HF183 singleplex) 9.39 ± 0.10 log10 CP/L (crAssphage singleplex), 9.44 ± 0.08 log10 CP/L (HF183 duplex), and 9.47 ± 0.06 log10 CP/L (crAssphage duplex) (Ahmed, Payyappat, et al., 2019). Therefore, their data suggested that those two targets had relatively the same concentration, but our study suggested crAsspahge had the highest concentration. Averaging the values at all sites, our data showed that crAssphage had an average concentration of $2.2x10^{8}$ CP/L and HF183 had an average concentration of $1.1x10^{8}$ CP/L with a p-value of $4.4x10^{-25}$ and a Pearson's correlation coefficient of 0.83. In another study, they tested for crAssphage, HF183, and Lachno3 in feces (Ahmed, Gyawali, et al., 2019). They found that the concentration of Lachno3 in feces was in the range of 4.59-9.37 log10 CP/g of human feces with an average of 7.33 log10 CP/g of human feces (Ahmed, Gyawali, et al., 2019). The concentration of crAssphage in feces was in the range of 6.30-8.92 log10 CP/g of human feces with an average of 7.35 log10 CP/g of human feces (Ahmed, Gyawali, et al., 2019). The concentration of HF183 in feces was in the range of 6.69-9.37 log10 CP/g of human feces with an average of 7.89 log10 CP/g of human feces (Ahmed, Gyawali, et al., 2019). Therefore, their work showed that HF183 had the highest average concentration in human feces while Lachno3 had the lowest average concentration in human feces. Another study focused on the Hawkesbury-Nepean River System to determine if there was potentially untreated sewage that flowed into it (Codello et al.). They checked for Lachno3 and HF183 in 27 different rivers (Codello et al.). A total of 9/27 rivers had detections for HF183 and 2/27 had detections for Lachno3 (Codello et al.). The concentration of HF183 was 3.6 times higher than Lachno3 in site industrial 03 and 4.6 times higher than Lachno3 in site eastern 01 (Codello et al.). In contrast, another study

found that Lachno3 (8.90 ± 0.17 log10 CP/L) had a higher concentration than both crAssphage (8.20 ± 0.24 log10 CP/L) and HF183 (8.50 ± 0.30 log10 CP/L) in untreated sewage samples (Ahmed, Gyawali, Hamilton, et al.). Overall, the other studies seem to favor HF183 over crAssphage and Lachno3. In our study Lachno3 was highly inferior to HF183 and crAssphage having the lowest concentration amongst all fecal indicators. Averaging the data from all sites, Lachno3 had a statistically lower concentration than PMMoV ($p=8.9x10^{-19}$), crAssphage ($p=1.3x10^{-42}$), and HF183 ($p=1.1x10^{-70}$). However, even though it was statistically the lowest, Lachno3 was more correlated to HF183 (0.66) and crAssphage (0.58) than PMMoV (0.09) which had the second lowest concentration. Several studies agree that crAssphage is found at higher concentrations than PMMoV. Holm et al. reported concentration ranges of 7.23×10^3 - 3.53×10^7 CP/mL (PMMoV) and 9.69×10^3 - 1.85×10^8 copies/ml (crAssphage) in sewage (Holm et al., 2022). While Nagarkar et al. also reported about a 10-fold difference in the upper limit when comparing CrAssphage (10^6 to 10^9 CP/L) and PMMoV (10^6 to 10^8 copies/L) concentration in the sewage (Nagarkar et al., 2022). They mentioned that HF183 had a concentration within 10^6 to 10^9 CP/L which was similar to crAssphage (Nagarkar et al., 2022). Lastly, crAssphage (5.95 ± 0.13 log10 CP/mL) again had a higher concentration than PMMoV (5.08 ± 0.34 log10 CP/mL) when analyzing composite samples of sewage (Ahmed, Bivins, et al., 2021). Therefore, these studies all agree that crAssphage was found in higher quantities than PMMoV which is comparable to our study. In our study crAssphage had a concentration of $2.2x10^8$ CP/L while PMMoV only had a concentration of $3.0x10^7$ CP/L when the data from all sites was averaged.

Lastly, COVID-19 cases were estimated based on the N1 and N2 SARS-CoV-2 gene targets. Another study also estimated the cases of COVID-19 with the formula we used. However, they used 126g of feces/(person*day) instead of 349g of feces/(person*day) and $10^{8.9}$ CP/g of feces instead of 10^{7} CP/g of feces (Li et al., 2022). They assessed three different water reclamation facilities and estimated their COVID-19 cases (Li et al., 2022). The actual cases and estimated cases (listed in parenthesis) were as follows: facility A had 44,996 (94,782 ± 7,583) cases of COVID-19, facility B had 2,645 (5,613 ± 449) cases of COVID-19, and facility C had 7,131 (7,987 ± 638) cases of COVID-19 (Li et al., 2022). Therefore, facility A and B estimated 2.1 times more cases of COVID-19, and facility C estimated 1.1 times more cases of COVID-19. Another study suggested that their model estimated that the amount of unreported cases were around 11-fold higher than the confirmed cases of COVID-19 (McMahan et al., 2021). A study by Kaplan et al. estimated with their sewage data that there was nearly three times the cases of COVID-19 in New Haven, Connecticut (United States) when compared to the reported cases (Kaplan et al., 2022, p. 2). They reported that they projected that 33.6% of the New Haven population had COVID-19 as of May 22nd, 2021 while clinically diagnosed cases of COVID-19 only suggested that 12% of the New Haven population had COVID-19 as of May 22nd, 2021 (Kaplan et al., 2022). Overall, it seems likely that the clinical data is underreporting cases of COVID-19 due to lack of participation, asymptomatic cases, and not reporting at home COVID-19 testing. One study estimated that it is expected that 253-409 cases of COVID-19 per 10,000 people is required for the wastewater to have a positive hit for COVID-19 (Hong et al., 2021). There were many data points in this study that had signal for the N1 and N2 gene targets despite having far fewer than 253-409

cases of COVID-19 reported clinically per 10,000 people. In fact, many data points in the three counties had less than 253 cases of COVID-19 in a population that was over 150,000 and still had a detection for the SARS-CoV-2 gene targets. Also, the sewage collected did not represent the whole county meaning that even less people would have COVID-19 than reported for the county. This suggests that there are many COVID-19 cases not reported. Interestingly, the Warren WWTP only collected sewage for approximately 20% of the county yet the sewage surveillance overestimated COVID-19 cases by over 3-fold. This suggests there was a high number of under reported cases of COVID-19 in Warren. Overall, Warren estimated 4.7 times (N1) and 4.1 times (N2) more cases of COVID-19 when compared to the confirmed cases, and 4.2 times (N1) and 3.6 times (N2) more cases of COVID-19 when compared to the total (confirmed and probable) cases of COVID-19. For Delhi, COVID-19 cases were estimated to be 1.6 times (N1) and 1.4 times (N2) higher than the confirmed cases of COVID-19 and 1.2 times (N1) and 1.0 times (N2) higher than the probable cases of COVID-19. Lastly for St. Clair, COVID-19 cases were estimated to be 1.9 times (N1) and 1.5 times (N2) higher than the confirmed cases of COVID-19 and 1.6 times (N1) and 1.3 times (N2) higher than the probable cases of COVID-19. One study looked at the results for 241 studies with 130,123 COVID-19 infected individuals and found that there were 31,411 asymptomatic individuals (Chen et al., 2021). There was 21.7% asymptomatic infections found after they excluded presymptomatic cases (Chen et al., 2021). In the end, they concluded that approximately one-fifth of COVID-19 cases are likely asymptomatic (Chen et al., 2021). Therefore, asymptomatic cases make up many cases of COVID-19 and will likely go unreported. With the asymptomatic cases not being reported, the lack

of individuals getting tested, and the underreporting of positive cases detected by at home tests, it is likely that a lot of cases of COVID-19 are going unreported. Therefore, the cases estimated with sewage surveillance may be a more accurate representation of the actual cases.

3.6 Conclusion

This research agrees with other studies that sewage surveillance is able to precede clinical cases of COVID-19. Our study suggested that sewage data preceded clinical data by 2.2 days on average, and for the nine peaks compared, sewage surveillance was found to precede clinical data by -3 to 11 days. Because the clinical data had daily data points and the sewage data had at most two data points a week, there may have been a built-in time shift. Shortening the lead. Overall, there is high variation throughout studies on how much of an early warning sewage data provides over clinical data. The study by Xiao et al. said that based on various studies sewage data was ahead of clinical data by 0-16 days (Xiao et al., 2022). Our data showed that the unaltered gene target concentration had the best correlation to the clinical cases of COVID-19. Normalization is supposed to enhance the correlation by considering the amount of feces in the sewage. Because the main source of SARS-CoV-2 in sewage should be feces, fluctuations in the amount of feces had the potential to negatively impact the signal. However, through the Pearson's correlation coefficients we showed that the normalized data did not show improved correlation with clinical cases. PMMoV provided the worst amount of correlation despite being widely recommended for sample normalization. The DNA normalized gene targets were able to provide a moderate positive correlation to the clinical cases. Normalizing with the flow rate provided a higher correlation than the fecal normalized gene targets in

2/3 sites with an overall moderate to high correlation to the clinical data for all three sites. The DNA based fecal indicators had a higher amount of correlation to clinical data than the RNA based fecal indicator (PMMoV). However, DNA fecal indicators required a different spin column workflow and assay mix increasing the cost and time. Overall, a significant amount of money would be saved from not normalizing the data, and these data show that there was no benefit to normalizing the data. It is unlikely, but possible, that the storage of the sewage pellets may have resulted in signal loss which could contribute to the lower correlations seen.

This study estimated the number of cases of COVID-19 and found that the estimated cases were statistically higher than the clinical cases in every case. It is likely that COVID-19 is under detected clinically due to patients not wanting to be tested, not reporting their at-home tests, or patients being asymptomatic and therefore not being tested for COVID-19. It has been suggested that 18-31% of COVID-19 cases are asymptomatic (Hata & Honda, 2020). Therefore, it is likely that many COVID-19 cases are not caught in clinical data meaning that the sewage data may provide a more realistic number of cases of COVID-19. Overall, there was a moderate to strong correlation between the confirmed cases and sewage estimated cases of COVID-19. Therefore, sewage monitoring allows tracking of COVID-19 trends while slightly preceding clinical data. With less clinical testing, sewage monitoring of SARS-CoV-2 may become the better tool for assessing COVID-19 trends and at a lower cost.

CHAPTER FOUR

DUPLEX CAPABILITY OF SARS-COV-2 SURROGATES PHI6 AND BCOV FOR INHIBITION AND RECOVERY CONTROL

4.1 Abstract

Coronavirus disease 2019 (COVID-19) is a highly contagious disease that typically causes symptomatic individuals to experience mild to moderate respiratory distress, but in too many cases may lead to hospitalization and even death. This disease is caused by an RNA virus known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Due to the fast-spreading nature of this virus it is vital that cases be identified quickly so appropriate steps can be taken to manage outbreaks of the virus. Due to the presence of SARS-CoV-2 in feces, COVID-19 infection can be efficiently monitored with sewage surveillance. Wastewater surveillance is beneficial because it can detect asymptomatic cases, cases from individuals that will not get tested, cases from individuals who take at-home tests and do not report their results, and quickly provide data on whole communities or large congregate facilities. In addition to testing sewage for SARS-CoV-2, it is also important to assess samples for inhibition and recovery to rule out any false negatives. Because the amount of inhibition or recovery from the SARS-CoV-2 PCR assays cannot be directly assessed, Phi6 and bovine coronavirus (BCoV) were evaluated as spiked recovery controls, i.e. surrogates, for SARS-CoV-2. A SARS-CoV-2 surrogate should have a similar structure to SARS-CoV-2. This similarity in structure should allow the surrogate to mimic how SARS-CoV-2 would react under the same sample preparation and assay. Both Phi6 and BCoV are enveloped RNA viruses with spike proteins which

are features of SARS-CoV-2. Both Phi6 and BCoV were found to have acceptable separation at an annealing temperature of 55°C allowing inhibition and SARS-CoV-2 gene targets to be tested simultaneously. BCoV required the quenchers ZEN™/Iowa Black® Fluorescent Quencher (ZEN/IABKFQ) for proper separation between the negative and positive droplets. Phi6 had reasonable separation with the quencher Black Hole Quencher®-1 (BHQ1). It was also determined that both BCoV and Phi6 can be duplexed without losing signal according to the Anova test. Lastly, it was discovered that BCoV had an average concentration that was 2-fold higher than Phi6.

4.2 Introduction

Coronavirus disease 2019 (COVID-19) is a disease that causes respiratory illness that is generally mild or moderate leading to minor symptoms such as coughing and fever (*Coronavirus*, n.d.). Although this disease is typically associated with mild symptoms, COVID-19 can lead to severe and even life-threatening illness. Due to the ability of this disease to be transmitted rapidly from person to person over 631 million people (November 14th, 2022) across the globe have been infected in less than three years of COVID-19's origin (*WHO Coronavirus (COVID-19) Dashboard*, n.d.). This disease is caused by a single-stranded and enveloped RNA virus known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Alexandersen et al., 2020). There are currently seven known coronaviruses (CoVs) that infect humans (Ye et al., 2020). Four of these CoVs commonly infect humans and are associated with respiratory illness that is mild such as the common cold (*Common Human Coronaviruses | CDC*, 2020). The other three of these CoVs are more dangerous and associated with more frequent deaths than the common human coronaviruses (HCoVs).

Generally, COVID-19 is diagnosed by completing reverse transcription-polymerase chain reaction (RT-PCR) on nasopharyngeal swabs which provides rapid results with high specificity and moderate sensitivity (Koskinen et al., 2021). SARS-CoV-2 can be detected in other human bodily fluids such as saliva, tears, semen, and feces (Kutti-Sridharan et al., 2020). Because SARS-CoV-2 can be detected in feces, sewage surveillance can be used to monitor COVID-19 trends in different buildings and areas. As urine samples do not generally test positive for SARS-CoV-2, feces will be the main source of SARS-CoV-2 in sewage (Mumm et al., 2021). Approximately half (48.1%) of fecal samples tested in COVID-19 infected patients test positive for SARS-CoV-2 (S. Park et al., 2021, p. 2) whereas 84.6% of nasopharyngeal tested positive for infected patients (Gopaul et al., 2020). However, enough viral RNA is shed in feces to make sewage surveillance suitable for tracking COVID-19 trends in regions or sewer sheds and buildings. Sewage surveillance has been used to monitor polioviruses in the sewage which has aided polio eradication efforts for many decades (O'Reilly et al., 2020). Acute flaccid paralysis is used to diagnose patients with polio based on their symptoms (Soltani et al., 2014). However, paralysis is not commonly seen (<1% of cases) in infected patients (Tuma, 2021). Thus, sewage surveillance has potential to detect more cases of polio than acute flaccid paralysis (Tuma, 2021). Sewage surveillance is advantageous because there may be hesitancy to testing due to cost or concern of medical procedures. Sewage surveillance can quickly monitor whole communities without permission from each individual. Because at-home testing kits are now available, COVID-19 cases may not be reported which could falsely suggest low community spread of COVID-19. Thus, sewage surveillance is becoming more critical despite SARS-CoV-2 not being detected

as frequently in the feces. Testing wastewater samples will provide data for whole communities non-invasively while providing rapid data including both symptomatic and asymptomatic individuals (CDC, 2022c).

Phi6 is a bacteriophage that has double-stranded RNA (Frilander et al., 1995). Phi6 is also an RNA virus that is enveloped which makes it a potential surrogate for other enveloped viruses such as Ebola virus and CoVs including SARS-CoV-2 (Bangiyev et al., n.d.). Phi6 is thought to be an optimal surrogate for CoVs due to it having spike proteins, having a similar size, and being an enveloped virus (Fedorenko et al., 2020). Bovine CoV (BCoV) is an RNA virus that causes respiratory illness in wild ruminants and cattle (Saif, 2010). This CoV is enveloped and has single-stranded RNA as opposed to Phi6 that has double-stranded RNA (Saif, 2010). As BCoV is a CoV and is single-stranded, it is logical to think that it should be a better surrogate for SARS-CoV-2 than Phi6. However, as these targets have both been used previously as surrogates for SARS-CoV-2, both Phi6 and BCoV were chosen in this study. If both surrogates suggest full inhibition, it is likely that the SARS-CoV-2 gene targets may be falsely reported as negative. Thus, it is important to assess the inhibition in each sample to take measures to reduce inhibition and to understand if the sample is truly negative. Sewage is highly variable in time and source. Therefore, the best surrogate for the SARS-CoV-2 gene targets might differ on a sample-by-sample basis.

All samples analyzed in this study use Droplet Digital polymerase chain reaction (ddPCR). Running a duplex reaction gives two sets of data within a single reaction well which significantly reduces cost. As two targets can be analyzed in a single well, time spent on pipetting, i.e. Plating, (both manually and on the automated droplet generator)

and on reading the plates can be reduced. Thus, it is highly desirable to run Phi6 and BCoV in a duplex. However, in order to run a duplex, we must verify that there is no significant signal loss. This study determined under what conditions Phi6 and BCoV could be run as a duplex reaction to save time and money while providing robust inhibition and recovery control.

4.3. Materials and Methods

4.3.1 Sample Delivery, Sample Selection, and Sample Concentration

Our workflow used the methods described by Flood et al (Flood et al., 2021) as a starting point. Sewage samples were delivered to the lab in 500mL portions. Samples were transported in a cooler with icepacks while in transit to our lab. Samples were processed immediately upon arrival. The 500mL bottle of sewage was mixed by inverting the bottle between 40 and 60 times to minimize foaming. A 45 mL homogenous sewage sample was decanted into a 50mL conical bottom centrifuge tube. Approximately 4.0g of polyethylene glycol 8000 (catalog number 97061-100 from VWR) and 0.59g sodium chloride (catalog number 470302-520 from VWR) were added to the sample yielding a concentration of 8% polyethylene glycol and 0.2M sodium chloride. After the solids and solutions were combined, the centrifuge tubes were transferred to a rotator where the tubes were rotated at 70rpm for 2hrs. The rotator was held at 4°C. After mixing, the samples were centrifuged at 3,650 rpm for 1 hr. Once centrifugation was completed, the supernatant was discarded, and a 1-2 mL pellet remained.

4.3.2 RNA Elution and Purification

The QIAamp Viral RNA Mini Kit (catalog number 52904) was purchased from QIAGEN to purify and elute the RNA from the sewage samples. The pellet (200µL) was pipetted

into a solution containing 800µL of Buffer AVL, 8µL of carrier RNA, 10µL of Phi6, and 10µL of BCoV (Bovilis® Coronavirus from Merrick Animal Health). Both Phi6 and BCoV had a concentration of 10^5 gene copies/mL. A 200µL portion of pellet was then pipetted into a 2mL cryovial containing the buffer and RNA mix. The cryovial was vortexed for 10s before incubation for 10mins at room temperature. After incubation, 200proof ethanol was added to the cryovial in an 800µL portion, and the cryovial was vortexed for 10s. At this point in the procedure there was 1.8mL solution in the cyrovial. This solution was run through the QIAamp mini spin column in order to obtain the extracted RNA. The 1.8mL solution was added in three 630µL quantities to the spin column. Three additions were done because the spin column could not hold enough solution for a single addition. Each 630µL portion of solution was added to the spin column and centrifuged sequentially. The flow through was not retained. The centrifugation settings were 1min at 8,000rpm. These centrifuge settings were the same for every addition except when the buffer AW2 was added. The waste collected in the collection tube under the spin column was discarded after every addition until the elution step (buffer AVE). After the whole 1.8mL solution was added to the spin column and centrifuged, the wash buffers were added. A 500µL portion of buffer AW1 was added to the spin column and centrifuged. Next, a 500µL portion of buffer AW2 was added to the spin column. At this point in the process (and only during this step) the centrifuge speed and duration was increased to 14,000rpm for 4min. After centrifugation, centrifugation settings were returned to normal, and the spin column was transferred to a 1.5mL microcentrifuge tube. The microcentrifuge tube was used to collect the RNA eluted from the spin column. Buffer AVE was added to the spin column in a 40µL portion and

incubated at room temperature for 1min. After incubation, the spin column was centrifuged. This step was repeated which resulted in a final volume of 80µL of RNA eluate.

4.3.3 Plating, Droplet Generation, Thermocycling, and Droplet Reading

The assay mix was made using the solutions obtained from the One-Step RT-ddPCR Advanced Kit for Probes (catalog number 1864021) produced by Bio-Rad. This kit provided everything needed for the assay mix except the water and the primer/probe mix. Table 4.1 shows the primer/probe sequences used to detect the SARS-CoV-2 gene targets. Table 4.2 shows the components that were used for a singleplex reaction and a duplex reaction.

Table 4.1: BCoV/Phi6 Primer and Probe Sequences

Sequence (5'-3')	BCoV	Phi6	References
Forward Primer	5'-CTGGAAG TTGGTGG AGTT-3'	5'-TGGCGGCGGTCAAGAGC-3'	(Gendron et al., 2010; Graham et al., 2020)
Reverse Primer	5'-ATTATCG GCCTAA CATACAT C-3'	5'-GGATGATTCTCCAGAAGCT GCTG-3'	(Gendron et al., 2010; Graham et al., 2020)
Probe	5'-HEX-CCTTCATA T-ZEN-CTATACACATCAAG TTGTT- IABKFQ-3'	5'- FAM-CGGTCGTCGCAGGTCTGAC ACTCGC-BHQ1-3'	(Gendron et al., 2010; Graham et al., 2020, p.)

Table 4.1: A table containing both primers and probes used for Phi6 and BCoV. The fluorophore and quencher chosen are not equivalent to the ones used in the referenced papers.

Table 4.2: Assay Mix Calculations for a Single Well

Reagent	Singleplex	Duplex
Supermix	5.5µL	5.5µL
Reverse transcriptase	2.2µL	2.2µL
Dithiothreitol	1.1µL	1.1µL
Primer/probe mix 1	3.3µL	3.3µL
Primer/probe mix 2	N/A	3.3µL
Water	4.4µL	1.1µL
Total	16.5µL	16.5µL

Table 4.2: The assay mix calculations utilized for a singleplex and a duplex reaction. These calculations were based on the amount of assay mix required for one reaction (one well).

The values in Table 4.2 were based on a single reaction. Singleplex and duplex reactions only varied based on the amount of water and the additional primer/probe mix which would be added for a duplex reaction. Samples and controls (negative, no template, and positive controls) were pipetted into a 96-well plate with the correct assay mix. Each well had 5.5µL of sample/control and 16.5µL of assay mix. Each sample and control were run in triplicate per assay mix. Once plating was finished, the plate was sealed with foil. This plate was then vortexed to mix the sample and the assay mix. Lastly, the plate was spun in a plate spinner to collect the liquid at the bottom of the 96-well plate. This plate was then placed into the Automated Droplet Generator (catalog number 1864101 Bio-Rad) to generate droplets. The Automated Droplet Generator generated droplets into another 96-well plate. This new 96-well plate was sealed with foil and placed into the thermocycler. The thermocycler settings are listed in Table 4.3. The thermocycling settings were based on the instructions for the 1-Step RT-ddPCR Advanced Kit for Probes for general settings, and the annealing temperatures from Flood et al. (*One-Step RT-DdPCR Advanced Kit for Probes*, n.d.; Pearson et al., 2021)

Table 4.3: Thermocycler Settings for BCoV and Phi6

Step	Temperature (°C)	Duration
1	25	3mins
2	50	60mins
3	95	10mins
4*	95	30secs
5*	55	1min
6	98	10mins
7	4	30mins
8	4	Infinity

Table 4.3: Thermocycling consists of 8 steps. Steps 4 and 5 are marked with an asterisk to symbolize those steps are repeated. Steps 4 and 5 are completed back-to-back, and this cycle is repeated 40 times. Once step 8 is reached the plate can be put onto the droplet reader at any point.

Once thermocycling was complete, the plate was held at room temperature for 15mins before reading the plate with the QX200 Droplet Reader (catalog number 1864003) from Bio-Rad. Once the plate was read, thresholds for positive and negative droplets were set and the plate was quantified.

4.3.4 Temperature Gradient Study

Thermocycling settings remained the same except step 5 listed in Table 4.3. Instead of running one temperature a series of annealing temperatures were ran. The range 52.0-62.0°C was chosen. Both Phi6 and BCoV were ran on a temperature gradient to assure proper separation was obtainable at 55.0°C. Specific temperatures between the minimum and maximum temperature were automatically chosen by the thermocycler.

4.3.5 Data Analysis

Gene copies of each target per liter of sewage (CP/L) were calculated using various constant values throughout the experiment, experimentally determined values, and values

obtained QuantaSoft™. Equation 4.1 was used to calculate both the gene copies of BCoV and Phi6 per liter of sewage.

(Equation 4.1)

$$\frac{CP}{L} = copies\ per\ uL \times eluate\ volume\ (80uL) \times \frac{\frac{final\ concentrate\ volume\ 45mL)}{extraction\ volume\ (0.2mL)}}{initial\ volume\ (45mL)} \times 1000$$

Equation 4.1: Gene Copies per Liter of Sewage Calculation
Gene copies per liter of sewage (CP/L) for BCoV and Phi6 can be calculated with this formula.

The gene copies of each target per liter of sewage of BCoV and Phi6 were then compared for the parallel singleplex and duplex reactions. A total of 36 samples were ran for Phi6 and BCoV in both singleplex and duplex forms to determine if there was a significant variation in concentration when experiments were carried out in a duplex instead of a singleplex reaction. Of these 36 samples, four of the samples were removed due to low droplets or extremely high values found in the singleplex reactions that were considered a potential machine error or sample contamination. For instance, the negative control was run to estimate the maximum BCoV and Phi6 signal that can be obtained. The negative control consisted of 200µL of water, 8µL of carrier RNA, 800µL of buffer AVL, 10µL of Phi6, and 10µL of BCoV. Because water was used instead of sewage, no inhibition should occur, and the maximum signal should be obtained. Two of the 36 samples had a signal for BCoV in the singleplex reaction that were 2.9x and 9.9x higher than the negative control which led to the conclusion that there was a machine error or sample

contamination, and these values were considered unreliable. Thus, these two samples were removed from both BCoV and Phi6 analysis.

To determine if the singleplex and duplex reactions had values that were statistically equivalent, an ANOVA: Two-Factor with Replication test was ran for BCoV and Phi6. This test provided an F-value, an F critical value, and a p-value. F-values lower than the F critical value follow the null hypothesis which supports that a duplex can be ran without losing BCoV or Phi6 signal. P-values greater than 0.05 indicate that a duplex can be run without losing signal. Paired t-tests were also run to test duplex capability. T-values between the positive and negative of the T critical two-tail value support that a duplex and singleplex provided similar results. The calculated gene copies of each target per liter of sewage were also compared between Phi6 and BCoV to determine if one target was detected at a higher concentration than the other target.

4.4 Results

When viewing the data with QuantaSoft™ Software, the separation between the positive droplets and negative droplets was assessed to determine if the quencher used in our lab was reliable for BCoV. Separation in this study is defined as the visible distinction between positive and negative droplets. Initially, Black Hole Quencher®-1 (BHQ1) was used to quench the HEX fluorophore on BCoV because the Phi6 primer/probe mix obtained from Michigan State University utilized BHQ1 as the quencher. Many labs in the Michigan PCR Network have done research with Phi6 to assess sample inhibition/recovery along with routine SARS-CoV-2 gene target testing. Separation is acceptable for Phi6 with the BHQ1 quencher leading to the selection of this quencher for BCoV. However, BCoV did not have reliable separation with the BHQ1 quencher unless

the droplets were held in the thermocycler overnight. However, even after the prolonged holding period on the thermocycler, separation was poor. The quencher ZEN™/Iowa Black® Fluorescent Quencher (ZEN/IABKFQ) was later used for BCoV, and a temperature gradient was run to determine the best annealing temperature. Concentration was kept constant. There was separation at all annealing temperatures ran (52.0-62.0°C). The best separation was obtained at the annealing temperatures between 52.0°Cand 56.0°C. For detection of the N1 and N2 gene targets (used to detect SARS-CoV-2) an annealing temperature of 55.0°C was used. Because there was reasonable separation at that temperature, both Phi6 and BCoV could run at an annealing temperature of 55.0°C. This would allow N1 and N2 to be run with the surrogates on the same plate as two duplexes. Figure 4.1 shows the temperature gradient for BCoV with both quenchers and Phi6.

BCoV had approximately an 1,800-amplitude (relative fluorescence) difference between the positive and negative droplets at 54.1°C and 56.0°C with the ZEN/IABKFQ quencher. When looking at the BHQ1 quencher at 54.1°C and 56.0°C, separation was minimal (approximately 50-amplitude difference). At the lowest temperatures (52.0°C and 52.7°C) there was an approximate amplitude difference of 100. Besides the large separation, when the BCoV probe had the quenchers ZEN/IABKFQ, plates could be put on the droplet reader just after thermocycling was finished with no separation issues. Thus, the ZEN/IABKFQ quenchers were chosen for BCoV. Phi6 did not have any issues with separation, so the quencher was not changed. Phi6 had approximately a 700-amplitude difference between positive and negative droplets regardless of annealing temperature.

Figure 4.1: Temperature Gradient with Different Probes

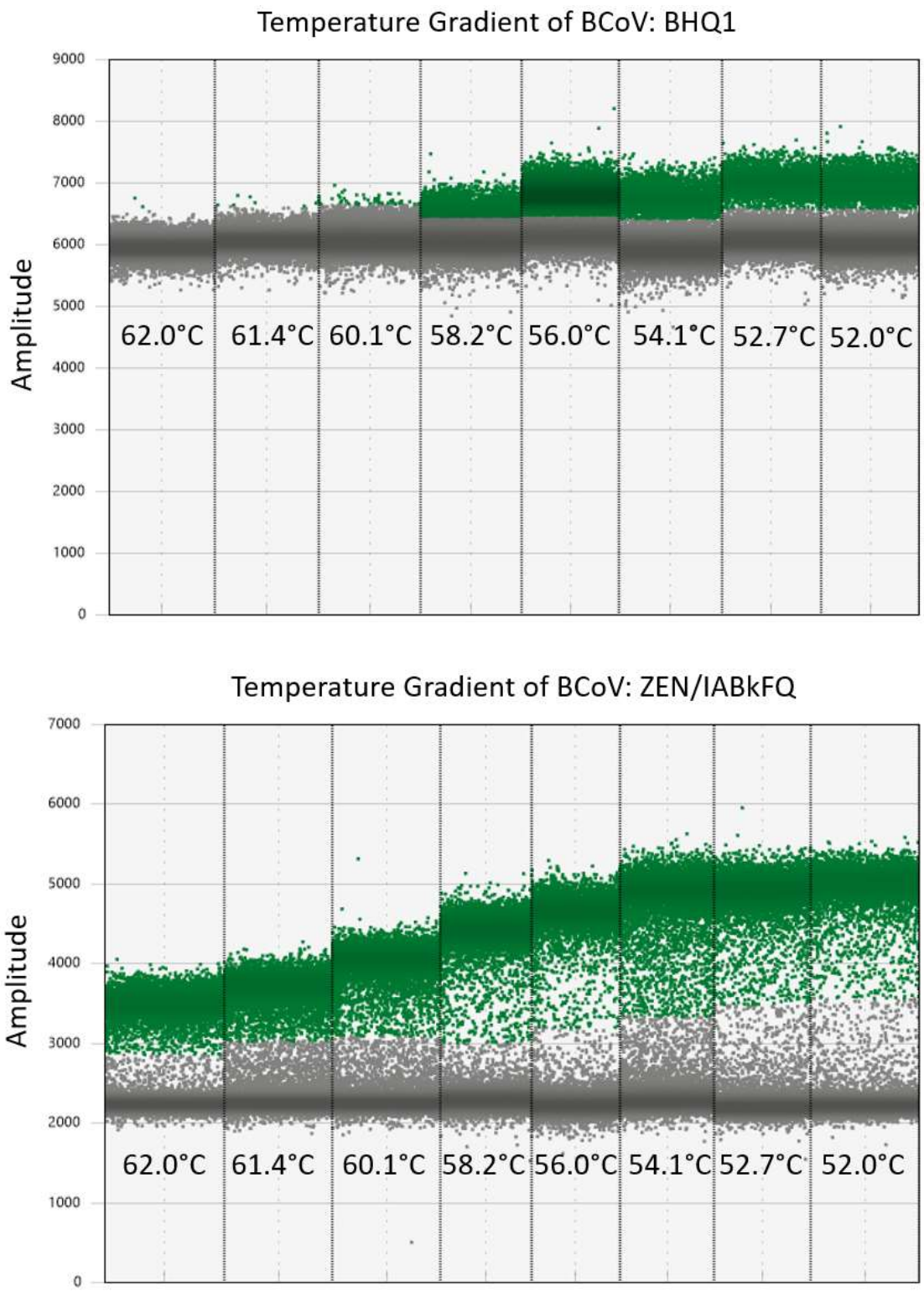


Figure 4.1: Temperature Gradient with Different Probes (Continued)

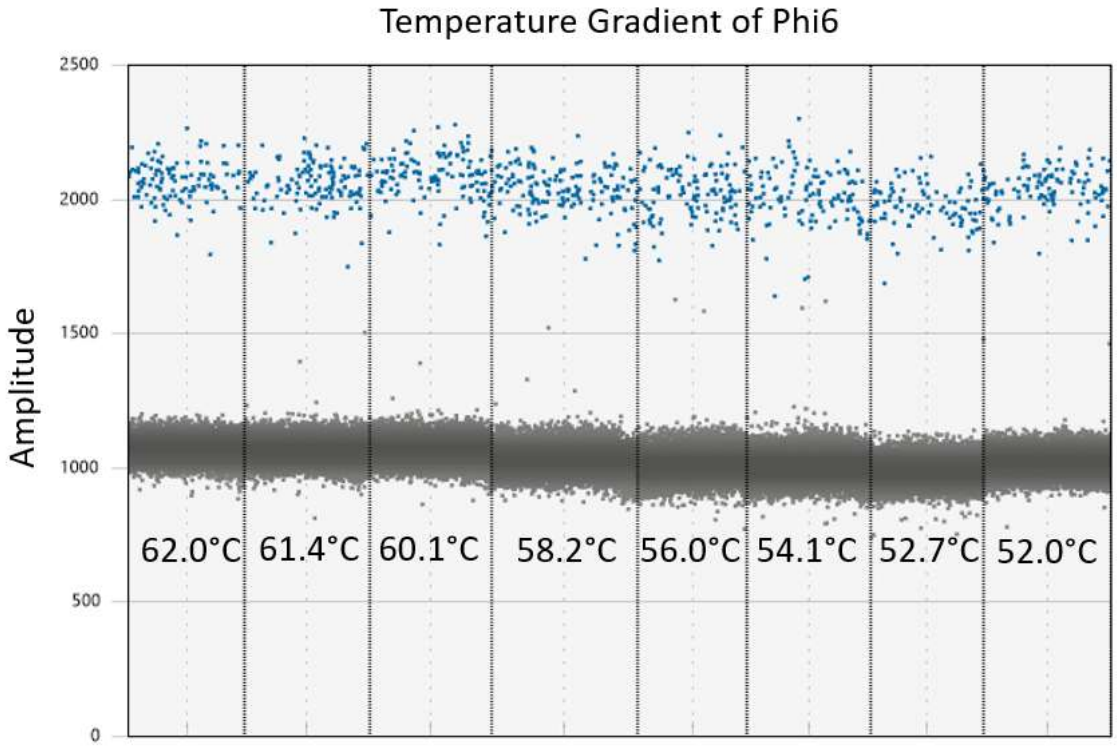


Figure 4.1: Temperature gradient of annealing temperatures for both BHQ1 and ZEN/IABKFQ quenchers on BCoV and Phi6. Green droplets indicate positive droplets, and gray droplets indicate negative droplets for BCoV. Blue droplets indicate positive droplets, and gray droplets indicate negative droplets for Phi6. The ZEN/IABKFQ quencher was able to drastically enhance the separation between positive and negative droplets. Phi6 only had one quencher tested because there was no issues with separation.

Using Excel, an ANOVA: Two-Factor with Replication test was run for both Phi6 and BCoV. A p-value of 0.5 indicated that singleplex and duplex reactions for BCoV have concentrations that are statistically equivalent. For Phi6 a p-value of 0.5 was obtained again indicating that singleplex and duplex reactions yield statistically equivalent concentrations. Thus, BCoV and Phi6 can be run as a duplex without compromising the concentration obtained.

Next, paired t-tests were calculated on a sample-by-sample basis for all 32 sewage samples when applicable. There was one non-detect for BCoV and seven non-detects for Phi6 which reduced the number of paired t-tests that could be run. For BCoV most of the samples (25/31 or 80.6%) supported that singleplex and duplex reactions produced concentrations that were statistically the same. For Phi6, most samples (22/25 or 88.0%) supported that singleplex and duplex reactions produced concentrations that were statistically the same.

Of the 32 samples, 31 of these samples had signal for either Phi6 or BCoV. However, there were six samples that only had a signal for BCoV indicating BCoV is more robust than Phi6 to inhibitors or compounds that interfere with recovery. Despite BCoV and Phi6 being spiked into the buffer AVL with the same concentration (10^5 gene copies/mL), BCoV always had a higher concentration than Phi6. Comparing BCoV's and Phi6's average concentration, BCoV had a 2.5times higher concentration. Figure 4.2 shows the differentiation between BCoV's and Phi6's concentration for all 32 samples.

