The Rapid Assessment of Aggregated Wastewater Samples for Genomic Surveillance of SARS-CoV-2 on a City-Wide Scale

Eric C. Rouchka [1,2,†], Julia H. Chariker [2,†], Kumar Saurabh [3], Sabine Waigel [4], Wolfgang Zacharias [4], Mei Zhang [5], Daymond Talley [6], Ian Santisteban [7,8], Madeline Puccio [9], Sarah Moyer [10,11], Rochelle H. Holm [9], Ray A. Yeager [9,12], Kevin J. Sokoloski [8,13,*], Joshua Fuqua [7,8,*], Aruni Bhatnagar [9] and Ted Smith [9]

1 Department of Biochemistry and Molecular Genetics, University of Louisville, 323 E. Chestnut St., Louisville, KY 40202, USA; eric.rouchka@louisville.edu
2 KY INBRE Bioinformatics Core, University of Louisville, 522 E. Gray St., Louisville, KY 40202, USA; julia.chariker@louisville.edu
3 Brown Cancer Center, University of Louisville, 530 S. Jackson St., Louisville, KY 40202, USA; saurabh_bt@hotmail.com
4 Department of Medicine, University of Louisville, 530 S. Jackson St., Louisville, KY 40402, USA; sabine.waigel@louisville.edu (S.W.); wolfgang.zacharias@louisville.edu (W.Z.)
5 Department of Neuroscience, University of Louisville, 505 S. Hancock St., Louisville, KY 40202, USA; mei.zhang@louisville.edu
6 Louisville/Jefferson County Metropolitan Sewer District, Morris Forman Water Quality Treatment Center, 4522 Algonquin Parkway, Louisville, KY 40211, USA; daymond.talley@louisvillemsd.org
7 Department of Pharmacology and Toxicology, University of Louisville, 505 S. Hancock St., Louisville, KY 40202, USA; ian.santisteban@louisville.edu
8 Center for Predictive Medicine, University of Louisville, 505 S. Hancock St., Louisville, KY 40202, USA
9 Christina Lee Brown Envirome Institute, University of Louisville, 302 E. Muhammad Ali Blvd., Louisville, KY 40202, USA; madeline.puccio@louisville.edu (M.P.); rochelle.holm@louisville.edu (R.H.H.); ray.yeager@louisville.edu (R.A.Y.); aruni.bhatnagar@louisville.edu (A.B.); ted.smith@louisville.edu (T.S.)
10 Department of Health Management and System Sciences, School of Public Health and Information Sciences, University of Louisville, 485 E. Gray St., Louisville, KY 40202, USA; sarah.moyer@louisvilleky.gov
11 Department of Public Health and Wellness, Louisville Metro Government, 400 E. Gray St., Louisville, KY 40202, USA
12 Department of Environmental and Occupational Health Sciences, School of Public Health and Information Sciences, University of Louisville, 485 E. Gray St., Louisville, KY 40202, USA
13 Department of Microbiology and Immunology, University of Louisville, 505 S. Hancock St., Louisville, KY 40202, USA
* Correspondence: kevin.sokoloski@louisville.edu (K.J.S.); j.fuqua@louisville.edu (J.F.); Tel.: +1-(502)-852-2149 (K.J.S.); +1-(502)-852-8413 (J.F.)
† These authors contributed equally to this work.

Abstract: Throughout the course of the ongoing SARS-CoV-2 pandemic there has been a need for approaches that enable rapid monitoring of public health using an unbiased and minimally invasive means. A major way this has been accomplished is through the regular assessment of wastewater samples by qRT-PCR to detect the prevalence of viral nucleic acid with respect to time and location. Further expansion of SARS-CoV-2 wastewater monitoring efforts to include the detection of variants of interest/concern through next-generation sequencing has enhanced the understanding of the SARS-CoV-2 outbreak. In this report, we detail the results of a collaborative effort between public health and metropolitan wastewater management authorities and the University of Louisville to monitor the SARS-CoV-2 pandemic through the monitoring of aggregate wastewater samples over a period of 28 weeks. Through the use of next-generation sequencing approaches the polymorphism signatures of Variants of Concern/Interest were evaluated to determine the likelihood of their prevalence within the community on the basis of their relative dominance within sequence datasets. Our data indicate that wastewater monitoring of water quality treatment centers and smaller neighborhood-scale catchment areas is a viable means by which the prevalence and genetic variation of SARS-CoV-2 within a metropolitan community of approximately one million individuals may be monitored, as our efforts detected the introduction and emergence of variants of concern in the city of Louisville. Importantly, these efforts confirm that regional emergence and

spread of variants of interest/concern may be detected as readily in aggregate wastewater samples as compared to the individual wastewater sheds. Furthermore, the information gained from these efforts enabled targeted public health efforts including increased outreach to at-risk communities and the deployment of mobile or community-focused vaccination campaigns.

1. Introduction

The clinical prevalence of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in the early stages of the pandemic led to the application of wastewater surveillance over a range of scales- from pooled sources covering large swaths of a community encompassing hundreds of thousands of individuals, to relatively isolated samples consisting of dozens to hundreds of individuals living in congregate settings [1–6]. These efforts were initially designed to monitor the trends of "true" incidence of SARS-CoV-2 infection within the targeted communities, as wastewater sampling represents a non-invasive means of quantitatively assessing the presence of SARS-CoV-2 nucleic acid derived from individuals participating and non-participating in clinical testing efforts [1,7–12]. Thus, the sampling of wastewater represents a means by which SARS-CoV-2 public health may be assessed on a large scale to determine the disease/infection burden [2,8,9,13–17].

The genetic diversity of SARS-CoV-2 is a major public health concern for several reasons. Foremost, genetic changes to the pathogen may alter the transmissibility and virulence leading to increasing threats to individual and public health. Second, alterations to the viral Spike (S) protein may contribute to the evasion of currently available vaccines, resulting in breakthrough infections and clinical illness in the vaccinated population [18,19]. This is evidenced by the impact of S protein mutations such as D614G and N501Y on viral infectivity and transmission [20,21]. The genetic diversity of SARS-CoV-2 has largely been determined via the sequencing of clinical samples obtained from infected individuals; however, changes in the viral genome have been shown to be detectable in wastewater samples [22–27]. Indeed, the presence of specific SARS-CoV-2 variants is often detectable in the wastewater of a community prior to detection in clinical samples [2,5,9,13,25,27,28]. As such, the genomic surveillance of SARS-CoV-2 through wastewater sampling represents a cost-effective and unbiased means by which the community prevalence and genetic diversity of SARS-CoV-2 infections may be assessed.

In this report we describe efforts undertaken by a partnership between the Center for Predictive Medicine and Emerging Infectious Diseases at the University of Louisville, the Kentucky IDeA Networks of Biomedical Research Excellence (KY INBRE), Christina Lee Brown Envirome Institute, the Louisville Metropolitan Sewer District, and the public health department of the city of Louisville, Kentucky to monitor SARS-CoV-2 via the assessment of aggregated wastewater samples. The findings of these studies have been instrumental to the public health response to SARS-CoV-2 by (i) providing an unbiased assessment of community prevalence, (ii) delineating the genetic diversity of SARS-CoV-2 in the community, and (iii) enabling the targeted deployment of public health resources and efforts. Thus, similar efforts may be applied by other communities to better understand community prevalence to aid in the development of proactive responses to SARS-CoV-2.

