

Estimating Relative Abundance of 2 SARS-CoV-2 Variants through Wastewater Surveillance at 2 Large Metropolitan Sites, United States

Alexander T. Yu,¹ Bridgette Hughes,¹ Marlene K. Wolfe, Tomas Leon, Dorothea Duong, Angela Rabe, Lauren C. Kennedy, Sindhu Ravuri, Bradley J. White, Krista R. Wigginton, Alexandria B. Boehm, Duc J. Vugia

Monitoring severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants of concern (VOCs) is critical for public health management of coronavirus disease. Sequencing is resource-intensive and incompletely representative, and not all isolates can be sequenced. Because wastewater SARS-CoV-2 RNA concentrations correlate with coronavirus disease incidence in sewersheds, tracking VOCs through wastewater is appealing. We developed digital reverse transcription PCRs to monitor abundance of select mutations in Alpha and Delta VOCs in wastewater settled solids, applied these to July 2020–August 2021 samples from 2 large US metropolitan sewersheds, and compared results to estimates of VOC abundance from case isolate sequencing. Wastewater measurements tracked closely with case isolate estimates (Alpha, rp 0.82–0.88; Delta, rp 0.97). Mutations were detected in wastewater even at levels <5% of total SARS-CoV-2 RNA and in samples available 1–3 weeks before case isolate results. Wastewater variant monitoring should be strategically deployed to complement case isolate sequencing.

By November 2021, the coronavirus disease (COVID-19) pandemic, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), had claimed >5 million lives worldwide, including

>700,000 in the United States (1–3). Since its emergence in late 2019, SARS-CoV-2 has mutated, resulting in some variants categorized by the World Health Organization as variants of concern (VOCs). VOCs have evidence of potential increased infectiousness, immune evasion, and clinical severity, and they have spread globally. Some VOCs, such as Alpha and Delta, have become the predominant strain at different times and regions (4,5). COVID-19 diagnostics, therapeutics, or vaccines may have decreased effectiveness against VOCs (6,7). As of November 2021, VOCs in the United States included the Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), and Delta (B.1.617.2) variants (3).

Monitoring for VOCs is critical for management of the ongoing COVID-19 pandemic, enabling public health officials to track their public health impact, implement control measures, and allocate resources effectively. Detection of SARS-CoV-2 variants occurs primarily through genomic sequencing of isolates collected for PCR-based diagnosis of persons with active COVID-19 infection. Sequencing is resource- and time-intensive and has limits on capacity because of equipment, reagents, and trained personnel (8). As such, complete and timely sequencing of case isolates is not feasible or practical, particularly when case numbers have been high. During January 2020–September 2021, <3% of COVID-19 cases in the United States had isolates that were sequenced and available on public repositories (3). Nonrandom selection of isolates for sequencing and nonuniform result reporting could make results susceptible to bias and not truly representative of circulating variants (4,8,9). Also,

Author affiliations: California Department of Public Health, Richmond, California, USA (A.T. Yu, T. Leon, A. Rabe, S. Ravuri, D.J. Vugia); California Department of Public Health, Sacramento, California, USA (A.T. Yu, T. Leon, A. Rabe, S. Ravuri, D.J. Vugia); Verily Life Sciences, South San Francisco, California, USA (B. Hughes, D. Duong, B.J. White); Stanford University, Stanford, California, USA (M.K. Wolfe, L.C. Kennedy, A.B. Boehm); Emory University, Atlanta, Georgia, USA (M.K. Wolfe); University of Michigan, Ann Arbor, Michigan, USA (K.R. Wigginton)

DOI: <https://doi.org/10.3201/eid2805.212488>

¹These first authors contributed equally to this article.

substantial delays can occur between isolate collection, sequencing and availability of results to public health (9). Given its timeliness, representativeness, and comparatively low costs, wastewater surveillance for VOCs can be a useful supplement to case-based sequencing surveillance (10–12).

Since early in the pandemic, wastewater samples have been collected and analyzed to quantify the amount of SARS-CoV-2 RNA in sewage. Estimates of viral RNA abundance in sewage correlate closely with reported COVID-19 case counts for the catchment area (sewershed) (13,14) and provide a comprehensive snapshot of real-time community transmission independent of individual care-seeking or testing behavior. Therefore, there is a strong interest in determining if wastewater can also provide useful information on circulating VOCs (15). Both sequencing and PCR assays targeting specific mutations have been proposed as methods to detect mutations and deletions in SARS-CoV-2 RNA in wastewater.

Variant monitoring using environmental samples presents technical challenges. Variants are characterized by the presence of multiple mutations on the same RNA genome, and some share ≥ 1 mutations (16). Unlike isolates from an individual case, which consist of a single genome, wastewater samples likely contain material from multiple SARS-CoV-2 variants shed from different persons, each variant at low concentrations and in various states of genomic integrity because of degradation (17). Therefore, because wastewater contains a complex mixture of SARS-CoV-2 RNA fragments, the presence of ≥ 1 variant mutation sequences does not alone prove that the variant is present in wastewater.