Figure 4.2: BCoV vs. Phi6 Concentration

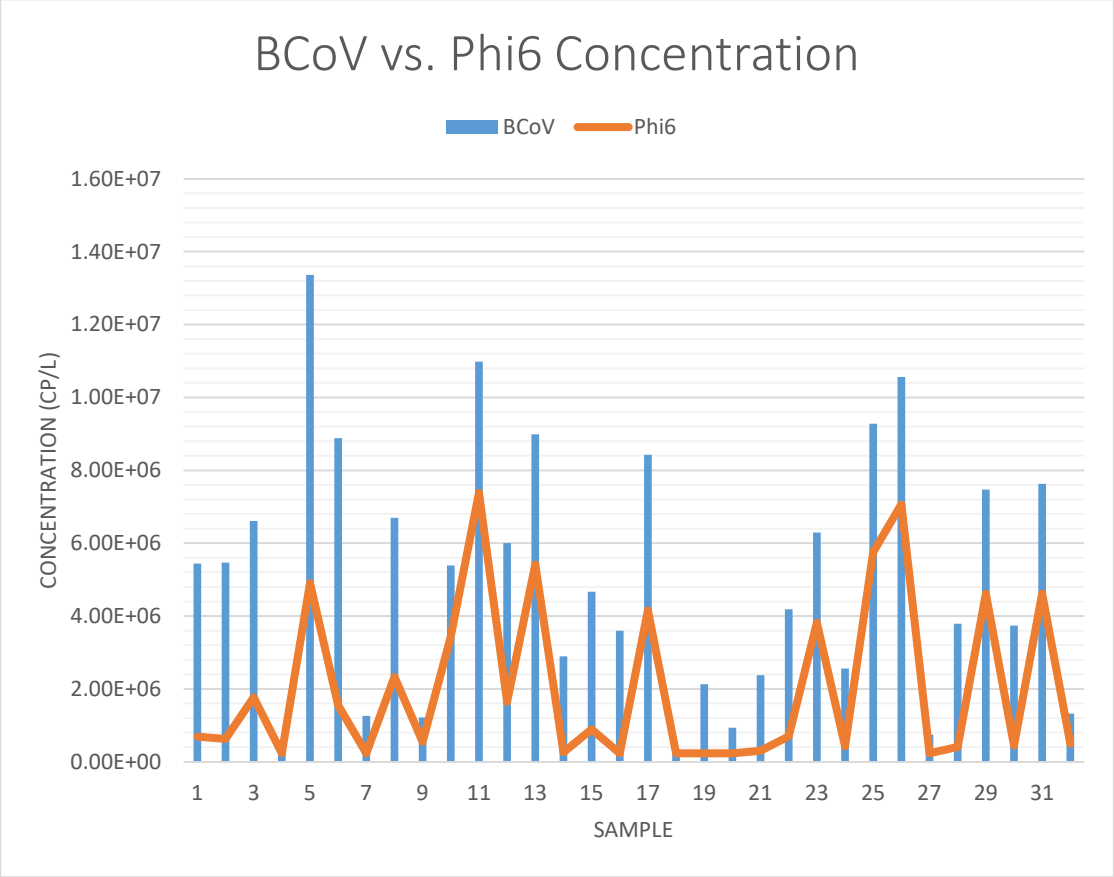


Figure 4.2: The average concentration of the six wells (three from singleplex and three from duplex) is shown as a blue bar for BCoV and an orange line for Phi6. BCoV always had a higher concentration than Phi6.

Figure 4.3: Graphical Comparison of Phi6 and BCoV Singleplex and Duplex

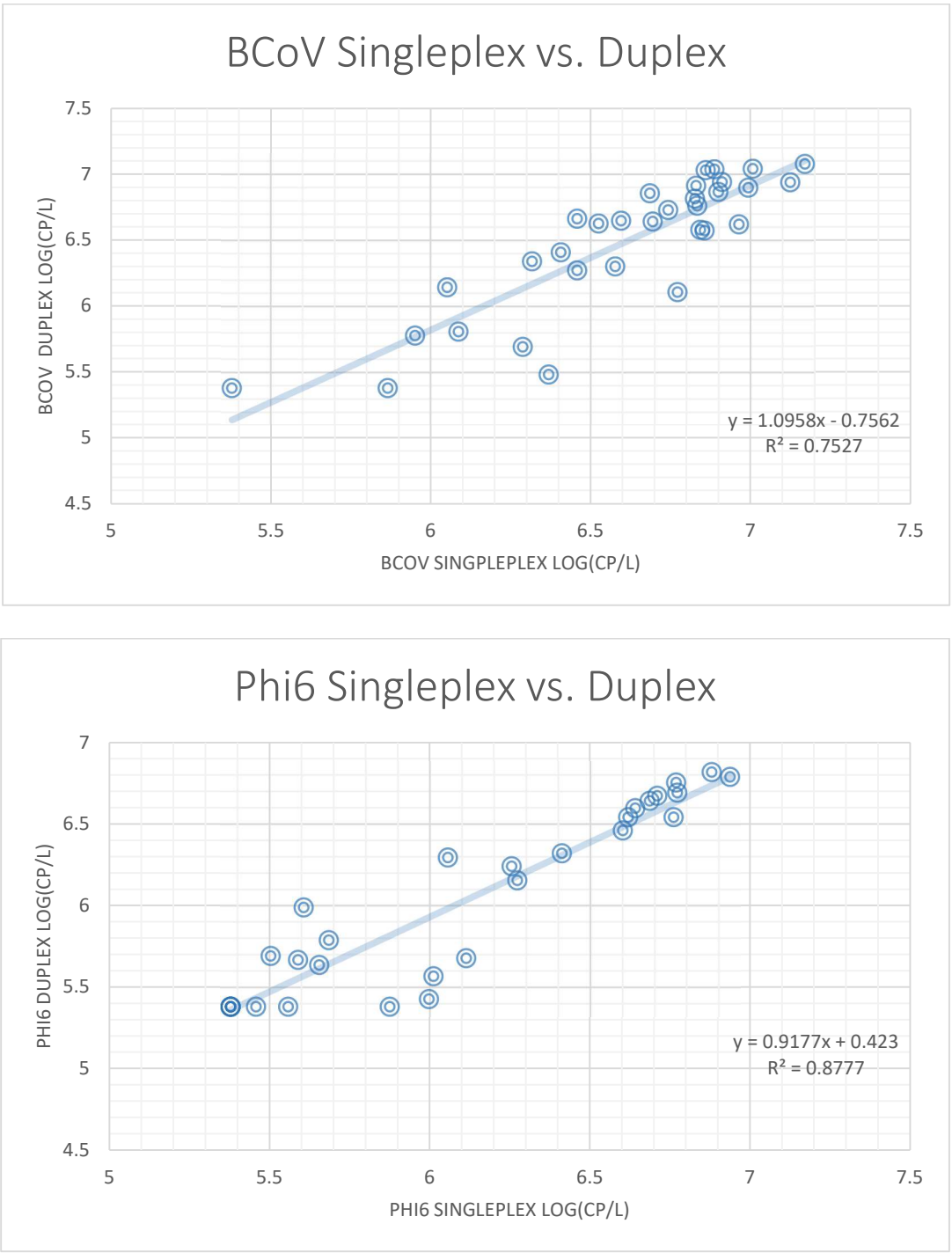


Figure 4.3: Graphs were adjusted to enlarge the trendline and points for clarity. The y-axis was set to 4.5 log(CP/L), and the x-axis was set to 5 log(CP/L). Phi6 had a higher R^2 value than BCoV showing there wass a stronger correlation between singleplex and duplex reactions for Phi6 than BCoV.

When the singleplex and duplex reactions were graphically compared, Phi6's data points had a higher similarity between the singleplex and duplex reactions than BCoV as shown in Figure 4.3. The R^2 value was 0.88 for Phi6 and 0.75 for BCoV. The Pearson's correlation coefficient was 0.97 for Phi6 and 0.83 for BCoV. Therefore, the data indicate the Phi6 has a stronger correlation between the concentrations obtained in singleplex and duplex reactions than BCoV.

Overall, BCoV and Phi6 can be duplexed while maintaining a concentration that is statistically the same. BCoV had a lower R^2 value and Pearson's correlation coefficient than Phi6, but both BCoV and Phi6 had a strong positive correlation.

4.5 Discussion

The goal of this study was to determine if Phi6 and BCoV could be duplexed to provide simultaneous detection of two targets for inhibition/spiked recovery control checks without compromising the ability to quantify. Running a duplex as opposed to two singleplex reactions reduces cost and time and allows inhibition for both surrogates to be assessed on the same plate as the SARS-CoV-2 gene targets. We found that on a sample-by-sample basis, 80.6% (BCoV) and 88.0% (Phi6) of the samples tested for BCoV supported that singleplex and duplex reactions had no statistical variation in concentration. An ANOVA test agreed that there was no statistical significance between the concentrations of the singleplex and duplex reactions. The SARS-CoV-2 surrogate BCoV was found to have an average concentration 2.5 times greater than Phi6, always had a higher concentration than Phi6, and was only fully inhibited (=MDL) in one sample compared to Phi6 that was fully inhibited in seven samples. These two findings suggest that BCoV is a better spike recovery control than Phi6. BCoV would have significantly

reduced the cost of reruns as only one sample was fully inhibited indicating it should be rerun compared to seven reruns for Phi6. There are instances in routine SARS-CoV-2 testing where either N1 or N2 SARS-CoV-2 gene targets have signal, but Phi6 does not have signal suggesting that Phi6 is not an ideal surrogate for SARS-CoV-2. BCoV is inexpensive and readily obtainable as a cattle vaccine, and BCoV has a higher resemblance to SARS-CoV-2 virus. Spiking samples with SARS-CoV-2 is impractical and would require a biosafety level 3-4 lab. (*Biosafety Level Requirements*, n.d.). Our lab is a biosafety level 2+ lab which is appropriate for only routine diagnostic testing of SARS-CoV-2 (Yeh et al., 2021). Thus, to minimize risk and cost, SARS-CoV-2 surrogates were used to estimate the amount of inhibition and recovery of SARS-CoV-2 would likely face under the same conditions.

Both Phi6 and BCoV are useful inhibition and spike recovery controls. Phi6 is cost efficient, has a similar structure to SARS-CoV-2, and has similar survival in sewage which makes Phi6 a good surrogate for SARS-CoV-2 and other CoVs (Gomes et al., 2022). BCoV is a good surrogate for SARS-CoV-2 because it is also a CoV, it is a single-stranded RNA virus, it has a similar size, it is enveloped, and the structure is very similar (Susan I. Gerber and John T. Watson, 2020). BCoV has been listed to require either biosafety level 1 or 2 labs (NR-445 Bovine Coronavirus, Mebus (Viruses), n.d.; NR-451 Bovine Coronavirus, Mebus, Chemically Inactivated (Antigen Preparations), n.d.).

Phi6 and BCoV have been used together in one study as SARS-CoV-2 to determine if ozone is optimal for decontaminating rooms with SARS-CoV-2 (Franke et al., 2021). They found that both Phi6 and BCoV behaved similarly (Franke et al., 2021).

Currently, there does not appear to be any studies that duplex Phi6 and BCoV. One target is needed to determine if a sample is inhibited or to detect recovery problems, but looking at two targets provides benefits that are similar to using both N1 and N2. Because BCoV is another CoV it is likely to be superior to Phi6 because Phi6 is double-stranded and has more structural deviation to SARS-CoV-2 than BCoV. If both surrogates are below the MDL this suggests full inhibition or recovery failure, and that there is high potential for false negative determination and the sample should be rerun with dilution. Besides interference with the PCR reaction, improper binding to the spin column, and competition between the target and other molecules to bind to the silica in the spin column can cause the sample to have no signal or low signal for the target or surrogate. Reduced retrieval is expected because sewage is likely to have many inhibiters that can degrade the RNA or inhibit ddPCR. Detecting SARS-CoV-2 in sewage is less sensitive than standard clinical testing because SARS-CoV-2 is not excreted as commonly in feces as respiratory specimens. It fluctuates, but generally SARS-CoV-2 is only found in the feces of 50% of patients. For instance, one study found that only 41% of the 44 patients had SARS-CoV-2 in their feces (Britton et al., 2021). Besides the diminished testing capability of feces in general, diluting it in sewage and the potential degradative molecules in the sewage could also greatly diminish the SARS-CoV-2 signal. In the sewage there are high amounts of substances that can degrade RNA such as ribonucleases (Panchal et al., 2021). There are PCR inhibitors in both sewage sludge which includes proteins, heavy metals, fats, and polyphenols and in wastewater which includes metal ions and polysaccharides (Schrader et al., 2012). PCR can be inhibited in various ways such as the inhibitor being precipitated along with the RNA leading to sequestering or degrading of the RNA, and

inhibitors binding to vital enzymes or RNA which could cause polymerases to become inhibited (Ahmed et al., 2022). Because the signal of SARS-CoV-2 is much lower in fecal samples, and is subject to inhibition and interference, two surrogates provide more information in evaluating false negatives and evaluating corrective measures.

This research determined that Phi6 and BCoV can be ran at an annealing temperature of 55.0°C. Phi6 and BCoV are generally ran at an annealing temperature of 60.0°C (Flood et al., 2021; Loeb et al., 2020). However, because the N1 and N2 gene targets are ran at an annealing temperature of 55.0°C, it is more practical to run BCoV and Phi6 at 55.0°C (Flood et al., 2021). We found that it was vital for the BCoV probe to have the quencher ZEN/IABKFQ for proper separation while Phi6 had proper separation with the quencher BHQ1. Our primer and probes were purchased from Integrated DNA Technologies. This company sells a quencher called ZEN which was stated to reduce background signals, and it is considered an internal quencher and comes paired with IABKFQ (*PrimeTime QPCR Probes*, n.d.). The company also mentions that ZEN generally is the main quencher, but IABKFQ additionally reduces background noise and enhances performance (*ZEN and IBFQ*, n.d.). As shown in our results, using this combination of quenchers drastically improved the amount of separation between the positive and negative droplets for BCoV. The BHQ1 quencher performed poorly and should not be used with sewage samples. The amplitude difference was approximately 2000 between positive and negative bands for the ZEN/IABKFQ quenchers at the comparable temperature of 52.0°C. At best Phi6 had an amplitude difference of 700 with the BHQ1 quencher. The separation with ZEN/IABKFQ was almost triple this value. Thus, it is likely that

switching Phi6's quencher to ZEN/IABKFQ would further enhance separation. However, since separation was reasonable, the quencher was not changed for Phi6.

4.6 Conclusion

It is necessary that every sample be assessed for inhibition and recovery to determine if there is potential for false negatives when analyzing the SARS-CoV-2 gene targets. Analyzing both BCoV and Phi6 provides more certainty if the SARS-CoV-2 gene targets may be fully inhibited. If BCoV and Phi6 are both fully inhibited, then the sample can be rerun after dilution or the sample modified by addition of chemicals such as bovine serum albumin. It is important that results are reported quickly to allow safety measures to be taken in certain areas if necessary. It was found that Phi6 and BCoV could be duplexed without losing signal. Thus, duplexing BCoV and Phi6 will allow for the SARS-CoV-2 gene targets, Phi6, and BCoV to be analyzed all together. This reduces time, cost, and allows samples to be rerun in a timely manner if the sample is fully inhibited according to Phi6 and BCoV. Lastly, to ensure optimal separation for BCoV, it is vital to use the quencher ZEN/IABKFQ for BCoV. BHQ1 is an appropriate quencher for N1, N2, and Phi6.

CHAPTER FIVE

ESTIMATION OF SARS-COV-2 INHIBITION AND COMPARISON OF THE FREQUENCY OF THE N1 AND N2 GENE TARGETS AND THE SARS-COV-2 SURROGATES PHI6 AND BCOV

5.1 Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an RNA virus which is responsible for coronavirus disease 2019 (COVID-19). Generally, clinical testing of individuals detects SARS-CoV-2 with a nasopharyngeal swab which can give details on COVID-19 infection on a person-by-person basis. Wastewater testing can be used alongside clinical testing to provide trends on the spread of COVID-19 within whole communities. Testing sewage by droplet digital PCR (ddPCR) provides cost-effective, rapid, COVID-19 data for entire communities allowing trends of COVID-19 increase to be identified earlier than clinical data. Sewage is complex and contains many aggressive chemicals, the SARS-CoV-2 gene targets are exposed to inhibitors that can interfere with PCR chemistry or the spin column clean-up procedure. It is important to detect inhibition to avoid false negative samples and allow measures to be taken to combat inhibition if necessary. Phi6 and bovine coronavirus (BCoV) can both be used as surrogates to assess if the sample is inhibited due to their similar structure to SARS-CoV-2. However, it is uncertain if BCoV and Phi6 are good surrogates for SARS-CoV-2 gene targets N1 and N2 in sewage. We found that BCoV (94.1%) was detected more frequently than Phi6 (81.4%) and that N1 (55.5%) is detected more frequently than N2 (51.7%) in sewage samples. Both BCoV ($p=5.9x10^{-84}$) and N1 ($p=6.3x10^{-12}$) had a statistically higher

concentration than Phi6 and N2, respectively. BCoV was found to be more consistent with the N1/N2 data only testing negative for 1.2% of the N1 and/or N2 positive samples while Phi6 missed 10.1% of the positive samples. When it came to BCoV and Phi6 retrieval, BCoV had statistically higher retrieval than Phi6 with an average amount of retrieval of 47.2% for BCoV and 27.8% for Phi6.

5.2 Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a betacoronavirus that is enveloped, single-stranded, and composed of RNA (Romano et al., 2020). Infection with SARS-CoV-2 causes individuals to get the disease named coronavirus disease 2019 (COVID-19) (CDC, 2020a). COVID-19 was first discovered in Wuhan, China in December of 2019 (CDC, 2022a). Soon after, the first case of COVID-19 was detected in Washington in January which was the first confirmed case in the United States (CDC, 2022a). In the same month (January 30, 2020) the World Health Organization (WHO) deemed COVID-19 a Public Health Emergency of International Concern and shortly after (March 11, 2020) the COVID-19 outbreak was called a global pandemic by the WHO (*CDC's Global Resources Pivot to Address COVID-19*, 2021). Most individuals that become ill due to COVID-19 experience symptoms that are mild to moderate (Shi et al., 2020, p. 19). One study looked at 44,415 individuals that had COVID-19 and found that most cases were mild (81%) followed by severe (14%) and critical (5%) cases. The most common symptoms associated with COVID-19 are fatigue, cough, and fever, but some individuals can also experience headache, runny nose, loss of smell, sore throat, or loss of taste (*Mild COVID-19 Symptoms*, 2021, p. 19).

COVID-19 is caused by one of the seven known human coronaviruses (HCoVs) (Kesheh et al., 2022b). There are two HCoVs that are associated with more critical disease than SARS-CoV-2 with both having higher mortality rates than SARS-CoV-2. Fortunately, those HCoVs are not as widespread as SARS-CoV-2. The diseases caused by these other HCoVs are called Middle East Respiratory Syndrome (MERS) and Severe acute respiratory syndrome (SARS), and they are associated with a 35% (MERS) and 9.6% (SARS) mortality rates (Peiris & Poon, 2021). COVID-19 has a drastically lower mortality rate (1.1% as of August 16th, 2022), but the number of cases is much higher (over 588 million as of August 16th, 2022) compared to SARS (8,098 cases as of the 2003 outbreak) and MERS (2,589 cases as of March 2022) (CSR, n.d.; *SARS | Basics Factsheet | CDC*, n.d.; *WHO Coronavirus (COVID-19) Dashboard*, n.d.). When testing for SARS-CoV-2, nasopharyngeal swabs are deemed to be the "gold standard" for diagnosing patients with COVID-19. Although nasopharyngeal or nasal swabs are more typical ways to test for COVID-19, testing wastewater for SARS-CoV-2 has also been implemented to detect SARS-CoV-2 within buildings or communities. Feces will likely be the number one source of SARS-CoV-2 in the sewage as SARS-CoV-2 is rarely detected in urine. Kim et al. determined that only 0.8% (2/247) of their urine samples were positive for SARS-CoV-2 when they detected it using reverse transcription polymerase chain reaction (RT-PCR) (Kim et al., 2020). The number of samples positive for SARS-CoV-2 when feces are tested greatly varies from study to study, but the detection of SARS-CoV-2 in feces is generally found to be much lower than detection with a nasopharyngeal swab or nasal swab. In the Kim et al. study, feces were also tested for SARS-CoV-2 and only 10.1% (13/129) of the samples were positive for SARS-CoV-

2 (Kim et al., 2020). In contrast, another study with a higher sample set found that 61.8% (173/280) of the fecal samples tested positive for SARS-CoV-2 (Lavania et al., 2022). Lavania et al. also determined that SARS-CoV-2 could be detected in the feces of asymptomatic, moderate, and severe/critical cases of COVID-19 (Lavania et al., 2022). The percentages of individuals with SARS-CoV-2 in their feces based on severity of disease were as follows: 67.4% (62/92) were asymptomatic cases, 55.5% (60/108) were moderate cases, and 63.7% were severe/critical cases of COVID-19 (Lavania et al., 2022). Consequently, detection of SARS-CoV-2 in wastewater may be difficult because it is not detected as frequently in feces, wastewater will dilute the fecal content, and the chemical and biological components of sewage may inhibit or interfere with SARS-CoV-2 assays. However, testing sewage for SARS-CoV-2 provides significant benefits such as one sample providing a whole snapshot of a community and wastewater testing can detect asymptomatic cases in addition to symptomatic cases (Kirby, 2021).

This study uses the SARS-CoV-2 gene targets (N1 and N2) recommend by the Centers for Disease Control and Prevention (CDC) (CDC, 2020c). These gene targets assess the concentration of SARS-CoV-2 in the sewage. However, due to the various components in the sewage matrix, polymerase chain reaction (PCR) could be inhibited and lead to samples being falsely labeled as negative for SARS-CoV-2. Additionally, the complex sewage matrix may have molecules that interfere with the spin column's ability to elute the target RNA leading to a reduced amount of recovery. It is important that inhibition be assessed on a sample-by-sample basis because the sewage matrix variability. To assess inhibition, SARS-CoV-2 surrogates were used. A surrogate is a virus that is similar structurally to the virus of interest. Thus, other coronaviruses (CoVs) can be surrogates

for SARS-CoV-2 because they all belong to the same family. In order to reduce the chance of illness, it would be more desirable to work with an animal CoV as opposed to a human CoV. Besides CoVs, bacteriophages can also be used as surrogates for SARS-CoV-2 (Machado et al., 2021). Bacteriophages and SARS-CoV-2 are of similar size, and they are not harmful to the environment or humans (Machado et al., 2021). This study will use the surrogates bovine CoV (BCoV) and Phi6. Both SARS-CoV-2 surrogates have been used together as surrogates in a previous study. These surrogates were used to determine how effective the room was decontaminated when an ozone based decontamination system was applied which would then estimate how SARS-CoV-2 would act under the same conditions (Franke et al., 2021). These surrogates were chosen based on their similarity to SARS-CoV-2, their low cost, and their ability to be delivered quickly. Although Phi6 and BCoV can be used as surrogates for SARS-CoV-2, there is no guarantee that they will provide an accurate depiction of SARS-CoV-2 under all circumstances. To determine if these surrogates could be reliable, the frequency of the SARS-CoV-2 gene targets and surrogates were assessed. If the surrogates and SARS-CoV-2 gene targets have a similar pattern of detection, the surrogates would minimally be reliable to determine if a sample is fully inhibited. The percentage of retrieval will also be assessed for BCoV and Phi6 which gives insight into the degree of SARS-CoV-2 inhibition. The aim of this study is to estimate the frequency of inhibition the SARS-CoV-2 gene targets and to determine if Phi6 and BCoV have a similar pattern of detection as the SARS-CoV-2 gene targets.

5.3 Materials and Methods

5.3.1 Sample Preparation and Centrifugation

Our workflow used the methods described by Flood et al (Flood et al., 2021) as a starting point. Sewage samples were delivered to the lab in 500mL portions. Samples were transported in a cooler with icepacks while in transit to our lab. Samples were processed immediately upon arrival. The 500mL bottle of sewage was mixed by inverting the bottle between 40 and 60 times to minimize foaming. A 45 mL homogenous sewage sample was decanted into a 50mL conical bottom centrifuge tube. Approximately 4.0g of polyethylene glycol 8000 (catalog number 97061-100 from VWR) and 0.59g sodium chloride (catalog number 470302-520 from VWR) were added to the sample yielding a concentration of 8% polyethylene glycol and 0.2M sodium chloride. After the solids and solutions were combined, the centrifuge tubes were transferred to a rotator where the tubes were rotated at 70rpm for 2hrs. The rotator was held at 4°C. After mixing, the samples were centrifuged at 3,650 rpm for 1 hr. Once centrifugation was completed, the supernatant was discarded, and a 1-2 mL pellet remained.

5.3.2 Sample Purification and Elution

The QIAamp Viral RNA Mini Kit (catalog number 52904) was purchased from QIAGEN to purify and elute the RNA from the sewage samples. The pellet (200µL) was pipetted into a solution containing 800µL of Buffer AVL, 8µL of carrier RNA, 10µL of Phi6, and 10µL of BCoV (Bovilis® Coronavirus from Merrick Animal Health). Both Phi6 and BCoV had a concentration of 10^5 gene copies/mL. A 200µL portion of pellet was then pipetted into a 2mL cryovial containing the buffer and RNA mix. The cryovial was vortexed for 10s before incubation for 10mins at room temperature. After incubation,

200proof ethanol was added to the cryovial in an 800µL portion, and the cryovial was vortexed for 10s. At this point in the procedure there was 1.8mL solution in the cyrovial. This solution was run through the QIAamp mini spin column in order to obtain the extracted RNA. The 1.8mL solution was added in three 630µL quantities to the spin column. Three additions were done because the spin column could not hold enough solution for a single addition. Each 630µL portion of solution was added to the spin column and centrifuged sequentially. The flow through was not retained. The centrifugation settings were 1min at 8,000rpm. These centrifuge settings were the same for every addition except when the buffer AW2 was added. The waste collected in the collection tube under the spin column was discarded after every addition until the elution step (buffer AVE). After the whole 1.8mL solution was added to the spin column and centrifuged, the wash buffers were added. A 500µL portion of buffer AW1 was added to the spin column and centrifuged. Next, a 500µL portion of buffer AW2 was added to the spin column. At this point in the process (and only during this step) the centrifuge speed and duration was increased to 14,000rpm for 4min. After centrifugation, centrifugation settings were returned to normal, and the spin column was transferred to a 1.5mL microcentrifuge tube. The microcentrifuge tube was used to collect the RNA eluted from the spin column. Buffer AVE was added to the spin column in a 40µL portion and incubated at room temperature for 1min. After incubation, the spin column was centrifuged. This step was repeated which resulted in a final volume of 80µL of RNA eluate. This solution was tested for the SARS-CoV-2 gene targets and surrogates.

5.3.3 RNA Amplification and Detection of SARS-CoV-2 Targets and Surrogates

This eluate was pipetted along with controls (negative, extraction, positive, and no template) in triplicate. The negative control provided the maximum Phi6 and BCoV concentration. This control was made with water instead of the pellet and eluted through the spin column as usual. The external control was like the negative control, but buffer AVE was used instead of water. Both reactions in this study were ran as a duplex. The N1 and N2 gene targets were ran as a duplex reaction, and Phi6 and BCoV were ran as a duplex reaction. Because both reactions were run in a duplex, the assay mix was made the same way. The assay mix was composed of three different solutions provided in the 1-Step RT-ddPCR Advanced Kit for Probes from Bio-Rad, primer/probe (p/p) solution made up from primers and probes purchased from Integrated DNA Technologies or p/p solution obtained from Michigan State University, and water. Many labs in the Michigan PCR Network commonly test Phi6, N1, and N2 and use the primer/probe solutions obtained for Michigan State University. The sequences of the primers and probes used in this study are listed in Table 5.1.

The volumes of these reagents for one reaction were as follows: 1.1µL dithiothreitol, 1.1µL water, 3.3µL p/p solution 1, 3.3µL p/p solution 2, 5.5µL Supermix, and 2.2µL reverse transcriptase. The final volume of assay mix was 16.5µL of assay mix per well. An amount of 5.5µL of sample or control was added to the assay mix to yield a final volume of 22µL per well.

Table 5.1: Primer and Probe Sequences

Gene Target	Sequence	References
BCoV	5'-CTGGAAGTTGGTGGAGTT-3' (Forward) 5'-ATTATCGGCCTAACATACATC-3' (Reverse) 5'-HEX-CCTTCATA T-ZEN-CTATACACATCAAGTTGTT-IABKFQ-3'	(Graham et al., 2020)
Phi6	5'-TGGCGGCGGTCAAGAGC-3'(Forward) 5'-GGATGATTCTCCAGAAGCTGCTG-3' (Reverse) 5'- FAM-CGGTCGTCGCAGGTCTGACACTCGC-BHQ1-3'(Probe)	(Gendron et al., 2010)
N1	5'-GACCCCAAAATCAGCGAAAT-3' (Forward) 5'-TCTGGTTACTGCCAGTTGAATCTG-3' (Reverse) 5'-FAM-ACCCCGCATTACGTTTGGTGGACC-BHQ1-3' (Probe)	(CDC, 2020c)
N2	5'-TTACAAACATTGGCCGCAAA-3'(Forward) 5'-GCGCGACATTCCGAAGAA-3' (Reverse) 5'-HEX-ACAATTTGCCCCCAGCGCTTCAG-BHQ1-3' (Probe)	(CDC, 2020c)

Table 5.1: The primer and probes sequences used for N1, N2, BCoV, and Phi6.

Phi6 and BCoV, and SARS-CoV-2 N1 and N2 targets were run in triplicate. Once the assay mix and sample/control were combined by pipetting both into a 96-well plate, the plate was sealed with foil and vortexed for 15s. The plate was then put into a plate spinner to gather the liquid to the bottom of the plate. The 96-well plate was then placed into the Automated Droplet Generator where a new 96-well plate was added to hold the

droplets generated by the machine. This new plate was then put into the thermocycler. The settings for the thermocycler were as follows (temperature, time): setting 1 (25°C, 3min), setting 2 (50°C, 60min), setting 3 (95°C, 10min), setting 4 (consists of 40 cycles of alternating between 95°C for 30s then 55°C for 1min), setting 5 (98°C, 10min), and setting 6 (4°C until ready to read the plate). Thermocycling settings for the general settings were based on the document found on Bio-Rad's website for 1-Step RT-ddPCR Advanced Kit for Probes, and annealing temperatures for N1 and N2 were found in a research article (*One-Step RT-DdPCR Advanced Kit for Probes*, n.d.; Pearson et al., 2021). The annealing temperatures for BCoV and Phi6 were chosen as in Chapter 4. After thermocycling was complete, the plate was read by the QX200 Droplet Reader gene copies per 20µL well computed.

5.3.4 Data Analysis

The droplet reader was paired with QuantaSoft™ Software to score the data. The data obtained from this software provided information on the number of droplets (accepted, positive, and negative), concentration, and copies per 20µL well. The positive droplets were assessed to determine if the sample was positive for SARS-CoV-2. Three or more positive droplets were considered a positive hit. Accepted droplets were assessed to determine the validity of the values obtained from the QuantaSoft™ Software. Wells that had greater than 10,000 accepted droplets were considered acceptable. The gene copies per 20µL well were used to calculate the concentration of the various gene targets. Equation 5.1 shows the formula utilized to calculate the concentration of the gene targets. The concentration was in units of gene copies per liter of sewage (CP/L).

(Equation 5.1)

$$\frac{CP}{L} = copies\ per\ \mu L \times \frac{pellet\ volume\ (mL)}{extraction\ volume\ (0.2mL)} \times \frac{eluate\ volume\ (80\mu L)}{initial\ volume\ (45mL)} \times \frac{1000mL}{1L}$$

Equation 5.1: Concentration Calculation for the SARS-CoV-2 Gene Targets
This formula shows how the concentrations for the N1, N2, and E gene targets and the surrogates Phi6 and BCoV were calculated. Pellet volume varied between 1-2mL for the SARS-CoV-2 gene targets. Pellet volume was 45mL for BCoV and Phi6.

The negative control provided the maximum concentration that could be obtained. The concentrations of Phi6 and BCoV on a sample-by-sample basis were divided by the concentration of the negative control to provide a percentage of Phi6 and BCoV retrieval.

5.4 Results

This study assessed the percentage of retrieval of the SARS-CoV-2 surrogates Phi6 and BCoV, and the frequency of these surrogates compared to the N1 and N2 gene targets. Phi6 was implemented into routine testing with the N1 and N2 gene targets prior to BCoV. Thus, there were more samples tested for Phi6 than BCoV. The sampling areas in this study included Eastern Michigan University (EMU), Oakland University (OU), Clinton Township (CT), Warren (WN), Michigan Health Department (MCHD), Delhi, Grosse Ile (GILE), South Lyon (Lyon), Mount Clemons (Mtclms), New Baltimore (Baltimore), Richmond, Romeo, St. Clair, and Wixom. Some areas had multiple sampling sites. For instance, EMU had 12 unique sampling sites within the university. The Phi6 inhibition data was collected from March 9th, 2021 until July 19th, 2022 with a total of 3,389 samples derived from 44 unique sites. The BCoV inhibition data was collected from March 8th, 2022 until July 19th, 2022 and analyzed 42 unique sites.

However, frozen pellets were thawed for the sites WN5, Delhi, and St. Clair and tested for inhibition. These frozen pellets dated back to September 21st, 2021.

The amount of Phi6/BCoV retrieved from each sample was determined to assess inhibition. The amount of retrieval was determined by dividing the concentration of the surrogate obtained from the sewage by the concentration of the negative control. The negative control was run similarly to the standard procedure, but water was used instead of 200µL of sewage. As water should not be inhibiting the signal, the negative control represents the maximum signal that can be obtained from spiking in the surrogate. Concentrations obtained for the surrogates that were equivalent to the method detection limit were fully inhibited. For the full sample set, Phi6 was fully inhibited in 419/3,389 (12.4%) of samples. Of the 44 sites analyzed, most of the sites (43/44) had minimally three samples that were fully inhibited. One site never had a fully inhibited sample, but it only had 13 samples. The number of samples collected varied based on site.

The Phi6 data suggested that there is a high level of inhibition. In fact, a total of 668/3,389 (19.7%) of samples had less than 5% of the spiked Phi6 retrieved. The number of samples that had more than 50% of the spiked in Phi6 retrieved were 560/3,389 (16.5%) of samples. Therefore, there were more samples that had less than 5% of Phi6 retrieved than samples that had more than 50% of Phi6 retrieved. The 10-20% retrieval interval was the most populated interval for Phi6. A total of 684/3,389 samples were in the 10-20%. The 90-100% retrieval interval had the lowest number of samples (31/3,389). Interestingly, there were 139 samples that had more than 100% retrieval. Figure 5.1 is a bar graph that groups Phi6's retrieval into various intervals.

Figure 5.1: Phi6 Retrieval from Sewage Samples

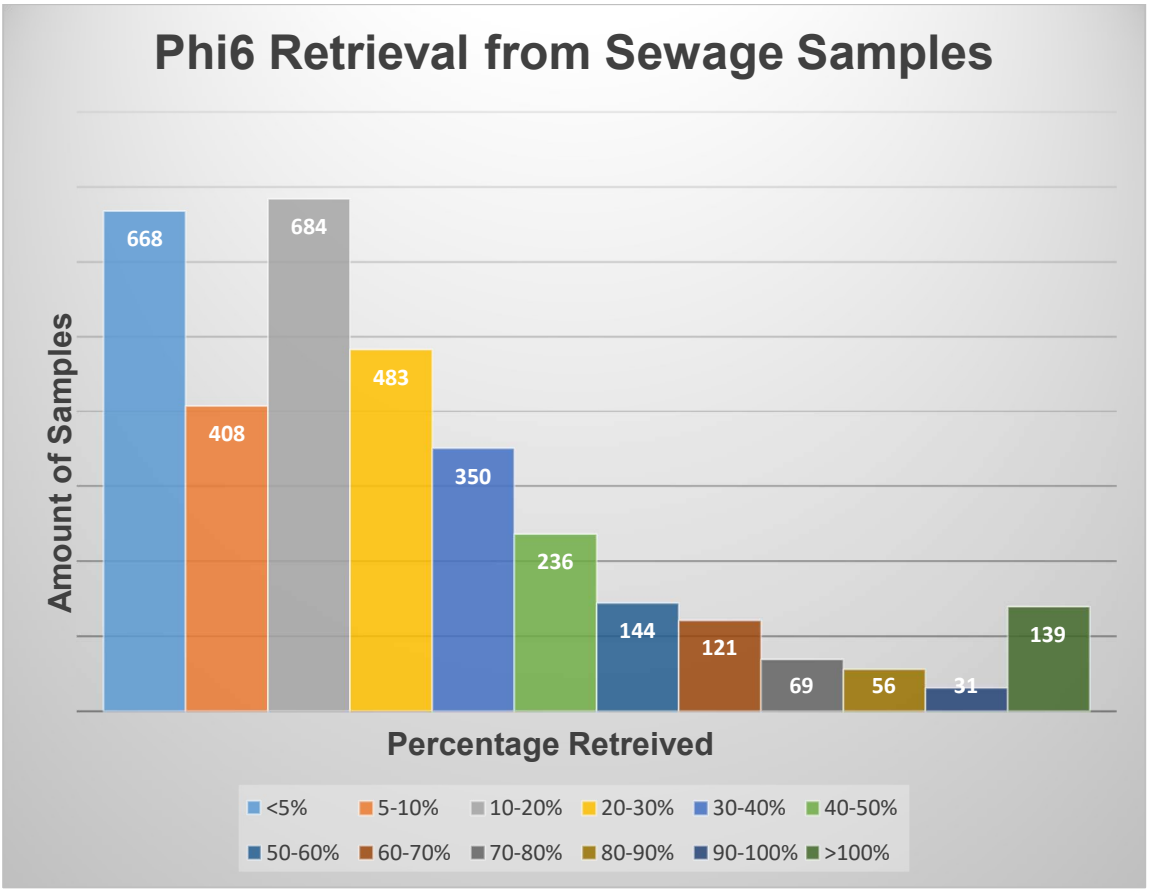


Figure 5.1: The percentage of Phi6 retrieved from 3,389 samples was analyzed. The 3,389 samples were comprised of 44 different sites. Each site had a various number of samples. The number of samples analyzed per site was based on the number of samples delivered to the lab throughout the duration of these experiments.