2. Results

2.1. Detection of SARS-CoV-2 RNA in Aggregate Wastewater Samples Correlates with Community Prevalence

As shown in Figure 1, the weekly incidence of SARS-CoV-2 infection in Louisville KY steadily increased over an approximately 3-month period, corresponding to the date range of October 2020 to January 2021, with peak incidence occurring during the first week of January 2021. During this time the analysis of aggregated wastewater samples obtained from the

five Water Quality Treatment Centers (WQTCs) that service the Louisville metropolitan area demonstrated the clear presence of SARS-CoV-2 nucleic acid as detected by qRT-PCR. As expected, increased detection of SARS-CoV-2 RNAs correlated with the observed increases in community prevalence as reflected by the available public health data (Spearman's r = −0.533; *p*-value = 0.0042). Nonetheless, the increases among the two measurements were temporally asynchronous, in that levels of viral RNA in the aggregated wastewater increased approximately one week earlier than the reported increases in incidence as determined via clinical testing. After reaching peak incidence in early January 2021, the weekly incidence of confirmed SARS-CoV-2 steadily decreased to levels below those observed in October 2020. Notably, despite the significant decrease in clinically confirmed SARS-CoV-2 cases over the next 3 months in the Louisville metropolitan area, the detection of SARS-CoV-2 RNA in aggregate wastewater samples remained at levels similar to those observed during the peak incidence period. This discrepancy remained until Week 24, whereafter the levels of SARS-CoV-2 nucleic acid detected via qRT-PCR began to steadily decrease back to levels consistent with those observed before the January 2021 peak.
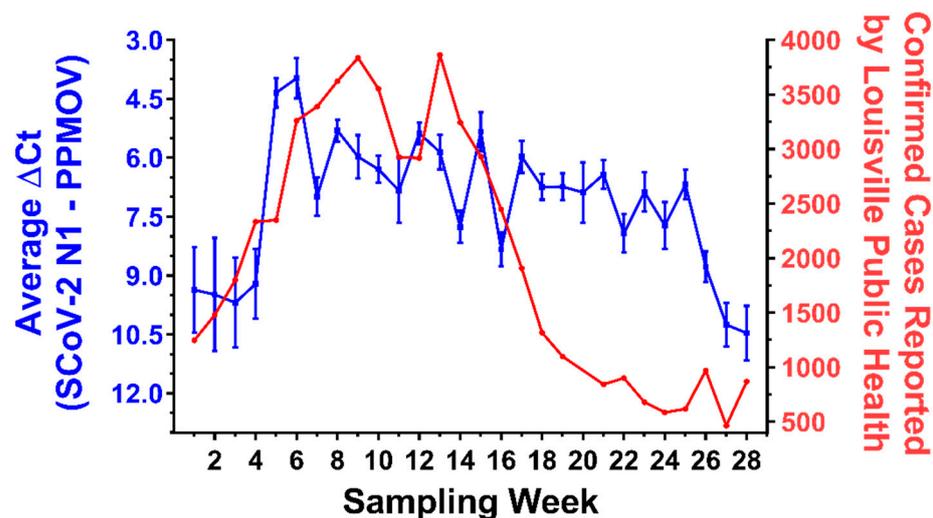

Figure 1. The Quantitative Detection of SARS-CoV-2 in the Aggregated Wastewater of Water Quality Treatement Plants, Louisville, Kentucky. On the left Y-axis (Blue) are the averaged ΔCt values obtained from the 5 water quality treatment centers (WQTCs) that service the Louisville metropolitan area. Metadata regarding the selection of the WQTCs and the populations they serve may be found in a companion publication Yeager et al., 2021 [29]. The ΔCt values represent the levels of SARS-CoV-2 N1 amplicon normalized to the PPMOV control amplicon. Note that the left Y-axis is plotted in an inverse manner. Individual data points represent the means of three technical replicates per WQTC per timepoint, with the error bars representing the standard deviation of the means. Plotted on the right Y-axis (Red) are the numbers of confirmed SARS-CoV-2 cases reported to the city of Louisville public health authorities (as per publically available health data) on a per week basis. The X-axis represents the particular week to which the data belongs, with week 1 pertaining to the first week of October 2020, and week 28 being the last week of April 2021. The week pertaining to that of the 3rd and 4th weeks of December are omitted from the above data set, as the underlying public health and wastewater data sets were incomplete for that period of time due to the holiday season. The relevant public health data may be found in the supplemental data files accompanying this manuscript.

The discrepancy between the SARS-CoV-2 nucleic acid detection and clinical incidence data sets could be due to a multitude of factors. Perhaps the most obvious potential reason for the loss of correlation between the public health data and the aggregate wastewater sample data during this period is the undercounting of infected individuals or the possibil-

ity of viral shedding that continues for weeks after symptomatic infection [30]. However, it also remained possible that aggregate wastewater sources exhibit reduced exchange rates relative to smaller wastewater tributaries leading to the "stagnation" of viral RNA levels. If the observations reported above for the aggregate wastewater sources were due to slower rates of exchange, then smaller wastewater sources with higher relative rates of exchange would be expected to exhibit contrary behaviors to the larger aggregate sites. Thus, to determine whether flow rate was impacting the aggregate wastewater observations we utilized qRT-PCR to assay smaller wastewater tributaries that fed into the two major WQTCs of Louisville, Morris Forman WQTC and Derek R. Guthrie WQTC. As exhibited by the data presented in Figure 2, the trends exhibited by these smaller wastewater tributaries closely match the observations obtained for the larger aggregate WQTCs into which they feed. Similar observations were made across other wastewater tributaries in the Louisville metropolitan area. Additionally, the observations for the smaller wastewater sheds are more or less consistent with the averaged measurements for the entire city of Louisville, as reported in Figure 1.
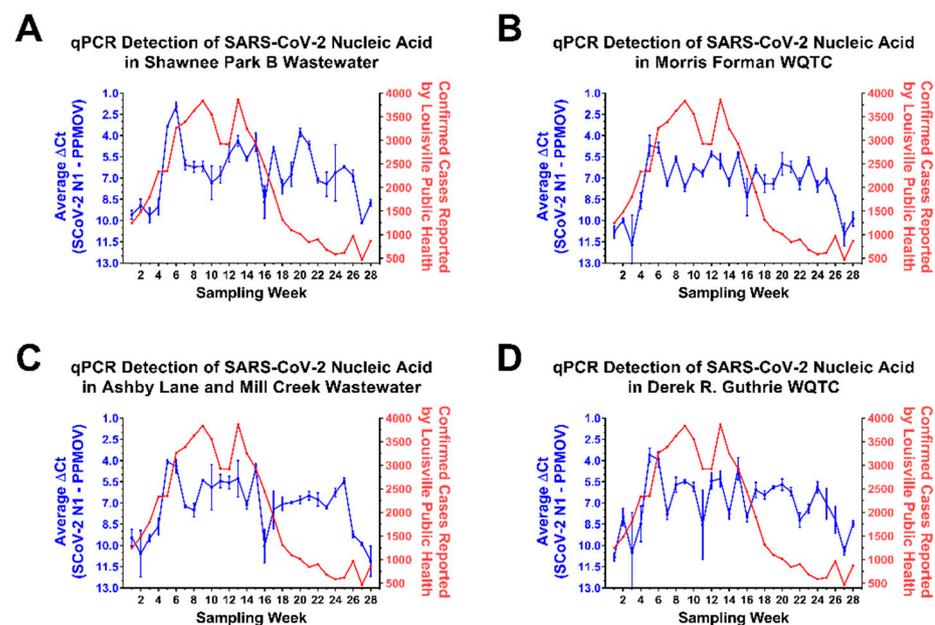

Figure 2. The Quantitative Detection of SARS-CoV-2 in Wastewater Tributaries of the Morris Forman and Derek R. Guthrie WQTCs. On the left Y-axes (Blue) are the average ΔCt values obtained from wastewater tributaries associated with the Shawnee Park B wastewatershed (**A**), and the aggregate wastewater treatment center to which it feeds- Morris Forman WQTC (**B**). Panels (**C**,**D**) are similar, but the areas represented are the Ashby Lane and Mill Creek tributary and the Derek R. Guthrie WQTC. The ΔCt values represent the levels of SARS-CoV-2 N1 amplicon normalized to the PPMOV control amplicon. Note that the left Y-axis is plotted in an inverse manner. Individual data points represent the means of three technical replicates per WQTC per timepoint, with the error bars representing the standard deviation of the means. Plotted on the right Y-axis (Red) are the numbers of confirmed SARS-CoV-2 cases reported to the city of Louisville public health authorities on a per week basis. The X-axis represents the particular week to which the data belongs, with week 1 pertaining to the first week of October 2020, and week 28 being the last week of April 2021. The week pertaining to that of the 3rd and 4th weeks of December are omitted from the above data set, as the underlying data sets were incomplete for that period of time.