We developed targeted digital reverse transcription PCR mutation assays to retrospectively and prospectively monitor wastewater settled solids for the presence and abundance of mutations present in the Alpha (B.1.1.7) and Delta (B.1.617.2) VOCs. We chose wastewater solids because they contain orders of magnitude higher concentrations of viral RNA than wastewater influent (18,19); previous work has documented a strong coupling between SARS-CoV-2 RNA concentrations in wastewater solids and incidence in the associated population contributing to the wastewater (19). We prospectively monitored wastewater solids of a large metropolitan sewershed in California (San Jose), USA, during July 2020–August 2021 for a deletion present in the Alpha variant. We then retrospectively measured the abundance of this deletion in a second large metropolitan area (Sacramento, CA, USA) where samples had been routinely collected. We also measured

concentrations of mutations suggestive of Delta in both sewersheds. We then compared these totals against estimates of Alpha and Delta abundance in each of these sewersheds by using COVID-19 case isolate sequencing data available to the California Department of Public Health (CDPH).

Methods

Mutation Assay Development for Alpha and Delta Variants

We developed assays *in silico* to target mutations present in Alpha (HV69-70) and Delta (Del156-157/R158G). We screened primers and probe sequences (Appendix Table, <https://wwwnc.cdc.gov/EID/article/28/5/21-2488-App1.pdf>) for specificity using BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), and then tested them *in vitro* against a wide range of viral genomes, including wild-type SARS-CoV-2 and SARS-CoV-2 VOCs, including Alpha and Delta. We further tested the sensitivity and specificity of the assays by diluting variant gRNA containing the mutations in no (0 copies), low (100 copies), and high (10,000 copies) background of wild-type gRNA (Appendix).

Wastewater Sample Collection

This study used samples from 2 publicly owned treatment works (POTWs) that serve ≈ 1.5 million residents of Santa Clara County, California, USA (San Jose), and Sacramento County, California, USA (Sacramento). Details of collection processes have been described (14).

We collected samples from the POTWs to span the period before and including the presumed emergence of Alpha and Delta variants in the communities. Before presumed emergence, sampling was 1–4 times per month; during the periods of suspected emergence, sampling was 3–7 times per week. At San Jose, 133 (HV69-70) and 48 (del156-157/R158G) samples and at Sacramento, 64 (HV69-70) and 48 (del156-157/R158G) samples were included for analyses of each mutation.

We extracted RNA from the settled solids and processed within 24 hours of sample collection to measure concentrations of the nucleoprotein (N) gene using digital droplet reverse transcription PCR (Appendix) (20). The N gene codes for the SARS-CoV-2 nucleocapsid; the specific region of the genome targeted by the assay is conserved on SARS-CoV-2 genomes. We included internal recovery controls. Thereafter, we stored RNA samples at -80°C for 0–300 days before analyzing them a second time for the N

gene and the Delta mutation (Del156–157/R158G) or the Alpha mutation (HV69-70), using digital droplet reverse transcription PCR. By comparing the N gene concentration in the samples before and after storage, we confirmed negligible RNA degradation. All wastewater data are publicly available (<https://doi.org/10.25740/zf117dn1545>).

Incident COVID-19 Cases and Case Isolate Sequences

Each POTW provided sewershed boundary shapefiles. We determined the number of PCR-confirmed COVID-19 cases reported to CDPH as a function of episode date (earliest of either specimen collection or symptom onset date) residing within each sewershed using methods reported previously (20) (Appendix).

COVID-19 case isolate whole-genome sequence data available to CDPH included data from the CDC and laboratory partners. We assigned sequence data to a sewershed on the basis of residential home postal code for the sample. We assigned the PANGO lineage based on the software version available at the time data was extracted; most recent results used pangoleARN and pango-designation version 1.2.66 (21).

We calculated VOC abundance estimates by dividing the number of sequences identified as Alpha or Delta (using the World Health Organization definition and including all PANGO sublineages Q.*, for Alpha, and AY.*, for Delta) by the total number of isolates sequenced from persons residing in the sewersheds over 14-day periods. To estimate time between isolate sample collection and sequence result and to measure the effect of that time delay on VOC estimates, we compared 14-day VOC abundance estimates over time against a final estimate generated on August 24, 2021. We chose a 14-day VOC window (versus a 7- or 28-day window) to balance timeliness of results and number of available case isolates sequenced within the window, given that fewer case isolates increase the uncertainty of estimates.

We performed Pearson correlations between the wastewater mutation and case isolate variant datasets, comparing the mean ratio of mutations in wastewater (HV69-70 and Del156-157/R158G to the N gene) to the proportion of case isolates sequenced and characterized as Alpha or Delta, each averaged over the previous 14 days. We used 0 as a replacement for samples where the measurement was below the limit of detection (nondetect); we repeated the analysis by using half the detection limit (500 copies/g), and the results were the same. We set statistical significance at $p < 0.05$ and performed analyses in R studio version 1.4.1106 (<https://www.rstudio.com>).

Results

Variant Mutation Assay Specificity and Sensitivity

In silico analysis indicated no cross-reactivity between the assays and deposited sequences in GenBank. When challenged against wild-type gRNA, the respiratory virus panel, and actual or synthetic variant gRNA, no cross-reactivity occurred. Positive controls and no-template controls run on the sample plate performed according to expectations. Variant mutation concentrations were measured in no, low, and high background of wild-type gRNA, which does not contain the mutations. Results of mutation assays in the presence of high and low background wild-type gRNA are similar to their results in the absence of background wild-type gRNA (Appendix Figure 1), indicating that the assays are sensitive and specific.