Overall, the percentage of Phi6 retrieved per interval was as follows: 668/3,389 (<5%), 408/3,389 (5-10%), 684/3,389 (10-20%), 483/3,389 (20-30%), 350/3,389 (30-40%), 236/3,389 (40-50%), 144/3,389 (50-60%), 121/3,389 (60-70%), 69/3,389 (70-80%), 56/3,389 (80-90%), 31/3,389 (90-100%), and 139/3,389 (>100%).

Table 5.2 shows the number of samples analyzed per site, the percentage of samples fully inhibited, and other site-specific data. Although all 44 sites were unique, some sites were collected from the same area. This means multiple sites can belong to the same city, township, or university. There were five areas that had more than one site, and nine areas that were only comprised of a single sampling site. Overall, EMU6 had the highest number of samples fully inhibited with full inhibition displayed in 49.2% of its samples. EMU12 had the highest percentage of samples with more than 50% of Phi6 retrieved (54.0%). This site also had the highest average percentage of Phi6 retrieval (65.1%). Overall, the average percentage of retrieval for all 3,389 samples was only 30.0% which indicated a high degree of inhibition across all sewage samples.

The percentage of BCoV retrieved was also assessed in this study. However, the amount of BCoV retrieved was only assessed in 1,072 samples opposed to the 3,389 samples analyzed for Phi6. The surrogate BCoV was only analyzed in 1,072 samples because it was introduced later in the study. This surrogate was tested over four months (March 8th, 2022 – July 19th, 2022). When BCoV retrieval was assessed, it was found that 62/1,072 (5.8%) of samples were fully inhibited.

Table 5.2: Site by Site Phi6 Retrieval

Site	Samples Fully Inhibited (%)	Samples with >50% Retrieval (%)	Average amount of Phi6 Retrieved (%)	Total Samples	City, Township, or University
CT1	4.72%	15.0%	30.1%	127	Township
CT2	6.15%	12.3%	26.3%	130	Township
CT3	12.0%	4.80%	18.5%	125	Township
CT4	19.4%	6.20%	14.5%	129	Township
CT5	8.50%	15.4%	26.4%	130	Township
CT6	30.5%	8.59%	15.4%	128	Township
CT7	3.92%	9.80%	21.3%	51	Township
EMU1	7.80%	32.8%	49.5%	64	University
EMU2	0.00%	38.5%	46.2%	13	University
EMU3	5.00%	21.7%	37.1%	60	University
EMU4	4.48%	29.9%	39.8%	67	University
EMU5	14.8%	9.84%	22.0%	61	University
EMU6	49.2%	4.62%	11.1%	65	University
EMU7	5.00%	33.3%	44.2%	60	University
EMU8	12.9%	24.2%	33.3%	62	University
EMU9	4.76%	31.7%	39.3%	63	University
EMU10	7.27%	9.10%	26.1%	55	University
EMU11	4.62%	30.8%	40.5%	65	University
EMU12	10.0%	54.0%	65.1%	50	University
OU1	8.22%	20.5%	40.6%	73	University
OU2	8.70%	20.3%	31.6%	69	University
OU3	9.72%	15.3%	30.9%	72	University
OU4	4.23%	22.5%	34.3%	71	University
OU5	21.9%	8.22%	18.5%	73	University
OU6	5.71%	8.57%	23.5%	70	University
OU7	8.47%	18.6%	30.7%	59	University
OU8	19.1%	10.6%	25.8%	47	University
WN1	9.10%	31.8%	52.5%	44	City
WN2	6.38%	27.7%	44.4%	47	City
WN3	8.33%	33.3%	55.9%	48	City
WN4	9.52%	26.2%	39.3%	42	City
WN5	6.96%	22.6%	38.8%	115	City
MCHD1	11.5%	24.1%	35.0%	87	City
MCHD2	10.8%	20.5%	28.5%	83	City
MCHD3	17.2%	9.20%	18.8%	87	City
Delhi	8.33%	7.29%	35.7%	96	Township
GILE	9.38%	9.38%	30.6%	32	Township

Table 5.2: Site by Site Phi6 Retrieval (Continued)

Site	Samples Fully Inhibited (%)	Samples with >50% Retrieval (%)	Average amount of Phi6 Retrieved (%)	Total Samples	City, Township, or University
Lyon	9.21%	13.2%	27.9%	76	City
Mtclms	14.9%	14.9%	29.2%	101	City
Baltimore	14.7%	9.47%	44.2%	95	City
Richmond	13.4%	15.5%	25.4%	97	City
Romeo	9.10%	19.3%	30.1%	88	City
St. Clair	16.0%	9.00%	22.4%	100	City
Wixom	36.5%	4.81%	12.6%	104	City

Table 5.2: This table shows the average percentage of retrieval, the number of samples analyzed per site, the percentage of samples fully inhibited, and the percentage of samples with more than 50% retrieval. All values were based on the addition of Phi6 to the 3,389 sewage samples. There were 44 sites analyzed from various areas. The sites that belonged to the same area varied numerically. Related areas include samples retrieved from the same city, township, or university.

Phi6 and BCoV were directly compared by removing the excess Phi6 data points and directly comparing the samples both surrogates had in common. The comparable Phi6 data had 198/1,072 (18.5%) samples that were fully inhibited. In contrast to Phi6, there were many sites that had no samples that were fully inhibited for BCoV. A total of 12/42 sites had no samples fully inhibited for BCoV. Only two sites had no samples that were fully inhibited for Phi6. The number of samples and the interval of Phi6 retrieval was as follows: 306/1,072 (<5%), 131/1,072 (5-10%), 192/1,072 (10-20%), 136/1,072 (10-20%), 100/1,072 (30-40%), 57/1,072 (40-50%), 36/1,072 (50-60%), 38/1,072 (60-70%), 14/1,072 (70-80%), 16/1,072 (80-90%), 9/1,072 (90-100%), and 37/1,072 (>100%).

Figure 5.2: BCoV Retrieval and Comparable Phi6 Retrieval from Sewage Samples

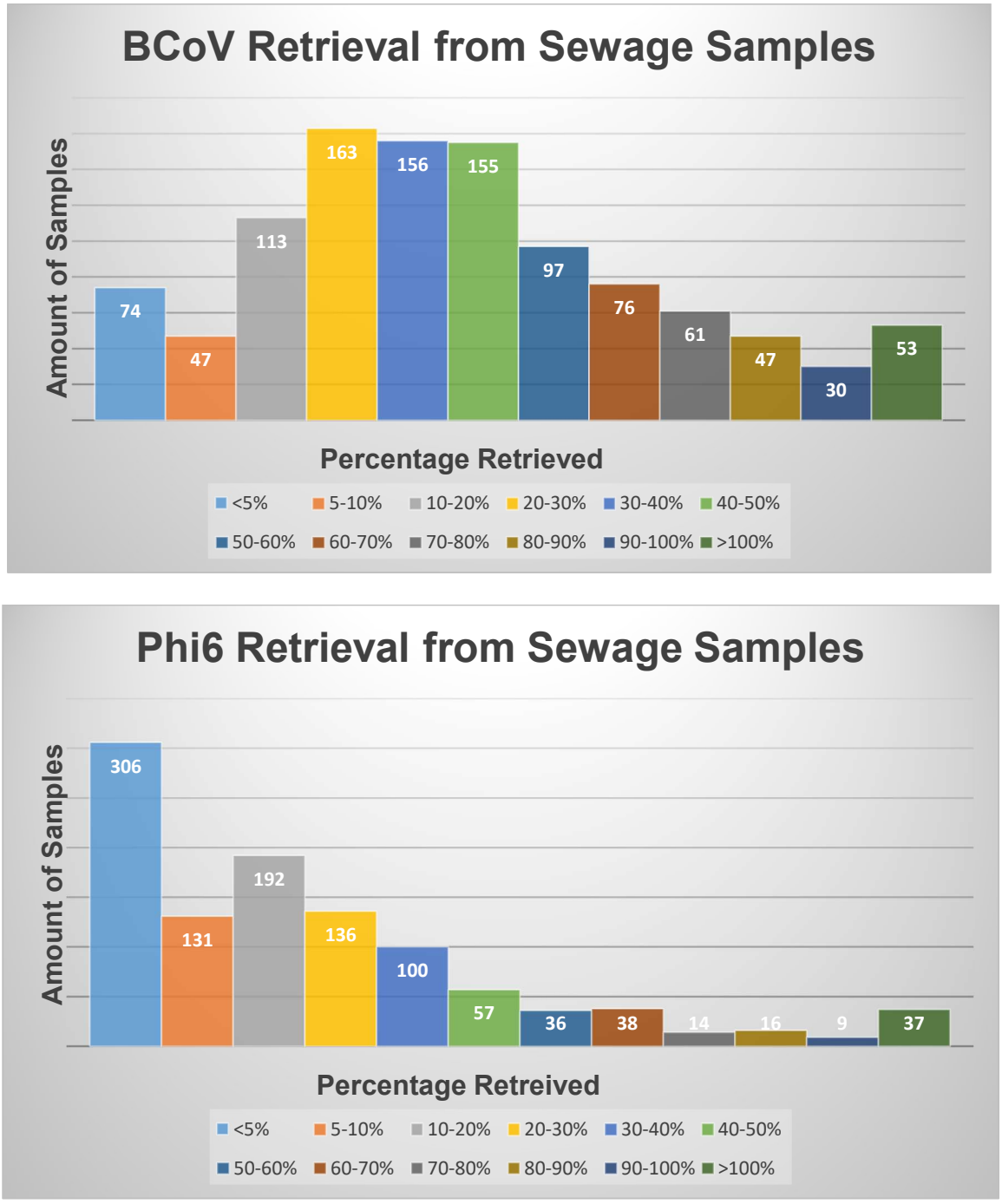


Figure 5.2: The percentage of BCoV and Phi6 retrieved from the 1,072 samples analyzed. The 1,072 samples were comprised of 42 different sites. The number of samples analyzed per site was based on what was delivered to the lab throughout the duration of these experiments.

The number of samples in each interval of retrieval were as follows for BCoV: 74/1,072 (<5%), 47/1,072 (5-10%), 113/1,072 (10-20%), 163/1,072 (20-30%), 156/1,072 (30-40%), 155/1,072 (40-50%), 97/1,072 (50-60%), 76/1,072 (60-70%), 61/1,072 (70-80%), 47/1,072 (80-90%), 30/1,072 (90-100%), and 53/1,072 (>100%). Figure 5.2 shows the graphical representation of this data. We found that only 6.9% (74/1,072) of samples had less than 5% of the spiked in BCoV retrieved with most of these samples (62/74) being fully inhibited. The interval with the highest number of samples was the 20-30% interval for BCoV (163/1,072 samples) and the <5% interval for Phi6 (306/1,072). Only 150/1,072 (14.0%) of the samples had over 50% Phi6 retrieval while 364/1,072 of BCoV's samples had more than 50% retrieval. Table 5.3 shows the amount of retrieval on a site-by-site basis for BCoV and the comparable Phi6 data.

The site with the most fully inhibited samples was EMU6 for both BCoV (41.2% of samples) and Phi6 (66.7% of samples). For BCoV and Phi6, the site with the highest number of samples over 50% retrieval (100% of samples) was GILE, but this site only had one sample. The next sites that had the highest number of samples with over 50% retrieval was WN1 for BCoV (90.1% of samples) and EMU12 for Phi6 (50.0%). The sites with the highest average retrieval were CT1 for BCoV (85.3%) and EMU12 for Phi6 (53.4%). Baltimore had the highest average Phi6 retrieval (70.1%), but it was excluded due to one sample having nearly 2000% retrieval with its removal leading to an average of 16.0% retrieval. Although the comparable Phi6 data had 68.4% less samples than the original Phi6 data, both the comparable and regular Phi6 data sets agreed that site EMU6 had the most fully inhibited samples, and site EMU12 had the highest amount of retrieval and samples over 50% Phi6 retrieval.

Table 5.3: Site by Site BCoV Retrieval and Comparable Phi6 Retrieval

Site	Samples Fully Inhibited (%)	Samples with >50% Retrieval (%)	Average Retrieval (%)	Total Samples	City, Township, or University
CT1	0.00% (8.11%)	37.8% (8.11%)	85.3% (24.0%)	37	Township
CT2	0.00% (10.5%)	28.9% (7.89%)	68.1% (21.6%)	38	Township
CT3	5.26% (18.4%)	15.8% (2.63%)	51.8% (13.4%)	38	Township
CT4	7.89% (18.4%)	10.5% (5.26%)	28.4% (12.2%)	38	Township
CT5	2.63% (15.8%)	36.8% (21.1%)	57.5% (24.3%)	38	Township
CT6	13.5% (27.0%)	16.2% (5.41%)	28.8% (12.2%)	37	Township
EMU1	0.00% (21.4%)	57.1% (28.6%)	64.7% (43.4%)	14	University
EMU3	18.2% (18.2%)	27.3% (9.10%)	35.4% (23.7%)	11	University
EMU4	7.14% (14.3%)	35.7% (28.6%)	57.5% (32.9%)	14	University
EMU5	7.14% (28.6%)	35.7% (0.00%)	65.0% (13.5%)	14	University
EMU6	41.2% (66.7%)	8.33% (8.33%)	14.1% (9.46%)	12	University
EMU7	7.69% (23.1%)	53.8% (23.1%)	59.4% (37.2%)	13	University
EMU8	27.3% (45.5%)	36.4% (9.10%)	40.8% (19.5%)	11	University
EMU9	7.69% (23.1%)	38.5% (30.8%)	48.7% (31.0%)	13	University
EMU10	0.00% (22.2%)	22.2% (11.1%)	39.0% (21.2%)	9	University
EMU11	7.14% (7.14%)	71.4% (21.4%)	71.7% (38.1%)	14	University
EMU12	7.14% (21.4%)	64.3% (50.0%)	63.0% (53.4%)	14	University
OU1	0.00% (13.3%)	40.0% (13.3%)	43.6% (27.2%)	15	University
OU2	7.69% (23.1%)	23.1% (15.4%)	42.5% (23.8%)	13	University

Table 5.3: Site by Site BCoV Retrieval and Comparable Phi6 Retrieval (Continued)

Site	Samples Fully Inhibited (%)	Samples with >50% Retrieval (%)	Average Retrieval (%)	Total Samples	City, Township, or University
OU3	6.67% (33.3%)	33.3% (0.00%)	33.1% (16.3%)	15	University
OU4	0.00% (0.00%)	42.9% (14.3%)	48.9% (26.1%)	14	University
OU5	18.8% (37.5%)	12.5% (12.5%)	26.6% (15.9%)	16	University
OU6	14.3% (14.3%)	28.6% (14.3%)	36.6% (20.9%)	13	University
OU7	6.25% (12.5%)	25.0% (12.5%)	39.9% (22.6%)	16	University
OU8	12.5% (25.0%)	37.5% (6.25%)	38.1% (23.7%)	16	University
WN1	0.00% (9.10%)	90.1% (36.4%)	76.5% (44.8%)	11	City
WN2	7.69% (7.69%)	69.2% (23.1%)	60.4% (47.7%)	13	City
WN3	14.3% (21.4%)	50.0% (28.6%)	51.8% (26.0%)	14	City
WN4	9.10% (18.2%)	62.7% (36.4%)	71.6% (44.9%)	11	City
WN5	4.80% (10.5%)	59.2% (25.0%)	57.9% (40.1%)	76	City
MCHD1	2.94% (11.8%)	35.3% (20.6%)	52.2% (28.2%)	34	City
MCHD2	5.71% (11.4%)	34.3% (20.0%)	47.3% (28.5%)	35	City
MCHD3	5.88% (20.6%)	17.6% (11.8%)	31.8% (20.0%)	34	City
Delhi	0.00% (10.5%)	30.3% (7.89%)	43.6% (38.9%)	38	Township
GILE	0.00% (0.00%)	100% (100%)	65.0% (195%)	1	Township
Lyon	0.00% (20.7%)	31.0% (6.89%)	51.9% (24.4%)	29	City
Mtclms	5.41% (21.6%)	37.8% (16.2%)	44.8% (29.0%)	37	City
Baltimore	0.00% (20.0%)	22.9% (5.71%)	37.3% (70.1%)	35	City

Table 5.3: Site by Site BCoV Retrieval and Comparable Phi6 Retrieval (Continued)

Site	Samples Fully Inhibited (%)	Samples with >50% Retrieval (%)	Average Retrieval (%)	Total Samples	City, Township, or University
Richmond	8.82% (20.6%)	26.5% (11.8%)	43.5% (22.5%)	34	City
Romeo	0.00% (13.8%)	31.0% (17.2%)	41.2% (30.0%)	29	City
St. Clair	1.23% (19.8%)	33.3% (9.88%)	42.9% (21.1%)	37	City
Wixom	21.1% (39.5%)	13.2% (7.89%)	26.3% (13.1%)	38	City

Table 5.3: This table shows various percentages based on BCoV retrieval and Phi6 retrieval in sewage samples. Phi6 data was listed in parentheses. Phi6 data was based on the same 1,072 samples tested for BCoV. There were 42 sites analyzed. Site CT7 and EMU2 were not sampled during this interval of time. Sites that have a relationship varied numerically. Related sites include samples retrieved from the same city, township, or university.

Overall, the average percentage of Phi6 and BCoV retrieval for all 1,072 samples was 27.8% (Phi6) and 47.2% (BCoV). Observing all 3,389 samples, Phi6 had an average percentage of retrieval of 30.0%.

Paired t-tests were run to compare the partial Phi6 data to BCoV. It was found that BCoV had a statistically higher number of samples that had more than 50% retrieval ($p=7.7x10^{-13}$) and a higher average percentage of retrieval ($p=3.5x10^{-4}$). As expected, Phi6 had a statistically higher percentage of samples that were fully inhibited ($p=4.4x10^{-14}$).

On a site-by-site basis it was shown that BCoV was favored to have more samples with over 50% retrieval (39/42 sites) and to have a higher average percentage of retrieval (41/42 sites). Phi6 was favored to have a higher percentage of samples fully inhibited (36/42 sites). The remaining sites had equivalent percentages for BCoV and Phi6 for all three categories excluding GILE that suggested Phi6 had a higher average percentage of

retrieval. That site only had one sample that was likely an error due to an unreasonably high amount of retrieval (nearly 200%). BCoV was found to be more robust than Phi6 due to the higher amount of retrieval and the lesser amount of inhibition.

The frequencies of Phi6, BCoV, N1, and N2 were also investigated. Phi6 was implemented earlier than BCoV and had more data points. Phi6 was compared to the N1 and N2 gene targets. There were 2,893 data points collected from 43 different sites for all three targets. This data was collected from June 1st, 2021 – July 19th, 2022. We found that 1,606/2,893 samples were positive for the N1 gene target, 1,495/2,893 samples were positive for the N2 gene target, and 2,512/2,893 samples were positive for Phi6. There were 1,732 detections for N1 and/or N2, and there were 2,661 detections for N1, N2, and/or Phi6. The average concentration of Phi6 and the SARS-CoV-2 gene targets were as follows: $7.3x10^4$ CP/L ($1.3x10^5$ CP/L without MDL values) for N1, $5.7x10^4$ CP/L ($1.0x10^5$ CP/L without MDL values) for N2, and $3.2x10^6$ CP/L ($3.7x10^6$ CP/L without MDL values) for Phi6. Table 5.4 shows the percentage of detections for each target at the various sites in this study.

A total of 12/43 sites had 100.0% of their detections positive for Phi6 meaning that Phi6 accounted for all the N1 and N2 positive detections in those 12 sites. Because Phi6 is an inhibition control used to give insight on the presence and the absence of the N1 and N2 gene targets, Phi6 should have a signal if the N1 and N2 gene targets are positive. In the remaining 31/43 sites Phi6 suggested that N1 and/or N2 were fully inhibited in at least one sample that was SARS-CoV-2 positive.

Table 5.4: Site by Site Frequency of N1, N2, and Phi6

Site	Percentage N1 Detection	Percentage N2 Detection	Percentage Phi6 Detection	Detections/Total Samples	City, Township, or University
CT1	79.8%	69.2%	96.2%	104/105	Township
CT2	78.8%	69.2%	95.2%	104/107	Township
CT3	80.2%	67.7%	89.6%	96/101	Township
CT4	74.5%	64.3%	88.8%	98/106	Township
CT5	74.3%	67.6%	93.3%	105/107	Township
CT6	70.0%	50.0%	78.9%	90/104	Township
CT7	25.9%	25.9%	100%	27/28	Township
EMU1	26.7%	31.1%	100%	45/50	University
EMU3	48.9%	48.9%	100%	45/47	University
EMU4	54.9%	51.0%	96.1%	51/52	University
EMU5	48.8%	46.5%	95.3%	43/49	University
EMU6	46.9%	40.6%	90.6%	32/50	University
EMU7	68.2%	65.9%	100%	44/47	University
EMU8	70.0%	65.0%	100%	40/48	University
EMU9	25.5%	31.9%	97.9%	47/49	University
EMU10	20.0%	27.5%	100%	40/43	University
EMU11	40.4%	42.6%	100%	47/50	University
EMU12	11.9%	11.9%	100%	42/47	University
OU1	44.2%	38.5%	98.1%	52/57	University
OU2	32.7%	34.7%	98.0%	49/53	University
OU3	36.5%	32.7%	96.2%	52/57	University
OU4	32.7%	30.8%	98.1%	52/54	University
OU5	56.9%	51.0%	88.2%	51/57	University
OU6	15.7%	19.6%	100%	51/55	University
OU7	35.2%	35.2%	100%	54/59	University
OU8	30.2%	30.2%	90.7%	43/48	University
WN1	54.8%	54.8%	93.5%	31/33	City
WN2	63.6%	54.5%	100%	33/36	City
WN3	15.2%	18.1%	100%	33/37	City
WN4	55.2%	62.1%	96.6%	29/32	City
WN5	86.5%	78.1%	94.8%	96/98	City
MCHD1	20.5%	21.8%	97.4%	78/86	City
MCHD2	52.6%	51.3%	96.1%	76/82	City
MCHD3	18.6%	21.4%	98.6%	70/84	City
Delhi	85.6%	80.0%	94.4%	90/93	Township
GILE	70.0%	70.0%	96.7%	30/32	Township
Lyon	84.7%	81.9%	93.1%	72/74	City
Mtclms	81.5%	81.5%	93.5%	92/100	City

Table 5.4: Site by Site Frequency of N1, N2, and Phi6 (Continued)

Site	Percentage N1 Detection	Percentage N2 Detection	Percentage Phi6 Detection	Detections/Total Samples	City, Township, or University
Baltimore	85.7%	79.1%	87.9%	91/94	City
Richmond	75.3%	75.3%	96.5%	85/94	City
Romeo	81.7%	76.8%	95.1%	82/86	City
St. Clair	92.1%	84.3%	92.1%	89/97	City
Wixom	80.8%	69.2%	82.1%	78/101	City

Table 5.4: This table shows how frequently the N1, N2, and Phi6 targets were present based on the number of detections not the total samples.

Overall, our data showed that Phi6 was not completely accurate at determining the presence and the absence of the N1 and N2 gene targets. However, for most sites Phi6 was able to account for over 90% of the detections meaning that Phi6 seemed to be a reasonable surrogate. Averaging 43 sites it was found that Phi6 accounted for 95.3% of the positive detections. However, the percentage of detections that Phi6 was able to detect ranged from 78.9-100.0%. Thus, Phi6 may be a better surrogate for the SARS-CoV-2 gene targets in some sites opposed to others. Phi6 was the worst surrogate for site CT6, and Phi6 accounted for all the detections for sites CT7, EMU1, EMU3, EMU7, EMU8, EMU10, EMU11, EMU12, OU6, OU7, WN2, and WN3. Overall, Phi6 estimated that 381/2,893 samples were fully inhibited. There was at least one positive target in 2,661/2,893 samples. The percentage of detections obtained (target detections/total detections) for each target were as follows: 60.3% (N1), 56.2% (N2), and 94.4% (Phi6). There were 929 (Phi6), 62 (N1), and 13 (N2) samples that were only positive for that specified target and no other target. There was a total of 149 (5.6%) samples out of 2,661 detections that were not positive for Phi6 but were positive for N1 and/or N2. As 1,732

samples were positive for the N1 and/or N2, Phi6 falsely labeled 8.6% of SARS-CoV-2 positive samples as fully inhibited.

The frequency of BCoV was assessed and compared to the frequency of the N1 and N2 gene targets. Again, Phi6 was directly compared to BCoV by removing additional data points that were not tested for BCoV. There was a total of 1,058 samples and 42 sites analyzed. This data was mainly collected from March 7th, 2022 to July 19th, 2022. Data starting from September 21st, 2021 was collected for WN5, Delhi, and St. Clair due to frozen pellets being thawed and tested for BCoV. For the comparable Phi6 data,we found that 927/1,058 samples were positive for either N1, N2, or Phi6. The BCoV data showed that 1,004/1,058 samples were positive for either N1, N2, or BCoV. Lastly, there were 646/1,058 samples that had a detection for either the N1 or N2 gene target. Overall, the number of detections per 1,058 samples were as follows: 619 (N1), 547 (N2), 862 (Phi6), and 996 (BCoV). Table 5.5 gives more detail on a site-by-site basis on the percentage of N1, N2, Phi6, and BCoV detections.

The data supported that BCoV was observed more frequently than Phi6. It was seen that BCoV was able to account for all the detections in 35/42 sites while only 21/42 sites had 100.0% Phi6 frequency. This shows that BCoV was more likely to be positive when the N1 and N2 gene targets were positive than Phi6. Averaging the sites, it was found that BCoV accounted for 99.2% of detections, and Phi6 accounted for 94.5% of detections. The percentage of detections ranged from 80.0-100% for Phi6 and 92.3-100.0% for BCoV. Site EMU6 had the lowest percentage of detection for Phi6, and site EMU9 had the lowest percentage detection for BCoV.

Table 5.5: Site by Site Frequency of N1, N2, BCoV, and Phi6

Site	Percentage N1 Detection	Percentage N2 Detection	Percentage BCoV (Phi6) Detection	Detections/Total Samples	City, Township, or University
CT1	77.8% (80.0%)	66.7% (68.6%)	100% (94.3%)	36/36 (35/36)	Township
CT2	86.5% (88.9%)	67.6% (69.4%)	100% (91.7%)	37/37 (36/37)	Township
CT3	91.7% (94.3%)	72.2% (74.3%)	97.2% (85.7%)	36/37 (35/37)	Township
CT4	70.6% (72.7%)	47.1% (48.5%)	100% (90.9%)	34/37 (33/37)	Township
CT5	75.0% (75.0%)	61.1% (61.1%)	100% (86.1%)	36/37 (36/37)	Township
CT6	72.7% (77.4%)	51.5% (54.8%)	93.9% (83.9%)	33/36 (31/36)	Township
EMU1	14.3% (18.2%)	14.3% (18.1%)	100% (100%)	14/14 (11/14)	University
EMU3	11.1% (11.1%)	11.1% (11.1%)	100% (100%)	9/11 (9/11)	University
EMU4	23.1% (23.1%)	15.4% (15.4%)	100% (92.3%)	13/14 (13/14)	University
EMU5	23.1% (30.0%)	15.4% (20.0%)	100% (100%)	13/14 (10/14)	University
EMU6	28.6% (40.0%)	0.00% (0.00%)	100% (80.0%)	7/12 (5/12)	University
EMU7	66.7% (80.0%)	50.0% (60.0%)	100% (100%)	12/13 (10/13)	University
EMU8	62.5% (83.3%)	37.5% (50.0%)	100% (100%)	8/11 (6/11)	University
EMU9	30.8% (36.4%)	23.1% (27.3%)	92.3% (90.1%)	13/13 (11/13)	University
EMU10	0.00% (0.00%)	0.00% (0.00%)	100% (100%)	9/9 (7/9)	University
EMU11	38.5% (38.5%)	30.8% (30.8%)	100% (100%)	13/14 (13/14)	University
EMU12	0.00% (0.00%)	7.69% (9.10%)	100% (100%)	13/14 (11/14)	University
OU1	33.3% (35.7%)	33.3% (35.7%)	100% (92.9%)	15/15 (14/15)	University
OU2	25.0% (30.0%)	33.3% (40.0%)	100% (100%)	12/13 (10/13)	University

Table 5.5: Site by Site Frequency of N1, N2, BCoV, and Phi6 (Continued)

Site	Percentage N1 Detection	Percentage N2 Detection	Percentage BCoV (Phi6) Detection	Detections/Total Samples	City, Township, or University
OU3	26.7% (36.4%)	20.0% (27.3%)	93.3% (90.1%)	15/15 (11/15)	University
OU4	28.6% (28.6%)	21.4% (21.4%)	100% (100%)	14/14 (14/14)	University
OU5	57.1% (66.7%)	35.7% (41.2%)	92.9% (83.3%)	14/16 (12/16)	University
OU6	33.3% (33.3%)	16.7% (16.7%)	100% (100%)	12/14 (12/14)	University
OU7	6.67% (7.14%)	13.3% (14.3%)	100% (100%)	15/16 (14/16)	University
OU8	7.14% (8.33%)	7.14% (14.3%)	100% (100%)	14/16 (12/16)	University
WN1	36.4% (40.0%)	36.4% (40.0%)	100% (100%)	11/11 (10/11)	City
WN2	66.7% (66.7%)	66.7% (66.7%)	100% (100%)	12/13 (12/13)	City
WN3	16.7% (18.2%)	25.0% (27.3%	100% (100%)	12/14 (11/14)	City
WN4	60.0% (60.0%)	80.0% (80.0%)	100% (90.0%)	10/11 (10/11)	City
WN5	94.6% (95.9%)	87.8% (89.0%	98.6% (93.2%)	74/75 (73/75)	City
MCHD1	9.38% (10.3%)	18.8% (20.7%)	100% (100%)	32/33 (29/33)	City
MCHD2	40.6% (43.3%)	40.6% (43.3%)	100% (100%)	32/34 (30/34)	City
MCHD3	9.67% (11.5%)	9.67% (11.5%)	100% (100%)	31/33 (26/33)	City
Delhi	89.3% (93.1%)	84.0% (87.5%)	100% (93.1%)	75/75 (72/75)	Township
GILE	100% (100%)	100% (100%)	100% (100%)	1/1 (1/1)	Township
Lyon	86.2% (92.6%)	79.3% (85.2%)	100% (85.2%)	29/29 (27/29)	City
Mtclms	77.8% (87.5%)	72.2% (81.3%)	97.2% (90.6%)	36/37 (32/37)	City
Baltimore	82.9% (90.6%)	74.3% (81.2%)	100.0% (87.5%)	35/35 (32/35)	City

Table 5.5: Site by Site Frequency of N1, N2, BCoV, and Phi6 (Continued)

Site	Percentage N1 Detection	Percentage N2 Detection	Percentage BCoV (Phi6) Detection	Detections/Total Samples	City, Township, or University
Richmond	61.3% (70.4%)	61.3% (70.4%)	100% (100%)	31/34 (27/34)	City
Romeo	62.1% (69.2%	55.2% (61.5%)	100% (96.2%)	29/29 (26/29)	City
St. Clair	88.5% (97.2%)	82.1% (90.1%)	100% (90.1%)	78/79 (71/79)	City
Wixom	79.3% (85.2%)	69.0% (72.1%)	100% (81.5%)	29/37 (27/37)	City

Table 5.5: The frequency of the N1, N2, BCoV and Phi6 targets present based on the detections and not the total samples. The BCoV data is not in parenthesis, and the comparable Phi6 data is in parenthesis.

We also compared BCoV and Phi6 side by side to determine which target was more likely to be present if only one surrogate was detected. Additionally, the same comparison was made between the SARS-CoV-2 gene targets. We also determined how many samples were positive for both surrogates at the same time or both SARS-CoV-2 gene targets at the same time. These data are listed in Figure 5.3.

It was found that BCoV was the only surrogate to have signal in 135 samples while only one sample was positive for Phi6 but negative for BCoV. Next, The N1 gene target was more likely to be the sole SARS-CoV-2 gene target present. There were 99 samples that only the N1 gene target was present, and 27 samples that the N2 gene target was the only target present. A high number of samples (861) were positive for both BCoV and Phi6 and around half of the samples (520) were positive for both SARS-CoV-2 gene targets.

Figure 5.3: Number of Positive Samples

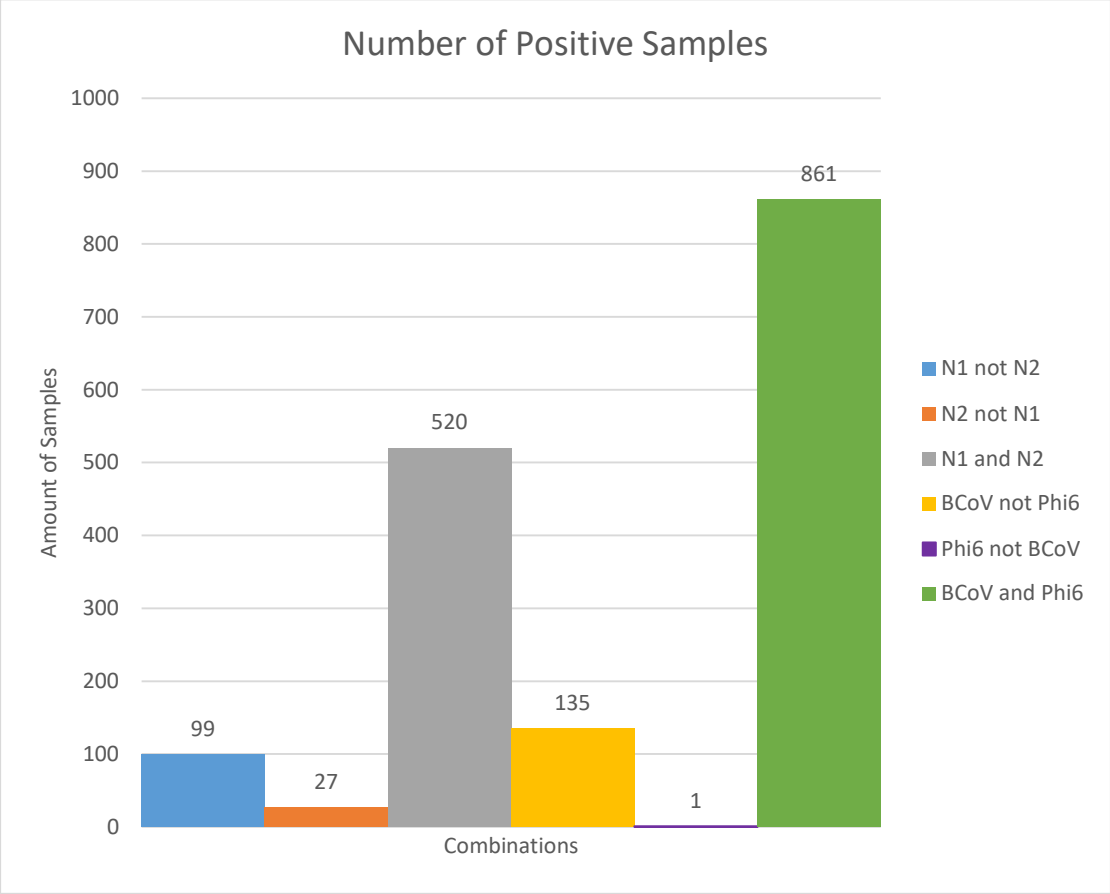


Figure 5.3: The number of samples positive for only BCoV, only Phi6, or Phi6 and BCoV and the number of samples positive for only N1, only N2 or N1 and N2. Surrogate and SARS-CoV-2 gene target data were treated as separate events. For example, positive for BCoV not Phi6 means that there is no Phi6 signal, but there could be N1 and/or N2 gene target signal.