Together these observations confirm that the analysis of aggregated wastewater samples obtained from treatment plants are capable of indicating the presence of SARS-CoV-2 nucleic acid. In addition, our analyses of wastewater tributaries and their corresponding individual treatment centers indicate that the detection of SARS-CoV-2 RNA is more or less equitable across local, regional, and metropolitan-wide wastewater sources. While

SARS-CoV-2 RNA levels correlated with increasing community prevalence they did not immediately correlate with a reported decrease community incidence, and the likely underlying causes of which are discussed in greater depth later.

2.2. Aggregate Wastewater Sources Can Provide a Rapid Means to Detect the Prevalence of Variants of Interest/Concern

While the detection of SARS-CoV-2 RNA levels in aggregate wastewater is useful in determining the presence and prevalence of SARS-CoV-2 within a community, it is also potentially a valuable resource by which the genetic diversity of SARS-CoV-2 in a community may be monitored [22,23,25–28]. To this end, the samples obtained from aggregated wastewater sources during an 8-week period corresponding to the 1st week of March to the final week of April 2021 were subjected to next-generation sequencing to determine the sequence diversity of the circulating SARS-CoV-2 populations. As exhibited in Figure 3, the depth of coverage across the SARS-CoV-2 genome obtained from the aggregate wastewater samples was more than sufficient to enable the detection of sequence polymorphisms. Moreover, the genome coverage and depth of sequencing correlated with the prevalence of SARS-CoV-2 within the sample (as per comparisons of Figures 1–3). Unsurprisingly, the depth of coverage obtained from the sequencing efforts was dependent on the concentration of the sample, as detected by qRT-PCR (Figure 3A).
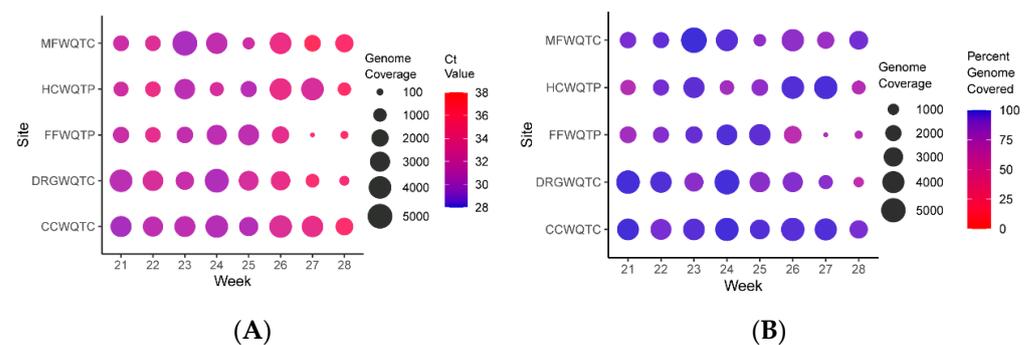

(A)   **(B)**

Figure 3. Quantitative Analysis of Per-Base Genome Coverage and Percent Genome Covered for the Aggregate Wastewater Sites. (**A**) Shown above is a circle plot depicting the depth of coverage per nucleotide (circle size) and the color of the datapoint corresponds to the Ct value of the associated sample, as detected by qRT-PCR. (**B**) Similar to that presented in A, but the circle size represents the depth of coverage, and the color of the circle represents the percentage of the genome covered by at least ten reads (circle color). The individual aggregate wastewater sites are listed on the left vertical axis, and the sampling week listed on the bottom horizontal axis.

The detection of SARS-CoV-2 variants has been a subject of extensive research and public health interest due to their potential impact on global public health. Collectively these efforts have led to the identification of numerous strains with potential consequences to transmission, virulence, and vaccine escapism. At the time of the preparation of this manuscript (25 May 2021) 4 specific variants have been identified by the CDC as variants of interest (VOIs)- the B.1.526 and B.1.526.1 variants first detected in New York, NY, USA, the B.1.525 variant first detected in UK and Nigeria, and the P.2 variant first detected in Brazil; and 5 strains have been identified as variants of concern (VOCs)- the B.1.1.7 variant first reported in UK, the P.1 variant first detected in Japan but of likely Brazilian origin, the B.1.351 variant first reported in South Africa, the B.1.427 and B.1.429 variants first reported in California, USA.

Due to the complex nature of wastewater samples, the isolation and detection of long contiguous sequence reads has been prohibitive. Thus, efforts to identify the genetic composition of the SARS-CoV-2 circulating within populations have utilized the detection of polymorphism markers associated with the VOIs and VOCs. While this approach does not allow for the specific and certain identification of a particular VOI or VOC it allows for

the determination of likelihood on a community scale that compliments traditional clinical surveillance strategies, especially when clinical surveillance has a limited scope.

To identify and establish an unbiased threshold of significance by which to gauge the presence or absence of a given genetic marker, the level of variation of laboratory cultured WA1/2020 was assessed in parallel to wastewater samples using next-generation sequencing. From these studies an average SNP noise rate of ~1% was detected, resulting in the setting of a conservative threshold for the detection of polymorphisms of >5%. Hence, any sequence polymorphism that meets or exceeds the threshold of 5% is considered to be a genuine detection event. As shown in Figure 4A,C, there are 89 nonsynonymous amino acid polymorphisms associated with one or more of the previously identified VOI/Cs. The detection and correlation of these polymorphism markers enables a broad determination of the circulating SARS-CoV-2 strains; however, it should be noted that this approach loses resolution between strains with close relatedness, such as the B.1.429 and B.1.427 VOCs.
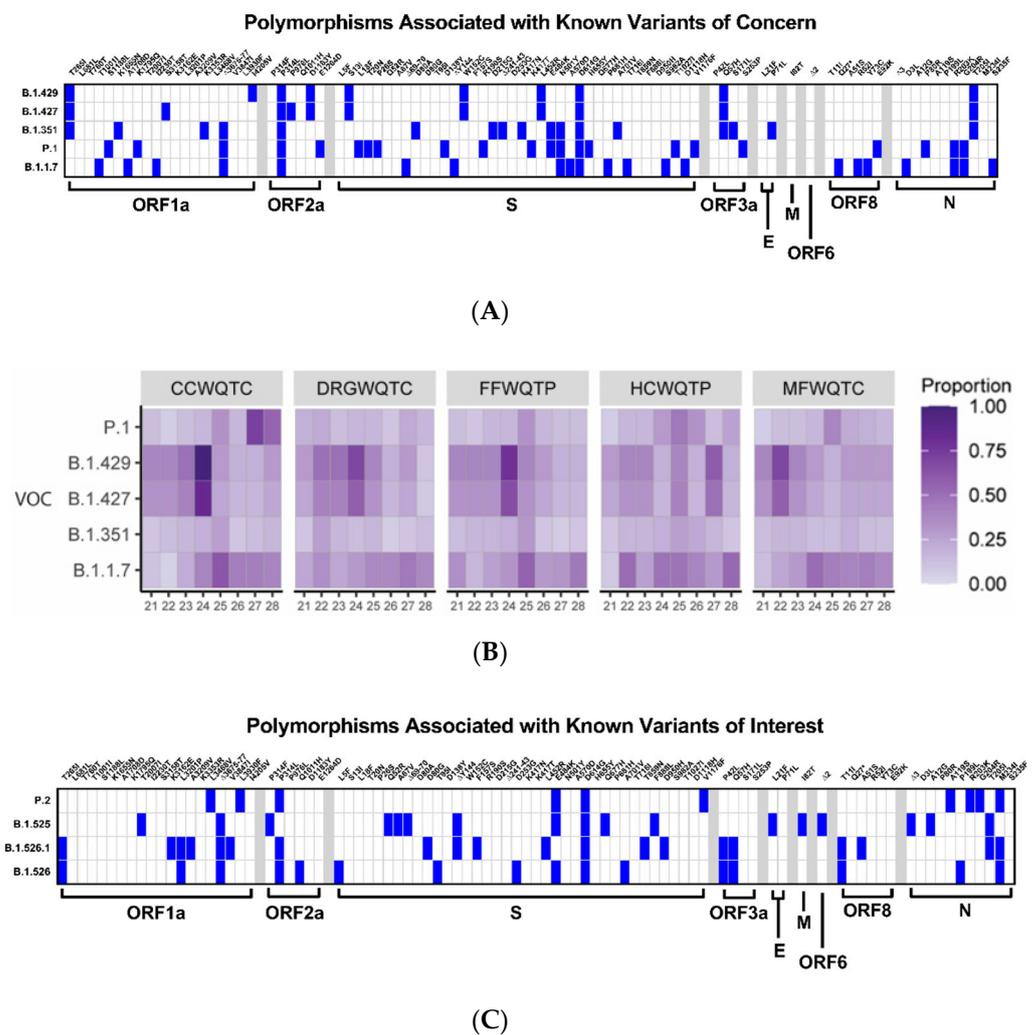
(**A**)