Variant Mutation Concentrations in Wastewater Solids

Results for positive and negative controls were as expected, and recovery controls indicated consistent RNA recovery from samples and lack of substantial inhibition (Appendix). We measured HV69-70 concentrations up to 1 time/day at San Jose and up to 3 times/week at Sacramento; concentrations ranged from not detected to $>10,000$ copies/g (Figure 1, 2). N and HV69-70 concentrations at San Jose before 15 Feb 2021 are not presented graphically; samples collected during July and September 2020 did not have measurable HV69-70. HV69-70 was measured for the first time in San Jose solids on November 25, 2020, at concentrations of ≈ 103 copy/g. We did not detect HV69-70 in Sacramento wastewater solids before late February 2021; results for samples collected in October 2020 (not shown in plot) and late January 2021 were nondetect for HV69-70. At both locations, the concentration of HV69-70 relative to the N gene (HV69-70/N ratio) increased over time beginning in early March 2021, peaked in early June 2021 at San Jose and May 2021 at Sacramento when HV69-70/N was ≈ 1 , and then fell until HV69-70 became undetectable at San Jose and present at very low relative concentrations at Sacramento (0.01) in late July 2021 (Figure 1, 2).

Del156–157/R158G concentrations were measured as frequently as three times per week at both San Jose and Sacramento and ranged from not detected to 100,000 copies/g (Figure 1, 2). We observed Del156-157/R158G nondetects in samples collected before early April 2021 at both sites, and then both sites experienced a small peak in Del156–157/R158G concentration in early to mid-May 2021 (Del156–157/

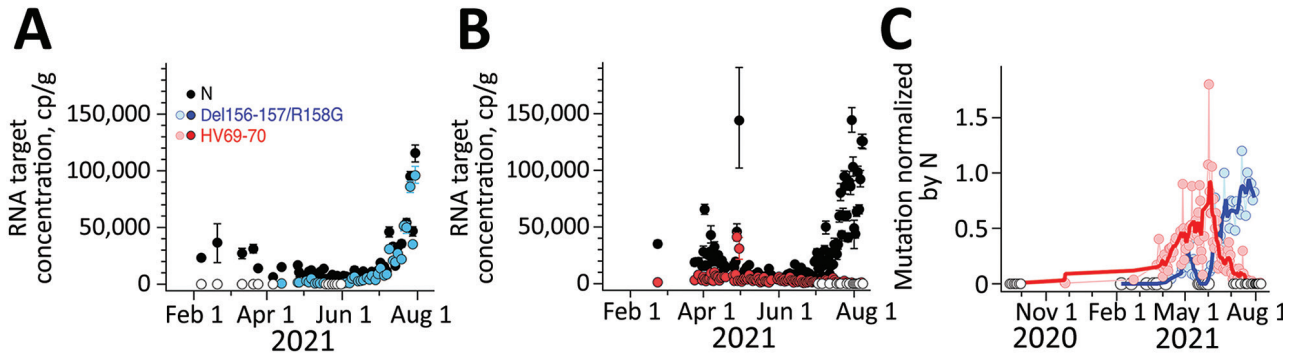


Figure 1. Measurements of severe acute respiratory syndrome coronavirus 2 variants of concern in wastewater solids, San Jose, California, USA. Concentrations of N gene and mutations found in Delta (Del156-157/R158G; panel A), and Alpha (HV69-70; panel B) variants in wastewater solids and their ratio (panel C). Error bars in panels A and B represent SDs derived from the 10 replicates run for each sample; open white circles are nondetects (below the limit of detection) and shown as 0. Errors include technical and replication errors. If error bars are not visible, then errors are smaller than the symbol. Panel C shows smoothed lines for visual reference for mutation ratios. For Del156-157/R158G/N ratio, the smoothed line is a 3-point running average, and for the HV69-70/N ratio, the smoothed line is a 7-point running average; each approximates a weekly average. The timescale for the HV69-70 data (B) is truncated for visualization; additional data on dates before February 15, 2021, are described in the article and shown (C), with the exception of data from July 14, 2020, which was nondetect. N, nucleoprotein.

R158G relative to N ≈ 0.2 – 0.3 at the 2 sites), followed by a decline to undetectable levels over ≈ 2 weeks, followed by a sharp increase until the end of the data series. During this time, N gene concentrations in wastewater increased contemporaneously. The concentration of Del156-157/R158G relative to N (Del156-157/R158G/N ratio) (Figure 1, 2) increased to ≈ 0.8 at the sites by the end of the data series.

Trends in Variants in Sequenced Case Isolates from Sewersheds

We analyzed trends in Alpha and Delta variants confirmed from case isolates collected from residents (case isolates) of the San Jose and Sacramento sewersheds from early February through late July 2021 (Figure 3). Alpha proportions increased in both sewersheds from early March, peaking in May–June and decreasing in early July. Delta was first identified in isolates in early April and by the end of July accounted for almost all sequenced isolates. In San Jose, a small peak in Delta occurred in May, before a large sustained increase in June; a similar peak is also evident, to a lesser extent, in the Sacramento Delta data. During this period, the 7-day average laboratory-confirmed incident COVID-19 cases ranged from 1 to 30/100,000 population (Appendix Figure 2) in each sewershed. Incident COVID-19 cases in each sewershed is positively and significantly correlated with N gene measurements in the settled solids (Pearson R [r_p] 0.8, df 46–131, $p < 10^{-10}$ for both San Jose and Sacramento N gene datasets, regardless of whether they were generated when measuring the Delta or Alpha mutation).