Overall, BCoV had a higher concentration than Phi6, and N1 had a higher concentration than N2. The surrogates and gene targets average concentrations were as follows: $7.2x10^4$ CP/L ($1.2x10^5$ CP/L without MDL values) for N1, $5.8x10^4$ CP/L ($1.0x10^5$ CP/L without MDL values) for N2, $5.9x10^6$ CP/L ($6.2x10^6$ CP/L without MDL values) for BCoV, and $3.4x10^6$ CP/L ($4.1x10^6$ CP/L without MDL values) for Phi6. A paired t-test was run to determine which surrogate and SARS-CoV-2 gene target had the highest concentration. BCoV had a statistically higher concentration than Phi6 ($5.9x10^{-84}$), and N1 had a statistically higher concentration than N2 ($p=6.3x10^{-12}$). There was a Pearson's correlation coefficient of 0.59 for Phi6 and BCoV demonstrating a moderate positive correlation, and a Pearson's correlation coefficient of 0.99 for N1 and N2 indicating an extremely high positive correlation. Next, we also found that BCoV was typically detected more frequently or at the same frequency as Phi6. For the SARS-CoV-2 gene targets, 24/43 sites favored the N1 gene target, and 10/43 sites favored the N2 gene target. Lastly, besides having a higher concentration than Phi6, BCoV was also more consistent with the N1 and N2 gene target data. The number of SARS-CoV-2 positive samples identified as fully inhibited were as follows: 149/1,732 (full Phi6 data set), 65/646 (comparable Phi6 data set), and 8/646 (BCoV). Additionally, there were 237/1,732 samples that were positive for N1 but not N2 and 126/1,732 samples positive for N2 but not N1, indicating both targets should be analyzed.

5.5 Discussion

Sewage surveillance is a crucial tool to detect SARS-CoV-2. The N1 and N2 gene targets can detect the presence of SARS-CoV-2, but these targets could wrongfully suggest that the sample is negative when compounds interfere with the targets ability to bind to the

spin column, or these compounds instead interfere with target amplification. PCR inhibitors detrimentally impact PCR when they inhibit the DNA polymerase, form a bond with the template DNA, and chelate magnesium which is vital to the DNA polymerase (Acharya et al., 2017). Certain compounds that can harmfully interfere with the PCR reaction are as follows: humic acid, Immunoglobulin G, calcium chloride, ferric chloride, proteins, bile, fats, EDTA, etc. (Acharya et al., 2017). Water contributions in the sewage dilute target RNA, and low SARS-CoV-2 detection in feces make the SARS-CoV-2 signal in sewage harder to detect than a regular clinical specimen. Therefore, PCR inhibitors in a sample with a low SARS-CoV-2 concentration could cause complete signal loss leading to false negatives.

The SARS-CoV-2 surrogates chosen in this study were BCoV and Phi6. Phi6 was chosen as a surrogate for SARS-CoV-2 because of its similarly sized spike proteins (~80-100nm), its similar structure to SARS-CoV-2, it's safe, it's cheap, and it has been used to model other viruses that are enveloped including SARS-CoV-2 (Gomes et al., 2022). Some studies have suggested that Phi6 is a good surrogate for SARS-CoV-2. For instance, String et al. concluded that Phi6 was inactivated with chlorine at a similar rate as SARS-CoV-2 and declared Phi6 can act as a conservative surrogate for the virus SARS-CoV-2 when disinfection via chlorine is analyzed (String et al., 2021). However, other studies have suggested Phi6 is not an appropriate surrogate for SARS-CoV-2. Weyersberg et al. found that Phi6 was more resistant to ultraviolet than SARS-CoV-2 regardless if it was ultraviolet A, B, or C (Weyersberg et al., 2022). They suggested that the single-stranded RNA of SARS-CoV-2 could be more sensitive to ultraviolet than the double stranded RNA of Phi6, and that SARS-CoV-2 is almost three times longer in size

(30kbp) than Phi6 (13.5kbp) which may contribute to Phi6 being more resilient against UV than SARS-CoV-2 (Weyersberg et al., 2022).

As SARS-CoV-2 is a CoV, other CoVs would be expected to be suitable surrogates for the various SARS-CoV-2 gene targets. Therefore, the second surrogate chosen for this study was a CoV. Interestingly, in the String et al. study, the CoV murine hepatitis virus was not a good surrogate while the bacteriophage Phi6 was a conservative surrogate for SARS-CoV-2. Therefore, murine hepatitis virus was not chosen for this study. Due to safety concerns, no HCoVs were chosen in this study. Instead, an animal CoV called BCoV was chosen. BCoV is low-cost, easy to obtain quickly, and has been used in other studies as a surrogate to SARS-CoV-2. One study found that BCoV only had 17.3% recovery in 40mL of sewage using the electronegative membrane filtration method (Juel et al., 2021). A study by Franke et al. used both BCoV and Phi6 as surrogates for SARS-CoV-2 to determine if ozone was able to effectively decontaminate rooms (Franke et al., 2021).

To estimate the degree of SARS-CoV-2 inhibition, BCoV and Phi6 retrieval were assessed. Retrieval was determined by dividing the surrogate concentration obtained from each sample by the maximum concentration of surrogate. The maximum value was the concentration of surrogate obtained when water was added instead of sewage pellet. There was a total of 1,072 samples that were derived from 42 different sites. Overall, Phi6 had the most samples fully inhibited (36/42) while BCoV had more samples with over 50% retrieval (39/42) and the highest average retrieval (43/44). Statistically, BCoV was superior to the Phi6 sample set for all these categories. The p-values were as follows: Phi6 had more samples fully inhibited ($p=4.4x10^{-14}$), BCoV had more samples over 50%

retrieval ($p=7.7x10^{-13}$), and BCoV had a higher average percentage of retrieval ($p=3.5x10^{-4}$). The interval of retrieval with the most samples was also different for Phi6 and BCoV with Phi6s most populated interval being <5% (306/1,072) and BCoVs most populated interval being 20-30% (163/1,072). The surrogate BCoV only had 6.9% (74/1,072) samples in the <5% range, and 83.8% of these samples were fully inhibited. Overall, Phi6 suggested a much higher degree of inhibition with only 14.0% of its samples having >50% retrieval compared to BCoV that had 34.0% of its samples with over 50% retrieval.

Samples that exceeded 100% retrieval did occur in this study. There were 37/1,072 samples for Phi6, and 53/1,072 samples for BCoV that exceeded 100%. This can be attributed to variation in the triplicate wells leading to percentages slightly over 100% retrieval. However, when sites suggest a percentage of retrieval that is significantly higher than 100%, it suggests that the negative control received less surrogate due to pipetting error, or the RNA did not properly bind to the spin column due to a faulty spin column. If the RNA has issues binding to the spin column due to production issues or degraded 200proof ethanol, the RNA yield will be significantly lower.

There are three distinct sample collection areas in this study – townships, cities, and universities. When averaging the values obtained from all the city sites, only 5.8% of samples were fully inhibited, 40.1% of samples had over 50% retrieval, and 49.2% was the average amount of BCoV retrieved. For Phi6 the values for the city were 17.8% (fully inhibited), 13.3% (samples over 50% retrieval), and 32.7% (average retrieval). The percentages obtained when the township values were averaged were as follows: 3.7% (fully inhibited), 34.5% (samples over 50%), and 53.6% (average retrieval) for BCoV and

13.6% (fully inhibited), 19.8% (samples over 50%), and 42.6% (average retrieval) for Phi6. The percentages obtained when the university values were averaged were as follows: 10.4% (fully inhibited), 36.5% (samples over 50%), and 45.7% (average retrieval) for BCoV and 23.7% (fully inhibited), 16.3% (samples over 50%), and 26.3% (average retrieval) for Phi6. Overall, the university sites had a higher percentage of fully inhibited samples than the city and township sites for both Phi6 and BCoV. The location for samples with more than 50% retrieval varied for BCoV (city) and Phi6 (township). For both BCoV and Phi6, the township sites had the highest average percentage of retrieval. Overall, there does not appear to be a big difference between areas. On a site-by-site basis the amount of inhibition faced by the surrogates varied. Both BCoV (41.2%) and Phi6 (66.7%) agreed that the site EMU6 had the highest number of samples fully inhibited. However, BCoV and Phi6 disagreed on which site had the most samples over 50% and the highest average percentage of retrieval. Phi6 supported that EMU12 had the highest number of samples over 50% retrieval (50.0%) and the highest overall retrieval (53.4%). BCoV supported that WN1 had the highest number of samples over 50% retrieval (90.1%), and CT1 the highest percentage of retrieval (85.3%). Overall, the amount of inhibition/retrieval varies on the sewage conditions which will fluctuate based on season, weather, etc. Human habits can cause the sewage matrix to drastically change. Also, weather and faulty sewers can lead to large quantities of water entering the sewage system which could reduce inhibition by reducing compounds interfering with the PCR reaction or competing to bind on the spin column silica.

There was more Phi6 data due to Phi6 being implemented earlier than BCoV. Comparing the full Phi6 data set to the partial Phi6 data set we found that the average percentage of

retrieval was 30.0% (27.8% for partial data set), and the most populated interval was 10-20% with 684/3,389 samples (<5% with 306/1,072 samples for partial data set). The <5% retrieval range comprised of 668 samples for the data full set, and the partial Phi6 data sets second highest interval was the 10-20% range. Additionally, both sets of Phi6 data agreed that EMU6 had the highest amount of fully inhibited samples, and EMU12 had the highest number of samples with more than 50% Phi6 retrieval and the highest average Phi6 retrieval. Therefore, even though there are more than 2,000 additional samples, both Phi6 data sets acted similarly indicating that behaviors impacting sewage content were somewhat consistent throughout the study.

Overall, the data in this study showed that Phi6 is more likely to be inhibited and inhibited to a higher degree than BCoV. BCoV was only fully inhibited in 62/1,072 samples while Phi6 was fully inhibited three times more than BCoV (198/1,072). The average amount of retrieval was 27.8% for Phi6 and 47.2% for BCoV. However, although BCoV is more resilient than Phi6 does not make it a more accurate surrogate. A reliable surrogate should have a similar pattern of frequency compared to the target of interest.

To determine if the chosen surrogates could be reliable surrogates for SARS-CoV-2, the frequency of the N1, N2, Phi6, and BCoV were observed. Of the 1,058 samples analyzed, there were 619 detections for N1, 547 detections for N2, 862 detections for Phi6, and 996 detections for BCoV. For any combination of N1, N2, or Phi6 there were 927 detections, but when any combination of BCoV, N1, and N2 were observed there was 1,004 detections. It is expected that the surrogates would always be present unless there is an extremely high degree of inhibition rendering the sample fully inhibited. Phi6 was

detected in 81.5% of the total samples, and BCoV was detected in 94.1% of the total samples. Thus, Phi6 suggested that nearly 13% more of samples were fully inhibited than BCoV did. Overall, BCoV accounted for 99.2% of detections, and Phi6 accounted for 94.5% of detections. Phi6 and BCoV were then compared to determine if one surrogate was more consistent with the frequency of the N1 and N2 gene targets. There were eight SARS-CoV-2 positive samples that BCoV suggested were fully inhibited, and there were 65 SARS-CoV-2 positive samples that Phi6 suggested were fully inhibited. There were 646 SARS-CoV-2 positive samples indicating that BCoV missed 1.2% of the SARS-CoV-2 positive samples while Phi6 missed 10.1% of the SARS-CoV-2 positive samples. The full set of Phi6 data suggested that Phi6 was slightly better at detecting SARS-CoV-2 positive samples missing 8.6% (149/1,732) of SARS-CoV-2 positive samples. Of the total 1,058 samples, there were 135 samples that were positive for BCoV and negative for Phi6, and only one sample that was positive for Phi6 and negative for BCoV.

The one sample positive for Phi6 and not for BCoV had no signal for the N1 or N2 gene targets. Therefore, it is possible that the sample was fully inhibited as suggested by BCoV. Overall, these data showed that BCoV was more consistent with the SARS-CoV-2 gene targets than Phi6, and that BCoV was a reliable surrogate for detecting full inhibition in the SARS-CoV-2 gene targets. Every sample that is fully inhibited should be modified to reduce inhibition and rerun to determine if the sample is truly negative. If only Phi6 was assessed for inhibition, significantly more samples would have been rerun due to the higher number of samples reported as fully inhibited. Phi6 and BCoV could be duplexed because this would not require additional time and would have minimal additional cost. However, Phi6 appears to perform poorly and is not recommended. One

reason that Phi6 may be inferior to BCoV could be that Phi6 has double-stranded RNA while SARS-CoV-2 and other CoVs have single-stranded RNA. Therefore, we suggest that CoVs surrogate should be assessed first for new studies revolving around SARS-CoV-2.

The frequency of each target was also assessed for township, city and university. The frequencies obtained at the city locations were as follows: 58.2% (N1), 57.2% (N2), 99.7% (BCoV), and 94.3% (Phi6). The township sites analyzed had the following frequencies: 83.0% (N1), 68.8% (N2), 98.9% (BCoV), and 90.7% (Phi6). The university sites had the following frequencies: 27.2% (N1), 20.3% (N2), 98.9% (BCoV), and 96.2% (Phi6). The township sites had the highest amount of both N1 and N2 gene target detections, the city sites had the highest amount of BCoV detections, and the university sites had the highest amount of Phi6 detections. Using the full data set the same trend was observed despite there being one additional township site. For the N1 gene target, the full data set showed that the township data had the highest frequency (71.1%) followed by the city (63.2%) and university sites (41.4%). Similarly, the N2 gene target had the highest frequency at the township sites (61.7%) followed by the city (60.7%) and the university sites (40.8%). Lastly, the university sites had the highest amount of Phi6 frequency (97.6%) followed by the city (94.5%), and the township (92.1%) sites. Regardless of the area analyzed, the N1 gene target was always found at a higher frequency than the N2 gene target with the largest difference in frequency found in the township data set with a >14% difference (comparable data set) or a >9% difference (full data set). This data also showed that the township locations were more likely to have individuals with COVID-19 while the university sites were least likely to have individuals with COVID-19. The

sample set with 1,058 samples showed that the university sites had over 40% less SARS-CoV-2 detection than the township sites and over 30% less SARS-CoV-2 detection than the city sites for both N1 and N2 gene targets. For the full sample set, there was typically around 20% or more detection in the city and township sites compared to the university sites. For both comparable and full data sets, the frequency of the surrogates was similar (within 6%) regardless of the site being a city, township, or university location.

The differences in detection for the SARS-CoV-2 gene targets amongst the different areas can be attributed to their location and precautions taken to prevent the spread of COVID-19. If sewage is taken from or near a COVID-19 hot spot, there will be a higher amount of SARS-CoV-2 gene target detected opposed to areas that are further away from COVID-19 hot spots. Samples taken from highly populated areas would be expected to have higher SARS-CoV-2 levels. In contrast, samples collected in more rural areas would be expected to have less SARS-CoV-2 in their sewage. Lastly, areas with stricter COVID-19 guidelines will be expected to have less SARS-CoV-2 signal. Areas with more businesses that impose mask wearing and with better air ventilation systems would likely have less SARS-CoV-2 in their sewage. The universities in this study had stricter guidelines than the other areas in Michigan. Both universities in this study had mask mandates which were imposed throughout the duration of this study, students on campus needed to be vaccinated, and decisions were made based off our data. If high levels of SARS-CoV-2 were found in the sewage of a particular dormitory, the individuals in that building would be tested and sick individuals would be quarantined. Therefore, the universities had less SARS-CoV-2 signal due to their stricter COVID-19 guidelines.

The N1 and N2 gene targets were also compared to determine if one gene target was superior in detecting SARS-CoV-2 than the other gene target. Interestingly, although they are derived from the same gene, the N1 gene target was favored in this study. Comparing the frequency at each site, the N1 gene target was favored at 24/43 sites, the N2 gene target was favored at 10/43 sites, and the remaining 9/43 sites had equivalent N1 and N2 frequency. Thus, this data shows that the N1 gene target was the favored gene target in the sewage samples tested in our lab. The composition of all sewage samples is unique meaning that a sample collected from the same site could have a different composition the next day it is sampled. There is potential to remove a target on a site-by-site basis if there is a trend of data supporting that one target is favored over the other and does not provide additional information or very limited benefits. However, removing a gene target does not change much experimentally. Both gene targets can be run in a duplex which will provide data on both the N1 and N2 gene targets for essentially the same cost and time**.** If the N2 gene target was removed from analysis, 7.3% of SARS-CoV-2 detection would be lost loss while removing the N1 gene target would result in 13.7% SARS-CoV-2 signal loss. The SARS-CoV-2 signal is subjected to a high amount of dilution from the sewage and subject to many inhibitors within the sewage. This can cause the actual concentration to be greatly diminished. Dilution in addition to the low frequency of SARS-CoV-2 in the feces can lead to a non-detectable concentration of SARS-CoV-2 in the sewage resulting in false negatives. Therefore, to have a higher probability of detecting SARS-CoV-2 in sewage, it is suggested that both targets remain in use. A different study also looked at the N1 and N2 gene targets, and their data also showed that the N1 gene target could have a higher concentration and a higher amount of

detection than the N2 gene target. Our study suggested that there was a high correlation between the N1 and N2 gene targets (Pearson correlation coefficient of 0.99) despite the N1 gene target having a higher concentration and frequency than the N2 gene target. Therefore, the SARS-CoV-2 gene target which is preferred will change based on what is sampled, the location of the sample, and collection date. Thus, it is suggested that both the N1 and N2 gene targets be implemented into SARS-CoV-2 analysis in the sewage.

5.6 Conclusion

Overall, the inhibition data suggested that the surrogate BCoV was superior to Phi6. The average amount of retrieval was 47.2% for BCoV and 27.8% for Phi6. Statistically, Phi6 had more samples that were fully inhibited ($p=4.4x10^{-14}$) while BCoV had a higher number of samples over 50% retrieval ($p=7.7x10^{-13}$) and a higher average amount of retrieval ($p=3.5x10^{-4}$). The surrogate Phi6 had 198/1,072 samples fully inhibited while BCoV only had 62/1,072 of its samples fully inhibited. The frequency data also suggested BCoV was superior to Phi6. As suggested by the inhibition data, BCoV was detected more frequently than Phi6. BCoV had a statistically higher frequency than Phi6 ($p=5.7x10^{-84}$). The frequency of the BCoV data was more consistent with the frequency of the SARS-CoV-2 gene targets. Out of the 2.893 samples, Phi6 was fully inhibited in 149/1,732 (8.6%) samples that were SARS-CoV-2 positive. The comparable Phi6 data had a similar trend indicating 65/646 (10.1%) of SARS-CoV-2 positive samples were fully inhibited. The surrogate BCoV only missed 8/646 (1.2%) of SARS-CoV-2 positive samples. It was also determined that the N1 gene target was superior to the N2 gene target for detecting SARS-CoV-2 in the sewage. As the gene targets can be duplexed, the only additional cost to analyze both gene targets would be the primer/probe solution.

Therefore, our data suggests that both SARS-CoV-2 gene targets should be analyzed along with BCoV.

CHAPTER SIX

DILUTION, PASTEURIZATION, AND CHEMICAL ADDITIONS TO SEWAGE TO REDUCE PCR INHIBITION WHEN DETECTING THE SARS-COV-2 N1 AND N2 GENE TARGETS

6.1 Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an RNA virus that has quickly spread across the globe infecting over 100 million people with a disease known as coronavirus disease 2019 (COVID-19) in a short period of time. Monitoring the virus is essential to avoid mass outbreaks and to reduce the spread in areas of high transmission by putting precautions into place. Clinical monitoring requires participation from each individual. Monitoring SARS-CoV-2 via the sewage does not require consent from each individual and is able to monitor specific buildings or areas within a single sample. Sewage surveillance signal often precedes data from a clinical setting by one to two weeks and provides data on asymptomatic cases or unreported cases. This work monitored SARS-CoV-2 by detecting the N1 and N2 gene targets. Samples were modified to assess if samples that were partially/fully inhibited according to the surrogate data could be uninhibited/have increased retrieval after minor changes to the typical procedure. This study tested the addition of bovine serum albumin (BSA), dilution, the addition of polyvinylpyrrolidone (PVP), and pasteurization to determine which method modifications would enhance the detection of the N1 and N2 gene targets. It was found that adding BSA or PVP to the assay mix did not enhance the N1 or N2 gene target signal. When modifying the initial 45mL of sewage, it was found that BSA was the only

alteration that enhanced the N1 gene target's concentration, while dilution, BSA, and pasteurization enhanced the N2 gene target's concentration. Therefore, the enhancements were more beneficial for the N2 gene target. The Phi6 and BCoV data were not consistently enhanced along with the SARS-CoV-2 data indicating that the surrogates may only be suited to determine if a sample is fully inhibited.

6.2. Introduction

Coronavirus disease 2019 (COVID-19) is a disease which can be deadly to humans. This disease is caused by a single-stranded and enveloped RNA virus known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (V'kovski et al., 2021). There are currently only seven other coronaviruses (CoVs) that infect humans with the first human CoV emerging in the 1960s (X. Zhao et al., 2020). There are also CoVs that infect various animals such as: pheasant CoV, bovine CoV (BCoV), rat CoV, canine CoV, murine hepatitis virus (MHV), feline CoV, turkey CoV, porcine transmissible gastroenteritis virus, and many more (Alluwaimi et al., 2020). There are three human CoVs (Middle East respiratory syndrome coronavirus (MERS-CoV), SARS-CoV-1, and SARS-CoV-2) that are deadly and highly pathogenic to humans while the other four human CoVs are associated with mild/moderate disease such as the common cold (*Common Human Coronaviruses | CDC*, 2020; Zhu et al., 2020). Despite COVID-19 being associated with a lower rate of mortality (1%) compared to MERS (35%) and SARS (13%), there is an increased number of deaths associated with COVID-19 than the other dangerous human CoVs (Azer, 2020; Pormohammad et al., 2020; *WHO Coronavirus (COVID-19) Dashboard*, n.d.). As of July 8th, 2022 there has been over 551 million cases and over 6 million deaths attributed to COVID-19 making it the most

severe health crisis since the Spanish flu in 1918 (Cascella et al., 2022b; *WHO Coronavirus (COVID-19) Dashboard*, n.d.).

Generally, COVID-19 is diagnosed by running reverse transcriptase-polymerase chain reaction (RT-PCR) on a nasopharyngeal swab collected from individuals (Zitek, 2020). Detection of SARS-CoV-2 is not limited to nasopharyngeal swabs. The SARS-CoV-2 virus has also been detected in stool, urine, sputum, blood, and saliva (Jones et al., 2020). However, samples taken from the respiratory tract are capable of detecting SARS-CoV-2 more frequently (70-100%) than other bodily fluids such as urine (<5%) and feces (30-60%) (Jones et al., 2020). Because SARS-CoV-2 can be detected in the feces, SARS-CoV-2 can be monitored in the sewage. Nevertheless, there are drawbacks to testing for SARS-CoV-2 in the sewage. The rate of SARS-CoV-2 detection in the feces is significantly less than respiratory specimens, and the SARS-CoV-2 signal will undergo a high rate of dilution from water contributed through inflow and infiltration, bathing, and cleaning. This could lead to some samples wrongly designated as negative for SARS-CoV-2. However, sewage surveillance is a valuable tool due to its ability to quickly and cost-effectively monitor a whole area with hundreds or thousands of people with a single sample. Wastewater testing can also detect asymptomatic cases of COVID-19 and detect COVID-19 cases attributed to individuals unwilling to be tested clinically. Sewage surveillance has become even more important due to the availability of at home COVID-19 testing kits which will yield results that will not be reported. Wastewater monitoring can also provide an early warning to a rise in cases that precedes general clinical testing (CDC, 2022c). One study suggested that sewage surveillance provides data that is one to two weeks ahead of clinical data which will allow health interventions to be taken earlier

if necessary (Kumar et al., 2021). However, the magnitude of the lead time from sewage surveillance varies by study.

SARS-CoV-2 is monitored by detecting a portion of the virus. This study used the N1 and N2 gene targets to identify two regions in the nucleocapsid gene of SARS-CoV-2. These gene targets have been recommended by the Centers for Disease Control and Prevention (CDC) (CDC, 2020c). For our study surrogates are added to samples as to assess inhibition. Surrogates must have a reasonable amount of structural similarity to the virus of interest to have a higher probability of mirroring that virus. CoVs (bovine coronavirus (BCoV), mouse hepatitis virus (MHV), etc), viruses in human stool (pepper mild mottle virus, etc), and bacteriophages (Phi6, MS2, etc) have been suggested as surrogates for SARS-CoV-2 (Sala-Comorera et al., 2021). One study used Phi6, transmissible gastroenteritis virus, and MHV as SARS-CoV-2 surrogates to determine how these surrogates would survive under various storage conditions on fish and meat over a 30 day period (Bailey et al., 2022). Our study uses Phi6 and BCoV. Phi6 was chosen to assess SARS-CoV-2 inhibition because its lysate which is spiked into each sample was based on a standard operating procedure from Michigan State University, and this virus has been used as a surrogate for SARS-CoV-2 in many other studies. This virus is highly similar to SARS-CoV-2 due to it being enveloped, similarly sized, and containing spike proteins (Fedorenko et al., 2020). BCoV was also chosen because it has already been used as a SARS-CoV-2 surrogate, and BCoV is a CoV meaning that it will have a similar structure to SARS-CoV-2. This virus is inexpensive and purchased as a modified live virus in a cattle vaccine. MHV has been used in studies as a SARS-CoV-2

surrogate, but it less widely available. The MHV surrogate is 10 times more expensive than the BCoV vaccine.

The purpose of this study was to determine if small changes to the regular procedure could enhance the SARS-CoV-2 signal by alleviating sample inhibition. This was accomplished by testing the SARS-CoV-2 surrogates with the same procedural changes in parallel. The most common way to reduce PCR inhibition in each sample is to dilute the samples, but this reduces the amount of detectable DNA which could lead to false negatives (Sidstedt et al., 2020). Two types of dilutions – dilution of the eluate and dilution of the pellet – were implemented to determine if one dilution method was superior to the other at enhancing the signal of SARS-CoV-2 gene targets. Because many of the samples regularly tested have a low concentration of SARS-CoV-2, diluting the sample could cause total signal loss. Thus, other method alterations which consisted of pasteurization and adding various chemicals to the sewage were also evaluated. Pasteurization was assessed because this required no alterations to the composition of the sewage. Pasteurization may be beneficial by reducing PCR inhibiters such as enzymes and microbes in the sewage which could interfere with the clean-up or PCR chemistry (Patil et al., 2019). Pasteurization in this study was carried out for two hours. Implementing pasteurization into the procedure would increase the time needed to process the samples. Bovine serum albumin (BSA) and polyvinylpyrrolidone (PVP) were added to determine if they could reduce sample inhibition. BSA has been utilized to reduce PCR inhibition caused by extracts from feces, iron chloride, fulvic acids, hemin, tannic acids, and other compounds (Giambernardi et al., 1998). PVP is also known to reduce PCR inhibition by interacting with phenols (Schrader et al., 2012). The ability of

PVP and BSA to reduce inhibition will be based on the presence of certain inhibitory molecules.

6.3 Materials and Methods

6.3.1 Sample Selection and Preparation

Our workflow used the methods described by Flood et al (Flood et al., 2021) as a starting point. Archived pellet samples were chosen that had a low to moderate concentrations of Phi6 and BCoV. Prior BCoV and Phi6 concentrations were available for previously tested samples. Both SARS-CoV-2 gene target concentrations and surrogate concentrations were assessed under conditions including normal sample preparation, diluting the sample, adding BSA (catalog number 97061-422 from VWR), adding PVP (catalog number 97061-818 from VWR), and pasteurizing the sample. All sewage samples were mixed by 40-60 inversions of the sewage stock. After mixing the sample, a 45 mL aliquot of sewage was mixed with 4.0g of polyethylene glycol 8,000 and 0.585g of sodium chloride. Sodium chloride (catalog number 470302-520) and polyethylene glycol 8000 (catalog number 97061-100) were both purchased from VWR. Four different sub samples were collected. One sample was completed adhering to the normal method for routine SARS-CoV-2 monitoring – adding only PEG and NaCl. This sample was called regular because there were no alternations to the routine procedure. This sample was also used later in the procedure for the two dilutions (dilution of the pellet and dilution of the eluate). For dilution of the pellet, 100µL of water and 100µL of pellet were added to a solution of buffer AVL, carrier RNA, Phi6, and BCoV which will be described more in depth later in the methods. The 2x dilution of eluate utilized 20µL of RNA eluted from the spin column and 20µL of water. For the pasteurized sample, 45mL

of sewage was poured into a conical bottom centrifuge tube and this conical bottom centrifuge tube was placed into a hot water bath at a temperature of 60°C. This sample was inverted 40-60 times every 20-30mins. After 2hrs of pasteurization, PEG and NaCl were added to this sample. The BSA experiments used BSA (0.25g), NaCl, and PEG. The amount of BSA was chosen based on a study that used 0.5% BSA and found that BSA was able to eliminate inhibitors (Oikarinen et al., 2009). The PVP experiments used 1.0g PVP, NaCl, and PEG. The amount of PVP chosen was based on a study that used 2% PVP for RNA extraction in sludge to help remove PCR inhibitors (Monpoeho et al., 2000). These four subsamples were held at 4°C while these samples were rotated at a speed of 70rpm. After 2hrs, these samples were centrifuged at 3,650rpm at a temperature of 4°C for 1hr.

6.3.2 Purification and Extraction of RNA

The QIAamp Viral RNA Mini Kit (catalog number 52904) was purchased from QIAGEN to purify and elute the RNA from the sewage samples. The pellet (200µL) was pipetted into a solution containing 800µL of Buffer AVL, 8µL of carrier RNA, 10µL of Phi6, and 10µL of BCoV (Bovilis® Coronavirus from Merrick Animal Health). Both Phi6 and BCoV had a concentration of 10^5 gene copies/mL. A 200µL portion of pellet was then pipetted into a 2mL cryovial containing the buffer and RNA mix. The cryovial was vortexed for 10s before incubation for 10mins at room temperature. After incubation, 200proof ethanol was added to the cryovial in an 800µL portion, and the cryovial was vortexed for 10s. At this point in the procedure there was 1.8mL solution in the cyrovial. This solution was run through the QIAamp mini spin column in order to obtain the extracted RNA. The 1.8mL solution was added in three 630µL quantities to the spin

column. Three additions were done because the spin column could not hold enough solution for a single addition. Each 630µL portion of solution was added to the spin column and centrifuged sequentially. The flow through was not retained. The centrifugation settings were 1min at 8,000rpm. These centrifuge settings were the same for every addition except when the buffer AW2 was added. The waste collected in the collection tube under the spin column was discarded after every addition until the elution step (buffer AVE). After the whole 1.8mL solution was added to the spin column and centrifuged, the wash buffers were added. A 500µL portion of buffer AW1 was added to the spin column and centrifuged. Next, a 500µL portion of buffer AW2 was added to the spin column. At this point in the process (and only during this step) the centrifuge speed and duration was increased to 14,000rpm for 4min. After centrifugation, centrifugation settings were returned to normal, and the spin column was transferred to a 1.5mL microcentrifuge tube. The microcentrifuge tube was used to collect the RNA eluted from the spin column. Buffer AVE was added to the spin column in a 40µL portion and incubated at room temperature for 1min. After incubation, the spin column was centrifuged. This step was repeated which resulted in a final volume of 80µL of RNA eluate.

6.3.3 Plating the Samples

Samples and controls (negative control, positive control, and no-template control) were all run in triplicate. Both reactions were run in duplex meaning that there were two reactions per well. BCoV and Phi6 were ran in a duplex reaction and N1 and N2 were ran in a duplex reaction. Each well received 5.5µL of RNA eluate and 16.5µL of assay mix. The assay mix was prepared with the reagents provided in Bio-Rad's One-Step RT-

ddPCR Advanced Kit for Probes, water, and primer/probe mix. The assay mix for one well duplex reaction included: 2.2µL reverse transcriptase, 3.3µL primer/probe mix 1, 1.1µL of water, 5.5µL Supermix, 1.1µL dithiothreitol, and 3.3µL primer/probe mix 2. The sequences of the primers and probes used in this study are listed in Table 6.1. Michigan State University supplied the N1 and N2 SARS-CoV-2 primers and probes and Phi6 to our lab and other labs in the Michigan PCR Network that also routinely tested for SARS-CoV-2 in sewage. After the initial data was observed, BSA and PVP were directly pipetted into the water of the assay mix to determine if inhibition could be reduced. This was mentioned by Oikarinen et al. (BSA) and Monpoeho et al. (PVP). In order to do this, 1.875g of BSA and 6.375g of PVP were weighed and placed into 25mL of water.

Table 6.1: Sequences for N1, N2, Phi6, and BCoV

Gene Target	Sequences	References
N1	5'-GACCCCAAAATCAGCGAAAT-3' (Forward) 5'-TCTGGTTACTGCCAGTTGAATCTG-3' (Reverse) 5'-FAM-ACCCCGCATTACGTTTGGTGGACC-BHQ1-3' (Probe)	(CDC, 2020c)
N2	5'-TTACAAACATTGGCCGCAAA-3' (Forward) 5'-GCGCGACATTCCGAAGAA-3' (Reverse) 5'-HEX-ACAATTTGCCCCCAGCGCTTCAG-BHQ1-3' (Probe)	(CDC, 2020c)
Phi6	5'-TGGCGGCGGTCAAGAGC-3' (Forward) 5'-GGATGATTCTCCAGAAGCTGCTG-3' (Reverse) 5'-FAM-CGGTCGTCGCAGGTCTGACACTCGC-BHQ1-3' (Probe)	(Gendron et al., 2010)
BCoV	5'-CTGGAAGTTGGTGGAGTT-3' (Forward) 5'-ATTATCGGCCTAACATACATC-3' (Reverse) 5'-HEX-CCTTCATAT-ZEN-CTATACACATCAAGTTGTT-IABKFQ-3' (Probe)	(Graham et al., 2020)

Table 6.1: The sequences for all of the primers (reverse and forward) and probes used in these experiments.

This water was added to the assay mix (calculated for duplex reaction) to give final concentrations of 0.5% (BSA) and 1.7% (PVP) (Monpoeho et al., 2000; Oikarinen et al., 2009).

After the samples were plated, the 96-well plate was placed into the Automated Droplet Generator (ADG). The ADG generated droplets into another 96-well plate which was sealed and put into thermocycler to convert RNA to DNA and to amplify the DNA. Thermocycler settings were as follows: step 1 – 25°C for 3min, step 2 – 50°C for 1hr, step 3 – 95°C for 10min, step 4 – 40 cycles composed of: 95°C for 30s followed by 60°C for 1min, step 5 – 98°C for 10min, and step 6 – hold at 4°C. Thermocycling settings for the general settings were based on the document found on Bio-Rad's website for 1-Step RT-ddPCR Advanced Kit for Probes, and annealing temperatures for N1 and N2 were found in a research article (*One-Step RT-DdPCR Advanced Kit for Probes*, n.d.; Pearson et al., 2021). The annealing temperatures for BCoV and Phi6 were chosen in Chapter 4. The plate was then placed onto the QX200 Droplet Reader.

6.3.4 Analyzing the Data

The droplet reader was paired with QuantaSoft™ Software to quantify the data. The positive droplets were assessed to determine if the sample was positive for SARS-CoV-2. Three or more positive droplets were considered a positive detection. Accepted droplets were assessed to determine the validity of the values obtained from the QuantaSoft™ Software. Wells that had greater than 10,000 accepted droplets were considered acceptable. The copies per 20µL well was used to calculate the concentration of the various SARS-CoV-2 gene targets. Equation 6.1 shows the formula utilized to calculate

the concentration of the SARS-CoV-2 gene targets. The concentration was in units of gene copies per liter of sewage (CP/L).

(Equation 6.1)

$$\frac{CP}{L} = copies\ per\ \mu L \times \frac{pellet\ volume\ (mL)}{extraction\ volume\ (0.2mL)} \times \frac{eluate\ volume\ (80\mu L)}{initial\ volume\ (45mL)} \times \frac{1000mL}{1L}$$

Equation 6.1: Concentration calculation for the SARS-CoV-2 Gene Targets
This formula shows how the concentrations for the SARS-CoV-2 gene targets and surrogates were calculated. The pellet should be between 1-2mL for the gene targets (N1, and N2) and 45mL for the surrogates (Phi6 and BCoV).

6.4 Results

Five different method modifications were evaluated to determine if any of these enhanced the signal and reduced inhibition. Addition of BSA or PVP to the 45mL of sewage, pasteurization, and 2x dilutions on either the eluate or pellet were assessed for 29 samples. For the 2x dilution, the concentration was multiplied by two to account for half of the sample composing of water. When comparing the concentrations of the N1 and N2 gene targets with the regular method and the enhanced methods, the gene target concentrations were lowest when the method was not enhanced. For the N1 gene target, the highest to lowest concentrations obtained for each alteration to the method or regular method were as follows: 2x dilution on the pellet>BSA>pasteurization> PVP>2x dilution on the eluate>regular method. The N2 gene target behaved similarly to the N1 gene target with the highest to lowest concentrations as follows: 2x dilution on pellet>BSA>pasteurization>2x dilution on eluate>PVP>regular method.