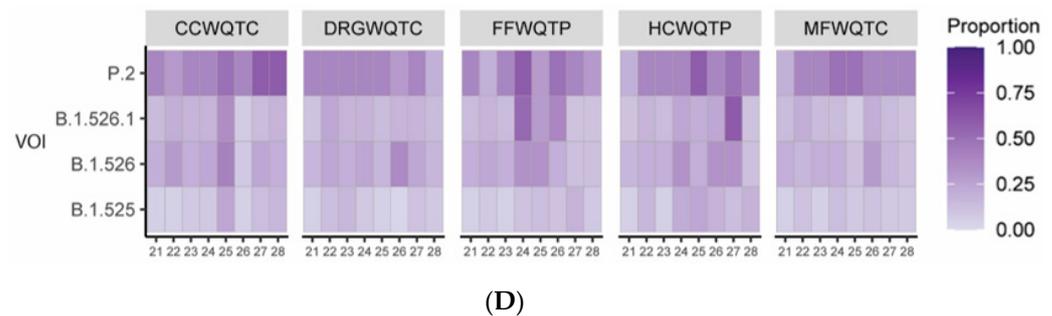(**B**)

(**C**)

Figure 4. *Cont.*

(D)

Figure 4. The Identification of Polymorphism Markers Informs About the Presence of Variants of Concern/Interest in the Community. Panels (**A,C**) detail the specific polymoprhisms associated with each of the currently identified VOCs and VOIs. Specific polymorphisms are found on the upper horizontal axis, and the SARS-CoV-2 gene to which they belong may be found on the lower horizonatal axis. Specific VOC/Is are listed on the left axis. Sequence analyses with respect to time for the aggregated wastewater samples obtained from each of the 5 water quality treatment centers (CC, Cedar Creek; DRG, Derek R. Guthrie; FF, Floyd's Fork; HC, Hite Creek; and MF, Morris Forman) for each of the identified VOCs (panel **B**) and VOIs (panel **D**). Data shown represents the proportion of polymorphism markers which meet or exceed the 5% prevalence threshold that were identified for each listed variant.

Interestingly, the assessment of the aggregated wastewater collection sites via next-generation sequencing revealed that the genetic composition of the community associated SARS-CoV-2 varied significantly during the period of time investigated. Specifically, as depicted in Figure 4B,D, the proportion of polymorphisms associated with a particular VOI or VOC that have met or exceeded the 5% threshold allowed for the determination of which strains were circulating with high likelihood in the community.

A primary observation that may be made from the aggregated wastewater analyses is that the genetic composition of the community associated SARS-CoV-2 of the Louisville metropolitan area varies significantly over a period of several weeks, but is, for the most part, highly consistent across the major wastewater watershed regions within any given week period. Nonetheless, the emergence and expansion of SARS-CoV-2 strains can be detected in the aggregated wastewater sequencing data. This is evidenced by the high prevalence detection of B.1.429/B.1.427 in the wastewater collected from the MFWQTC at week 4, and the emergence of the B.1.429/B.1.427 VOCs in the wastewater collected from CCWQTC, DRGWQTC, and FFWQTC. It should be noted that the time period of movement in this instance strongly correlates with the incubation period of SARS-CoV-2 [31–33].

In addition to detecting the geographic/spatial emergence of the B.1.429/B.1.427 VOCs, the introduction of new variants to the Louisville metropolitan area was also detected during this time by clinical sampling and aggregate wastewater analysis. In early March 2021 the P.1 variant was identified in a clinical sample taken from an individual in the Louisville metropolitan area on 5 March 2021. As supported via all available public data, and to the best of our knowledge, this represents the first instance of the P.1 VOC in the state of Kentucky. Contact tracing indicated the likely site of exposure/infection was in Southern Indiana, where the presence of P.1 had recently been detected. During this period of time the markers associated with the P.1 VOC were not detected in the aggregate wastewater of the Louisville metropolitan area indicating likely limited dissemination at that time; however, several weeks later markers associated with the P.1 VOC were detected in the treatment center which services the infected individual's domicile, MFWQTC. This observation was further confirmed through the analysis of the wastewater tributary associated with the individual's domicile, namely Shawnee Park B. The local sewershed demonstrated a signal associated with the presence of P1 three weeks before detection in the aggregate WQTC and as with B.1.1.7 predicted the emergence of the variant within the community before it became a stable part of the viral strains present in the population.

With the P1 variant, public health efforts including contact tracing were undertaken, but as evidenced by the increased detection of P.1 markers in the Shawnee Park B wastewater samples several weeks after the initial clinical detection of the P.1 VOC, the P.1 variant was not sufficiently contained (Figure 5). Furthermore, as with the spatial emergence of the B.1.429/B.1.427 strain, the P.1 strain was later found to have spread to other sites within the city as evidenced by the detection of the P.1 associated VOC markers in the CCWQTC. At this point in time continued spread of the P.1 variant beyond the regions currently identified is expected.
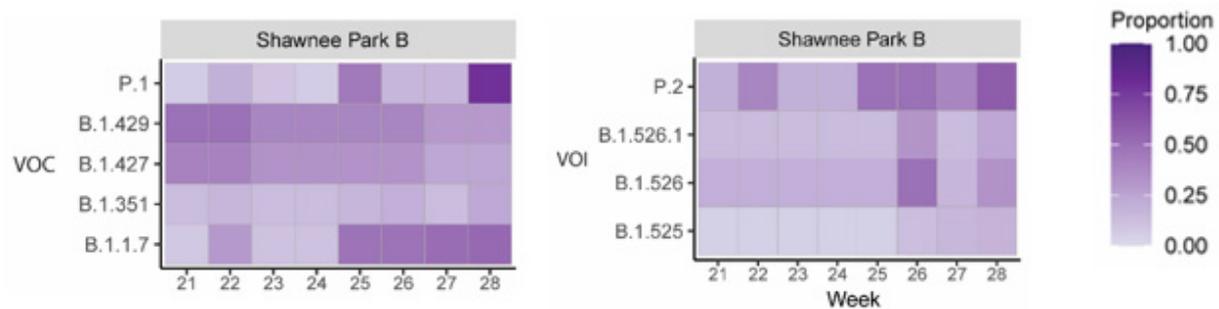

Figure 5. Realtime Identification of the Emergence of the P.1 Variant in a Northwestern Louisville Community. Sequence analyses with respect to time for the wastewater tributary from the Shawnee Park neighborhood in Louisville, KY. Data shown represents the proportion of polymorphism markers which meet or exceed the 5% prevalence threshold that were identified for each listed variant. Week 21 corresponds to the week period/date of the initial detection of the P.1 variant in a clinical sample.

The appearance and emergence of B.1.1.7 is also detected during this period, with the markers associated with this variant becoming prominent concurrent with the rise of B.1.429/B.1.427, and prior to the emergence of markers associated with the P.1 variant. Whether this variant becomes the dominant strain in the Louisville metropolitan area remains to be seen, but the markers associated with the B.1.1.7 VOC are more prevalent and persist to a higher degree than those for the P.1 VOC over the reported time period.

Altogether these observations confirm that wastewater sampling, even on an aggregated scale covering several hundred thousand persons, is capable of detecting the emergence of variants within a community in real time.