Relationship between Proportion of Alpha and Delta Variants in Case Isolates and Wastewater Mutation Data

We compared ratios of HV69-70 and Del156-157/R158G mutations to the N gene (VOC abundance estimates based on wastewater) against the proportion of all case isolates sequenced and identified as Alpha and Delta variant (VOC abundance estimates based on case isolate sequencing) from each sewershed (Figure 3). Trends of wastewater VOC abundance estimates follow closely and temporally the trends of case isolate sequencing VOC abundance estimates during this period at both sewersheds, including features such as an early peak in Delta in May. Alpha and Delta mutation gene ratios from wastewater were strongly correlated with the corresponding ratios of each VOC from case isolates sequenced: r_p 0.82 ($p < 10^{-5}$, df 19 in San Jose) and 0.88 ($p < 10^{-7}$, df 21 in Sacramento) for Alpha; r_p 0.97 ($p < 10^{-15}$, df 23 for both San Jose and Sacramento) for Delta. When compared to the opposing variant, the mutation gene ratios were not correlated ($p > 0.05$ for all).

Completeness of and Delays in Receiving SARS-CoV-2 Isolate Sequence Data

During February 1–August 1, 2021, the total number of case isolates sequenced over a 14-day period in our 2 sewersheds varied (2–520 median for 8% of all sequenced case isolates from San Jose and 6% from Sacramento). Earliest isolate sequencing results were available to CDPH ≈ 5 days after sample collection date. Approximately 75% of all sequenced isolate results in our dataset were available within 2–3 weeks. As more isolate sequencing data were received, estimated

proportions of VOCs changed over time. Around 3 weeks were required for 95% of VOC estimates (14-day window) to be within 10% of the final estimate.

Discussion

Our results show that the HV69-70 and Del156-157/R158G mutation assays as used for wastewater settled solids were sensitive and specific. By using these PCR mutation assays, we found strong correlation between wastewater estimates and case isolate sequencing-derived estimates of circulating Alpha and Delta in 2 large metropolitan communities in California, USA. Mutations were detected in wastewater samples collected 1–3 weeks earlier than when Alpha and Delta variant estimates generated by case-isolate sequencing were available and reliable. Targeted mutation assays applied to SARS-CoV-2 RNA extracted from wastewater solids can be a rapid, efficient, and reliable way to monitor VOCs introduced to and circulating in a community. Monitoring for VOCs using wastewater may provide earlier complementary surveillance data than from case isolate sequencing data, if mutation assays are or can be developed for new and existing VOCs and put into use in a timely manner.

Use of PCRs targeting characteristic mutations thought to be particular to a SARS-CoV-2 variant may concurrently detect other SARS-CoV-2 strains that carry the same mutations. Targeting a single mutation in wastewater, as was done in our study for Alpha, carries an increased potential risk for mischaracterization. For example, on September 8, 2021, according to GISAID (<https://www.gisaid.org>), a global repository of case isolate sequence data, 1,043,561 (97%) of

the 1,077,360 Alpha (B.1.1.7 and Q sublineages) sequences contained the HV69-70 mutation. However, HV69-70 was also present in other variants, such as B.1.258.19, where it was present in all 141 B.1.258.19 sequences in GISAID, and B.1.617.2, where it was present in 647 (0.2%) of 402,038 sequences. Targeting multiple mutations, as was done in our study with Delta, can increase specificity. Of the 937,570 sequences in GISAID classified as Delta (B.1.617.2 and AY sublineages), 842,354 (90%) have the Del156–157/R158G mutations (referred to as E156G/del157–158 in GISAID). Although this combination of mutations can be present in other variants, it is rarer; the non-Delta variant with the highest percentage of sequences with these mutations is B.1.617.3, for which there were 266 isolates in the global GISAID database and only 77 (29%) possessing these mutations. The non-Delta variant with these mutations for which there are the largest number of isolates in GISAID is B.1.1.7, for which 6 (0.0006%) of the >1,053,637 million sequences have these mutations.

Our findings show that use of mutation assays (HV69-70 for Alpha, Del156-157/R158G for Delta) to estimate circulating variants in wastewater correlated well with estimates from case isolate sequencing data. Wastewater estimates for Alpha, based on a single deletion assay, were robust over time in 2 large municipalities over 8 months (r_p 0.82, $p < 10^{-5}$ in San Jose; r_p 0.88, $p < 10^{-7}$ in Sacramento), including periods of high (tail of 2020 winter, 2021 summer) and low (2021 spring) community SARS-CoV-2 transmission. Similarly, estimates for Delta, based on multiple mutations, correlated highly with estimates from sequenced case isolates (r_p 0.97, $p < 10^{-15}$ for both San Jose