Table 6.2: Average Concentration for Each Enhanced Method

Method Enhancement	N/A	2x Eluate Dilution	2x Pellet Dilution	PVP	Pasteurization	BSA
N1 Concentration (CP/L)	3.98×10^4	4.25×10^4	12.2×10^4	4.61×10^4	5.36×10^4	5.59×10^4
N2 Concentration (CP/L)	2.57×10^4	3.72×10^4	9.90×10^4	3.01×10^4	4.03×10^4	4.65×10^4

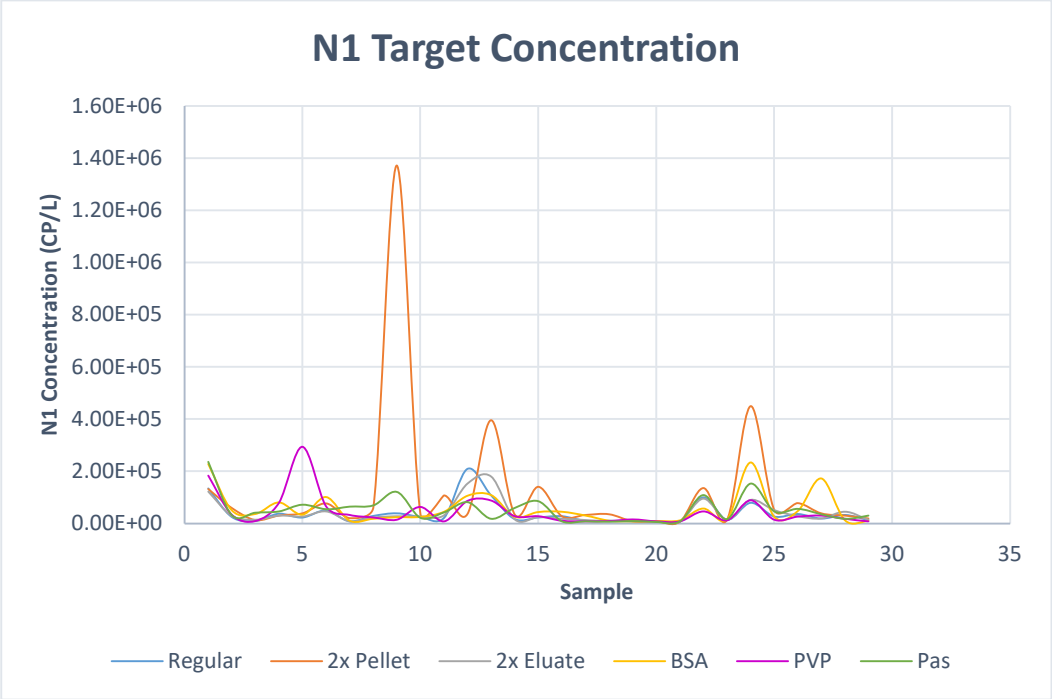
Table 6.2: Concentration in copies per liter obtained when the samples were pasteurized, diluted, unaltered, or had BSA or PVP added. Concentrations of the diluted samples were multiplied by two to account for the 2x dilution.

The actual concentrations (copies/liter – CP/L) are listed in Table 6.2. The method detection limit was set to the same value for all non-detects to avoid skewing the average. The PVP and BSA in this portion of the study were only added to the initial 45mL of sewage and not into the assay mix. A paired t-test was run to compare the unaltered method to each of the five adjusted methods. The N1 gene target only had a statistically different value ($p<0.05$) when BSA was added. The N2 gene target was enhanced by4/5 of the modification having a statistically greater value ($p<0.05$) when compared to the unaltered samples. The signal was not enhanced for either target when PVP was added. The N1 gene target had an average concentration (3.98×10^4 CP/L) that was statistically higher ($p<0.05$) than the N2 gene target (2.57×10^4 CP/L). As they are both derived from the nucleocapsid gene, it would be expected that their concentration would be similar. The number of detections was compared amongst the modified samples for the N1 and N2 gene targets. The unaltered method had five samples for the N1 gene target that had no signal and seven samples for the N2 gene target that had no signal. For the N1 gene target, the number of non-detect samples for each method enhancement were as follows:

seven (pasteurized), six (2x dilution on eluate), six (PVP), five (unaltered), five (BSA), and four (2x dilution on pellet). Thus, pasteurizing the sewage led to the most non-detects, and diluting the pellet led to the highest number of SARS-CoV-2 detections for the N1 gene target. For the N2 gene target, the number of non-detect samples for each method enhancement were as follows: eleven (PVP), eight (2x dilution on eluate), seven (unaltered), seven (pasteurized), five (2x dilution on pellet), and five (BSA). Thus, adding PVP led to the most non-detects, and diluting the pellet and adding BSA to the sewage led to the highest number of SARS-CoV-2 detections for the N2 gene target. Phi6 and BCoV were also observed for each of the method enhancements to determine if these surrogates would have a similar trend as the N1 and N2 gene targets. In general, Phi6 and BCoV are used to estimate the amount of inhibition that SARS-CoV-2 would likely face under the same conditions. Usually, when BCoV and Phi6 are used to estimate inhibition, the sample concentration of Phi6 or BCoV is divided by the negative control to get a percentage of retrieved surrogate. For that calculation the negative control would be the highest amount of sample that can be obtained, and each sample would be divided by the negative control to get the percentage of retrieval. For this study the concentrations were solely observed because the same concentrations of Phi6 and BCoV were spiked into all the samples, and the only interest was to determine if there was an increase or decrease in their signal. The concentration obtained for the 2x dilution of eluate was multiplied by two to make up for the signal loss due to dilution. The concentration for the 2x dilution of the pellet was not multiplied by two because Phi6 and BCoV were not diluted. When the concentrations of BCoV were ranked based on the various method enhancements, the following order was obtained: regular<BSA<2x dilution on eluate<2x

dilution on pellet<PVP<pasteurization. The rankings for Phi6 were as follows: PVP<2x dilution on pellet<regular<pasteurization<2x dilution on eluate<BSA. BCoV concentration was statistically enhanced ($p<0.05$) when the eluate was diluted, or the sewage was pasteurized. The concentration of Phi6 had a statistically significant increase ($p<0.05$) when the eluate was diluted or BSA was added. The number of fully inhibited samples were assessed for each adjustment to the method. The number of samples that had BCoV samples that were fully inhibited were as follows: two (regular), one (2x dilution on eluate), one (BSA), one (PVP), zero (pasteurization), and zero (2x dilution on pellet). The number of samples that had Phi6 fully inhibited were as follows: six (PVP), four (regular), three (BSA), two (2x dilution on elute), two (pasteurization), and one (2x dilution on pellet). Figure 6.1 shows the concentration of each SARS-CoV-2 gene target and surrogate under the various method alternations. We also assessed if adding BSA and PVP to the assay mix alone could increase the signal. Again, this was tested for 29 samples, but less than half of the samples had a signal for the N1 and N2 gene targets. The samples with BSA have a slightly higher average concentration for both N1 and N2 gene targets compared to the unaltered data. In contrast, PVP had the lowest concentration for both N1 and N2 targets. Nevertheless, there was no statistical difference between the regular samples or the samples with BSA or PVP added to the assay mix when observing the N1 and N2 gene target data. The number of samples that were positive for the N1 gene target were as follows: 14 (BSA), 13 (regular), and 11 (PVP). The number of samples that were positive for the N2 gene target were as follows: 7 (regular), 6 (BSA), and 3 (PVP).

Figure 6.1: Concentration of SARS-CoV-2 Gene Targets and SARS-CoV-2 Surrogates Based on Method Variations



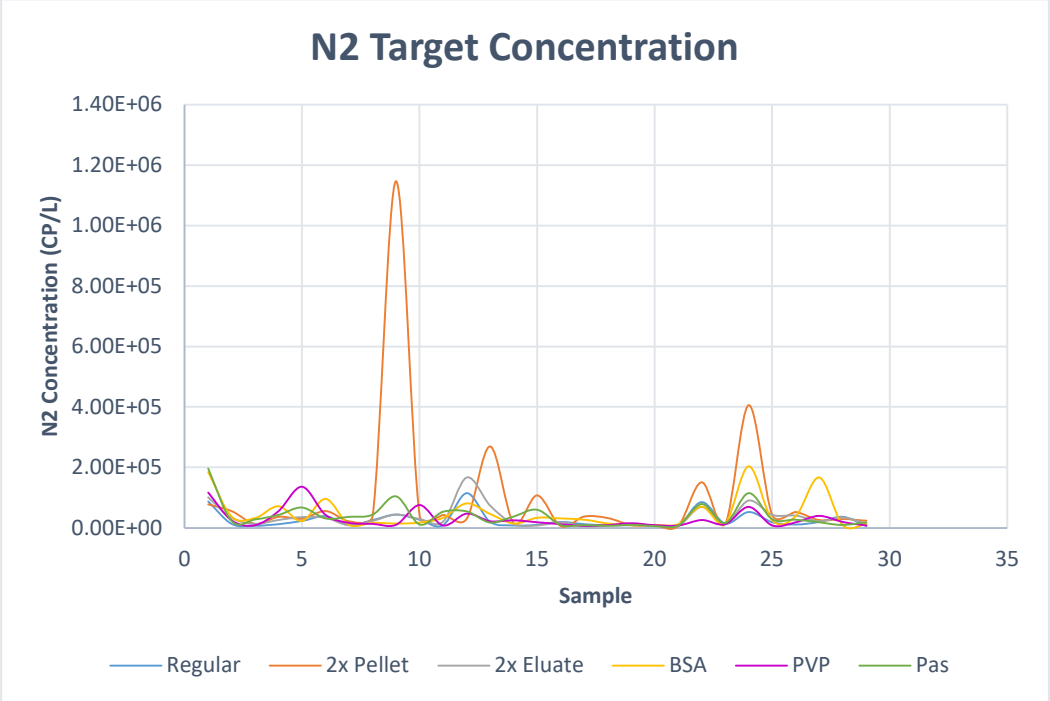
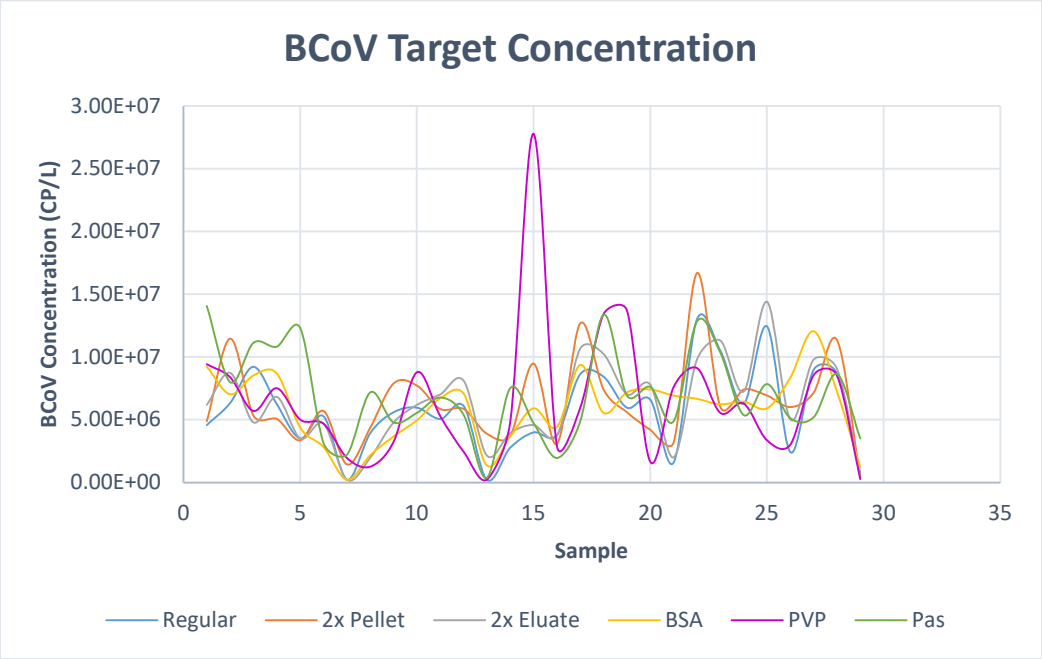


Figure 6.1: Concentration of SARS-CoV-2 Gene Targets and SARS-CoV-2 Surrogates Based on Method Variations (Continued)



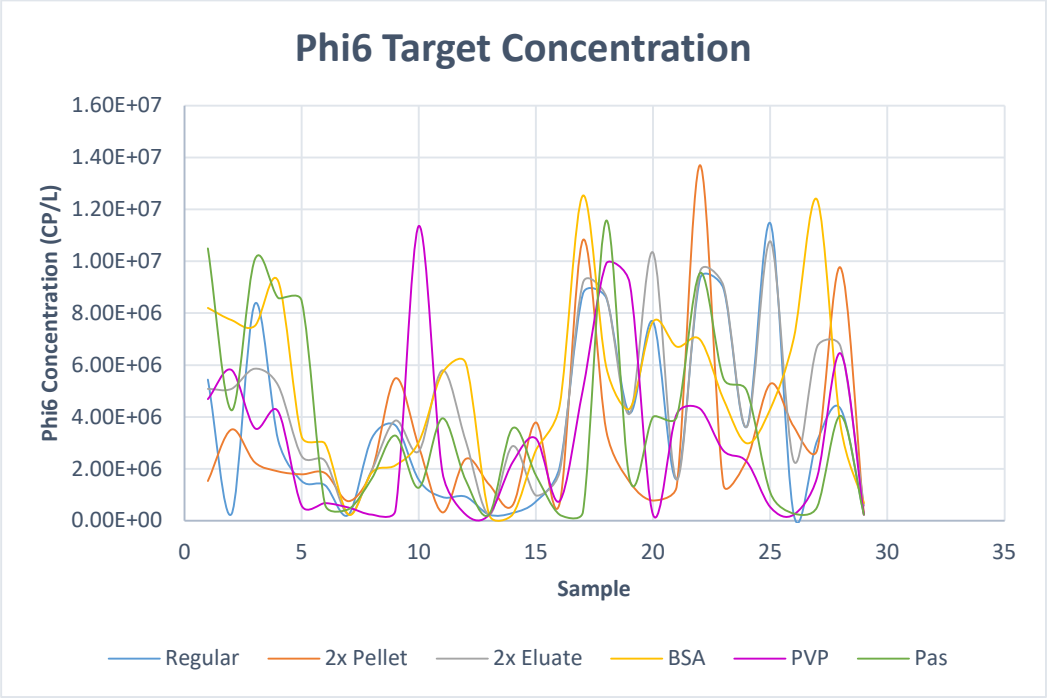


Figure 6.1: These figures show the concentration of the SARS-CoV-2 gene targets N1 and N2 and the SARS-CoV-2 surrogates Phi6 and BCoV when the method was altered. Pas is short for pasteurization. Concentration was observed for 29 samples.

Next, the SARS-CoV-2 surrogates BCoV and Phi6 were analyzed. According to the SARS-CoV-2 surrogates, the regular samples had a statistically higher concentration compared to the concentrations obtained for either of the assay mix alterations (BCoV) or when BSA was added to the assay mix (Phi6). As suggested by the p-values, the regular method always had a higher average concentration of surrogate compared to the concentrations obtained from the samples with altered assay mix. All samples, excluding one Phi6 sample that had BSA added, had signal for both BCoV and Phi6 indicating that no samples were fully inhibited.

6.5 Discussion

The first set of experiments addressed what would happen to the concentration of the N1 and N2 gene targets and SARS-CoV-2 surrogates (Phi6 and BCoV) when the samples had method adjustments to the initial 45mL of sewage. The regular (unaltered) method always had the lowest concentration for the N1 and N2 gene targets when comparing the average concentration to the average concentrations obtained from the altered methods. A paired t-test demonstrated that the addition of BSA to the initial 45mL of sewage was the only method alteration that significantly enhanced the concentration of both SARS-CoV-2 gene targets. No other method alteration could statistically enhance the concentration of the N1 gene target. Interestingly, the N2 gene target was shown to have its concentration enhanced for all method alterations to the initial 45mL of sewage except when PVP was added despite both gene targets arising from the same gene. A paired-t-test showed that the N1 target has a significantly higher concentration than the N2 gene target. The N2 gene target was also detected less or the same number of times as the N1 gene target regardless of the method alteration. Many studies use both the N1 and N2 gene targets

sequences recommended by the CDC (CDC, 2020c). In general, there are many studies that have a similar finding that the N1 gene target is superior to the N2 gene target. One study mentioned that the N1 gene target had a higher amount of sensitivity than the N2 gene target (Vogels et al., 2020). Another study analyzed the SARS-CoV-2 gene targets in sewage, and found that when these samples were tested, 91.2% were positive for N1 and 85.3% were positive for N2 which again indicates N1 is detected more frequently than N2 (Pérez-Cataluña et al., 2021). However, all studies do not agree that the N1 gene target is superior for SARS-CoV-2 detection. A study by Klein et al. mentioned that the N1 and N2 gene targets provide a comparable amount of fitness to detect SARS-CoV-2 with no differences observed when comparing genome copies and cycle threshold values (Klein et al., 2020). However, Nalla et al. concluded that the N2 gene target provided more sensitivity when detecting samples than the N1 gene target (Nalla et al., 2020). Therefore, it is feasible that N1 may not always be the most suitable gene target despite support from this study and other studies. Using two gene targets is preferred because it will provide a higher amount of certainty that the sample is truly negative for SARS-CoV-2. In some studies, the discrepancy between the amount of gene target detected could be attributed to the sample type. Each gene target may degrade differently. Based on our prior study the gene targets degraded similarly in sewage, so it is unlikely that they are degrading at different rates. For our study, the N2 gene target appears to be more prone to inhibition than the N1 gene target because the concentration of the N2 gene target was enhanced with four method alterations while N1's concentration was only enhanced with one method alteration. The method changes may have reduced competition for binding sites on the spin column or reduced interference with PCR

amplification allowing more N2 gene target to be detected. N1 appears to be less prone to having interference with spin column binding or PCR inhibition because the method enhancements were not as beneficial to N1.

From the average concentrations, the BCoV data suggested that all methods enhanced the signal, and the Phi6 data suggested that all methods, but PVP enhanced the signal. Paired t-tests indicated that pasteurizing the sample and diluting the eluate increased BCoV's concentration ($p<0.05$), and the t-test also indicated ($p<0.05$) that BSA added to the sewage and diluting the eluate enhanced the concentration for Phi6. These observations are somewhat consistent with the N2 gene target which showed statistical significance for many of the method alternations. BCoV and Phi6 do not appear to behave similarly to the N1 and N2 gene targets due to inconsistencies in which methods are significant. This inconsistency was also observed when the assay mix was altered. The N1 and N2 data suggested that neither the addition of BSA nor PVP caused any statistical difference in the amount of SARS-CoV-2 gene target obtained. The Phi6 and BCoV data provided evidence that there was a statistical difference ($p<0.05$) that negatively impacted the concentration when BSA and PVP were added to the assay mix. Thus, the surrogates Phi6 and BCoV in this study did not reliably predict how the N1 and N2 gene targets would react when different methods of enhancement were selected. The trend observed for the surrogates may vary from the N1 and N2 gene targets because the Phi6/BCoV concentration could not be normalized because problems with the negative control. The percentage of retrieval was not assessed. Assessing the percentage of retrieval instead of the concentrations could have strengthened the trend between the SARS-CoV-2 gene

targets and surrogate data. Unfortunately, there was insufficient sample to repeat the experiment and insufficient funding.

PVP was used in sludge samples for removing PCR inhibitors (Monpoeho et al.). It was mentioned that PVP is supposed to reduce inhibition by binding to phenolic compounds in order to form complexes with them (Monpoeho et al.). Organic compounds such as phenol, humic acids, bile salts, polysaccharides, and urea are inhibitory to PCR (Schrader et al.). In this study, the addition of PVP was not able to enhance the N1 or N2 gene target signal when added to the sewage or assay mix. In fact, although there was no statistical difference in the concentration, the addition of PVP created one false negative for N1 and four false negatives for N2. When PVP was added to the assay mix two samples were wrongly identified as negative for N1 and four samples were wrongly identified as negative for N2. Thus, PVP was detrimental to reducing sample inhibition of the SARS-CoV-2 gene targets in sewage.

Pasteurization was another method used to enhance the N1 and N2 gene target signal. Pasteurization was found to statistically enhance ($p=0.010$) only the concentration of the N2 gene target even though the average concentration was higher for the N1 gene target ($p=0.056$). Pasteurization is a technique which applies high temperature over a period of time which reduces microorganisms by several orders of magnitude if they are susceptible to heat, if there is long enough contact time, and if the temperature is high enough (Islam & Johnston, 2006). Part of the issue with pasteurization is that RNA is heat sensitive which could lead to the RNA degrading causing false negatives to occur. It has been mentioned that DNA is stable even at high temperatures over 100°C, but RNA is much less stable at high temperatures and temperatures above 65°C should be avoided

(*Precautions for Handling of RNA*, n.d.). Pasteurization was carried out at 60°C for 2hrs in this study so it would not negatively impact the RNA. The paired t-test did not show that pasteurization negatively impacted the N1 and N2 gene targets. Overall, three samples for N1 and one sample for N2 had a false negative signal. Thus, pasteurization appears to be slightly detrimental. Overall, pasteurization enhanced concentration, but it also led to some to some false negatives. Because there was generally a signal increase after pasteurization, it is even possible that some samples were false positive as two of the samples only had a signal for one target. Pasteurization is promising, but it would need to be tested for each sample site to see if it is beneficial. It would be assumed that the sewage matrix would have some similarity if it were taken from the same site. In general, pasteurization is cost-effective, but a more time-consuming way to enhance the concentration of the N2 gene target.

Dilution is another common way to reduce inhibition. It has been said that as the concentration increases so does the amount of inhibitors in the sample (McCord et al., 2014). Thus, although there is a reduction in the intensity of the signal, the amplification should overall be enhanced due to there being less inhibitors in the sample (McCord et al., 2014). Dilution has been proven to be a useful tool in some studies. In one study a five-fold dilution was carried out on DNA extract which lead to increased DNA quantification (3.3-fold) and increased quantitative PCR sensitivity by 25% (Acharya et al., 2017). However, dilution is not always beneficial. A different article mentioned that they tried using both inhibitor removal kit columns and diluting samples, but they were unable to increase eDNA yields, reduce filter clogging, or decrease PCR inhibitors (Hunter et al., 2019). Thus, dilution is not always favorable and will need to be assessed

to determine if it is the right process before implementing it into the normal workflow. Dilution is appropriate when the concentration is generally found in high amounts. However, when the target of interest is generally found at a low concentration, the target DNA could be diluted to the point where the sample becomes negative for the target of interest (Al-Soud et al., 2005). The average concentration of the N1 and N2 gene targets in our study were found to be higher for both the dilution of the eluate and the pellet (concentration multiplied by two to account for the 2x dilution). Diluting the pellet was the method enhancement that yielded the highest average concentration for both N1 and N2 gene targets. However, the N1 gene target was not shown to have statistical significance when the pellet underwent a 2x dilution ($p=0.052$) or when the eluate underwent a 2x dilution ($p=0.383$). However, the N2 gene target was shown to be enhanced in both cases with a p-values of 0.041 (pellet) and 0.004 (eluate). A 2x dilution was chosen to minimally dilute the sample because in many cases the concentrations obtained from each sample is low. Therefore, a small dilution was chosen in order to try to reduce the chance of false negatives while reducing inhibition. The 2x dilution of the eluate increased the number of false negative samples by one for each gene target. The 2x dilution on the pellet reduced inhibition leading to one additional sample becoming positive for N1 and two additional samples becoming positive samples for N2. Therefore, diluting the pellet should be chosen over diluting the eluate. This alteration was beneficial because it led to three more SARS-CoV-2 positive detections.

BSA was used to decrease sample inhibition. BSA was chosen based on the study by Oikarinen et al. who used BSA to reduce PCR inhibition in feces (Oikarinen et al., 2009). BSA enhancement was investigated as an addition to the assay mix and as an addition

into the initial 45mL portion of sewage. When BSA was added to the initial 45mL of sewage there was a statistically significant improvement to the N1 and N2 signal. The addition of BSA to the sewage was the only alteration to the method that significantly enhanced the concentration of both the N1 and N2 gene targets. Besides the N1 and N2 gene targets having a statistically higher concentration, two N2 samples were uninhibited when BSA was added to the 45mL of sewage. There was no statistical increase in signal when BSA was added to the assay mix. Overall, a similar number of samples had detection for the regular samples (13 detections for N1 and 6 detections for N2) as compared to samples with BSA added to the assay mix (14 detections for N1 and 6 detections for N2). Interestingly, despite having a similar number of detections, the samples that were positive varied bringing into question the use of this enhancement. Six samples that were non-detects for N1, and three samples that were non-detect for N2 became detectable. Farell et al. mentioned that BSA is temperature sensitive meaning that the high temperatures that are involved in PCR could be causing BSA to lose its ability to enhance the SARS-CoV-2 gene target's concentrations (Farell & Alexandre, 2012). The study by Oikarinen et al. found that 13/108 of their stool samples were fully inhibited, but after addition of BSA all samples had a signal (Oikarinen et al., 2009). Their study found that BSA did cause a reduction in sample sensitivity. However, BSA eliminated all false negatives which occurred due to PCR inhibitors (Oikarinen et al., 2009). Thus, they determined that overall BSA was a beneficial addition to stool samples because 12% more samples had signal despite the reduction in sensitivity (Oikarinen et al., 2009). The authors suggested that the PCR inhibition in stool samples may have a dietary origin (Oikarinen et al., 2009). BSA might reduce sample inhibition by hydrogen bonding with

phenolic compounds present from industrial discharges (Carol Kreader, 1995). Also, BSA through hydrophobic interaction may bind to lipids, and its high lysine content may allow it to bind to anions (Carol Kreader, 1995).

6.6 Conclusion

This research suggests that the N2 gene target may be more susceptible to PCR inhibition than the N1 gene target based on the statistically lower concentration observed in the N2 gene target. This idea is additionally supported based on the observation that N2 concentrations increased for most method alterations. The N2 signal was enhanced for all method enhancements with the exception of PVP addition. The N1 signal was enhanced for BSA addition to the sewage pellet. PVP along with BSA are thought to reduce PCR inhibition by interacting with phenolic compounds, but PVP had no benefit in our study. In contrast, BSA was found to enhance the concentration of both the N1 and N2 gene targets when added to the sewage. Therefore, BSA may be reducing inhibition by binding to other compounds that are not phenol based.

Adding BSA to the assay mix did not significantly impact the concentration. BSA may be more beneficial when added to the pellet because BSA was mentioned to be temperature sensitive in another study which means that BSA may not have any inhibitor reducing abilities during PCR amplification. This is likely why adding BSA to the assay mix is unable to further enhance the signal. Overall, BSA infused sewage was the only method enhancement that was able to provide a statistically enhanced concentration for both N1 and N2 gene targets. The N2 gene target concentration was enhanced by pasteurization, BSA addition to the sewage, and dilution of eluate or pellet. The Phi6 and BCoV data did not correlate with the N1 and N2 gene targets. This was likely due to low to our inability

to normalize these data by the faulty negative control. This part of the study could not be repeated making our surrogate data not definitive.

CHAPTER SEVEN

COMPARISON OF THE PERCENTAGE OF RETRIEVED SARS-COV-2 GENE TARGETS E, N1, AND N2 TO THE SARS-COV-2 SURROGATES PHI6, BCOV, AND MHV

7.1 Abstract

Coronavirus disease 2019 (COVID-19) is a fast spreading and potentially lethal virus caused by an RNA virus called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Due to the highly transmissible nature of this virus, hundreds of millions of people have had this virus within less than three years of the virus's origin. Monitoring cases of this disease is vital to allow precautions to be taken to decrease the rapid spread of COVID-19. Sewage monitoring has been implemented to rapidly screen whole communities for COVID-19. This allows both asymptomatic and symptomatic cases to be detected cost effectively and unobtrusively. Sewage is a difficult matrix to analyze by PCR techniques and sample inhibition can lead to excessive false negatives. This study compares three different SARS-CoV-2 surrogates to three common SARS-CoV-2 gene targets to determine if the surrogates provide reasonable estimates of inhibition. The surrogates chosen – murine hepatitis virus (MHV), Phi6, and bovine coronavirus (BCoV) – have been used as SARS-CoV-2 surrogates in other studies. Currently, our lab uses both BCoV and Phi6. We found that the N1 and N2 gene targets have a similar amount of inhibition and are resilient with only one sample being fully inhibited. In general, the E gene target was detected less frequently and at lower concentrations than the other gene targets. The coronavirus surrogates were beneficial with no statistical difference observed

between them and the SARS-CoV-2 gene targets (N1 and N2) when comparing the percentage of retrieval. However, Phi6 was statistically different than all the SARS-CoV-2 gene targets and falsely identified 12 samples as fully inhibited. Our study suggests either BCoV and MHV could be used as surrogates, but BCoV is superior because of its low cost and ready availability.

7.2 Introduction

Coronavirus disease 2019 (COVID-19) is generally associated with mild disease such as cough, loss of smell, body aches, fever, sore throat, and headache, but some patients do not have symptoms and some patients may die from this disease (Subbarao & Mahanty, 2020). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a single-stranded and enveloped RNA virus which causes the disease COVID-19 (Alexandersen et al., 2020). There are seven coronaviruses (CoVs) known to infect humans which include human CoV-NL63 (HCoV-NL63), HCoV-HKU1, HCoV-229E, HCoV-OC43, Middle East respiratory syndrome coronavirus (MERS-CoV), and SARS-CoV-1 and 2 (Kesheh et al., 2022a). Most HCoVs are betacoronaviruses, but HCoV-NL63 and HCoV-229E are alphacoronaviruses (Susan I. Gerber and John T. Watson, 2020). In general HCoVs are mostly associated with mild virus symptoms similar to the common cold (Coerdt & Khachemoune, 2021). As the twenty-first century has progressed, three HCoVs which are correlated with a higher risk of mortality have emerged – MERS-CoV, SARS-CoV-1, and SARS-CoV-2. These HCoVs are associated with a much higher fatality rate than the other four HCoVs. These HCoVs cause diseases that have the following fatality rates 1.2% (COVID-19), 9.6% (SARS), and 35% (MERS) (Peiris & Poon, 2021; *WHO Coronavirus (COVID-19) Dashboard*, n.d.). The common cold is

typically not associated with death and is less likely to result in death than the flu virus which has a 0.1% fatality rate (Health Desk, 2020).

COVID-19 is generally diagnosed with naso-oropharyngeal swabs (Rodrigues et al., 2021). This virus is also present in other bodily fluids including blood, semen, feces, serum, and occasionally urine (Johnson et al., 2021). Sewage surveillance can monitor SARS-CoV-2 mostly due to the feces present in the sewage. However, SARS-CoV-2 is only present in less than half of fecal samples (Johnson et al., 2021). Sewage is 99.9% water (Buchanan, n.d.). Therefore, between the inherently low concentration of SARS-CoV-2 in feces and the major dilution in sewage, the SARS-CoV-2 signal in wastewater often low. This study analyzes three different gene targets that detect SARS-CoV-2. The World Health Organization (WHO) mentions the N1, N2 and E gene targets can identify SARS-CoV-2 (*Whoinhouseassays.Pdf*, n.d.). All three of these gene targets were used to determine if one target was superior. However, testing two gene targets is optimal because two gene targets can be readily duplexed with little additional cost. Practical considerations of the droplet digital PCR platform make triplex assays difficult. It is very important to monitor COVID-19 quickly and with high accuracy to allow public health measures when they can still be meaningful. Analyzing two targets offers higher certainty leading to fewer false negatives. Also, mutations may impact the reliability of one or multiple of the SARS-CoV-2 gene targets analyzed. Thus, using two targets should allow more reliable detection and quantitation.

It is also important to determine which samples have no sample retrieved to rule out incorrectly designating a sample as negative. If a sample is known to be fully inhibited, measures can be taken to reduce inhibition. A common way to reduce inhibition is to

dilute the sample (Schrader et al., 2012). However, it is not optimal to preemptively dilute samples because target concentrations can be low especially when cases are less frequent. Thus, diluting the sample could cause the analyst to incorrectly identify the sample as negative. Surrogates can be used to estimate the amount of inhibition and identify samples that should be rerun with dilution. Inhibition of the SARS-CoV-2 gene targets cannot be directly assessed for each sample without modifying the results for routine SARS-CoV-2 testing. To test for inhibition, a surrogate is directly spiked into the sample and ran through the entire process of RNA purification and extraction. Therefore, the surrogate should not be naturally present in sewage, and the surrogate must be structurally similar virus to the virus of interest. It has been suggested that bacteriophages such as MS2 and Phi6 and other CoVs can be used as surrogates for SARS-CoV-2 (Boegel et al., 2021; String et al., 2021). This work directly compares Phi6, bovine coronavirus (BCoV), and murine hepatitis virus (MHV) for use as surrogates. These RNA viruses have been used as surrogates for SARS-CoV-2 in other studies. However, due to structural variance some surrogates may not mimic SARS-CoV-2 in all sewage matrices. This study used Droplet Digital polymerase chain reaction (ddPCR) to find surrogates that best represented how SARS-CoV-2 reacted in various sewage samples. Because every sewage matrix is different and varies over time and seasonally, multiple surrogates might be required for the highest accuracy. This work aims to identify the best surrogate/surrogates for estimating the amount of inhibition the SARS-CoV-2 gene targets would face in wastewater samples. When inhibition is accurately assessed, samples will be correctly identified as fully inhibited allowing effective procedures to be implemented to reduce inhibition.

7.3 Materials and Methods

7.3.1 Choosing a Sample/Pellet Composition

Our workflow used the methods described by Flood et al (Flood et al., 2021) as a starting point. Samples chosen were tested prior for SARS-CoV-2 during routine sewage monitoring and found to have no signal for the N1, N2, and E SARS-CoV-2 gene targets. Sewage samples were delivered to the lab in 500mL portions. Samples were transported in a cooler with icepacks while in transit to our lab. Samples were processed immediately upon arrival. The 500mL bottle of sewage was mixed by inverting the bottle between 40 and 60 times to minimize foaming. A 45 mL homogenous sewage sample was decanted into a 50mL conical bottom centrifuge tube. Approximately 4.0g of polyethylene glycol 8000 (catalog number 97061-100 from VWR) and 0.59g sodium chloride (catalog number 470302-520 from VWR) were added to the sample yielding a concentration of 8% polyethylene glycol and 0.2M sodium chloride. After the solids and solutions were combined, the centrifuge tubes were transferred to a rotator where the tubes were rotated at 70rpm for 2hrs. The rotator was held at 4°C. After mixing, the samples were centrifuged at 3,650 rpm for 1 hr. Once centrifugation was completed, the supernatant was discarded, and a 1-2 mL pellet remained. The supernatant was removed, and the pellet was saved. Part of the pellet was immediately analyzed for routine testing which was later used to determine which samples were negative for SARS-CoV-2. The remaining pellet was frozen and later used for this study.

7.3.2 RNA Extraction and Purification

In order to keep biosafety level requirements to a biosafety level two lab, SARS-CoV-2 eluate from a previous sewage sample that averaged around 3000 positive droplets was

spiked into each negative sample. This allowed each sample to be directly assessed for inhibition of the SARS-CoV-2 gene targets N1, N2, and E. The QIAamp Viral RNA Mini Kit from QIAGEN (catalog number 52906) was used to purify and eluate RNA for analysis. For each sample, 800µL of buffer AVL, 8µL of carrier RNA, 10µL of Phi6 (10^5 gene copies/mL), 10µL of BCoV (10^5 gene copies/mL – Bovilis® Coronavirus from Merrick Animal Health), 10µL of MHV (10^5 gene copies/mL – MHV Stock from ZeptoMetrix), and 25µL of SARS-CoV-2 eluate were added. A bulk solution of this mix was prepared, and 800µL was pipetted into each 2mL cryovial.

A 200µL portion of pellet was then pipetted into a 2mL cryovial containing the buffer and RNA mix. The cryovial was vortexed for 10s before incubation for 10mins at room temperature. After incubation, 200proof ethanol was added to the cryovial in an 800µL portion, and the cryovial was vortexed for 10s. At this point in the procedure there was 1.8mL solution in the cyrovial. This solution was run through the QIAamp mini spin column in order to obtain the extracted RNA. The 1.8mL solution was added in three 630µL quantities to the spin column. Three additions were done because the spin column could not hold enough solution for a single addition. Each 630µL portion of solution was added to the spin column and centrifuged sequentially. The flow through was not retained. The centrifugation settings were 1min at 8,000rpm. These centrifuge settings were the same for every addition except when the buffer AW2 was added. The waste collected in the collection tube under the spin column was discarded after every addition until the elution step (buffer AVE). After the whole 1.8mL solution was added to the spin column and centrifuged, the wash buffers were added. A 500µL portion of buffer AW1 was added to the spin column and centrifuged. Next, a 500µL portion of buffer AW2 was

added to the spin column. At this point in the process (and only during this step) the centrifuge speed and duration was increased to 14,000rpm for 4min. After centrifugation, centrifugation settings were returned to normal, and the spin column was transferred to a 1.5mL microcentrifuge tube. The microcentrifuge tube was used to collect the RNA eluted from the spin column. Buffer AVE was added to the spin column in a 40µL portion and incubated at room temperature for 1min. After incubation, the spin column was centrifuged. This step was repeated which resulted in a final volume of 80µL of RNA eluate. This solution was analyzed to determine the percentage of inhibition in each sample.