2.3. The Analysis of Aggregate Wastewater Sources Can Provide Broad Genetic Information on a Community-Level

As demonstrated above, the analysis of aggregated wastewater allows for rapid broad detection of genetic markers associated with SARS-CoV-2 variants within a community. Further analysis of the sequencing data outside of the lens of known VOI/Cs revealed the presence of other polymorphisms with potential public health ramifications within the Louisville community. As a major focus of SARS-CoV-2 surveillance is potential vaccine escapism we elected to focus our efforts on identifying highly prevalent polymorphisms of the SARS-CoV-2 S protein [34]. As with our previous efforts to detect polymorphisms associated with VOI/Cs, in order to remain conservative in our analyses a prevalence threshold of 5% was established. As shown in Figure 6, 96 polymorphisms exceeded the threshold of 5% prevalence, however relatively few specific polymorphisms reached saturation or persisted for several weeks.

The most notable dominant polymorphism is the D614G mutation, which was universally detected, as was expected as this mutant lineage is currently dominant globally. Additionally, several other polymorphisms not associated with known VOI/Cs exhibit prolonged detection and significant prevalence- these include a pair of polymorphisms associated with the C-terminal region of the S protein, C1254S and C1254*. Curiously, these mutations are located in the cytoplasmic tail of the S2 fragment of the S protein, and the functions and potential importance of a stop codon in this region are discussed later. In addition to the appearance of polymorphisms associated with the aforementioned

VOI/Cs, there is one region of the SARS-CoV-2 S protein that exhibits prolonged detection with moderate levels of sequence prevalence. These include the polymorphisms spanning the SD1 domain of the S1 component of the SARS-CoV-2 S protein. As with the C1254 polymorphisms, the implications of these mutants are discussed below.
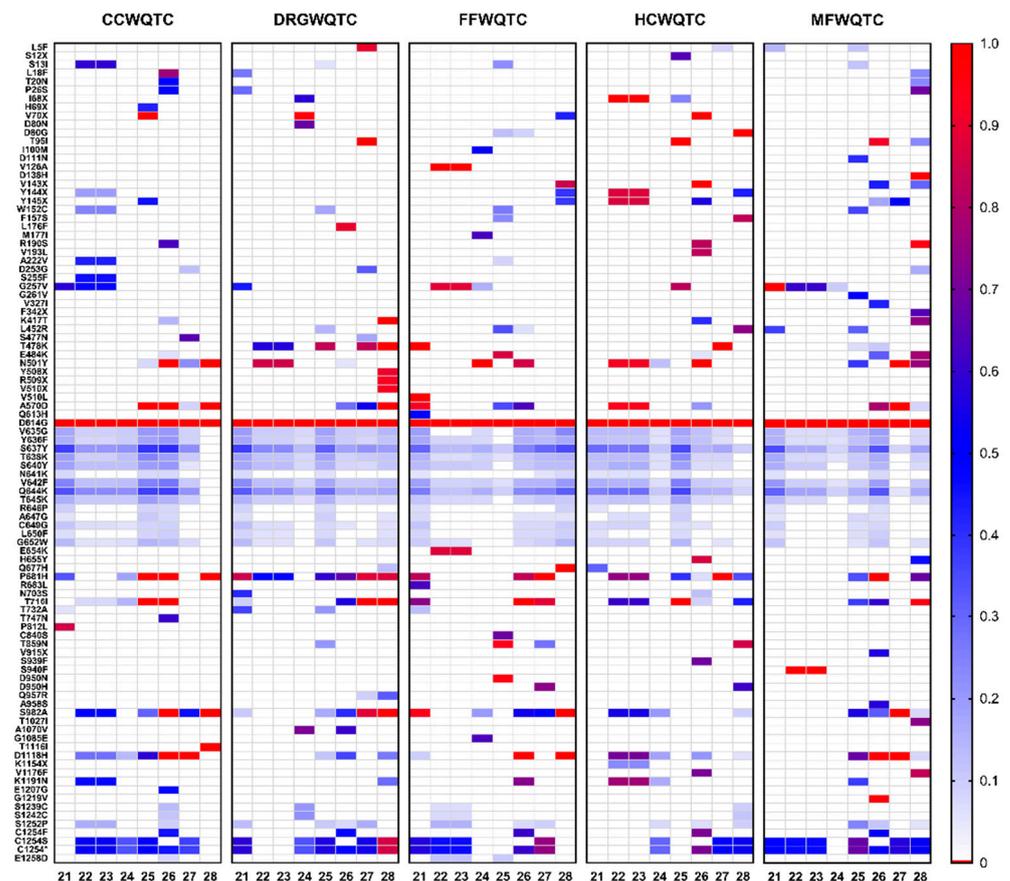

Figure 6. S Protein Polymorphisms Detected in Aggregate Wastewater Over an 8 Week Period. Above are heatmaps for each of the individual wastewater treatment facilities that service the Louisville metropolitan area. Time in weeks is listed on the horizontal axis. Individual polymorphisms that met or exceeded the threshold of 5% prevalence are listed on the left axis. Prevalence, as determined by the number of reads associated with a given polymorphism over total is indicated above in the figure. Silent mutations are not shown.

3. Discussion

3.1. Aggregate Wastewater Analyses Are Indicative of Prevalence/Genetic Variation on a Community-Level

In this study we report the quantitative detection of SARS-CoV-2 viral nucleic acid from aggregated wastewater samples in a metropolitan setting of approximately 1 million individuals. As reported in Figure 1, the detection of SARS-CoV-2 nucleic acid in the aggregated wastewater of Louisville, KY closely mirrored the observed public health data for the corresponding location and time period. Interestingly, a significant lag was observed between the decrease in confirmed cases and the corresponding decrease in SARS-CoV-2 in the aggregate wastewater. The gravity-fed combined wastewater/stormwater nature of the Louisville sewer management system represented a major limitation towards determining the underlying mechanism(s) behind the observed lag. A potential explanation was that sewer system conditions suffered reduced flow during the period in question, however several lines of evidence indicate that this was not the case. Foremost, smaller and larger wastewater tributaries exhibited similar levels of SARS-CoV-2 nucleic acid, and levels

of PPMOV did not fluctuate during the period in question. Moreover, genetic analyses revealed that while the overall levels of SARS-CoV-2 nucleic acid did not appreciably change, the genetic diversity of the SARS-CoV-2 polymorphisms in the wastewater varied to a great extent. Collectively, these observations support the conclusion that the public health data failed to account for all current cases during the period of weeks 14 to 26, and that SARS-CoV-2 prevalence likely remained highly similar to that reported during the peak of the infection. The precise reasons underlying the decrease in the number of confirmed cases detected during this period in time are unknown. One potential reason is a generalized scale back of testing efforts in the Louisville community, which resulted in decreased detection/reporting as asymptomatic individuals were unlikely to seek out testing if not readily available or clinically ill. It should be noted that this period of time corresponds to roughly when large scale vaccination efforts started in Louisville, KY, but the speed and magnitude of the decrease is in excess of the vaccination rate during this time. Regardless, altogether, these data reported here confirm the utility and versatility of assessing aggregated wastewater sources for the presence of SARS-CoV-2 nucleic acid.

In addition to readily being able to monitor the disease prevalence within a community, aggregate wastewater can be assessed using next-generation sequencing for the detection of polymorphisms [22,23,25–28]. As shown in Figures 3–5 the assessment of viral nucleic acid extracted from aggregated wastewater samples is capable of indicating the presence of variants of concern. When comparing aggregate sewersheds to samples of smaller wastewater tributaries there seems to be a lag in the aggregate setting related to resolution likely related to dilution of the virus, however; the virus is generally resolved within a couple of weeks as it spreads within the community. Depending on the public health goals and the planned response the total number of samples needed to cover a community can vary, for an observational program the use of aggregated wastewater collected from regional treatment centers should be adequate, but for an early response program to mitigate spread a sewershed level surveillance program may increase resolution and reaction times. Importantly, as shown in Figure 4, relative changes in strain prevalence can be readily detected, and the emergence of variants within a community can be monitored in real time using aggregated wastewater analyses. Obtaining this information can be informative to the public health response by aiding in general response preparations; however, as the precision with which aggregated, wastewater can identify specific individuals infected with particular variants is, obviously, nil, the continued analyses of clinical specimens is warranted.