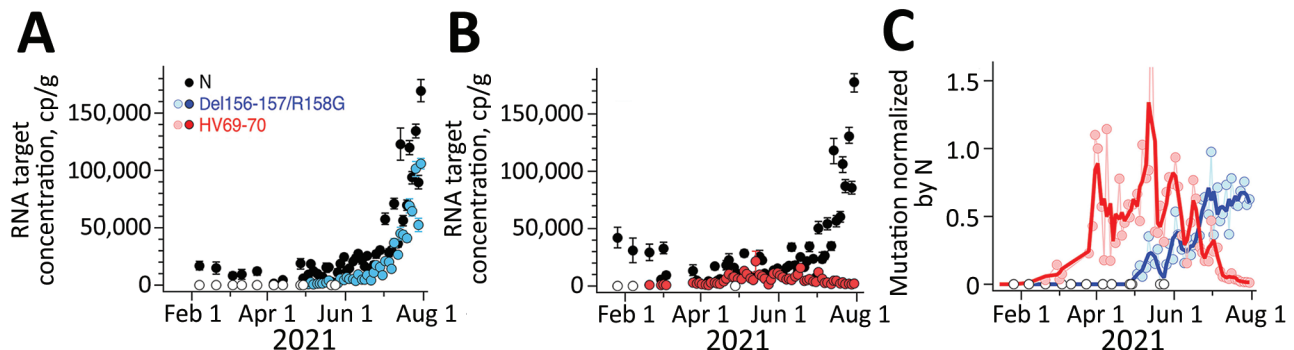
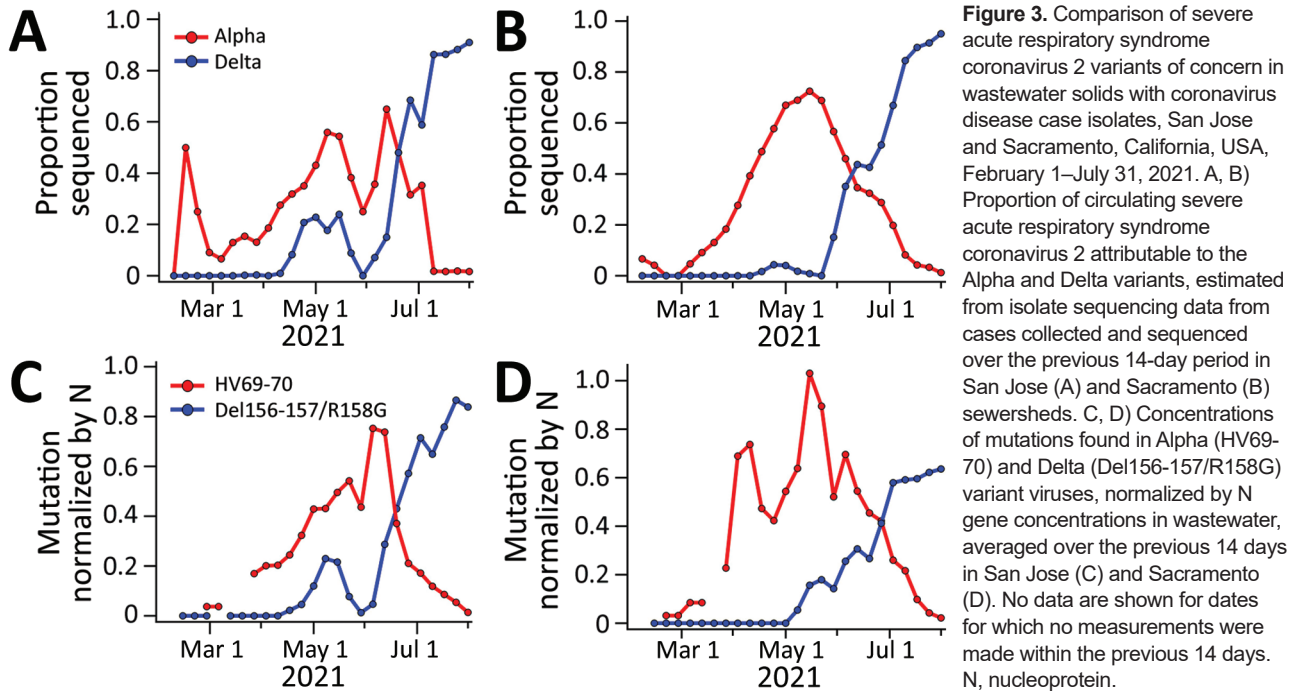


Figure 2. Measurements of severe acute respiratory syndrome coronavirus 2 variants of concern in wastewater solids, Sacramento, California, USA. Concentrations of N gene and mutations found in Delta (Del156-157/R158G; panel A), and Alpha (HV69-70; panel B) severe acute respiratory syndrome coronavirus 2 in wastewater solids and their ratio (panel C). Error bars in panels A and B represent SDs derived from the 10 replicates run for each sample; open white circles are nondetects (below the limit of detection) and shown as 0. Errors include technical and replication errors. If error bars are not visible, then errors are smaller than the symbol. For Del156-157/R158G/N ratio, the smoothed line is a 3-point running average, and for the HV69-70/N ratio, the smoothed line is a 7-point running average; each approximates a weekly average. The timescale for the HV69-70 data (B) is truncated for visualization; additional data on dates before January 15, 2021, are described in the article and were nondetects. One data point is located beyond the upper bound of the y-axis (C): the value for HV69-70/N on May 14, 2021, was 2.4. N, nucleoprotein.



and Sacramento). Concurrent monitoring of VOCs in both wastewater and case isolates can confirm whether targeted mutation assays used are correlated with the VOCs being monitored and mitigate risks for misinterpreting wastewater results. Discrepant or divergent estimates between the 2 datasets should be noticeable within weeks and would suggest another variant with the same mutations circulating at abundance, prompting investigation if unexpected.