7.3.3 Sample Plating

Samples and controls were run in triplicate. Each 96-well plate contained a negative control, no-template control, and positive control. The negative control used a 200µL portion of water, instead of the pellet, that had polyethylene glycol 8,000, and sodium chloride were added as a typical sample. Sequences for the SARS-CoV-2 gene targets and surrogates are listed in Table 7.1.

The primer/probe sequences for MHV are proprietary (Zeptometrix) and have not been published. The assay mix was made using the solutions from the One-Step RT-ddPCR Advanced Kit for Probes (catalog number 1864021) produced by Bio-Rad. The assay mix composition varied for duplex and simplex reactions by the amount of water and primer and probe mix. For a simplex reaction the volumes and reagents for the assay mix were as follows for one reaction: 5.5µL Supermix, 2.2µL reverse transcriptase, 1.1µL dithiothreitol, 3.3µL primer/probe mix, and 4.4µL of water.

Table 7.1: SARS-CoV-2 Gene Target and SARS-CoV-2 Surrogates Sequences

Gene Target	Sequences (5'-3')	Reference
E	ACAGGTACGTTAATAGTTAATAGCGT (Forward) ATATTGCAGCAGTACGCACACA (Reverse) FAM-ACACTAGCCATCCTTACTGCGCTTCG-BHQ1 (Probe)	(Corman et al., 2020b)
N1	GACCCCAAAATCAGCGAAAT (Forward) TCTGGTTACTGCCAGTTGAATCTG (Reverse) FAM-ACCCCGCATTACGTTTGGTGGACC-BHQ1 (Probe)	(CDC, 2020c)
N2	TTACAAACATTGGCCGCAAA (Forward) GCGCGACATTCCGAAGAA (Reverse) HEX-ACAATTTGCCCCCAGCGCTTCAG-BHQ1 (Probe)	(CDC, 2020c)
BCoV	CTGGAAGTTGGTGGAGTT (Forward) ATTATCGGCCTAACATACATC (Reverse) HEX-CCTTCATAT-ZEN-CTATACACATCAAGTTGTT-IABKFQ (Probe)	(Graham et al., 2020)
Phi6	TGGCGGCGGTCAAGAGC (Forward) GGATGATTCTCCAGAAGCTGCTG (Reverse) FAM-CGGTCGTCGCAGGTCTGACACTCGC-BHQ1 (Probe)	(Gendron et al., 2010)

Table 7.1: Primers (forward and reverse) and probes used in this study excluding MHV. MHV primer and probe sequences were provided by the company Zeptometrix and are proprietary.

For a duplex reaction the volumes and reagents for the assay mix were as follows for one reaction: 5.5µL Supermix, 2.2µL reverse transcriptase, 1.1µL dithiothreitol, 3.3µL primer/probe mix 1, 3.3µL primer/probe mix 2, and 1.1µL of water. In this study, BCoV and Phi6 were duplexed, and N1 and N2 were duplexed. Both MHV and E were run in simplex reactions. Once samples and assay mix were plated, the plate was put into the Automated Droplet Generator where droplets were generated into another 96-well plate. This new plate was put into the thermocycler at the following settings in the specified order: 25°C for 3min (step 1), 50°C for 1hr (step 2), 95°C for 10min (step 3), 40 cycles

that consisted of: 95°C for 30s then 60°C for 1min (step 4), 98°C for 10min (step 5), and the plate was held at 4°C until the plate was ready be placed onto the QX200 Droplet Reader (step 6). Thermocycling settings for the general settings were based on the document found on Bio-Rad's website for 1-Step RT-ddPCR Advanced Kit for Probes, and annealing temperatures for N1 and N2 were found in a research article (*One-Step RT-DdPCR Advanced Kit for Probes*, n.d.; Pearson et al., 2021). The annealing temperatures for BCoV and Phi6 were chosen in Chapter 4. The annealing temperature chosen for MHV was determined experimentally and found to have proper separation at 55°C. Therefore, that annealing temperature was used for all the targets and surrogates in this study.

7.3.4 Data Analysis

The droplet reader was paired with QuantaSoft™ Software to calculate gene copies per 20µL well. The positive droplets were assessed to determine if the sample was positive for SARS-CoV-2. Three or more positive droplets were considered a positive detection. Accepted droplets were assessed to determine the validity of the values obtained from the QuantaSoft™ Software. Wells that had greater than 10,000 accepted droplets were considered acceptable. Equation 7.1 shows the formula utilized to calculate the concentration of the SARS-CoV-2 gene targets. The concentration had units of gene copies per liter of sewage (CP/L). The copies per liter of sewage of each surrogate gene target was obtained for each sample and the negative control. The negative control represented the maximum signal of the surrogate.

(Equation 7.1)

$$\frac{CP}{L} = copies\ per\ \mu L \times \frac{pellet\ volume\ \ (mL)}{extraction\ volume\ (0.2mL)} \times \frac{eluate\ volume\ (80\mu L)}{initial\ volume\ (45mL)} \times \frac{1000mL}{1L}$$

Equation 7.1: Concentration Calculation for the SARS-CoV-2 Gene Targets
This formula shows how the concentrations for the N1, N2, and E gene targets were calculated.

Therefore, the concentration obtained for each sample was divided by the concentration of the negative control and multiplied by 100 to obtain the percentage of RNA retrieved. The percentage of RNA retrieved from the SARS-CoV-2 surrogates, and the SARS-CoV-2 gene targets were compared. An ANOVA: two-factor with replication test was ran to determine if there was statistically a difference between the percentage of RNA retrieved for the SARS-CoV-2 surrogates and the SARS-CoV-2 gene targets.

7.4 Results

A total of 52 samples were ran to assess how closely related the SARS-CoV-2 gene targets E, N1, and N2 were related to the SARS-CoV-2 surrogates Phi6, BCoV, and MHV. Unfortunately, N2 was unable to separate on one plate. Therefore, only 39 samples were analyzed for the N2 gene target. No eluate remained to rerun those samples for the N2 gene target. The percentage of retrieval of each RNA target was assessed by dividing the copies per liter of sample by the maximum copies per liter obtained. This maximum concentration was obtained by running water through the spin columns opposed to sewage. This was known as negative control. Thus, no inhibiters should be present, and the signal should be the maximum concentration that can be obtained.

Figure 7.1: Percentage of Gene Target Retrieved per Sample

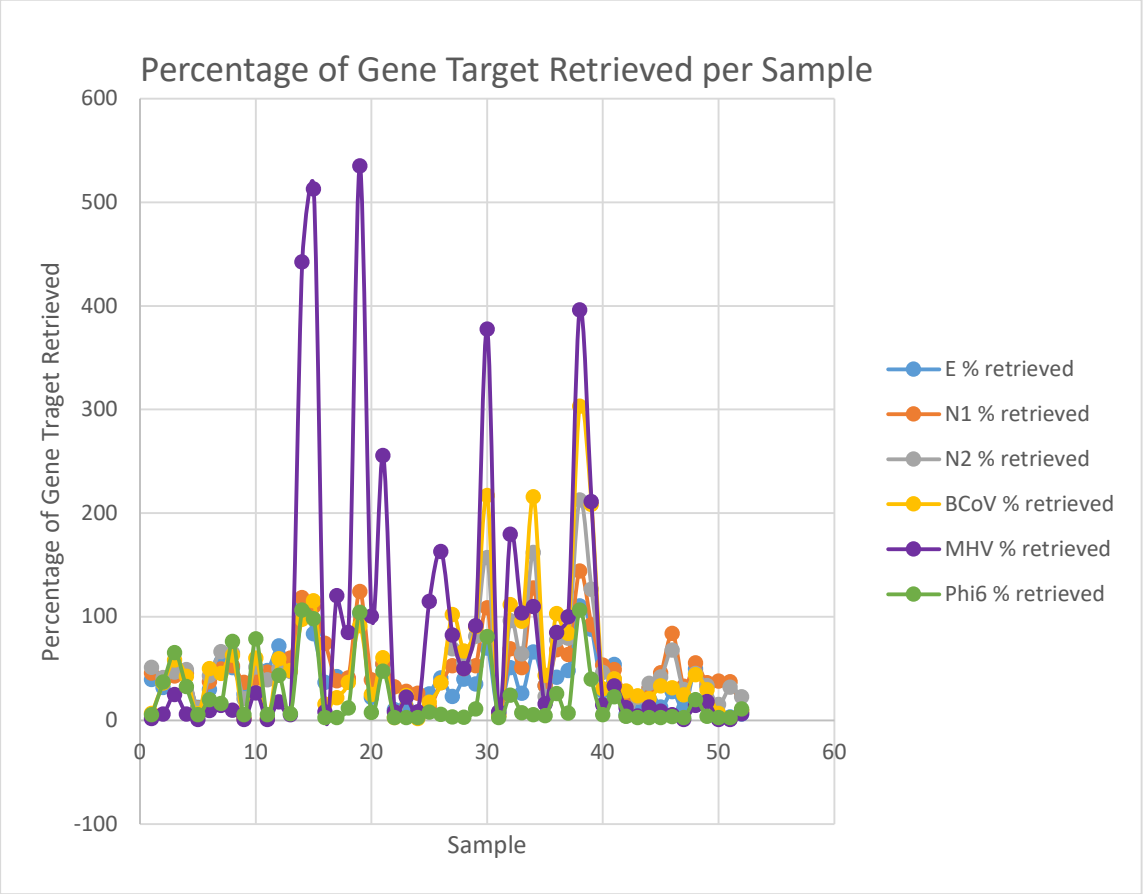


Figure 7.1: This graph compares the percentage of retrieved target for each sample.

A total of 24/52 of these sewage samples came from sites with varying location while the remaining samples came from one of those 24 locations and only varied based on sampling date. However, the sewage composition could change daily meaning that the sewage composition from the same site could vary greatly over time. Figure 7.1 shows the percentage of target RNA retrieved for both SARS-CoV-2 gene targets and surrogates on a sample-by-sample basis.

The E gene target was the SARS-CoV-2 gene target with the lowest percentage of retrieval. The N1 and N2 gene targets had a similar amount of retrieval. On average the gene targets had percentages of retrieval as follows: 39.9% (E), 52.5% (N1), and 57.3% (N2). Those averages included all 52 samples for E and N1. However, if we observe the same 39 samples that N2 had ample separation for, the percentages change slightly (E had 37.4% retrieval and N1 had 51.1% retrieval). The N2 gene target shows the highest retrieval.

Comparing the SARS-CoV-2 surrogates, Phi6 had the lowest retrieval at 23.3%. BCoV and MHV having more than double the amount of RNA retrieved when compared to Phi6 (57.7% for BCoV and 85.5% for MHV). Table 7.2 shows the percentages of each SARS-CoV-2 gene target and SARS-CoV-2 surrogate retrieved for both the full 52 samples and for 39 samples to have a better comparison to the N2 data. MHV had the highest percentage of retrieval with more than 85% of MHV retrieved. Although MHV had a high amount of retrieval, it was inconsistent with the retrieval seen in the SARS-CoV-2 gene targets with its retrieval suggesting nearly 30% more RNA was retrieved than observed in both N gene targets. Notably, MHV was more comparable to both N gene targets when the samples observed were reduced to the same 39 samples tested for N2.

Table 7.2: Percentage of RNA Target Retrieved

Gene Target	E	N1	N2	BCoV	MHV	Phi6
% Retrieved (39)	37.4%	51.1%	57.3%	62.9%	53.1%	20.7%
% Retrieved (52)	39.9%	52.5%	N/A	57.7%	85.5%	23.3%

Table 7.2: The percentage of the SARS-CoV-2 gene targets and the SARS-CoV-2 surrogates retrieved. Percentages are shown for all 52 samples and for 39 samples.

BCoV had a similar amount of retrieval compared to both N1 and N2 gene targets. Lastly, Phi6's percentage of retrieval was much lower compared to the other target's retrieval. The number of samples fully inhibited per target were also observed. The E gene target had the most samples fully inhibited (4/52 samples) while the remaining two N gene targets only had one sample fully inhibited. Of the SARS-CoV-2 surrogates, BCoV had the fewest number of samples that were fully inhibited (4/52). Of these fully inhibited samples, three of these samples correctly identified samples that had at least one of the SARS-CoV-2 gene targets fully inhibited. Of these three samples inhibited, two of these samples were only inhibited according to the E gene target, and the remaining sample showed that all three gene targets were inhibited. BCoV was the most consistent surrogate with the SARS-CoV-2 gene target data. The other two SARS-CoV-2 surrogates were not as accurate indicating more of the samples were negative for the SARS-CoV-2 gene targets. The MHV data had nine fully inhibited samples. This surrogate was able to correctly identify all four samples that were inhibited for at least one SARS-CoV-2 gene target, but it falsely identified five samples as fully inhibited. Some of these samples that MHV data showed to be fully inhibited had around 40% retrieval. Therefore, some

samples that MHV deemed as fully inhibited had a reasonable amount of retrieval. In contrast, BCoV did have low retrieval (6.8%) when strictly the E gene target was inhibited. Phi6 was the most unreliable surrogate and identified 16 of the samples were fully inhibited. Phi6 was able to identify all of the samples that were inhibited for at least one SARS-CoV-2 gene target, but the number of false negatives outweighed the benefit of Phi6's ability to identify fully inhibited samples. Table 7.3 summarizes how many samples were correctly identified as fully inhibited for at least one SARS-CoV-2 gene target, and how many times the surrogates wrongfully identified a sample as fully inhibited.

On a sample-by-sample basis each SARS-CoV-2 gene target was compared to the SARS-CoV-2 surrogates to determine which surrogate provided the closest amount of retrieval to the SARS-CoV-2 gene targets. There was a total of 143 target and sample combinations (52 N1, 52 E, and 39 N2). Overall, BCoV had the highest likelihood to match the SARS-CoV-2 gene target's percentage of retrieval. It was the superior surrogate in 84/143 combinations. These 84 samples can further be broken down by the SARS-CoV-2 gene targets as follows: 34/52 (N1), 24/52 (E), and 26/39 (N2). Phi6 was the best surrogate in 37 of the samples and more specifically 19/52 samples for E, 11/52 samples for N1, and 7/39 samples for N2. The remaining 22 combinations showed that MHV was the best surrogate for 9/52 samples for E, 7/52 samples for N1, and 6/39 samples for N2. As expected, when comparing the SARS-CoV-2 gene targets to the surrogates the worst surrogate was MHV followed by Phi6.

Table 7.3: SARS-CoV-2 Surrogates Ability to Identify Fully Inhibited Samples

SARS-CoV-2 Surrogates	BCoV	MHV	Phi6
Correctly Identified Fully Inhibited Samples	3/4	4/4	4/4
False Negatives	1/52	5/52	12/52

Table 7.3: The number of fully inhibited samples observed for the SARS-CoV-2 surrogates BCoV, MHV, and Phi6. There were four samples that were fully inhibited for one or more of the SARS-CoV-2 gene targets out of the total 52 samples.

BCoV was rarely the worst surrogate. In fact, BCoV was only the worst surrogate in 11/143 target and sample combinations (8 for E, 1 for N1, and 2 for N2). Phi6 was the worst surrogate in 51/143 combinations (14 for E, 20 for N1, and 17 for N2). MHV was found to be the least reliable surrogate in 81/143 combinations (30 for E, 31 for N1, and 20 for N2). These findings are summarized in Table 7.4.

Table 7.4: Favored and Least-Favored SARS-CoV-2 Surrogates

Gene Target	E	N1	N2	Total	Best/Worst Surrogate
BCoV	24	34	26	84	Best
Phi6	19	11	7	37	Best
MHV	9	7	6	22	Best
BCoV	8	1	2	11	Worst
Phi6	14	20	17	51	Worst
MHV	30	31	20	81	Worst

Table 7.4: A comparison of the percentage of retrieval of the SARS-CoV-2 surrogates BCoV, Phi6, and MHV to the percentage of retrieval that each SARS-CoV-2 gene target was found to have. When the surrogate had the closest amount of retrieval to a SARS-CoV-2 gene target it was considered the best surrogate. When there was the biggest disagreement between the percentages of retrieval, the surrogate was the worst. This table outlines the specific number of times each surrogate was considered the best or the worst for the three different SARS-CoV-2 surrogates.

ANOVA: two-factor with replication test was ran. For each SARS-CoV-2 gene target, the percentage retrieved was compared to all of the surrogates to determine if there was statistical difference between the SARS-CoV-2 target and the surrogate.

For the E gene target, the ANOVA test determined that the percentage of retrieval for both E and BCoV were statistically equivalent (p-value greater than 0.05). Both MHV and Phi6 had statistically different percentages of retrieval when compared with the E gene target (p-values less than 0.05). For the N1 gene target, both BCoV and MHV had percentages of retrieval that were statistically equivalent to the N1 gene target's percentage of retrieval (p-values greater than 0.05). Phi6 had a p-value lower than 0.05 indicating that there was a statistical difference between the percentage of retrieval obtained for Phi6 and N1. For the N2 gene target, both BCoV and MHV had percentages of retrieval that were statistically the same as N2's percentage for retrieval (p-values greater than 0.05). Again, Phi6 had a statistically different percentage for retrieval than N2's percentage of retrieval (p-value less than 0.05). Thus, Phi6 always had a statistically different percentage of retrieval, and BCoV always had statistically the same percentage retrieval as the SARS-CoV-2 gene targets.

7.5 Discussion

This study determined the amount of inhibition that the SARS-CoV-2 gene targets and surrogates faced in various sewage matrices. PCR may be inhibited due to various compounds in the sewage such as phenolic compounds, humic acids, and heavy metals (Botes et al., 2013). PCR is inhibited due to inhibitory molecules inactivating the polymerase or binding with other components in the PCR reaction such as the primers and DNA (Eilert & Foran, 2009). It is vital that inhibition be estimated to accurately

depict if the sample is fully inhibited to allow reruns when necessary and to avoid rerunning samples that are not fully inhibited. Samples need to be rerun if they are fully inhibited to determine if the sample is truly negative for the SARS-CoV-2 gene targets. Rerunning samples that are not fully inhibited will cause an unnecessary waste of resources. One common way to reduce inhibition is to dilute the sample (H. Wang et al., 2017). Another study suggested a combination of bovine serum albumin, sample dilution, additional Taq polymerase, and Thiopropyl Sepharose beads could aid in overcoming polymerase inhibition (Makuch et al., 1996). In general, the amount of inhibition undergone by the SARS-CoV-2 gene targets cannot be directly assessed. SARS-CoV-2 inhibition cannot be directly assessed because assessing inhibition requires that the target be spiked directly into the buffer solution that enters the spin column. This would cause nearly every sample tested to become positive for the SARS-CoV-2 gene targets, and the spiked in signal could not be differentiated from the actual signal of SARS-CoV-2 in the sewage. Therefore, surrogates were used in place of SARS-CoV-2 for our inhibition experiments. A SARS-CoV-2 surrogate is an RNA virus that has a similar structure to SARS-CoV-2. Phi6 has been mentioned as a potential surrogate for SARS-CoV-2 due to its similar size, its spike proteins, and its envelop to protect its genetic material which are all traits that SARS-CoV-2 exhibits (Fedorenko et al., 2020). Phi6 is also safe to work with (Prussin et al., 2018). SARS-CoV-2 can be handled in a biosafety level two lab for routine diagnostic testing (CDC, 2020c). Studies have previously utilized Phi6 as a surrogate for SARS-CoV-2. One of these studies used Phi6 to suggest that SARS-CoV-2 irradiation at a wavelength of 455nm could disinfect both surfaces and air without posing health risks to humans (Vatter et al., 2021).

MHV and BCoV are both CoVs that have been used as surrogates for SARS-CoV-2. Because they are the same type of virus as SARS-CoV-2, they both have a similar structure to SARS-CoV-2 despite infecting different hosts. Both BCoV and MHV only require a biosafety level two lab (*NR-445 Bovine Coronavirus, Mebus (Viruses)*, n.d.; Torii et al., 2022). One recovery study used Phi6, MHV, and pepper mild mottle virus as SARS-CoV-2 surrogates to assess five alternative polyethylene glycol precipitation methods in wastewater (Torii et al., 2022). BCoV has also been used to estimate the recovery of SARS-CoV-2 (Al-Duroobi et al., 2021). Thus, all these viruses have been used in prior experiments as surrogates for SARS-CoV-2. In this study Phi6, MHV, and BCoV were used as surrogates to estimate the amount of retrieval of the SARS-CoV-2 gene targets in sewage. Two different genes (the envelope and nucleocapsid genes) of SARS-CoV-2 were targeted with the N1, N2, and E gene targets. The World Health Organization has listed all three of these gene targets for SARS-CoV-2 detection (*Whoinhouseassays.Pdf*, n.d.).

Our study found that the percentages of retrieval of the SARS-CoV-2 gene targets were as follows: 39.8% for E (37.4% for 39 samples), 52.5% for N1 (51.1% for 39 samples), and 57.3% for N2. The SARS-CoV-2 surrogates had percentages of retrieval as follows: 23.3% for Phi6 (20.7% for 39 samples), 57.7% for BCoV (62.9% for 39 samples), and 85.5% for MHV (53.1% for 39 samples). These data showed that the E gene target was more prone to false negative than the N1 and N2 gene targets due to its average retrieval being over 10% less. The N1 and N2 gene targets had a similar amount of retrieval (6.2% difference). It is likely that the N1 and N2 gene targets had a closer amount of retrieval due to both gene targets arising from the nucleocapsid gene. The E gene target is derived

from a different gene known as the envelope gene. The data suggests that E may be more prone to inhibition than N1 and N2.

We found that the number of fully inhibited samples varied for the gene targets with the E gene target being most susceptible to inhibition labeling two samples as negative that were positive for the N1 and N2 gene targets. Interestingly, The CDC does not list the E gene target for SARS-CoV-2 detection, but it does list the N1 and N2 gene targets (*List-of-Acceptable-Commercial-Primers-Probes.Pdf*, n.d.). We suggest that the E gene target be removed from analysis due to the low amount of retrieval and because the E gene target falsely labeled two samples as negative for SARS-CoV-2. Removing the E gene target from analysis would reduce time and cost and allow one reaction to be run because N1 and N2 can be duplexed. Based on this data, it is likely that no SARS-CoV-2 detection or minimal SARS-CoV-2 detection would be lost if the E gene target was removed from analysis.

We also compared the samples with the surrogates fully inhibited samples to the samples that had at least one SARS-CoV-2 gene target inhibited. We found each surrogate suggested the following number of samples were fully inhibited that had signal for at least one SARS-CoV-2 gene target: 12 (Phi6), five (MHV), and one (BCoV).Thus, BCoV is a much better surrogate than MHV and Phi6 for detecting fully inhibited SARS-CoV-2 samples. We suggest that minimally BCoV be run to prevent unnecessary reruns.

The N1 and N2 gene targets were the superior SARS-CoV-2 gene targets. These data support that the N1 and N2 gene targets are highly resilient, and that these targets are typically able to withstand full inhibition in the harsh sewage environments. Nevertheless, the composition of the sewage is always changing, and based on the

sewage matrix, the number of samples fully inhibited could increase or decrease. Certain contaminants in the sewage may cause the RNA not to bind to the silica membrane which could lead to false negatives. One study found that a pH between 6 and 8 caused a sharp decrease in the amount of DNA that adhered to the silica in the spin column (Vandeventer et al., 2012). Another article mentioned that adding ethanol can further improve the binding of RNA to the silica membrane of the spin column (*Review of DNA Extraction Methodologies and Guidelinesfor Protocol Development*, n.d.). A different study revealed that incorrect binding to the spin column may be due to old low-quality ethanol, and it is necessary to use 200proof ethanol for proper binding of RNA to the silica membrane in the spin column (*How DNA Extraction Kits Work in 5 Simple Steps*, 2021).

Phi6 had a statistically different percentage of retrieval than the N1, N2, and E gene targets indicating that it is not an optimal surrogate. BCoV's percentage of retrieval was statistically equivalent to the E, N1, and N2 gene target's percentage of retrieval indicating that it was the best surrogate. MHV's percentage of retrieval was statistically equivalent to the N1 and N2 gene targets percentage of retrieval. As expected, both surrogates that were CoVs were able to more accurately depict the percentage of retrieval of the SARS-CoV-2 gene targets than the bacteriophage Phi6. There is a higher similarity amongst CoVs because they belong to the same genus which is called CoV (Pal et al., n.d.). Also, Phi6 has double-stranded RNA while CoVs have single-stranded RNA (Alexandersen et al., 2020; Butcher et al., 1997). These differences are likely responsible for Phi6's inability to accurately estimate the amount of retrieval of the SARS-CoV-2 gene targets. However, due to the similarities, low risk level, and low costs associated

with using Phi6 as a surrogate, it has been used commonly as a surrogate for CoVs.. Generally, it is desirable to duplex two targets because there is higher certainty that the sample is negative for SARS-CoV-2 or that the sample is truly fully inhibited. Also, if a mutation arises in the portion of gene that is targeted, one of the targets analyzed could become obsolete. Regarding the surrogates, Phi6 did not provide any additional useful data. Phi6 suggested a percentage of retrieval that was much lower than any of the SARS-CoV-2 gene targets, and it deemed many positive samples to be fully inhibited. Thus, Phi6 should not be used. MHV could be duplexed with BCoV. However, MHV was not able to provide any additional information in this study. Every sample that BCoV deemed fully inhibited MHV deemed fully inhibited as well as several other samples that were not fully inhibited. BCoV is readily available for purchase as a cattle vaccine for less than $50, and it is highly concentrated (~10^8 CP/mL). Therefore, one bottle of the vaccine can be used for a long duration. In contrast, MHV was available from one supplier, only available as a dilute solution (~10^5 CP/L), and was costly. Also, signal for MHV was only obtained when it was run through the QIAamp Viral RNA Mini Kit from QIAGEN meaning that the initial solution could not be used for a positive control unless it was run through a spin column. BCoV produced the least number of false negatives, it is cost-effective, it was not statistically different from any of the SARS-CoV-2 gene targets, and it can be used as a positive control without running it through a spin column. Because the spiking solution will always be the same, there will be no concern of mutations. Therefore, although two targets could be duplexed there does not appear to be any benefit to duplexing the surrogates because the other surrogates were unable to provide any additional data that was not provided by BCoV. Thus, current data suggest that BCoV

alone could be used to assess inhibition. However, if mutations in SARS-CoV-2 lead to the selection of new SARS-CoV-2 gene targets, the surrogates would need to be reassessed.

7.6 Conclusion

Phi6 was initially thought to be a desirable surrogate for SARS-CoV-2 due to its similar structure to SARS-CoV-2, low safety requirements (biosafety level one lab), and low cost. These traits made this virus of high interest in this study and in many other studies. Unfortunately, these data support that Phi6 is not a good surrogate for the SARS-CoV-2 gene targets selected in this study. The Phi6 data suggested that the percentage of gene target retrieved was much lower than the true values, and that the SARS-CoV-2 gene targets were much more susceptible to inhibitors. As expected, the CoV surrogates in this study were able to more accurately depict how the SARS-CoV-2 gene targets would act under various conditions. BCoV was found to be the superior surrogate positively identifying 3/4 fully inhibited samples. Also, BCoV was shown to have no statistical difference between its percentage of retrieval and the SARS-CoV-2 gene target's percentage of retrieval. MHV was comparable with the N1 and N2 gene targets, but the difference was significant between the percentages of retrieval for E and MHV. BCoV and MHV could be duplexed and used to assess inhibition of SARS-CoV-2 gene targets. However, the expense of MHV and its generally poor performance does not warrant its use. BCoV could be used as a singular target to monitor SARS-CoV-2 inhibition. Because BCoV's retrieval is statistically the same as the SARS-CoV-2 gene targets, the amount of inhibition (100%/percentage of retrieval) could be multiplied by the SARS-

CoV-2 concentration to obtain a more accurate depiction of the SARS-CoV-2 gene target concentration in sewage.

CHAPTER EIGHT

HANDLING A NEW VIRUS OUTBREAK

This study provides insight into what steps researchers should assess when developing a new sewage monitoring workflow as new viral diseases emerge. The primary information required will be the sequences needed to amplify a unique region of the new virus in order to detect the presence or absence of the virus in a sample. It is likely that the CDC and the WHO will be the first organizations to provide sequences that are targeted for various regions of the virus. These sequences will include primer and probe sequences which target specific regions on the virus of interest. These gene targets will then need to be assessed to determine their rates of degradation. Minimally, three different temperatures should be observed to allow an Arrhenius plot to be graphed with three data points. The degradation rate and half-life can then be calculated at any temperature. It is important to understand how quickly the gene targets are degrading to ensure that they are stable enough to persist in their environment long enough to be detected. In a clinical setting, this would not be a concern. However, when testing sewage, degradation rate is vital to ensure that the virus will be detectable. In our study all three gene targets had a half-life value near 100hrs in sewage. Average travel times for the wastewater to leave the home and enter the WWTP would be around two hours (*Frequently Asked Questions for the Water Reclamation Plant*, n.d.). The half-lives confirm that there will be minimal degradation of the gene targets in the sewage, during collection, and shipping to the lab. We found that the E, N1, and N2 gene targets were reliable gene targets to assess SARS-CoV-2 in the sewage. It is recommended that any new virus that emerges should be

monitored in the sewage if possible. However, if the virus is not present in the feces or urine of patients, the virus is unlikely to be present at high concentrations in sewage. Unfortunately, all viruses cannot be monitored in the sewage. Therefore, the applicability of wastewater testing for new viruses will need to be assessed on a virus-by-virus basis. Once the reliability of the targets has been assessed by determining that they degrade slowly enough to persist in the sewage, the frequency of the selected targets should be assessed. At least two targets should be evaluated and ideally, they can should be duplexed so that they can be tested for approximately the same cost and time as assessing a singular target. In this study we analyzed the N1, N2, and E gene targets. It was determined that the E gene target only added 2.3% more SARS-CoV-2 detection. Removal of either the N1 or N2 gene targets would result in approximately 10% SARS-CoV-2 signal loss. Therefore, the E gene target was removed, and the N1 and N2 gene targets were retained for ongoing monitoring and research. Based on our data, if a new CoV arose we would suggest that the nucleocapsid gene should be tested prior to the envelope gene if funding does not support both. However, as the virus would be different, the E gene has the potential to be the favored target for detection of the new virus. The N gene may not always be the best gene to target when analyzing new CoVs. Once the best targets have been identified for the new virus, the other gene targets may be eliminated from analysis.

Sewage flow is subject to dilution and large variations. Therefore, normalization can be implemented to in an attempt to improve the correlation between the viral signal and clinical cases. We found that the unnormalized signal always provided the best correlation to the clinical data. Therefore, we suggest that the data should remain

unnormalized at least initially. Normalization may prove beneficial, but it should not be an initial priority. A new virus may have an enhanced correlation to the clinical data if the signal is normalized. The virus's concentration may be normalized in sewage with flow if available, but the cost of flow meters would not be generally justified. If flow data are unavailable, which is typical when samples are not collected from a wastewater treatment plants, fecal indicators can be used instead of flow. However, the RNA based fecal indicator in our study (PMMoV) is dietary based and confounded the data. PMMoV in the sewage appears to be sporadic because it is not always excreted in feces. Some individuals may excrete more PMMoV or have no PMMoV in their feces based on their diet. Also, human feces are not the only source of PMMoV. PMMoV is a plant virus. Currently, PMMoV is recommended as the best RNA fecal indicator, but the DNA fecal indicators normalized the SARS-CoV-2 signal better than PMMoV. Overall, new viruses can be normalized in the sewage with flow rate and with fecal indicators, but there is a chance the unaltered data may provide the best correlation to the clinical data. Inhibition on a sample-by-sample basis must also be assessed. Typically, inhibition is assessed by spiking the sewage with surrogates that structurally resemble the virus of interest. If the new virus is another CoV, the surrogate BCoV would be suggested as a surrogate to determine the degree of sample inhibition. BCoV was shown to be reliable surrogate to assess inhibition for the N1, N2, and E gene targets in our study. Phi6 was determined to have a statistically different percentage of retrieval than all three SARS-CoV-2 gene targets in our study but performed adequately. A virus that is structurally similar should be chosen, but a virus that belongs to the same family of viruses is a good starting place because they will be structurally related to each other. However, it is

necessary to test the reliability of the surrogates. To test the reliability of a surrogate, previously determined virus negative samples should be spiked with the viral RNA from previous positive samples if possible. Then, the percentage of retrieval of the surrogate and the target can be compared directly to determine if their retrieval is statistically the same or different. In the case of SARS-CoV-2, the biosafety level in our lab was not high enough to utilize live SARS-CoV-2 virus. Therefore, the eluate of a sample with an extremely high concentration was spiked into the SARS-CoV-2 negative samples. In situations where the biosafety level is not high enough to work with a live virus, eluate from a highly concentrated sample can be utilized to determine which surrogate is the best match to the gene targets of interest. If the surrogate and the virus do not statistically vary in the amount of sample retrieved, the percentage of retrieval could be factored in to normalize the concentration of the virus. Overall, our study suggested that the SARS-CoV-2 gene targets underwent a moderate amount of inhibition in the sewage. The surrogate BCoV suggested that on average samples only had 47.2% of their concentration retrieved. The surrogate Phi6 suggested that the average amount of retrieval was 27.8%. When the surrogates suggest full inhibition, the sample should be modified to reduce inhibition. All methods listed in this study are used to combat PCR inhibition. Therefore, all our method adjustments could be assessed to reduce inhibition that the new virus may face. In our study, BSA was the only method enhancement that was able to statistically increase the concentration for both SARS-CoV-2 gene targets. However, pasteurization and dilution were also able to enhance the concentration of the N2 gene target. Therefore, it is suggested that pasteurization, addition of BSA to the initial sewage, and dilution could reduce PCR inhibition for a new virus. Reducing inhibition will be extremely

important in the case of a new virus. When a new virus arises, the concentration of virus that can be detected in the sewage will be low. As the virus spreads the concentration will increase and spread to new areas. It is important that the samples with low concentration be detected quickly to identify where the virus is spreading. Depending on the severity of the disease, precautions may need to be taken immediately. Using sewage surveillance will provide data in a timely fashion for whole communities allowing public health measures to be put in place for maximum effect.

Overall, our research supports the idea that monitoring viruses with sewage surveillance is useful and cost effective. If sewage surveillance was implemented sooner, SARS-CoV-2 the data could have allowed decisions on safety precautions to be taken sooner which may have greatly reduced the spread of the virus during the beginning of the pandemic. Implementing sewage surveillance early in the outbreak of a new disease could help reduce the spread of the virus while scientists research the virus. COVID-19 has caused many deaths, especially in the beginning of the pandemic. The next new virus that arises could be much deadlier. Therefore, it is important to reduce the spread of the virus at the beginning of the virus's outbreak rather than combating the disease once it has exponentially spread. Sewage surveillance can be a big part of monitoring a new virus. Wastewater surveillance is particularly important in the beginning of a new virus outbreak when the virus needs to be tracked and when the panic of the new virus has worn off leading to individual's letting their guards down. Currently, sewage surveillance is the best way that SARS-CoV-2 can be monitored at the community level. As time has passed individuals have become complacent and participation in a clinical setting has waned. This means that clinically there will be fewer reported cases due to individuals

with mild symptoms opting out of testing or using at-home testing kits and not reporting their results. Sewage surveillance will likely continue to be a major part of assessing how COVID-19 is spreading and when COVID-19 cases are escalating. We suggest that sewage surveillance should be implemented as soon as possible with new viruses as it is not reliant on participation and can be useful for monitoring specific areas allowing safety precautions to be placed where needed instead of large-scale closures.

APPENDIX

INSTITUTIONAL BIOSAFETY COMMITTEE APPROVAL LETTER

OAKLAND UNIVERSITY

Institutional Biosafety Committee

September 24, 2020

Associate Professor David Szlag
Department of Chemistry

Reference: IBC Application # 4570, "Community Monitoring of COVID-19 Infections via Innovative Sewage Surveillance using ddPCR"

Dear Dr. Szlag:

The Institutional Biosafety Committee (IBC), responsible for the review of research involving materials of human origin, recombinant DNA, and biohazardous, infectious or toxic materials, has reviewed the original proposal, noted the revisions provided by you upon committee's request, and approved the revised project referenced above.

The proposed activities were approved at BSL-2+ for a period of three years starting September 24, 2020 and ending September 23, 2023; and assigned the approval # 4570-20.

Any unanticipated problems, changes in, or halting of the activities proposed in the project, as described, must be promptly reported to the IBC or the biosafety officer.

In addition, all faculty, staff and students involved in the project must be trained and qualified to conduct the project in a responsible manner. Please note, in this regard, that you, the principal investigator, are fully responsible at all times for limiting or restricting access to laboratories and for assessing, in each circumstances, who may enter or work in a laboratory when biohazardous materials are in use.

Please let me know if you have any questions. Thank you.