Collectively, the unbiased assessment of SARS-CoV-2 prevalence and genetic diversity enables the development of a coordinated public health response towards SARS-CoV-2 through the actions of partnerships between the University of Louisville, the Department of Public Health and Wellness of the Louisville Metro Government, and the Louisville/Jefferson County Metropolitan Sewer District. Understanding the community prevalence through wastewater analysis represents a means by which potential burdens of SARS-CoV-2 infections on public health systems may be estimated in advance of the development of clinical illness. Moreover, identifying which specific communities or regions within a larger metropolitan area are experiencing enhanced circulation/transmission enables the targeted deployment of public health resources for maximum effect. Specific examples of efforts employed in Louisville, KY in response to wastewater monitoring data have included enhanced outreach to at-risk communities with high prevalence trends and the deployment of mobile vaccination clinics. Together these tangible responses to the information gleaned through wastewater analysis have contributed towards the city-wide efforts to combat SARS-CoV-2 infection.

3.2. Sequence Analysis of SARS-CoV-2 Genetic Diversity Reveals Potential Biological Insights

As discussed above, aggregate wastewater can be readily assessed using polymorphism markers associated with VOI/Cs to determine the presence and relative prevalence of specific variants within a community. However, limiting the assessment of sequence diversity to the known VOI/Cs limits the potential of wastewater analyses by seeking to

quantify specific markers ignores the primary utility of the next-generation sequencing approach. Wholistic examination of the sequences obtained from the sequencing of aggregate wastewater has the potential to reveal biological insight into the genetic changes driving the continued emergence of SARS-CoV-2. For the purpose of this study, we have limited the region of specific interest to the S protein of SARS-CoV-2. The data in Figure 6 reveals two interesting phenomena worth further discussion. First, is the apparent significant sequence variation in the SD2 domain of the S1 fragment. This region of S1 is associated with the flexibility with which the Receptor Binding Domain (RBD) can transition between the "up" and "down" conformations, which has been previously associated with receptor binding and entry [35–37]. Curiously, none of the polymorphisms associated with the SD2 domain (which encompasses amino acids 589-677) achieve high prevalence, indicating that perhaps the cost–benefit of these variations is not significantly capable of driving selective pressure. It is worth noting that many of these polymorphisms are not currently associated with any VOI/Cs and are relatively uncommon in the SARS-CoV-2 sequence repositories and have not been previously reported in the state of Kentucky. The second is the C1254* mutation which introduces a premature stop codon resulting in the C-terminal truncation of the S protein by 19 amino acids. This particular mutation was amongst the more prevalent and consistently identified polymorphisms in the Louisville metropolitan area, with a prevalence rate of approximately 60% across several weeks. Interestingly, the deletion of the C-terminal 19 amino acids has the potential to interfere with particle assembly, as the C-terminal domain includes a dibasic motif (KxHxx) involved in ER membrane retrieval via interaction with COPI [38]. This region has been demonstrated to be important to particle assembly in the related betacoronavirus SARS-CoV-1. Thus, the apparent partial selection of a variant with potentially reduced particle production is interesting and may have ramifications on transmission and pathogenesis.

Any discussion of SARS-CoV-2 polymorphism analysis in a complex sample mixture would be remiss if the disclaimer that the sequence diversity observed in wastewater is driven by several factors. One such factor is clearly the number of individuals infected with a particular variant at any given time. Yet, another equally important factor is the degree to which particular strains may be shed into wastewater. Thus, the use of wastewater, aggregate or not, to determine the genetic diversity of SARS-CoV-2 in a community is likely not providing a complete understanding of the diversity and selective pressures within a community.

4. Conclusions

To conclude, the analysis of aggregated wastewater has proven capable of quantitatively assessing the level of SARS-CoV-2 virus within a community and is often capable of predicting public health trends. In addition, the monitoring of genetic diversity through the use of polymorphism marker analysis can improve the understanding of the wastewater trends and provide further information on which public health decisions may be formed. Finally, these efforts have the potential to identify genetic polymorphisms with potential biological implications without the need for extensive clinical surveillance.

5. Materials and Methods

5.1. Wastewater Sample Prep

Wastewater samples were collected through a composite sampler with 24-h composite raw wastewater samples collected into an aseptic 125 mL polyethylene terephthalate bottles. Aggregated wastewater was collected via the sampling of wastewater sources at a rate of 30 mL over 15-min intervals, and the collected materials were homogenized prior to aliquoting and further assessment. Viral particles were precipitated using PEG-8000 (Millipore-Sigma, 89510, Tokyo, Japan). Throughout the process, samples were maintained on ice and processed within 12 h after collection. For each sample, 40 mL of chilled wastewater was passed through a 70 μm cell strainer (VWR, 76327-100, Radnor, PA, USA) and PEG-8000 and NaCl were added to a final concentration of 12.5 mM and 210 mM, respectively. Samples were refrigerated overnight and then centrifuged at 15,000× g for 30 min at 4 °C. The

pellet was resuspended with 1.1 mL TRIzol^{TM} (Thermo Scientific # 15596018, Waltham, MA, USA) and transferred to a sterile microfuge tube. The TRIzol^{TM} sample was then incubated for 5 min at room temperature and then centrifuged at 12,000× *g* for 5 min at 4 °C. The sample was then divided into two 500 µL samples, one for isolation and one for archiving at −80 °C. The sample for isolation had an additional 500 µL of TRIzol added and 900 µL of 100% Ethanol. Samples were vortexed and the RNA was isolated using a Direct-zol™ 96 MagBead RNA kit (Zymo Research, R2102, Irvine, CA, USA) with RNA eluted in 100 µL of DNAse/RNAse Free Water. RNA cleanup was done using the Zymo RNA Clean & Concentrator^{TM}-5 (Zymo #R1016) according to the manufacturer's instructions with RNA eluted in 60 µL of DNAse/RNAse Free Water. Purified RNA was inspected for yield and quality using a NanoDrop 1000. The number of viral copies in each sample was determined using a probe-based RT-qPCR on a QuantStudio 3 (Applied Biosystems, Waltham, MA, USA) real-time PCR system using Taq 1-Step Multiplex Master Mix (Thermo Fisher #A28527). The primer and probe sequences are shown in Table S1 with 5 primer/probe sets used for each sample and all samples ran in triplicate. 4 µL of sample was used for each 20 µL reaction. PCR cycling conditions were 25 °C for 2 min, 50 °C for 10 min, 95 °C for 2 min and 45 cycles of 95 °C for 2 s and 60 °C for 30 s. We generated a standard curve for each primer-probe set used and fit the Ct values to extrapolate copies per mL of wastewater. For this publication we are only reporting on the N1 and PMMoV Ct values generated from this methodology. The ΔCt between N1 and PMMoV was used to normalize the N1 concentrations for potential dilution and are graphed over the sampling period. A detailed protocol may be found in the Supplementary File S1 accompanying this study.

5.2. cDNA Synthesis

The Superscript$^®$ IV First-Strand Synthesis System (Thermo Fisher #18091050) was used to generate cDNA with random hexamer primers. The RT reaction was mixed according to manufacturer's instructions with a final reaction volume of 20 µL and 5 µL of our template RNA added to the mixture. The reverse transcriptase incubation step was performed with sequential incubation at 23 °C for 10 min, 50 °C for 30 min, and 80 °C for 10 min, according to the manufacturer's protocol with adjustment of the incubation times recommended by Swift Biosciences SNAP low input protocol.