Emergence of the Omicron VOC in November 2021 (3) provides an excellent example of the importance of interpreting wastewater mutation assay data in the context of case isolate sequencing data. Omicron also includes the HV69-70 mutation. At the time this study was conducted, the HV69-70 mutation, as noted previously, was rarely circulating in non-Alpha variants, suggesting that positive detections likely represented Alpha. However, Alpha disappeared from California circulation by end of summer 2021 and by December 2021, public health concern was for Omicron. With zero Alpha case isolates detected in either sewershed during September 1–December 1, the HV69-70 mutation assay was deployed on wastewater to screen for presence of Omicron, a more likely VOC to emerge in California than Alpha, while a more specific assay was developed (22).

For validated assays deployed in established wastewater sites, wastewater surveillance for VOCs could be an important adjunctive estimate of variant circulation. Because cost and limited genomic

testing capacity make sequencing all COVID-19 isolates impractical, especially during times of high case incidence, health departments and decision-makers extrapolate information from relatively small numbers or proportions of sequenced isolates, which may be biased and unrepresentative. For our case dataset, 14-day VOC estimates were derived from as few as 2–20 total case isolates and <1% of total cases sequenced.

Wastewater variant monitoring can overcome biases and delays seen with case isolate sequencing. Because everyone living in a sewershed contributes waste to the system, wastewater monitoring is independent of testing and care accessibility biases and results are more representative of cases in that sewershed. In addition, wastewater mutation assay results are available in a shorter time than VOC estimates from sequencing of case isolates. In our monitored sewersheds, the total average turnaround time from wastewater collection to testing results was <8 hours. In contrast, for our 2 sewersheds, it took 2–3 weeks after sample collection date for 75% of case isolate sequence results to be received and 3 weeks for most 14-day VOC estimates to be within 10% of their final estimate. These delays do not include the additional delay between case symptom onset and test taking that could further accentuate time advantages of wastewater variant monitoring.

Several limitations exist for using wastewater (solids or liquids) for SARS-CoV-2 variant monitoring.

Laboratory limits of detection for SARS-CoV-2 RNA in wastewater and for targeted mutations may result in no detection, especially at times with lower community COVID-19 case counts and consequent lower overall concentrations of SARS-CoV-2 RNA in wastewater. However, even in mid-May 2021, when case counts in these 2 sewersheds were as low as 1–2 cases/100,000 population, both SARS-CoV-2 RNA levels and variant abundance could still be measured and accurately estimated. Estimates of circulating Alpha and Delta mutations were also able to be consistently detected even at levels <5% of total SARS-CoV-2 RNA. Limits of detection, both for SARS-CoV-2 and for different mutations associated with variants, are likely to vary depending on laboratory methods used and which mutation is targeted; delineating these limits for each laboratory, sewershed, and assay is important for interpreting what a nondetect result implies about variant circulation.

Because newly identified variants to be monitored require new mutation assays to be designed, the time needed to design, test, and begin using assays is a crucial consideration (Appendix Figure 3). Although the time to design an assay *in silico* (<1 day) and test its sensitivity and specificity *in vitro* (3–5 days) is short, the time to receive reagents, including synthesized oligos and positive control RNA, from vendors can take 4–6 weeks because of supply chain issues and increased demand during the pandemic. In addition, before an assay can be designed, variant sequences and mutations must be accurately characterized, which can delay the assay design process. Efforts to develop assays before variants become VOCs and proactively order reagents can help ensure assays are available when needed for public health response.

Monitoring for VOCs will continue to be an important public health function and a need that will become more salient if SARS-CoV-2 testing of cases and sequencing resources or utilization decrease over time. Difficulty in surveillance based on case isolate sequencing, including difficulties attributable to non-representative sampling and delayed results, mean that complementary variant surveillance methods are needed. Detection and monitoring of variants in wastewater has been proposed as an adjunct methodology, and our experiences monitoring for 2 VOCs in 2 large California municipalities support the use of targeted PCR mutation assays as a useful method to estimate abundance of circulating VOCs and inform public health. In conjunction with continued COVID-19 case isolate sequencing, wastewater variant monitoring can be strategically deployed as an adjunct public health surveillance tool.

Acknowledgments

We thank the many people who contributed to wastewater sample collection, including Srividhya Ramamoorthy, Michael Cook, Ursula Bigler, James Noss, Lisa C. Thompson, Payak Sarkar, Noel un Enoki, and Amy Wong. We also thank the many groups involved in the sequencing of case isolates, including California, Santa Clara, and Sacramento County Public Health departments and the many public, private, and academic laboratories who have dedicated substantial resources towards the understanding of SARS-CoV-2 variants. In addition, we thank the California Department of Public Health COVID-19 Epidemiology and Data teams for their help with COVID-19 data.

Parts of this study were performed at Stanford University, which sits on the ancestral and unceded lands of the Muwekma Ohlone people. We pay our respects to them and their Elders, past and present, and are grateful for the opportunity to live and work here.

This work is supported in part by a gift from the CDC Foundation. This study was supported in part by the Epidemiology and Laboratory Capacity for Infectious Diseases Cooperative Agreement (no. 6NU50CK000539-03-02) from CDC. B.W., D.D., and B.H. are employees of Verily Life Sciences.

About the Author

Dr. Yu is a public health medical officer at the California Department of Public Health, where he works on enteric and waterborne diseases surveillance, as well as on the COVID-19 pandemic response. His research interests include enhanced and environmental surveillance methods for infectious diseases and climate change-related infectious diseases.