Sincerely,

Domenic Luongo, CHMM
Director - Laboratory Safety and Compliance

The Research Office

371 Wilson Blvd, Rochester, MI 48309-4486
(248) 370-4898 | Fax: (248) 370-2973 | oakland.edu

REFERENCES

Acharya, K. R., Dhand, N. K., Whittington, R. J., & Plain, K. M. (2017). PCR Inhibition of a Quantitative PCR for Detection of Mycobacterium avium Subspecies Paratuberculosis DNA in Feces: Diagnostic Implications and Potential Solutions. *Frontiers in Microbiology*, *8*. https://www.frontiersin.org/article/10.3389/fmicb.2017.00115

Ahmed, W., Bertsch, P. M., Angel, N., Bibby, K., Bivins, A., Dierens, L., Edson, J., Ehret, J., Gyawali, P., Hamilton, K. A., Hosegood, I., Hugenholtz, P., Jiang, G., Kitajima, M., Sichani, H. T., Shi, J., Shimko, K. M., Simpson, S. L., Smith, W. J. M., … Mueller, J. F. (2020). Detection of SARS-CoV-2 RNA in commercial passenger aircraft and cruise ship wastewater: A surveillance tool for assessing the presence of COVID-19 infected travellers. *Journal of Travel Medicine*, *27*(5), taaa116. https://doi.org/10.1093/jtm/taaa116

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. (2020). Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology. *Environmental Research*, *191*, 110092. https://doi.org/10.1016/j.envres.2020.110092

Ahmed, W., Bivins, A., Bertsch, P. M., Bibby, K., Gyawali, P., Sherchan, S. P., Simpson, S. L., Thomas, K. V., Verhagen, R., Kitajima, M., Mueller, J. F., & Korajkic, A. (2021). Intraday variability of indicator and pathogenic viruses in 1-h and 24-h composite wastewater samples: Implications for wastewater-based epidemiology. *Environmental Research*, *193*, 110531. https://doi.org/10.1016/j.envres.2020.110531

Ahmed, W., Gyawali, P., Feng, S., & McLellan, S. L. (2019). Host Specificity and Sensitivity of Established and Novel Sewage-Associated Marker Genes in Human and Nonhuman Fecal Samples. *Applied and Environmental Microbiology*, *85*(14), e00641-19. https://doi.org/10.1128/AEM.00641-19

Ahmed, W., Payyappat, S., Cassidy, M., & Besley, C. (2019). A duplex PCR assay for the simultaneous quantification of Bacteroides HF183 and crAssphage CPQ_056 marker genes in untreated sewage and stormwater. *Environment International*, *126*, 252–259. https://doi.org/10.1016/j.envint.2019.01.035

Ahmed, W., Payyappat, S., Cassidy, M., Besley, C., & Power, K. (2018). Novel crAssphage marker genes ascertain sewage pollution in a recreational lake receiving urban stormwater runoff. *Water Research*, *145*, 769–778. https://doi.org/10.1016/j.watres.2018.08.049

Ahmed, W., Sidhu, J. P. S., Smith, K., Beale, D. J., Gyawali, P., & Toze, S. (2016). Distributions of Fecal Markers in Wastewater from Different Climatic Zones for Human Fecal Pollution Tracking in Australian Surface Waters. *Applied and Environmental Microbiology*, *82*(4), 1316–1323. https://doi.org/10.1128/AEM.03765-15

Ahmed, W., Simpson, S. L., Bertsch, P. M., Bibby, K., Bivins, A., Blackall, L. L., Bofill-Mas, S., Bosch, A., Brandão, J., Choi, P. M., Ciesielski, M., Donner, E., D'Souza, N., Farnleitner, A. H., Gerrity, D., Gonzalez, R., Griffith, J. F., Gyawali, P., Haas, C. N., … Shanks, O. C. (2022). Minimizing errors in RT-PCR detection and quantification of SARS-CoV-2 RNA for wastewater surveillance. *Science of The Total Environment*, *805*, 149877. https://doi.org/10.1016/j.scitotenv.2021.149877

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. *The Science of the Total Environment*, *761*, 144216. https://doi.org/10.1016/j.scitotenv.2020.144216

Al-Duroobi, H., Moghadam, S. V., Phan, D. C., Jafarzadeh, A., Matta, A., & Kapoor, V. (2021). Wastewater surveillance of SARS-CoV-2 corroborates heightened community infection during the initial peak of COVID-19 in Bexar County, Texas. *FEMS Microbes*, *2*, xtab015. https://doi.org/10.1093/femsmc/xtab015

Alene, M., Yismaw, L., Assemie, M. A., Ketema, D. B., Mengist, B., Kassie, B., & Birhan, T. Y. (2021). Magnitude of asymptomatic COVID-19 cases throughout the course of infection: A systematic review and meta-analysis. *PLoS ONE*, *16*(3), e0249090. https://doi.org/10.1371/journal.pone.0249090

Alexandersen, S., Chamings, A., & Bhatta, T. R. (2020). SARS-CoV-2 genomic and subgenomic RNAs in diagnostic samples are not an indicator of active replication. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-19883-7

Alluwaimi, A. M., Alshubaith, I. H., Al-Ali, A. M., & Abohelaika, S. (2020). The Coronaviruses of Animals and Birds: Their Zoonosis, Vaccines, and Models for SARS-CoV and SARS-CoV2. *Frontiers in Veterinary Science*, *7*. https://www.frontiersin.org/article/10.3389/fvets.2020.582287

Al-Soud, W. A., Ouis, I.-S., Li, D.-Q., Ljungh, Å., & Wadström, T. (2005). Characterization of the PCR inhibitory effect of bile to optimize real-time PCR detection of Helicobacter species. *FEMS Immunology & Medical Microbiology*, *44*(2), 177–182. https://doi.org/10.1016/j.femsim.2004.12.004

Artesi, M., Bontems, S., Göbbels, P., Franckh, M., Maes, P., Boreux, R., Meex, C., Melin, P., Hayette, M.-P., Bours, V., & Durkin, K. (2020). A Recurrent Mutation at Position 26340 of SARS-CoV-2 Is Associated with Failure of the E Gene Quantitative Reverse Transcription-PCR Utilized in a Commercial Dual-Target Diagnostic Assay. *Journal of Clinical Microbiology*, *58*(10), e01598-20. https://doi.org/10.1128/JCM.01598-20

Azer, S. A. (2020). COVID-19: Pathophysiology, diagnosis, complications and investigational therapeutics. *New Microbes and New Infections*, *37*, 100738. https://doi.org/10.1016/j.nmni.2020.100738

Baek, Y. M., Jang, K.-J., Lee, H., Yoon, S., Baek, A., Lee, K., & Kim, D.-E. (2019). The bacterial endoribonuclease RNase E can cleave RNA in the absence of the RNA chaperone Hfq. *The Journal of Biological Chemistry*, *294*(44), 16465–16478. https://doi.org/10.1074/jbc.RA119.010105

Bailey et al. (2022). *Persistence of Coronavirus Surrogates on Meat and Fish Products during Long-Term Storage | Applied and Environmental Microbiology*. https://journals.asm.org/doi/10.1128/aem.00504-22

Bangiyev, R., Chudaev, M., Schaffner, D. W., & Goldman, E. (n.d.). Higher Concentrations of Bacterial Enveloped Virus Phi6 Can Protect the Virus from Environmental Decay. *Applied and Environmental Microbiology*, *87*(21), e01371-21. https://doi.org/10.1128/AEM.01371-21

Barati rashvanlou, R., Rezaee, A., Farzadkia, M., Gholami, M., & Kermani, M. (n.d.). Effect of micro-aerobic process on improvement of anaerobic digestion sewage sludge treatment: Flow cytometry and ATP assessment. *RSC Advances*, *10*(59), 35718–35728. https://doi.org/10.1039/d0ra05540a

Bartas, M., Volná, A., Beaudoin, C. A., Poulsen, E. T., Červeň, J., Brázda, V., Špunda, V., Blundell, T. L., & Pečinka, P. (2022). Unheeded SARS-CoV-2 proteins? A deep look into negative-sense RNA. *Briefings in Bioinformatics*, *23*(3), bbac045. https://doi.org/10.1093/bib/bbac045

Beattie, R. E., Blackwood, A. D., Clerkin, T., Dinga, C., & Noble, R. T. (2022). Evaluating the impact of sample storage, handling, and technical ability on the decay and recovery of SARS-CoV-2 in wastewater. *PLOS ONE*, *17*(6), e0270659. https://doi.org/10.1371/journal.pone.0270659

Bechhofer, D. H., & Deutscher, M. P. (2019). Bacterial ribonucleases and their roles in RNA metabolism. *Critical Reviews in Biochemistry and Molecular Biology*, *54*(3), 242–300. https://doi.org/10.1080/10409238.2019.1651816

Bernhardt, H. S., & Tate, W. P. (2012). Primordial soup or vinaigrette: Did the RNA world evolve at acidic pH? *Biology Direct*, *7*, 4. https://doi.org/10.1186/1745-6150-7-4

Beyene, G. T., Alemu, F., Kebede, E. S., Alemayehu, D. H., Seyoum, T., Tefera, D. A., Assefa, G., Tesfaye, A., Habte, A., Bedada, G., Tegene, B., Yeshambaw, M., Wassie, L., Mihret, A., Abdissa, A., & Mulu, A. (2021). Saliva is superior over nasopharyngeal swab for detecting SARS-CoV2 in COVID-19 patients. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-02097-2

Biosafety Level Requirements. (n.d.). Retrieved April 5, 2022, from https://www.phe.gov/s3/BioriskManagement/biocontainment/Pages/BSL-Requirements.aspx

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

Bland, J., Kavanaugh, A., Hong, L. K., & Kadkol, S. S. (2021). Development and validation of viral load assays to quantitate SARS-CoV-2. *Journal of Virological Methods*, *291*, 114100. https://doi.org/10.1016/j.jviromet.2021.114100

blog, L. B. P. (n.d.). *The SARS-CoV-2 Variant and its Impact on Diagnostic Testing*. Lab-Best-Practice. Retrieved January 26, 2022, from https://health.ucdavis.edu/blog/lab-best-practice/the-sars-cov-2-variant-and-its-impact-on-diagnostic-testing/2021/01

Boegel, S. J., Gabriel, M., Sasges, M., Petri, B., D'Agostino, M. R., Zhang, A., Ang, J. C., Miller, M. S., Meunier, S. M., & Aucoin, M. G. (2021). Robust Evaluation of Ultraviolet-C Sensitivity for SARS-CoV-2 and Surrogate Coronaviruses. *Microbiology Spectrum*, *9*(2), e00537-21. https://doi.org/10.1128/Spectrum.00537-21

Boonchai, W., & Iamtharachai, P. (2010). The pH of commonly available soaps, liquid cleansers, detergents and alcohol gels. *Dermatitis: Contact, Atopic, Occupational, Drug*, *21*(3), 154–156.

Botes, M., de Kwaadsteniet, M., & Cloete, T. E. (2013). Application of quantitative PCR for the detection of microorganisms in water. *Analytical and Bioanalytical Chemistry*, *405*(1), 91–108. https://doi.org/10.1007/s00216-012-6399-3

Brant, A. C., Tian, W., Majerciak, V., Yang, W., & Zheng, Z.-M. (2021). SARS-CoV-2: From its discovery to genome structure, transcription, and replication. *Cell & Bioscience*, *11*(1), 136. https://doi.org/10.1186/s13578-021-00643-z

Brisco, M. J., & Morley, A. A. (2012). Quantification of RNA integrity and its use for measurement of transcript number. *Nucleic Acids Research*, *40*(18), e144. https://doi.org/10.1093/nar/gks588

Britton, G. J., Chen-Liaw, A., Cossarini, F., Livanos, A. E., Spindler, M. P., Plitt, T., Eggers, J., Mogno, I., Gonzalez-Reiche, A. S., Siu, S., Tankelevich, M., Grinspan, L. T., Dixon, R. E., Jha, D., van de Guchte, A., Khan, Z., Martinez-Delgado, G., Amanat, F., Hoagland, D. A., … Faith, J. J. (2021). Limited intestinal inflammation despite diarrhea, fecal viral RNA and SARS-CoV-2-specific IgA in patients with acute COVID-19. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-92740-9

Buchanan, J. R. (n.d.). *Wastewater Basics 101.* 38.

Butcher, S. J., Dokland, T., Ojala, P. M., Bamford, D. H., & Fuller, S. D. (1997). Intermediates in the assembly pathway of the double-stranded RNA virus phi6. *The EMBO Journal*, *16*(14), 4477–4487. https://doi.org/10.1093/emboj/16.14.4477

Cao, Y., Raith, M. R., Smith, P. D., Griffith, J. F., Weisberg, S. B., Schriewer, A., Sheldon, A., Crompton, C., Amenu, G. G., Gregory, J., Guzman, J., Goodwin, K. D., Othman, L., Manasjan, M., Choi, S., Rapoport, S., Steele, S., Nguyen, T., & Yu, X. (2017). Regional Assessment of Human Fecal Contamination in Southern California Coastal Drainages. *International Journal of Environmental Research and Public Health*, *14*(8), 874. https://doi.org/10.3390/ijerph14080874

Carol Kreader. (1995). *Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein*. https://doi.org/10.1128/aem.62.3.1102-1106.1996

Cascella, M., Rajnik, M., Aleem, A., Dulebohn, S. C., & Di Napoli, R. (2022a). Features, Evaluation, and Treatment of Coronavirus (COVID-19). In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK554776/

Cascella, M., Rajnik, M., Aleem, A., Dulebohn, S. C., & Di Napoli, R. (2022b). Features, Evaluation, and Treatment of Coronavirus (COVID-19). In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK554776/

CDC. (2020a, February 11). *COVID-19 and Your Health*. Centers for Disease Control and Prevention. https://www.cdc.gov/coronavirus/2019-ncov/your-health/about-covid-19/basics-covid-19.html

CDC. (2020b, February 11). *Nucleic Acid Amplification Tests (NAATs)*. Centers for Disease Control and Prevention. https://www.cdc.gov/coronavirus/2019-ncov/lab/naats.html

CDC. (2020c, February 11). *Real-time RT-PCR Primers & Probes (Research Use Only)*. Centers for Disease Control and Prevention. https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html

CDC. (2022a, January 5). *CDC Museum COVID-19 Timeline*. Centers for Disease Control and Prevention. https://www.cdc.gov/museum/timeline/covid19.html

CDC. (2022b, January 26). *Wastewater Surveillance Testing Methods*. https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/testing-methods.html

CDC. (2022c, March 21). *National Wastewater Surveillance System*. Centers for Disease Control and Prevention. https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/wastewater-surveillance.html

CDC. (2022d, June 14). *Wastewater Surveillance Data Reporting and Analytics*. Centers for Disease Control and Prevention. https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/data-reporting-analytics.html

CDC's Global Resources Pivot to Address COVID-19. (2021, February 8). https://www.cdc.gov/globalhealth/resources/reports/annual/2021/global-resources-pivot.html

Cerrada-Romero, C., Berastegui-Cabrera, J., Camacho-Martínez, P., Goikoetxea-Aguirre, J., Pérez-Palacios, P., Santibáñez, S., José Blanco-Vidal, M., Valiente, A., Alba, J., Rodríguez-Álvarez, R., Pascual, Á., Oteo, J. A., Miguel Cisneros, J., Pachón, J., Casas-Flecha, I., Cordero, E., Pozo, F., & Sánchez-Céspedes, J. (2022). Excretion and viability of SARS-CoV-2 in feces and its association with the clinical outcome of COVID-19. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-11439-7

Chen, X., Huang, Z., Wang, J., Zhao, S., Wong, M. C.-S., Chong, K. C., He, D., & Li, J. (2021). Ratio of asymptomatic COVID-19 cases among ascertained SARS-CoV-2 infections in different regions and population groups in 2020: A systematic review and meta-analysis including 130 123 infections from 241 studies. *BMJ Open*, *11*(12), e049752. https://doi.org/10.1136/bmjopen-2021-049752

Chu, M. A., Jang, Y. Y., Lee, D. W., Kim, S. H., Ryoo, N., Park, S., Lee, J. H., & Chung, H. L. (2021). Viral load and rebound in children with coronavirus disease 2019 during the first outbreak in Daegu city. *Clinical and Experimental Pediatrics*, *64*(12), 652–660. https://doi.org/10.3345/cep.2020.02033

Codello, A., McLellan, S. L., Steinberg, P., Potts, J., Scanes, P., Ferguson, A., Hose, G. C., Griffith, M., Roguet, A., Lydon, K. A., Maher, W. A., Krikowa, F., & Chariton, A. (2021). A weight-of-evidence approach for identifying potential sources of untreated sewage inputs into a complex urbanized catchment. *Environmental Pollution*, *275*, 116575. https://doi.org/10.1016/j.envpol.2021.116575

Coerdt, K. M., & Khachemoune, A. (2021). Corona viruses: Reaching far beyond the common cold. *African Health Sciences*, *21*(1), 207–213. https://doi.org/10.4314/ahs.v21i1.27

Colton, H., Ankcorn, M., Yavuz, M., Tovey, L., Cope, A., Raza, M., Keeley, A. J., State, A., Poller, B., Parker, M., de Silva, T. I., & Evans, C. (2021). Improved sensitivity using a dual target, E and RdRp assay for the diagnosis of SARS-CoV-2 infection: Experience at a large NHS Foundation Trust in the UK. *The Journal of Infection*, *82*(1), 159–198. https://doi.org/10.1016/j.jinf.2020.05.061

Common Human Coronaviruses | CDC. (2020, May 27). https://www.cdc.gov/coronavirus/general-information.html

Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. K., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M. L., Mulders, D. G., Haagmans, B. L., van der Veer, B., van den Brink, S., Wijsman, L., Goderski, G., Romette, J.-L., Ellis, J., Zambon, M., … Drosten, C. (2020a). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*, *25*(3), 2000045. https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045

Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. K., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M. L., Mulders, D. G., Haagmans, B. L., van der Veer, B., van den Brink, S., Wijsman, L., Goderski, G., Romette, J.-L., Ellis, J., Zambon, M., … Drosten, C. (2020b). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*, *25*(3), 2000045. https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045

Coronavirus. (n.d.). Retrieved March 28, 2022, from https://www.who.int/westernpacific/health-topics/coronavirus

COVID-19, cold, allergies and the flu: What are the differences? (n.d.). Mayo Clinic. Retrieved September 18, 2022, from https://www.mayoclinic.org/diseases-conditions/coronavirus/in-depth/covid-19-cold-flu-and-allergies-differences/art-20503981

CSR. (n.d.). *MERS outbreaks*. World Health Organization - Regional Office for the Eastern Mediterranean. Retrieved February 16, 2022, from http://www.emro.who.int/health-topics/mers-cov/mers-outbreaks.html

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. *The Science of the Total Environment*, *770*, 145319. https://doi.org/10.1016/j.scitotenv.2021.145319

Daughton, C. G. (2020). Wastewater surveillance for population-wide Covid-19: The present and future. *The Science of the Total Environment*, *736*, 139631. https://doi.org/10.1016/j.scitotenv.2020.139631

DdPCR Supermix for Probes (No dUTP). (n.d.). Retrieved March 8, 2023, from https://www.bio-rad.com/sites/default/files/webroot/web/pdf/lsr/literature/10026868.pdf

Deng, Y., Zheng, X., Xu, X., Chui, H., Lai, W., Li, S., Tun, H. M., Poon, L. L. M., Ding, J., Peiris, M., Leung, G. M., & Zhang, T. (2022). Use of Sewage Surveillance for COVID-19: A Large-Scale Evidence-Based Program in Hong Kong. *Environmental Health Perspectives*, *130*(5), 057008. https://doi.org/10.1289/EHP9966

Dhama, K., Khan, S., Tiwari, R., Sircar, S., Bhat, S., Malik, Y. S., Singh, K. P., Chaicumpa, W., Bonilla-Aldana, D. K., & Rodriguez-Morales, A. J. (2020). Coronavirus Disease 2019–COVID-19. *Clinical Microbiology Reviews*, *33*(4), e00028-20. https://doi.org/10.1128/CMR.00028-20

Eilert, K. D., & Foran, D. R. (2009). Polymerase Resistance to Polymerase Chain Reaction Inhibitors in Bone*. *Journal of Forensic Sciences*, *54*(5), 1001–1007. https://doi.org/10.1111/j.1556-4029.2009.01116.x

Fabre, A.-L., Colotte, M., Luis, A., Tuffet, S., & Bonnet, J. (2014). An efficient method for long-term room temperature storage of RNA. *European Journal of Human Genetics*, *22*(3), Article 3. https://doi.org/10.1038/ejhg.2013.145

Farell, E. M., & Alexandre, G. (2012). Bovine serum albumin further enhances the effects of organic solvents on increased yield of polymerase chain reaction of GC-rich templates. *BMC Research Notes*, *5*, 257. https://doi.org/10.1186/1756-0500-5-257

Farkas, K., Hillary, L. S., Malham, S. K., McDonald, J. E., & Jones, D. L. (2020). Wastewater and public health: The potential of wastewater surveillance for monitoring COVID-19. *Current Opinion in Environmental Science & Health*, *17*, 14–20. https://doi.org/10.1016/j.coesh.2020.06.001

Fedorenko, A., Grinberg, M., Orevi, T., & Kashtan, N. (2020). Survival of the enveloped bacteriophage Phi6 (a surrogate for SARS-CoV-2) in evaporated saliva microdroplets deposited on glass surfaces. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-79625-z

Feng, S., Bootsma, M., & McLellan, S. L. (n.d.). Human-Associated Lachnospiraceae Genetic Markers Improve Detection of Fecal Pollution Sources in Urban Waters. *Applied and Environmental Microbiology*, *84*(14), e00309-18. https://doi.org/10.1128/AEM.00309-18

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

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

Fordyce, S. L., Kampmann, M.-L., van Doorn, N. L., & Gilbert, M. T. P. (2013). Long-term RNA persistence in postmortem contexts. *Investigative Genetics*, *4*(1), 7. https://doi.org/10.1186/2041-2223-4-7

Franke, G., Knobling, B., Brill, F. H., Becker, B., Klupp, E. M., Belmar Campos, C., Pfefferle, S., Lütgehetmann, M., & Knobloch, J. K. (2021). An automated room disinfection system using ozone is highly active against surrogates for SARS-CoV-2. *The Journal of Hospital Infection*, *112*, 108–113. https://doi.org/10.1016/j.jhin.2021.04.007

Frequently Asked Questions for the Water Reclamation Plant. (n.d.). Retrieved December 5, 2022, from https://www.rochestermn.gov/home/showpublisheddocument/4333/635609934479530000

Frilander, M., Poranen, M., & Bamford, D. H. (1995). The large genome segment of dsRNA bacteriophage phi6 is the key regulator in the in vitro minus and plus strand synthesis. *RNA*, *1*(5), 510–518.

Galani, A., Aalizadeh, R., Kostakis, M., Markou, A., Alygizakis, N., Lytras, T., Adamopoulos, P. G., Peccia, J., Thompson, D. C., Kontou, A., Karagiannidis, A., Lianidou, E. S., Avgeris, M., Paraskevis, D., Tsiodras, S., Scorilas, A., Vasiliou, V., Dimopoulos, M.-A., & Thomaidis, N. S. (2022). SARS-CoV-2 wastewater surveillance data can predict hospitalizations and ICU admissions. *The Science of the Total Environment*, *804*, 150151. https://doi.org/10.1016/j.scitotenv.2021.150151

Gandhi, R. T., Lynch, J. B., & del Rio, C. (2020). Mild or Moderate Covid-19. *New England Journal of Medicine*, *383*(18), 1757–1766. https://doi.org/10.1056/NEJMcp2009249

Gendron, L., Verreault, D., Veillette, M., Moineau, S., & Duchaine, C. (2010). Evaluation of Filters for the Sampling and Quantification of RNA Phage Aerosols. *Aerosol Science and Technology*, *44*(10), 893–901. https://doi.org/10.1080/02786826.2010.501351

Geraghty, E. (2022, May 21). How GIS Brings Wastewater Surveillance to Life. *Industry Blogs*. https://www.esri.com/en-us/industries/blog/articles/how-gis-brings-wastewater-surveillance-to-life/

Gerrity, D., Papp, K., Stoker, M., Sims, A., & Frehner, W. (2020). Early-pandemic wastewater surveillance of SARS-CoV-2 in Southern Nevada: Methodology, occurrence, and incidence/prevalence considerations. *Water Research X*, *10*, 100086. https://doi.org/10.1016/j.wroa.2020.100086

Giambernardi, T. A., Rodeck, U., & Klebe, R. J. (1998). Bovine serum albumin reverses inhibition of RT-PCR by melanin. *BioTechniques*, *25*(4), 564–566. https://doi.org/10.2144/98254bm03

Gomes, M., Bartolomeu, M., Vieira, C., Gomes, A. T. P. C., Faustino, M. A. F., Neves, M. G. P. M. S., & Almeida, A. (2022). Photoinactivation of Phage Phi6 as a SARS-CoV-2 Model in Wastewater: Evidence of Efficacy and Safety. *Microorganisms*, *10*(3), 659. https://doi.org/10.3390/microorganisms10030659

Gopaul, R., Davis, J., Gangai, L., & Goetz, L. (2020). Practical Diagnostic Accuracy of Nasopharyngeal Swab Testing for Novel Coronavirus Disease 2019 (COVID-19). *Western Journal of Emergency Medicine*, *21*(6), 1–4. https://doi.org/10.5811/westjem.2020.8.48420

Graham, K. E., Loeb, S. K., Wolfe, M. K., Catoe, D., Sinnott-Armstrong, N., Kim, S., Yamahara, K. M., Sassoubre, L. M., Mendoza Grijalva, L. M., Roldan-Hernandez, L., Langenfeld, K., Wigginton, K. R., & Boehm, A. B. (2020). SARS-CoV-2 RNA in Wastewater Settled Solids Is Associated with COVID-19 Cases in a Large Urban Sewershed. *Environmental Science & Technology*, acs.est.0c06191. https://doi.org/10.1021/acs.est.0c06191

Greenwald, H. D., Kennedy, L. C., Hinkle, A., Whitney, O. N., Fan, V. B., Crits-Christoph, A., Harris-Lovett, S., Flamholz, A. I., Al-Shayeb, B., Liao, L. D., Beyers, M., Brown, D., Chakrabarti, A. R., Dow, J., Frost, D., Koekemoer, M., Lynch, C., Sarkar, P., White, E., … Nelson, K. L. (2021). Tools for interpretation of wastewater SARS-CoV-2 temporal and spatial trends demonstrated with data collected in the San Francisco Bay Area. *Water Research X*, *12*, 100111. https://doi.org/10.1016/j.wroa.2021.100111

Haque, R., Moe, C. L., Raj, S. J., Ong, L., Charles, K., Ross, A. G., Shirin, T., Raqib, R., Sarker, P., Rahman, M., Rahman, M. Z., Amin, N., Mahmud, Z. H., Rahman, M., Johnston, D., Akter, N., Khan, T. A., Hossain, Md. A., Hasan, R., … Bhattacharya, P. (2022). Wastewater surveillance of SARS-CoV-2 in Bangladesh: Opportunities and challenges. *Current Opinion in Environmental Science & Health*, *27*, 100334. https://doi.org/10.1016/j.coesh.2022.100334

Haramoto, E., Kitajima, M., Kishida, N., Konno, Y., Katayama, H., Asami, M., & Akiba, M. (2013). Occurrence of Pepper Mild Mottle Virus in Drinking Water Sources in Japan. *Applied and Environmental Microbiology*, *79*(23), 7413–7418. https://doi.org/10.1128/AEM.02354-13

Harris-Lovett, S., Nelson, K., 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., Pagsuyoin, S., Prasek, S., Radniecki, T. S., Ram, J. L., Roper, D. K., Safford, H., … Korfmacher, K. S. (2021). Wastewater surveillance for SARS-CoV-2 on college campuses: Initial efforts, lessons learned and research needs. *MedRxiv*, 2021.02.01.21250952. https://doi.org/10.1101/2021.02.01.21250952

Hata, A., & Honda, R. (2020). Potential Sensitivity of Wastewater Monitoring for SARS-CoV-2: Comparison with Norovirus Cases. *Environmental Science & Technology*, acs.est.0c02271. https://doi.org/10.1021/acs.est.0c02271

Health Desk. (2020, September 18). *Are you as likely to die from a common cold as COVID-19?* [Health Desk.org]. https://health-desk.org/articles/are-you-as-likely-to-die-from-a-common-cold-as-covid-19

Healthline. (2020). *What is a Coronavirus?* Healthline. https://www.healthline.com/health/coronavirus-types

Hernandez Acosta, R. A., Esquer Garrigos, Z., Marcelin, J. R., & Vijayvargiya, P. (2022). COVID-19 Pathogenesis and Clinical Manifestations. *Infectious Disease Clinics of North America*, *36*(2), 231–249. https://doi.org/10.1016/j.idc.2022.01.003

Hillwig, M. S., Rizhsky, L., Wang, Y., Umanskaya, A., Essner, J. J., & MacIntosh, G. C. (2009). Zebrafish RNase T2 genes and the evolution of secretory ribonucleases in animals. *BMC Evolutionary Biology*, *9*, 170. https://doi.org/10.1186/1471-2148-9-170

Hiscott, J., Alexandridi, M., Muscolini, M., Tassone, E., Palermo, E., Soultsioti, M., & Zevini, A. (2020). The global impact of the coronavirus pandemic. *Cytokine & Growth Factor Reviews*, *53*, 1–9. https://doi.org/10.1016/j.cytogfr.2020.05.010

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

Holm, R. H., Nagarkar, M., Yeager, R. A., Talley, D., Chaney, A. C., Rai, J. P., Mukherjee, A., Rai, S. N., Bhatnagar, A., & Smith, T. (2022). Surveillance of RNase P, PMMoV, and CrAssphage in wastewater as indicators of human fecal concentration across urban sewer neighborhoods, Kentucky. *FEMS Microbes*, *3*, xtac003. https://doi.org/10.1093/femsmc/xtac003

Hong, P.-Y., Rachmadi, A. T., Mantilla-Calderon, D., Alkahtani, M., Bashawri, Y. M., Al Qarni, H., O'Reilly, K. M., & Zhou, J. (2021). Estimating the minimum number of SARS-CoV-2 infected cases needed to detect viral RNA in wastewater: To what extent of the outbreak can surveillance of wastewater tell us? *Environmental Research*, *195*, 110748. https://doi.org/10.1016/j.envres.2021.110748

How DNA Extraction Kits Work in 5 Simple Steps. (2021, October 27). https://bitesizebio.com/25018/how-dna-extraction-kits-work-in-the-lab/

How Sewage Pollution Ends Up In Rivers. (n.d.). American Rivers. Retrieved September 18, 2022, from https://www.americanrivers.org/threats-solutions/clean-water/sewage-pollution/

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

Hunter, M. E., Ferrante, J. A., Meigs-Friend, G., & Ulmer, A. (2019). Improving eDNA yield and inhibitor reduction through increased water volumes and multi-filter isolation techniques. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-40977-w

Islam, M. F., & Johnston, R. B. (2006). Household Pasteurization of Drinking-water: The Chulli Water-treatment System. *Journal of Health, Population, and Nutrition*, *24*(3), 356–362.

Jamal, H. (n.d.). *Physical Characteristics of Sewage (Waste Water) | Characteristics of Fresh Sewage | Temperature, Color, Odor, Solids*. Retrieved January 28, 2022, from https://www.aboutcivil.org/characteristics-sewage-physical

Jennings, W. C., Gálvez-Arango, E., Prieto, A. L., & Boehm, A. B. (2020). CrAssphage for fecal source tracking in Chile: Covariation with norovirus, HF183, and bacterial indicators. *Water Research X*, *9*, 100071. https://doi.org/10.1016/j.wroa.2020.100071

Johnson, H., Garg, M., Shantikumar, S., Thachil, J., Rai, B., Aboumarzouk, O. M., Hashim, H., & Philip, J. (2021). COVID-19 (SARS-CoV-2) in Non-Airborne body fluids: A systematic review & Meta-analysis. *Turkish Journal of Urology*, *47*(2), 87–97. https://doi.org/10.5152/tud.2021.20586

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. *The Science of the Total Environment*, *749*, 141364. https://doi.org/10.1016/j.scitotenv.2020.141364

Juel, M. A. I., Stark, N., Nicolosi, B., Lontai, J., Lambirth, K., Schlueter, J., Gibas, C., & Munir, M. (2021). Performance evaluation of virus concentration methods for implementing SARS-CoV-2 wastewater based epidemiology emphasizing quick data turnaround. *The Science of the Total Environment*, *801*, 149656. https://doi.org/10.1016/j.scitotenv.2021.149656

Kaplan, E. H., Zulli, A., Sanchez, M., & Peccia, J. (2022). Scaling SARS-CoV-2 wastewater concentrations to population estimates of infection. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-07523-7

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), Article 7925. 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. (n.d.). Rapid, Large-Scale Wastewater Surveillance and Automated Reporting System Enable Early Detection of Nearly 85% of COVID-19 Cases on a University Campus. *MSystems*, *6*(4), e00793-21. https://doi.org/10.1128/mSystems.00793-21

Kashi, A. H., Fallah-karkan, M., Amini, E., & Vaezjalali, M. (2020). *The Presence of COVID-19 in Urine: A Systematic Review and Meta-analysis of the Literature* (p. 2020.05.15.20094920). medRxiv. https://doi.org/10.1101/2020.05.15.20094920

Kesheh, M. M., Hosseini, P., Soltani, S., & Zandi, M. (2022a). An overview on the seven pathogenic human coronaviruses. *Reviews in Medical Virology*, *32*(2), e2282. https://doi.org/10.1002/rmv.2282

Kim, J.-M., Kim, H. M., Lee, E. J., Jo, H. J., Yoon, Y., Lee, N.-J., Son, J., Lee, Y.-J., Kim, M. S., Lee, Y.-P., Chae, S.-J., Park, K. R., Cho, S.-R., Park, S., Kim, S. J., Wang, E., Woo, S., Lim, A., Park, S.-J., … Yoo, C. K. (2020). Detection and Isolation of SARS-CoV-2 in Serum, Urine, and Stool Specimens of COVID-19 Patients from the Republic of Korea. *Osong Public Health and Research Perspectives*, *11*(3), 112–117. https://doi.org/10.24171/j.phrp.2020.11.3.02

Kirby, A. E. (2021). Using Wastewater Surveillance Data to Support the COVID-19 Response—United States, 2020–2021. *MMWR. Morbidity and Mortality Weekly Report*, *70*. https://doi.org/10.15585/mmwr.mm7036a2

Kitajima, M., Sassi, H. P., & Torrey, J. R. (2018). Pepper mild mottle virus as a water quality indicator. *Npj Clean Water*, *1*(1), Article 1. https://doi.org/10.1038/s41545-018-0019-5

Klein et al. (2020, December 8). *Optimizing SARS-CoV-2 molecular diagnostic using N gene target: Insights about reinfection | medRxiv*. https://www.medrxiv.org/content/10.1101/2020.12.06.20244905v1.full

Koczera, P., Martin, L., Marx, G., & Schuerholz, T. (2016). The Ribonuclease A Superfamily in Humans: Canonical RNases as the Buttress of Innate Immunity. *International Journal of Molecular Sciences*, *17*(8), 1278. https://doi.org/10.3390/ijms17081278

Koskinen, A., Tolvi, M., Jauhiainen, M., Kekäläinen, E., Laulajainen-Hongisto, A., & Lamminmäki, S. (2021). Complications of COVID-19 Nasopharyngeal Swab Test. *JAMA Otolaryngology–Head & Neck Surgery*, *147*(7), 672–674. https://doi.org/10.1001/jamaoto.2021.0715

Kumar, M., Joshi, M., Patel, A. K., & Joshi, C. G. (2021). Unravelling the early warning capability of wastewater surveillance for COVID-19: A temporal study on SARS-CoV-2 RNA detection and need for the escalation. *Environmental Research*, *196*, 110946. https://doi.org/10.1016/j.envres.2021.110946

Kutti-Sridharan, G., Vegunta, R., Vegunta, R., Mohan, B. P., & Rokkam, V. R. P. (2020). SARS-CoV2 in Different Body Fluids, Risks of Transmission, and Preventing COVID-19: A Comprehensive Evidence-Based Review. *International Journal of Preventive Medicine*, *11*, 97. https://doi.org/10.4103/ijpvm.IJPVM_255_20

Larsen, D. A., & Wigginton, K. R. (2020). Tracking COVID-19 with wastewater. *Nature Biotechnology*, *38*(10), 1151–1153. https://doi.org/10.1038/s41587-020-0690-1

Lavania, M., Joshi, M. S., Ranshing, S. S., Potdar, V. A., Shinde, M., Chavan, N., Jadhav, S. M., Sarkale, P., Mohandas, S., Sawant, P. M., Tikute, S., Padbidri, V., Patwardhan, S., & Kate, R. (2022). Prolonged Shedding of SARS-CoV-2 in Feces of COVID-19 Positive Patients: Trends in Genomic Variation in First and Second Wave. *Frontiers in Medicine*, *9*. https://www.frontiersin.org/articles/10.3389/fmed.2022.835168

Lee, K. W., Yusof Khan, A. H. K., Ching, S. M., Chia, P. K., Loh, W. C., Abdul Rashid, A. M., Baharin, J., Inche Mat, L. N., Wan Sulaiman, W. A., Devaraj, N. K., Sivaratnam, D., Basri, H., & Hoo, F. K. (2020). Stroke and Novel Coronavirus Infection in Humans: A Systematic Review and Meta-Analysis. *Frontiers in Neurology*, *11*. https://www.frontiersin.org/articles/10.3389/fneur.2020.579070