5.3. Library Preparation

Libraries were prepared using the Swift Biosciences SNAP low input protocol for SARS-CoV-2 (Swift Bioscience, Ann Arbor, MI, USA, Cat # COSG1V2-96, SN-5X296). 10 µL of cDNA was combined with 20 µL of reaction mix and proceeded with multiplex PCR according to protocol. The PCR product was cleaned up using SPRIselect beads (Beckman Coulter, Brea, CA, USA, Cat. No. B23318) at a 1.0X ratio. The purified sample/beads mix was resuspended in 17.4 µL of TE buffer provided in the post-PCR kit. Samples were indexed through PCR with the SNAP Unique Dual Indexing Primers (Swift Bioscience, Ann Arbor, MI, USA, Cat. # SN91096-1-PLATE). The indexing PCR product was further cleaned up and eluted from the beads using a 0.65X PEG NaCl clean-up. The purified libraries were then eluted in 22 µL of TE buffer and transferred to fresh tubes and stored at −20 °C. 1 additional cycle was added to the multiplex PCR and 2 additional cycles were added to the indexing PCR to obtain higher library yields. Library normalization was performed according to SwiftBio's Normalase 2 nM final pool protocol. 5 µL of Normalase I Master Mix were added to each 20 µL library eluate for a final pool of 2 nM and thoroughly mixed. Samples were placed in the thermocycler to incubate at 30 °C for 15 min. 5 µL of each library were pooled, and 1 µL of Normalase II Master Mix per library was added and thoroughly mixed. The library pool was placed in the thermocycler to incubate at 37 °C for 15 min. 0.2 µL of Reagent X1 per library was added to the pool to inactivate Normalase II at 95 °C for 2 min and held at 4 °C.

5.4. Next-Generation Sequencing

Library pool and PhiX were denatured and diluted following Illumina's directions. Libraries with 1% PhiX spike-in were sequenced at read length 2 × 149 bp using the NextSeq 500/550 Mid Output Kit v2.5 300 Cycles (Illumina, San Diego, CA, USA, Cat# 20024905), targeting 1–5 M reads per library.

5.5. Data Analysis

Sequencing reads were analyzed using a custom bioinformatics pipeline. Low quality bases were trimmed using Trimmomatic v0.38 (1),and were then aligned to the NC_045512.2 reference genome using bwa mem v 0.7.17-r1188 (2). Single nucleotide variants (SNVs) relative to the reference were detected using bcftools mpileup (3). SNVs occurring in at least 5% of the reads with at least five separate supporting instances were marked for further interrogation. SNVs occurring at locations of interest as they relate to specific SARS-CoV-2 variants (B.1.1.7, B.1.351, B.1.526, P.1, and B.1.429) were reported for all of the samples (Supplementary Table S1).

5.6. Correlation/Statistical Data Analysis

Public health data and the SARS-CoV-2 nucleic acid levels detected in wastewater samples with respect to time were analyzed using Spearman's Correlation test to determine the Spearman's r and *p*-Values via Graphpad Prism 7.05.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/pathogens10101271/s1, Table S1: Public Health Information Reported for the City of Louisville, current as of April 2021; File S1: a detailed protocol for the detection of SARS-CoV-2 nucleic acid in wastewater, including all primer sequences.

References

1. Achak, M.; Alaoui Bakri, S.; Chhiti, Y.; M'Hamdi Alaoui, F.E.; Barka, N.; Boumya, W. SARS-CoV-2 in hospital wastewater during outbreak of COVID-19: A review on detection, survival and disinfection technologies. *Sci. Total Environ.* **2021**, *761*, 143192. [CrossRef]

2. Scott, L.C.; Aubee, A.; Babahaji, L.; Vigil, K.; Tims, S.; Aw, T.G. Targeted wastewater surveillance of SARS-CoV-2 on a university campus for COVID-19 outbreak detection and mitigation. *Environ. Res.* **2021**, *200*, 111374. [CrossRef] [PubMed]

3. Spurbeck, R.R.; Minard-Smith, A.; Catlin, L. Feasibility of neighborhood and building scale wastewater-based genomic epidemiology for pathogen surveillance. *Sci. Total Environ.* **2021**, *789*, 147829. [CrossRef]

4. Fernandez-Cassi, X.; Scheidegger, A.; Banziger, C.; Cariti, F.; Tunas Corzon, A.; Ganesanandamoorthy, P.; Lemaitre, J.C.; Ort, C.; Julian, T.R.; Kohn, T. Wastewater monitoring outperforms case numbers as a tool to track COVID-19 incidence dynamics when test positivity rates are high. *Water Res.* **2021**, *200*, 117252. [CrossRef]

5. Roka, E.; Khayer, B.; Kis, Z.; Kovacs, L.B.; Schuler, E.; Magyar, N.; Malnasi, T.; Oravecz, O.; Palyi, B.; Pandics, T.; et al. Ahead of the second wave: Early warning for COVID-19 by wastewater surveillance in Hungary. *Sci. Total Environ.* **2021**, *786*, 147398. [CrossRef] [PubMed]

6. Colosi, L.M.; Barry, K.E.; Kotay, S.M.; Porter, M.D.; Poulter, M.D.; Ratliff, C.; Simmons, W.; Steinberg, L.I.; Wilson, D.D.; Morse, R.; et al. Development of Wastewater Pooled Surveillance of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) from Congregate Living Settings. *Appl. Environ. Microbiol.* **2021**, *87*, e0043321. [CrossRef] [PubMed]

7. Ahmed, W.; Angel, N.; Edson, J.; Bibby, K.; Bivins, A.; O'Brien, J.W.; Choi, P.M.; Kitajima, M.; Simpson, S.L.; Li, J.; et al. First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: A proof of concept for the wastewater surveillance of COVID-19 in the community. *Sci. Total Environ.* **2020**, *728*, 138764. [CrossRef]

8. Bivins, A.; North, D.; Ahmad, A.; Ahmed, W.; Alm, E.; Been, F.; Bhattacharya, P.; Bijlsma, L.; Boehm, A.B.; Brown, J.; et al. Wastewater-Based Epidemiology: Global Collaborative to Maximize Contributions in the Fight Against COVID-19. *Environ. Sci. Technol.* **2020**, *54*, 7754–7757. [CrossRef] [PubMed]

9. Li, X.; Kulandaivelu, J.; Zhang, S.; Shi, J.; Sivakumar, M.; Mueller, J.; Luby, S.; Ahmed, W.; Coin, L.; Jiang, G. Data-driven estimation of COVID-19 community prevalence through wastewater-based epidemiology. *Sci. Total Environ.* **2021**, *789*, 147947. [CrossRef]

10. Xu, X.; Zheng, X.; Li, S.; Lam, N.S.; Wang, Y.; Chu, D.K.W.; Poon, L.L.M.; Tun, H.M.; Peiris, M.; Deng, Y.; et al. The first case study of wastewater-based epidemiology of COVID-19 in Hong Kong. *Sci. Total Environ.* **2021**, *790*, 148000. [CrossRef]

11. Hillary, L.S.; Farkas, K.; Maher, K.H.; Lucaci, A.; Thorpe, J.; Distaso, M.A.; Gaze, W.H.; Paterson, S.; Burke, T.; Connor, T.R.; et al. Monitoring SARS-CoV-2 in municipal wastewater to evaluate the success of lockdown measures for controlling COVID-19 in the UK. *Water Res.* **2021**, *200*, 117214. [CrossRef] [PubMed]

12. Jafferali, M.H.; Khatami, K.; Atasoy, M.; Birgersson, M.; Williams, C.; Cetecioglu, Z. Benchmarking virus concentration methods for quantification of SARS-CoV-2 in raw wastewater. *Sci. Total Environ.* **2021**, *755*, 142939. [CrossRef] [PubMed]

13. Weidhaas, J.; Aanderud, Z.T.; Roper, D.K.; VanDerslice, J.; Gaddis, E.B.; Ostermiller, J.; Hoffman, K.; Jamal, R.; Heck, P.; Zhang, Y.; et al. Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds. *Sci. Total Environ.* **2021**, *775*, 145790. [CrossRef] [PubMed]

14. Westhaus, S.; Weber, F.A.; Schiwy, S.; Linnemann, V.; Brinkmann, M.; Widera, M.; Greve, C.; Janke, A.; Hollert, H.; Wintgens, T.; et al. Detection of SARS-CoV-2 in raw and treated wastewater in Germany—Suitability for COVID-19 surveillance and potential transmission risks. *Sci. Total Environ.* **2021**, *751*, 141750. [CrossRef] [PubMed]

15. Wu, F.; Zhang, J.; Xiao, A.; Gu, X.; Lee, W.L.; Armas, F.; Kauffman, K.; Hanage, W.; Matus, M.; Ghaeli, N.; et al. SARS-CoV-2 Titers in Wastewater Are Higher than Expected from Clinically Confirmed Cases. *Msystems* **2020**, *5*, e00614-20. [CrossRef]