References

1. State of California. Tracking COVID-19 in California [cited 2021 Sep 28]. <https://covid19.ca.gov/state-dashboard>
2. World Health Organization. COVID-19 dashboard [cited 2021 Sep 28]. <https://covid19.who.int>
3. Centers for Disease Control and Prevention. COVID data tracker [cited 2021 Sep 28]. <https://covid.cdc.gov/covid-data-tracker>
4. Campbell F, Archer B, Laurenson-Schafer H, Jinnai Y, Konings F, Batra N, et al. Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. *Euro Surveill.* 2021;26:2100509. <https://doi.org/10.2807/1560-7917.ES.2021.26.24.2100509>
5. Twohig KA, Nyberg T, Zaidi A, Thelwall S, Sinnathamby MA, Aliabadi S, et al.; COVID-19 Genomics UK (COG-UK) consortium. Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: a cohort study. *Lancet Infect Dis.* 2022;22:35–42. [https://doi.org/10.1016/S1473-3099\(21\)00475-8](https://doi.org/10.1016/S1473-3099(21)00475-8)

6. Sanders RW, de Jong MD. Pandemic moves and countermoves: vaccines and viral variants. *Lancet*. 2021;397:1326–7. [https://doi.org/10.1016/S0140-6736\(21\)00730-3](https://doi.org/10.1016/S0140-6736(21)00730-3)
7. World Health Organization Regional Office for Africa. Weekly bulletin on outbreak and other emergencies: week 39: 21–27 September 2020. 2020 Sep 28 [cited 2021 Sep 28]. <https://apps.who.int/iris/handle/10665/335723>
8. Walensky RP, Walke HT, Fauci AS. SARS-CoV-2 variants of concern in the United States – challenges and opportunities. *JAMA*. 2021;325:1037–8. <https://doi.org/10.1001/jama.2021.2294>
9. World Health Organization Regional Office for Europe. Methods for the detection and identification of SARS-CoV-2 variants, March 2021 [cited 2021 Sep 28]. <https://apps.who.int/iris/handle/10665/340067>
10. Crits-Christoph A, Kantor RS, Olm MR, Whitney ON, Al-Shayeb B, Lou YC, et al. Genome sequencing of sewage detects regionally prevalent SARS-CoV-2 variants. *MBio*. 2021;12:e02703–20. <https://doi.org/10.1128/mBio.02703-20>
11. Izquierdo-Lara R, Elsinga G, Heijnen L, Munnink BBO, Schapendonk CME, Nieuwenhuijse D, et al. Monitoring SARS-CoV-2 circulation and diversity through community wastewater sequencing, the Netherlands and Belgium. *Emerg Infect Dis*. 2021;27:1405–15. <https://doi.org/10.3201/eid2705.204410>
12. Bar-Or I, Weil M, Indenbaum V, Bucris E, Bar-Ilan D, Elul M, et al. Detection of SARS-CoV-2 variants by genomic analysis of wastewater samples in Israel. *Sci Total Environ*. 2021;789:148002. <https://doi.org/10.1016/j.scitotenv.2021.148002>
13. Kirby AE, Walters MS, Jennings WC, Fugitt R, LaCross N, Mattioli M, et al. Using Wastewater surveillance data to support the COVID-19 response – United States, 2020–2021. *MMWR Morb Mortal Wkly Rep*. 2021;70:1242–4. <https://doi.org/10.15585/mmwr.mm7036a2>
14. Wolfe MK, Archana A, Catoe D, Coffman MM, Dorevich S, Graham KE, et al. Scaling of SARS-CoV-2 RNA in settled solids from multiple wastewater treatment plants to compare incidence rates of laboratory-confirmed COVID-19 in their sewersheds. *Environ Sci Technol Lett*. 2021;8:393–404. <https://doi.org/10.1021/acs.estlett.1c00184>
15. Peccia J, Zulli A, Brackney DE, Grubaugh ND, Kaplan EH, Casanovas-Massana A, et al. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nat Biotechnol*. 2020;38:1164–7. <https://doi.org/10.1038/s41587-020-0684-z>
16. Chakraborty C, Bhattacharya M, Sharma AR. Present variants of concern and variants of interest of severe acute respiratory syndrome coronavirus 2: their significant mutations in S-glycoprotein, infectivity, re-infectivity, immune escape and vaccines activity. *Rev Med Virol*. 2021 Jun 27 [Epub ahead of print]. <https://doi.org/10.1002/rmv.2270>
17. Wurtzer S, Waldman P, Ferrier-Rembert A, Frenois-Veyrat G, Mouchel JM, Boni M, et al.; OBEPINE consortium. Several forms of SARS-CoV-2 RNA can be detected in wastewaters: Implication for wastewater-based epidemiology and risk assessment. *Water Res*. 2021;198:117183. <https://doi.org/10.1016/j.watres.2021.117183>
18. Li B, Di DYW, Saingam P, Jeon MK, Yan T. Fine-scale temporal dynamics of SARS-CoV-2 RNA abundance in wastewater during A COVID-19 lockdown. *Water Res*. 2021;197:117093. <https://doi.org/10.1016/j.watres.2021.117093>
19. Graham KE, Loeb SK, Wolfe MK, Catoe D, Sinnott-Armstrong N, Kim S, et al. SARS-CoV-2 RNA in wastewater settled solids is associated with COVID-19 cases in a large urban sewershed. *Environ Sci Technol*. 2021;55:488–98. <https://doi.org/10.1021/acs.est.0c06191>
20. Wolfe MK, Topol A, Knudson A, Simpson A, White B, Vugia DJ, et al. High-frequency, high-throughput quantification of SARS-CoV-2 RNA in wastewater settled solids at eight publicly owned treatment works in northern California shows strong association with COVID-19 incidence. *mSystems*. 2021;6:e0082921. <https://doi.org/10.1128/mSystems.00829-21>
21. Rambaut A, Holmes EC, O’Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol*. 2020;5:1403–7. <https://doi.org/10.1038/s41564-020-0770-5>
22. KQED News. Scientists find traces of Omicron in sewage from Sacramento and Merced [cited 2021 Dec 8]. <https://www.kqed.org/news/11898161/scientists-find-traces-of-omicron-in-sewage-from-sacramento-and-merced>