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. *The Science of the Total Environment*, *817*, 152958. https://doi.org/10.1016/j.scitotenv.2022.152958

Liotti, F. M., Menchinelli, G., Marchetti, S., Morandotti, G. A., Sanguinetti, M., Posteraro, B., & Cattani, P. (2021). Evaluation of three commercial assays for SARS-CoV-2 molecular detection in upper respiratory tract samples. *European Journal of Clinical Microbiology & Infectious Diseases*, *40*(2), 269–277. https://doi.org/10.1007/s10096-020-04025-0

List-of-Acceptable-Commercial-Primers-Probes.pdf. (n.d.). Retrieved May 22, 2022, from https://www.cdc.gov/coronavirus/2019-ncov/downloads/List-of-Acceptable-Commercial-Primers-Probes.pdf

Loeb, S., Graham, K., Catoe, D., Wolfe, M., Boehm, A., & Wigginton, K. (2020, August 6). *One-Step RT-ddPCR for Detection of SARS-CoV-2, Bovine Coronavirus, and PMMoV RNA in RNA Derived from Wastew...* Protocols.Io. https://www.protocols.io/view/one-step-rt-ddpcr-for-detection-of-sars-cov-2-bovi-bi6vkhe6

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

Ma, Q., Liu, J., Liu, Q., Kang, L., Liu, R., Jing, W., Wu, Y., & Liu, M. (2021). Global Percentage of Asymptomatic SARS-CoV-2 Infections Among the Tested Population and Individuals With Confirmed COVID-19 Diagnosis. *JAMA Network Open*, *4*(12), e2137257. https://doi.org/10.1001/jamanetworkopen.2021.37257

Maal-Bared, R., Qiu, Y., Li, Q., Gao, T., Hrudey, S. E., Bhavanam, S., Ruecker, N. J., Ellehoj, E., Lee, B. E., & Pang, X. (2023). Does normalization of SARS-CoV-2 concentrations by Pepper Mild Mottle Virus improve correlations and lead time between wastewater surveillance and clinical data in Alberta (Canada): Comparing twelve SARS-CoV-2 normalization approaches. *Science of The Total Environment*, *856*, 158964. https://doi.org/10.1016/j.scitotenv.2022.158964

Machado, G. T., Pinto, C. R. de C., da Fonseca, L. A. V., Ramos, T. C. dos S., Paggi, T. F. P., & Spira, B. (2021). Bacteriophages as surrogates for the study of viral dispersion in open air. *Archives of Microbiology*, *203*(7), 4041–4049. https://doi.org/10.1007/s00203-021-02382-8

Mahlknecht, J. (2022). Presence and persistence of SARS-CoV-2 in aquatic environments: A mini-review. *Current Opinion in Environmental Science & Health*, *29*, 100385. https://doi.org/10.1016/j.coesh.2022.100385

Makuch, J., Brucker, E., Schmid, R., Mynahan, P., & Owen, T. (1996). Removal of PCR Inhibitors from Forensic Samples and Enhanced Sensitivity for D1S80 using Fluorescent Scanning. *Canadian Society of Forensic Science Journal*, *29*(1), 29–32. https://doi.org/10.1080/00085030.1996.10757045

Mallapaty, S. (2021). Closest known relatives of virus behind COVID-19 found in Laos. *Nature*, *597*(7878), 603–603. https://doi.org/10.1038/d41586-021-02596-2

Mardian, Y., Kosasih, H., Karyana, M., Neal, A., & Lau, C.-Y. (2021). Review of Current COVID-19 Diagnostics and Opportunities for Further Development. *Frontiers in Medicine*, *8*. https://www.frontiersin.org/articles/10.3389/fmed.2021.615099

Markt, R., Mayr, M., Peer, E., Wagner, A. O., Lackner, N., & Insam, H. (2021). Detection and Stability of SARS-CoV-2 Fragments in Wastewater: Impact of Storage Temperature. *Pathogens*, *10*(9), Article 9. https://doi.org/10.3390/pathogens10091215

McClary-Gutierrez, J. S., Mattioli, M. C., Marcenac, P., Silverman, A. I., Boehm, A. B., Bibby, K., Balliet, M., de los Reyes, F. L., Gerrity, D., Griffith, J. F., Holden, P. A., Katehis, D., Kester, G., LaCross, N., Lipp, E. K., Meiman, J., Noble, R. T., Brossard, D., & McLellan, S. L. (2021). SARS-CoV-2 Wastewater Surveillance for Public Health Action. *Emerging Infectious Diseases*, *27*(9), e210753. https://doi.org/10.3201/eid2709.210753

McCord, B., Pionzio, A. M., & Thompson, R. E. (2014). *Analysis of the Effect of a Variety of PCR Inhibitors on the Amplification of DNA using Real Time PCR, Melt Curves and STR Analysis*. https://www.semanticscholar.org/paper/Analysis-of-the-Effect-of-a-Variety-of-PCR-on-the-McCord-Pionzio/5340b93affd813573a789f2c2a322f17ecb4ecaa

McMahan, C. S., Self, S., Rennert, L., Kalbaugh, C., Kriebel, D., Graves, D., Colby, C., Deaver, J. A., Popat, S. C., Karanfil, T., & Freedman, D. L. (2021). COVID-19 wastewater epidemiology: A model to estimate infected populations. *The Lancet Planetary Health*, *5*(12), e874–e881. https://doi.org/10.1016/S2542-5196(21)00230-8

Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., & Brouwer, A. (2020). Presence of SARS-Coronavirus-2 RNA in Sewage and Correlation with Reported COVID-19 Prevalence in the Early Stage of the Epidemic in The Netherlands. *Environmental Science & Technology Letters*, acs.estlett.0c00357. https://doi.org/10.1021/acs.estlett.0c00357

Melvin, R. G., Chaudhry, N., Georgewill, O., Freese, R., & Simmons, G. E. (2021). Predictive power of SARS-CoV-2 wastewater surveillance for diverse populations across a large geographical range. *MedRxiv*, 2021.01.23.21250376. https://doi.org/10.1101/2021.01.23.21250376

Michigan.gov. (n.d.-a). *Coronavirus*. Retrieved September 3, 2022, from https://www.michigan.gov/coronavirus

Michigan.gov. (n.d.-b). *Michigan Data*. Retrieved September 3, 2022, from https://www.michigan.gov/coronavirus/stats

Mild COVID-19 Symptoms: Timeline, Progression, Contagiousness. (2021, February 16). Healthline. https://www.healthline.com/health/mild-covid-symptoms

Mohapatra, R. K., Pintilie, L., Kandi, V., Sarangi, A. K., Das, D., Sahu, R., & Perekhoda, L. (2020). The recent challenges of highly contagious COVID-19, causing respiratory infections: Symptoms, diagnosis, transmission, possible vaccines, animal models, and immunotherapy. *Chemical Biology & Drug Design*, 10.1111/cbdd.13761. https://doi.org/10.1111/cbdd.13761

Mollaei, H. R., Afshar, A. A., Kalantar-Neyestanaki, D., Fazlalipour, M., & Aflatoonian, B. (2020). Comparison five primer sets from different genome region of COVID-19 for detection of virus infection by conventional RT-PCR. *Iranian Journal of Microbiology*, *12*(3), 185–193.

Monpoeho, S., Dehée, A., Mignotte, B., Schwartzbrod, L., Marechal, V., Nicolas, J. C., Billaudel, S., & Férré, V. (2000). Quantification of enterovirus RNA in sludge samples using single tube real-time RT-PCR. *BioTechniques*, *29*(1), 88–93. https://doi.org/10.2144/00291st03

Mumm, J.-N., Ledderose, S., Ostermann, A., Rudelius, M., Hellmuth, J. C., Münchhoff, M., Munker, D., Scherer, C., Volz, Y., Ebner, B., Giessen-Jung, C., Lampert, C., Vilsmaier, T., Schneider, S., Gapp, M., Milger-Kneidinger, K., Behr, J., von Bergwelt-Baildon, M., Keppler, O. T., … Rodler, S. (2021). Dynamics of urinary and respiratory shedding of Severe acute respiratory syndrome virus 2 (SARS-CoV-2) RNA excludes urine as a relevant source of viral transmission. *Infection*. https://doi.org/10.1007/s15010-021-01724-4

Nagarkar, M., Keely, S. P., Jahne, M., Wheaton, E., Hart, C., Smith, B., Garland, J., Varughese, E. A., Braam, A., Wiechman, B., Morris, B., & Brinkman, N. E. (2022). SARS-CoV-2 monitoring at three sewersheds of different scales and complexity demonstrates distinctive relationships between wastewater measurements and COVID-19 case data. *Science of The Total Environment*, *816*, 151534. https://doi.org/10.1016/j.scitotenv.2021.151534

Nalla, A. K., Casto, A. M., Huang, M.-L. W., Perchetti, G. A., Sampoleo, R., Shrestha, L., Wei, Y., Zhu, H., Jerome, K. R., & Greninger, A. L. (2020). Comparative Performance of SARS-CoV-2 Detection Assays Using Seven Different Primer-Probe Sets and One Assay Kit. *Journal of Clinical Microbiology*, *58*(6), e00557-20. https://doi.org/10.1128/JCM.00557-20

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

Nicole Acosta, María A. Bautista, Jordan Hollman, JanineMcCalder, Alexander Buchner Beaudet, & Lawrence Man. (2021). *A multicenter study investigating SARS-CoV-2 in tertiary-care hospital wastewater. Viral burden correlates with increasing hospitalized cases as well as hospital-associated transmissions and outbreaks—ScienceDirect*. *201*, 117369.

NR-445 Bovine coronavirus, Mebus (Viruses). (n.d.). Retrieved April 19, 2022, from https://www.beiresources.org/Catalog/animalViruses/NR-445.aspx

Ohyama, H., Sakai, T., Agari, Y., Fukui, K., Nakagawa, N., Shinkai, A., Masui, R., & Kuramitsu, S. (2014). The role of ribonucleases in regulating global mRNA levels in the model organism Thermus thermophilus HB8. *BMC Genomics*, *15*(1), 386. https://doi.org/10.1186/1471-2164-15-386

Oikarinen, S., Tauriainen, S., Viskari, H., Simell, O., Knip, M., Virtanen, S., & Hyöty, H. (2009). PCR inhibition in stool samples in relation to age of infants. *Journal of Clinical Virology*, *44*(3), 211–214. https://doi.org/10.1016/j.jcv.2008.12.017

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

One-Step RT-ddPCR Advanced Kit for Probes. (n.d.). Retrieved March 8, 2023, from https://www.bio-rad.com/sites/default/files/webroot/web/pdf/lsr/literature/10049226.pdf

O'Reilly, K. M., Allen, D. J., Fine, P., & Asghar, H. (2020). The challenges of informative wastewater sampling for SARS-CoV-2 must be met: Lessons from polio eradication. *The Lancet Microbe*, *1*(5), e189–e190. https://doi.org/10.1016/S2666-5247(20)30100-2

Pabbaraju, K., Wong, A. A., Douesnard, M., Ma, R., Gill, K., Dieu, P., Fonseca, K., Zelyas, N., & Tipples, G. A. (2021). Development and validation of RT-PCR assays for testing for SARS-CoV-2. *Official Journal of the Association of Medical Microbiology and Infectious Disease Canada*, *6*(1), 16–22. https://doi.org/10.3138/jammi-2020-0026

Pal, M., Berhanu, G., Desalegn, C., & Kandi, V. (n.d.). Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update. *Cureus*, *12*(3), e7423. https://doi.org/10.7759/cureus.7423

Panchal, D., Prakash, O., Bobde, P., & Pal, S. (2021). SARS-CoV-2: Sewage surveillance as an early warning system and challenges in developing countries. *Environmental Science and Pollution Research*, *28*(18), 22221–22240. https://doi.org/10.1007/s11356-021-13170-8

Park, G. W., Ng, T. F. F., Freeland, A. L., Marconi, V. C., Boom, J. A., Staat, M. A., Montmayeur, A. M., Browne, H., Narayanan, J., Payne, D. C., Cardemil, C. V., Treffiletti, A., & Vinjé, J. (2020). CrAssphage as a Novel Tool to Detect Human Fecal Contamination on Environmental Surfaces and Hands. *Emerging Infectious Diseases*, *26*(8), 1731–1739. https://doi.org/10.3201/eid2608.200346

Park, S., Lee, C.-W., Park, D.-I., Woo, H.-Y., Cheong, H. S., Shin, H. C., Ahn, K., Kwon, M.-J., & Joo, E.-J. (2021). Detection of SARS-CoV-2 in Fecal Samples From Patients With Asymptomatic and Mild COVID-19 in Korea. *Clinical Gastroenterology and Hepatology*, *19*(7), 1387-1394.e2. https://doi.org/10.1016/j.cgh.2020.06.005

Paterson, B. J., & Durrheim, D. N. (2022). Wastewater surveillance: An effective and adaptable surveillance tool in settings with a low prevalence of COVID-19. *The Lancet Planetary Health*, *6*(2), e87–e88. https://doi.org/10.1016/S2542-5196(22)00009-2

Patil, S., Ananthan, A., Nanavati, R. N., Nataraj, G., & Prasad, P. (2019). Effect of different methods of pasteurization on bactericidal action of human milk: A prospective observational study. *The Indian Journal of Medical Research*, *150*(5), 504–507. https://doi.org/10.4103/ijmr.IJMR_600_18

Pauline A. Bariola, Gustavo C. MacIntosh, & Pamela J. Green. (1999). Regulation of S-Like Ribonuclease Levels in Arabidopsis. Antisense Inhibition of RNS1 orRNS2 Elevates Anthocyanin Accumulation. *Plant Physiology*, *119*(1), 331–342. https://doi.org/10.1104/pp.119.1.331

Pearson, J. D., Trcka, D., Lu, S., Hyduk, S. J., Jen, M., Aynaud, M.-M., Hernández, J. J., Peidis, P., Barrios-Rodiles, M., Chan, K., Woodgett, J., Mazzulli, T., Attisano, L., Pelletier, L., Cybulsky, M. I., Wrana, J. L., & Bremner, R. (2021). Comparison of SARS-CoV-2 indirect and direct RT-qPCR detection methods. *Virology Journal*, *18*, 99. https://doi.org/10.1186/s12985-021-01574-4

Peiris, M., & Poon, L. L. M. (2021). Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) (Coronaviridae). *Encyclopedia of Virology*, 814–824. https://doi.org/10.1016/B978-0-12-814515-9.00138-7

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. *The Science of the Total Environment*, *758*, 143870. https://doi.org/10.1016/j.scitotenv.2020.143870

Pormohammad, A., Ghorbani, S., Khatami, A., Farzi, R., Baradaran, B., Turner, D. L., Turner, R. J., Bahr, N. C., & Idrovo, J. (2020). Comparison of confirmed COVID-19 with SARS and MERS cases - Clinical characteristics, laboratory findings, radiographic signs and outcomes: A systematic review and meta-analysis. *Reviews in Medical Virology*, e2112. https://doi.org/10.1002/rmv.2112

Precautions for Handling of RNA. (n.d.). Retrieved June 22, 2022, from https://lifescience.roche.com/en_us/articles/precautions-for-handling-of-rna.html

PrimeTime qPCR Probes. (n.d.). Integrated DNA Technologies. Retrieved April 8, 2022, from https://www.idtdna.com/pages/products/crispr-genome-editing/alt-r-genome-editing-detection-kit

Prussin, A. J., Schwake, D. O., Lin, K., Gallagher, D. L., Buttling, L., & Marr, L. C. (2018). Survival of the Enveloped Virus Phi6 in Droplets as a Function of Relative Humidity, Absolute Humidity, and Temperature. *Applied and Environmental Microbiology*, *84*(12), e00551-18. https://doi.org/10.1128/AEM.00551-18

Randazzo, W., Truchado, P., Cuevas-Ferrando, E., Simón, P., Allende, A., & Sánchez, G. (2020). SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. *Water Research*, *181*, 115942. https://doi.org/10.1016/j.watres.2020.115942

Review of DNA Extraction Methodologies and Guidelinesfor Protocol Development. (n.d.). Retrieved October 3, 2022, from https://globalsnowleopard.org/wp-content/uploads/2021/07/Review-of-DNA-Extraction-Methodologies-and-Guidelines-for-Protocol-Development_Final.pdf

Rodrigues, J., Gouveia, C., Santos, M. A., Costa, O., Côrte-Real, R., & Brito, M. J. (2021). Comparison of nasopharyngeal samples for SARS-CoV-2 detection in a paediatric cohort. *Journal of Paediatrics and Child Health*, 10.1111/jpc.15405. https://doi.org/10.1111/jpc.15405

Roldan-Hernandez, L., Graham, K. E., Duong, D., & Boehm, A. B. (2022). Persistence of Endogenous SARS-CoV-2 and Pepper Mild Mottle Virus RNA in Wastewater-Settled Solids. *ACS ES&T Water*. https://doi.org/10.1021/acsestwater.2c00003

Romano, M., Ruggiero, A., Squeglia, F., Maga, G., & Berisio, R. (2020). A Structural View of SARS-CoV-2 RNA Replication Machinery: RNA Synthesis, Proofreading and Final Capping. *Cells*, *9*(5), 1267. https://doi.org/10.3390/cells9051267

Ross, M. P., & Spiller, H. A. (1999). Fatal ingestion of sodium hypochlorite bleach with associated hypernatremia and hyperchloremic metabolic acidosis. *Veterinary and Human Toxicology*, *41*(2), 82–86.

Safford, H. R., Shapiro, K., & Bischel, H. N. (2022). Wastewater analysis can be a powerful public health tool—If it's done sensibly. *Proceedings of the National Academy of Sciences*, *119*(6), e2119600119. https://doi.org/10.1073/pnas.2119600119

Saif, L. J. (2010). Bovine Respiratory Coronavirus. *The Veterinary Clinics of North America. Food Animal Practice*, *26*(2), 349–364. https://doi.org/10.1016/j.cvfa.2010.04.005

Sala-Comorera, L., Reynolds, L. J., Martin, N. A., O'Sullivan, J. J., Meijer, W. G., & Fletcher, N. F. (2021). Decay of infectious SARS-CoV-2 and surrogates in aquatic environments. *Water Research*, *201*, 117090. https://doi.org/10.1016/j.watres.2021.117090

SARS | Basics Factsheet | CDC. (n.d.). Retrieved July 18, 2022, from https://www.cdc.gov/sars/about/fs-sars.html

Schrader, C., Schielke, A., Ellerbroek, L., & Johne, R. (2012). PCR inhibitors – occurrence, properties and removal. *Journal of Applied Microbiology*, *113*(5), 1014–1026. https://doi.org/10.1111/j.1365-2672.2012.05384.x

Sedlak, R. H., Kuypers, J., & Jerome, K. R. (2014). *A multiplexed droplet digital PCR assay performs better than qPCR on inhibition prone samples*. *80*(4), 285–286. https://doi.org/10.1016/j.diagmicrobio.2014.09.004

Shereen, M. A., Khan, S., Kazmi, A., Bashir, N., & Siddique, R. (2020). COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research*, *24*, 91–98. https://doi.org/10.1016/j.jare.2020.03.005

Shi, Y., Wang, G., Cai, X., Deng, J., Zheng, L., Zhu, H., Zheng, M., Yang, B., & Chen, Z. (2020). An overview of COVID-19. *Journal of Zhejiang University. Science. B*, *21*(5), 343–360. https://doi.org/10.1631/jzus.B2000083

Sidova, M., Tomankova, S., Abaffy, P., Kubista, M., & Sindelka, R. (2015). Effects of post-mortem and physical degradation on RNA integrity and quality. *Biomolecular Detection and Quantification*, *5*, 3–9. https://doi.org/10.1016/j.bdq.2015.08.002

Sidstedt, M., Rådström, P., & Hedman, J. (2020). PCR inhibition in qPCR, dPCR and MPS—mechanisms and solutions. *Analytical and Bioanalytical Chemistry*, *412*(9), 2009–2023. https://doi.org/10.1007/s00216-020-02490-2

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

Sludge, farmer's friend or toxic slime? (2009, May 5). Grist. https://grist.org/article/2009-05-05-sludge-fertilizer-sewage/

S.m.k, H.-A., & S.d, H. (2000). *Determination Of The Mean Daily Stool Weight, Frequency Of Defecation And Bowel Transit Time: Assessment Of 1000 Healthy Subjects*. *3*(4), 0–0.

Soltani, J., Esmailnasab, N., Roshani, D., Karimi, M., & Amjadi, M.-J. (2014). Acute Flaccid Paralysis and Its Differential Diagnosis in in Kurdistan Province, Western Iran; an 11-Year Surveillance. *Iranian Journal of Pediatrics*, *24*(2), 131–139.

Song et al. (2022). *SARS-CoV-2: The Monster Causes COVID-19*. https://www.frontiersin.org/articles/10.3389/fcimb.2022.835750/full

Stachler, E., Kelty, C., Sivaganesan, M., Li, X., Bibby, K., & Shanks, O. C. (2017). Quantitative CrAssphage PCR Assays for Human Fecal Pollution Measurement. *Environmental Science & Technology*, *51*(16), 9146–9154. https://doi.org/10.1021/acs.est.7b02703

String, G. M., White, M. R., Gute, D. M., Mühlberger, E., & Lantagne, D. S. (2021). Selection of a SARS-CoV-2 Surrogate for Use in Surface Disinfection Efficacy Studies with Chlorine and Antimicrobial Surfaces. *Environmental Science & Technology Letters*, acs.estlett.1c00593. https://doi.org/10.1021/acs.estlett.1c00593

Subbarao, K., & Mahanty, S. (2020). Respiratory Virus Infections: Understanding COVID-19. *Immunity*, *52*(6), 905–909. https://doi.org/10.1016/j.immuni.2020.05.004

Susan I. Gerber and John T. Watson. (2020). PRE-2019 Coronaviruses. In *Goldman-Cecil Medicine* (pp. 2164–2167).

Szymczak, W. A., Goldstein, D. Y., Orner, E. P., Fecher, R. A., Yokoda, R. T., Skalina, K. A., Narlieva, M., Gendlina, I., & Fox, A. S. (2020). Utility of Stool PCR for the Diagnosis of COVID-19: Comparison of Two Commercial Platforms. *Journal of Clinical Microbiology*, *58*(9), e01369-20. https://doi.org/10.1128/JCM.01369-20

Tahan, S., Parikh, B. A., Droit, L., Wallace, M. A., Burnham, C.-A. D., & Wang, D. (n.d.). SARS-CoV-2 E Gene Variant Alters Analytical Sensitivity Characteristics of Viral Detection Using a Commercial Reverse Transcription-PCR Assay. *Journal of Clinical Microbiology*, *59*(7), e00075-21. https://doi.org/10.1128/JCM.00075-21

Takeuchi, K., Yanagisawa, H., Kurosawa, Y., Iida, Y., Kawai, K., & Fujimaki, S. (2022). Degradation of SARS-CoV-2 specific ribonucleic acid in samples for nucleic acid amplification detection. *PLoS ONE*, *17*(3), e0264541. https://doi.org/10.1371/journal.pone.0264541

Tan, S. H., Allicock, O., Armstrong-Hough, M., & Wyllie, A. L. (2021). Saliva as a gold-standard sample for SARS-CoV-2 detection. *The Lancet. Respiratory Medicine*, *9*(6), 562–564. https://doi.org/10.1016/S2213-2600(21)00178-8

Tarun, J., Susan, J., Suria, J., Susan, V. J., & Criton, S. (2014). Evaluation of pH of Bathing Soaps and Shampoos for Skin and Hair Care. *Indian Journal of Dermatology*, *59*(5), 442–444. https://doi.org/10.4103/0019-5154.139861

Technical-bulletin_sars-cov-2-variants-mutations.pdf. (n.d.). Retrieved January 25, 2022, from https://www.primetech.co.jp/Portals/0/db/product/LGC/biosearch/technical-bulletin_sars-cov-2-variants-mutations.pdf

The Basics: RNase Control - US. (n.d.). Retrieved September 18, 2022, from https://www.thermofisher.com/us/en/home/references/ambion-tech-support/nuclease-enzymes/general-articles/the-basics-rnase-control.html

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*, 116181. https://doi.org/10.1016/j.watres.2020.116181

Torii, S., Oishi, W., Zhu, Y., Thakali, O., Malla, B., Yu, Z., Zhao, B., Arakawa, C., Kitajima, M., Hata, A., Ihara, M., Kyuwa, S., Sano, D., Haramoto, E., & Katayama, H. (2022). Comparison of five polyethylene glycol precipitation procedures for the RT-qPCR based recovery of murine hepatitis virus, bacteriophage phi6, and pepper mild mottle virus as a surrogate for SARS-CoV-2 from wastewater. *The Science of the Total Environment*, *807*, 150722. https://doi.org/10.1016/j.scitotenv.2021.150722

Tsujimoto, Y., Terada, J., Kimura, M., Moriya, A., Motohashi, A., Izumi, S., Kawajiri, K., Hakkaku, K., Morishita, M., Saito, S., Takumida, H., Watanabe, H., Tsukada, A., Morita, C., Yamaguchi, Y., Katsuno, T., Kusaba, Y., Sakamoto, K., Hashimoto, M., … Sugiyama, H. (n.d.). Diagnostic accuracy of nasopharyngeal swab, nasal swab and saliva swab samples for the detection of SARS-CoV-2 using RT-PCR. *Infectious Diseases (London, England)*, 1–9. https://doi.org/10.1080/23744235.2021.1903550

Tuma, J. N. (2021). Surveillance to Track Progress Toward Polio Eradication—Worldwide, 2019–2020. *MMWR. Morbidity and Mortality Weekly Report*, *70*. https://doi.org/10.15585/mmwr.mm7018a2

Uddin, M. K. M., Shirin, T., Hossain, M. E., Alam, A. N., Ami, J. Q., Hasan, R., Miah, M., Shaly, N. J., Ahmed, S., Rahman, S. M. M., Rahman, M., & Banu, S. (n.d.). Diagnostic Performance of Self-Collected Saliva Versus Nasopharyngeal Swab for the Molecular Detection of SARS-CoV-2 in the Clinical Setting. *Microbiology Spectrum*, *9*(3), e00468-21. https://doi.org/10.1128/Spectrum.00468-21

U.S. Census Bureau QuickFacts: Ingham County, Michigan. (n.d.-a). Retrieved October 18, 2022, from https://www.census.gov/quickfacts/inghamcountymichigan

U.S. Census Bureau QuickFacts: Macomb County, Michigan. (n.d.-b). Retrieved October 18, 2022, from https://www.census.gov/quickfacts/macombcountymichigan

U.S. Census Bureau QuickFacts: St. Clair County, Michigan. (n.d.-c). Retrieved October 18, 2022, from https://www.census.gov/quickfacts/stclaircountymichigan

Vadde, K. K., Al-Duroobi, H., Phan, D. C., Jafarzadeh, A., Moghadam, S. V., Matta, A., & Kapoor, V. (2022). Assessment of Concentration, Recovery, and Normalization of SARS-CoV-2 RNA from Two Wastewater Treatment Plants in Texas and Correlation with COVID-19 Cases in the Community. *ACS ES&T Water*. https://doi.org/10.1021/acsestwater.2c00054

van Doorn, A. S., Meijer, B., Frampton, C. M. A., Barclay, M. L., & de Boer, N. K. H. (2020). Systematic review with meta-analysis: SARS-CoV-2 stool testing and the potential for faecal-oral transmission. *Alimentary Pharmacology & Therapeutics*, *52*(8), 1276–1288. https://doi.org/10.1111/apt.16036

Vanaerschot, M., Mann, S. A., Webber, J. T., Kamm, J., Bell, S. M., Bell, J., Hong, S. N., Nguyen, M. P., Chan, L. Y., Bhatt, K. D., Tan, M., Detweiler, A. M., Espinosa, A., Wu, W., Batson, J., Dynerman, D., Consortium, C., Wadford, D. A., Puschnik, A. S., … Crawford, E. D. (2020). *Identification of a polymorphism in the N gene of SARS-CoV-2 that adversely impacts detection by a widely-used RT-PCR assay* (p. 2020.08.25.265074). bioRxiv. https://doi.org/10.1101/2020.08.25.265074

Vandeventer, P. E., Lin, J. S., Zwang, T. J., Nadim, A., Johal, M. S., & Niemz, A. (2012). Multiphasic DNA Adsorption to Silica Surfaces under Varying Buffer, pH, and Ionic Strength Conditions. *The Journal of Physical Chemistry. B*, *116*(19), 5661–5670. https://doi.org/10.1021/jp3017776

Vatter, P., Hoenes, K., & Hessling, M. (2021). Blue light inactivation of the enveloped RNA virus Phi6. *BMC Research Notes*, *14*, 187. https://doi.org/10.1186/s13104-021-05602-y

V'kovski, P., Kratzel, A., Steiner, S., Stalder, H., & Thiel, V. (2021). Coronavirus biology and replication: Implications for SARS-CoV-2. *Nature Reviews Microbiology*, *19*(3), Article 3. https://doi.org/10.1038/s41579-020-00468-6

Vogels, C. B. F., Brito, A. F., Wyllie, A. L., Fauver, J. R., Ott, I. M., Kalinich, C. C., Petrone, M. E., Casanovas-Massana, A., Catherine Muenker, M., Moore, A. J., Klein, J., Lu, P., Lu-Culligan, A., Jiang, X., Kim, D. J., Kudo, E., Mao, T., Moriyama, M., Oh, J. E., … Grubaugh, N. D. (2020). Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT–qPCR primer–probe sets. *Nature Microbiology*, *5*(10), Article 10. https://doi.org/10.1038/s41564-020-0761-6

Wang, H., Qi, J., Xiao, D., Wang, Z., & Tian, K. (2017). A re-evaluation of dilution for eliminating PCR inhibition in soil DNA samples. *Soil Biology and Biochemistry*, *106*, 109–118. https://doi.org/10.1016/j.soilbio.2016.12.011

Wang, W., Xu, Y., Gao, R., Lu, R., Han, K., Wu, G., & Tan, W. (2020). Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA*, *323*(18), 1843–1844. https://doi.org/10.1001/jama.2020.3786

Waste Wattage: Cities Aim to Flush Heat Energy Out of Sewers. (2012, December 11). Science. https://www.nationalgeographic.com/science/article/121211-sewage-heat-recovery

Wei, X., Viadero, R. C., & Bhojappa, S. (2008). Phosphorus removal by acid mine drainage sludge from secondary effluents of municipal wastewater treatment plants. *Water Research*, *42*(13), 3275–3284. https://doi.org/10.1016/j.watres.2008.04.005

Weidhaas, J., Aanderud, Z. T., Roper, D. K., VanDerslice, J., Gaddis, E. B., Ostermiller, J., Hoffman, K., Jamal, R., Heck, P., Zhang, Y., Torgersen, K., Laan, J. V., & LaCross, N. (2021). Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds. *The Science of the Total Environment*, *775*, 145790. https://doi.org/10.1016/j.scitotenv.2021.145790

Weyersberg, L., Klemens, E., Buehler, J., Vatter, P., Hessling, M., Weyersberg, L., Klemens, E., Buehler, J., Vatter, P., & Hessling, M. (2022). UVC, UVB and UVA susceptibility of Phi6 and its suitability as a SARS-CoV-2 surrogate. *AIMS Microbiology*, *8*(3), Article microbiol-08-03-020. https://doi.org/10.3934/microbiol.2022020

WHO Coronavirus (COVID-19) Dashboard. (n.d.). Retrieved January 8, 2022, from https://covid19.who.int

Whoinhouseassays.pdf. (n.d.). Retrieved January 30, 2022, from https://www.who.int/docs/default-source/coronaviruse/whoinhouseassays.pdf

Wolfe, M. K. (n.d.). Invited Perspective: The Promise of Wastewater Monitoring for Infectious Disease Surveillance. *Environmental Health Perspectives*, *130*(5), 051302. https://doi.org/10.1289/EHP11151

Wölfel, R., Corman, V. M., Guggemos, W., Seilmaier, M., Zange, S., Müller, M. A., Niemeyer, D., Jones, T. C., Vollmar, P., Rothe, C., Hoelscher, M., Bleicker, T., Brünink, S., Schneider, J., Ehmann, R., Zwirglmaier, K., Drosten, C., & Wendtner, C. (2020). Virological assessment of hospitalized patients with COVID-2019. *Nature*, *581*(7809), Article 7809. https://doi.org/10.1038/s41586-020-2196-x

Worobey, M., Levy, J. I., Malpica Serrano, L., Crits-Christoph, A., Pekar, J. E., Goldstein, S. A., Rasmussen, A. L., Kraemer, M. U. G., Newman, C., Koopmans, M. P. G., Suchard, M. A., Wertheim, J. O., Lemey, P., Robertson, D. L., Garry, R. F., Holmes, E. C., Rambaut, A., & Andersen, K. G. (2022). The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic. *Science*, *377*(6609), 951–959. https://doi.org/10.1126/science.abp8715

Wu et al. (2020). *SARS-CoV-2 Titers in Wastewater Are Higher than Expected from Clinically Confirmed Cases | mSystems*. https://journals.asm.org/doi/10.1128/mSystems.00614-20

Wu, F., Lee, W. L., Chen, H., Gu, X., Chandra, F., Armas, F., Xiao, A., Leifels, M., Rhode, S. F., Wuertz, S., Thompson, J., & Alm, E. J. (2022). Making waves: Wastewater surveillance of SARS-CoV-2 in an endemic future. *Water Research*, *219*, 118535. https://doi.org/10.1016/j.watres.2022.118535

Wu, Y.-C., Chen, C.-S., & Chan, Y.-J. (2020). The outbreak of COVID-19: An overview. *Journal of the Chinese Medical Association*, *83*(3), 217–220. https://doi.org/10.1097/JCMA.0000000000000270

Wu, Z., Greaves, J., Arp, L., Stone, D., & Bibby, K. (2020). Comparative fate of CrAssphage with culturable and molecular fecal pollution indicators during activated sludge wastewater treatment. *Environment International*, *136*, 105452. https://doi.org/10.1016/j.envint.2019.105452

Xiao, A., Wu, F., Bushman, M., Zhang, J., Imakaev, M., Chai, P. R., Duvallet, C., Endo, N., Erickson, T. B., Armas, F., Arnold, B., Chen, H., Chandra, F., Ghaeli, N., Gu, X., Hanage, W. P., Lee, W. L., Matus, M., McElroy, K. A., … Alm, E. J. (2022). Metrics to relate COVID-19 wastewater data to clinical testing dynamics. *Water Research*, *212*, 118070. https://doi.org/10.1016/j.watres.2022.118070

Ye, Z.-W., Yuan, S., Yuen, K.-S., Fung, S.-Y., Chan, C.-P., & Jin, D.-Y. (2020). Zoonotic origins of human coronaviruses. *International Journal of Biological Sciences*, *16*(10), 1686–1697. https://doi.org/10.7150/ijbs.45472

Yeh, K. B., Tabynov, K., Parekh, F. K., Mombo, I., Parker, K., Tabynov, K., Bradrick, S. S., Tseng, A. S., Yang, J.-R., Gardiner, L., Olinger, G., & Setser, B. (2021). Significance of High-Containment Biological Laboratories Performing Work During the COVID-19 Pandemic: Biosafety Level-3 and -4 Labs. *Frontiers in Bioengineering and Biotechnology*, *9*. https://www.frontiersin.org/article/10.3389/fbioe.2021.720315

ZEN and IBFQ. (n.d.). Integrated DNA Technologies. Retrieved April 8, 2022, from https://www.idtdna.com/pages/support/faqs/if-zen-is-positioned-so-close-to-fam-why-is-the-3-quencher-required-

Zhao, X., Ding, Y., Du, J., & Fan, Y. (2020). 2020 update on human coronaviruses: One health, one world. *Medicine in Novel Technology and Devices*, *8*, 100043. https://doi.org/10.1016/j.medntd.2020.100043

Zhao, Y., Sun, J., Li, Y., Li, Z., Xie, Y., Feng, R., Zhao, J., & Hu, Y. (2020). The Strand-biased Transcription of SARS-CoV-2 and Unbalanced Inhibition by Remdesivir. *BioRxiv*, 2020.10.15.325050. https://doi.org/10.1101/2020.10.15.325050

Zhou, Y., Pei, F., Ji, M., Wang, L., Zhao, H., Li, H., Yang, W., Wang, Q., Zhao, Q., & Wang, Y. (2020). Sensitivity evaluation of 2019 novel coronavirus (SARS-CoV-2) RT-PCR detection kits and strategy to reduce false negative. *PLoS ONE*, *15*(11), e0241469. https://doi.org/10.1371/journal.pone.0241469

Zhu, Z., Lian, X., Su, X., Wu, W., Marraro, G. A., & Zeng, Y. (2020). From SARS and MERS to COVID-19: A brief summary and comparison of severe acute respiratory infections caused by three highly pathogenic human coronaviruses. *Respiratory Research*, *21*(1), 224. https://doi.org/10.1186/s12931-020-01479-w

Zitek, T. (2020). The Appropriate Use of Testing for COVID-19. *Western Journal of Emergency Medicine*, *21*(3), 470–472. https://doi.org/10.5811/westjem.2020.4.47370