16. Gonzalez, R.; Curtis, K.; Bivins, A.; Bibby, K.; Weir, M.H.; Yetka, K.; Thompson, H.; Keeling, D.; Mitchell, J.; Gonzalez, D. COVID-19 surveillance in Southeastern Virginia using wastewater-based epidemiology. *Water Res.* **2020**, *186*, 116296. [CrossRef] [PubMed]

17. Fuqua, J.L.; Rouchka, E.C.; Waigel, S.; Sokoloski, K.; Chung, D.; Zacharias, W.; Zhang, M.; Chariker, J.; Talley, D.; Santisteban, I.; et al. A rapid assessment of wastewater for genomic surveillance of SARS-CoV-2 variants at sewershed scale in Louisville, KY. *medRxiv* **2021**. [CrossRef]

18. Yao, H.; Lu, X.; Chen, Q.; Xu, K.; Chen, Y.; Cheng, M.; Chen, K.; Cheng, L.; Weng, T.; Shi, D.; et al. Patient-derived SARS-CoV-2 mutations impact viral replication dynamics and infectivity in vitro and with clinical implications in vivo. *Cell Discov.* **2020**, *6*, 76. [CrossRef] [PubMed]

19. Li, Q.; Wu, J.; Nie, J.; Zhang, L.; Hao, H.; Liu, S.; Zhao, C.; Zhang, Q.; Liu, H.; Nie, L.; et al. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell* **2020**, *182*, 1284–1294.e1289. [CrossRef] [PubMed]

20. Zhang, L.; Jackson, C.B.; Mou, H.; Ojha, A.; Peng, H.; Quinlan, B.D.; Rangarajan, E.S.; Pan, A.; Vanderheiden, A.; Suthar, M.S.; et al. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat. Commun.* **2020**, *11*, 6013. [CrossRef]

21. Plante, J.A.; Liu, Y.; Liu, J.; Xia, H.; Johnson, B.A.; Lokugamage, K.G.; Zhang, X.; Muruato, A.E.; Zou, J.; Fontes-Garfias, C.R.; et al. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **2021**, *592*, 116–121. [CrossRef] [PubMed]

22. Izquierdo-Lara, R.; Elsinga, G.; Heijnen, L.; Munnink, B.B.O.; Schapendonk, C.M.E.; Nieuwenhuijse, D.; Kon, M.; Lu, L.; Aarestrup, F.M.; Lycett, S.; et al. Monitoring SARS-CoV-2 Circulation and Diversity through Community Wastewater Sequencing, the Netherlands and Belgium. *Emerg. Infect. Dis.* **2021**, *27*, 1405–1415. [CrossRef]

23. La Rosa, G.; Mancini, P.; Bonanno Ferraro, G.; Veneri, C.; Iaconelli, M.; Lucentini, L.; Bonadonna, L.; Brusaferro, S.; Brandtner, D.; Fasanella, A.; et al. Rapid screening for SARS-CoV-2 variants of concern in clinical and environmental samples using nested RT-PCR assays targeting key mutations of the spike protein. *Water Res.* **2021**, *197*, 117104. [CrossRef] [PubMed]

24. Bi, C.; Ramos-Mandujano, G.; Tian, Y.; Hala, S.; Xu, J.; Mfarrej, S.; Esteban, C.R.; Delicado, E.N.; Alofi, F.S.; Khogeer, A.; et al. Simultaneous Detection and Mutation Surveillance of SARS-CoV-2 and co-infections of multiple respiratory viruses by Rapid field-deployable sequencing. *Med* **2021**, *2*, 689–700.e4. [CrossRef] [PubMed]

25. Fontenele, R.S.; Kraberger, S.; Hadfield, J.; Driver, E.M.; Bowes, D.; Holland, L.A.; Faleye, T.O.C.; Adhikari, S.; Kumar, R.; Inchausti, R.; et al. High-throughput sequencing of SARS-CoV-2 in wastewater provides insights into circulating variants. *Water Res.* **2021**. [CrossRef]

26. Crits-Christoph, A.; Kantor, R.S.; Olm, M.R.; Whitney, O.N.; Al-Shayeb, B.; Lou, Y.C.; Flamholz, A.; Kennedy, L.C.; Greenwald, H.; Hinkle, A.; et al. Genome Sequencing of Sewage Detects Regionally Prevalent SARS-CoV-2 Variants. *MBio* **2021**, *12*, e02703-20. [CrossRef]

27. Martin, J.; Klapsa, D.; Wilton, T.; Zambon, M.; Bentley, E.; Bujaki, E.; Fritzsche, M.; Mate, R.; Majumdar, M. Tracking SARS-CoV-2 in Sewage: Evidence of Changes in Virus Variant Predominance during COVID-19 Pandemic. *Viruses* **2020**, *12*, 1144. [CrossRef] [PubMed]

28. Nemudryi, A.; Nemudraia, A.; Wiegand, T.; Surya, K.; Buyukyoruk, M.; Cicha, C.; Vanderwood, K.K.; Wilkinson, R.; Wiedenheft, B. Temporal Detection and Phylogenetic Assessment of SARS-CoV-2 in Municipal Wastewater. *Cell Rep. Med.* **2020**, *1*, 100098. [CrossRef]

29. Yeager, R.; Holm, R.H.; Saurabh, K.; Fuqua, J.L.; Talley, D.; Bhatnagar, A.; Smith, T. Wastewater Sample Site Selection to Estimate Geographically Resolved Community Prevalence of COVID-19: A Sampling Protocol Perspective. *Geohealth* **2021**, *5*, e2021GH000420. [CrossRef]

30. Cevik, M.; Tate, M.; Lloyd, O.; Maraolo, A.E.; Schafers, J.; Ho, A. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: A systematic review and meta-analysis. *Lancet Microbe* **2021**, *2*, e13–e22. [CrossRef]

31. Guan, W.J.; Ni, Z.Y.; Hu, Y.; Liang, W.H.; Ou, C.Q.; He, J.X.; Liu, L.; Shan, H.; Lei, C.L.; Hui, D.S.C.; et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N. Engl. J. Med.* **2020**, *382*, 1708–1720. [CrossRef]

32. Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K.S.M.; Lau, E.H.Y.; Wong, J.Y.; et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* **2020**, *382*, 1199–1207. [CrossRef]

33. Lauer, S.A.; Grantz, K.H.; Bi, Q.; Jones, F.K.; Zheng, Q.; Meredith, H.R.; Azman, A.S.; Reich, N.G.; Lessler, J. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann. Intern. Med.* **2020**, *172*, 577–582. [CrossRef] [PubMed]

34. Yuan, M.; Huang, D.; Lee, C.D.; Wu, N.C.; Jackson, A.M.; Zhu, X.; Liu, H.; Peng, L.; van Gils, M.J.; Sanders, R.W.; et al. Structural and functional ramifications of antigenic drift in recent SARS-CoV-2 variants. *Science* **2021**, *373*, 818–823. [CrossRef]

35. Ou, J.; Zhou, Z.; Dai, R.; Zhang, J.; Zhao, S.; Wu, X.; Lan, W.; Ren, Y.; Cui, L.; Lan, Q.; et al. V367F mutation in SARS-CoV-2 spike RBD emerging during the early transmission phase enhances viral infectivity through increased human ACE2 receptor binding affinity. *J. Virol.* **2021**, *95*, JVI-00617. [CrossRef] [PubMed]

36. Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, *581*, 215–220. [CrossRef] [PubMed]

37. Huang, Y.; Yang, C.; Xu, X.F.; Xu, W.; Liu, S.W. Structural and functional properties of SARS-CoV-2 spike protein: Potential antivirus drug development for COVID-19. *Acta Pharmacol. Sin.* **2020**, *41*, 1141–1149. [CrossRef] [PubMed]

38. McBride, C.E.; Li, J.; Machamer, C.E. The cytoplasmic tail of the severe acute respiratory syndrome coronavirus spike protein contains a novel endoplasmic reticulum retrieval signal that binds COPI and promotes interaction with membrane protein. *J. Virol.* **2007**, *81*, 2418–2428. [CrossRef] [PubMed]