Address for correspondence: Alexander T. Yu, California Department of Public Health, 850 Marina Bay Pkwy, Building P, Richmond, CA 94804, USA; email: alexander.yu@cdph.ca.gov

Estimating Relative Abundance of 2 SARS-CoV-2 Variants through Wastewater Surveillance at 2 Large Metropolitan Sites, United States

Appendix

Supplementary Methods

Sampling information

Samples were collected by POTW staff using sterile technique in clean, labeled bottles. Approximately 50 ml of settled solids were collected each study day from each POTW. Settled solids samples were grab samples at SAC. It should be noted that “grab” samples from the sludge line of the primary clarifier represent composites of solids settled over the residence time of the clarifier. At SJ, POTW staff manually collected a 24 h composite sample. Samples were immediately stored at 4°C and transported to the lab and processed (within 6 hours of collection).

Mutation assay design

Assays were developed in silico using Primer3Plus (<https://primer3plus.com/>) to target mutations present in Alpha and Delta variants of concern. Mutation sequences were obtained from outbreak.info/compare-lineages. The parameters used in the development are provided in the list following this paragraph. Primers and probe sequences are provided in Appendix Table and schematic of the assays are provided in Appendix Figure 4 and Appendix Figure 5.

Parameters used with primer design software.

- Product size ranges: 60–200
- Primer size: min 15, opt 20, max 36
- Primer melting temperature: min 50°C, optimal 60°C, max 65°C

- GC% content: min 40%, optimal 50%, high 60%
- concentration of divalent cations = 3.8 mM
- concentration of dNTPs needs to be 0.8 mM
- Internal Oligo: size min 15, optimal 20, max 27
- Internal Oligo: Melting temp min 62°C, optimal 63°C, max 70°C
- Internal Oligo: GC% min 30%, optimal 50%, max 80%

Specificity Screening Against Other Targets

Primers and probe sequences were screened for specificity tested in vitro against a respiratory virus panel (NATtrol Respiratory Verification Panel, Zeptomatrix) that includes several influenza and coronavirus viruses, “wild-type” gRNA from SARS-CoV-2 strain 2019-nCoV/USA-WA1/2020 (ATCC® VR-1986D) which does not contain the mutations (hereafter referred to as WT-gRNA), heat inactivated SARS-CoV-2 strain Alpha (SARS-CoV-2 variant B.1.1.7, ATCC® VR-3326HK), and synthetic gRNA from Twist Biosciences (South San Francisco, California, USA) for Beta (Twist control 16), Gamma (Twist control 17), Kappa (Twist control 18), and Delta (Twist control 23) variants. RNA was extracted from the virus panel and whole viruses using the Chemagic Viral DNA/RNA 300 Kit H96 for the Perkin Elmer Chemagic 360. RNA was used undiluted as template using digital droplet RT-PCR with mutation primer and probes in singleton assays in single wells (see further details on digital PCR below). The concentration of targets used in the in vitro specificity testing was ≈ 275 copies per well.

Sensitivity and Specificity

For sensitivity and specificity testing, single wells were run using the HV69–70 assay for a subset of no and low background samples, but otherwise run in three wells. The number of copies of variant mutation sequences input to each well was estimated using a dilution series of variant gRNA in no background; the vendor specified concentration of the variant gDNA or gene blocks was scaled by the slope of the curve relating the measured ddRT-PCR concentration and the calculated input concentration based on the vendor estimates. Additional details of these experiments are provided in the main text of the paper.

Solids Sample Processing

The solids were dewatered by centrifugation at 24,000 x g for 30 minutes at 4°C. The supernatant was aspirated and discarded. A 0.5 - 1 g aliquot of the dewatered solids was dried at 110°C for 19–24 hrs to determine its dry weight. Bovine coronavirus (BCoV) was used as a positive recovery control. Each day, attenuated bovine coronavirus vaccine (PBS Animal Health, Calf-Guard Cattle Vaccine) was spiked into DNA/RNA shield solution (Zymo Research) at a concentration of 1.5 µL /mL. Dewatered solids were resuspended in the BCoV-spiked DNA/RNA shield to a concentration of 75 mg/mL. This concentration of solids was chosen as previous work titrated solutions with varying concentrations of solids to identify a concentration at which inhibition of the SARS-CoV-2 assays was minimized (5). Between five and ten 5/32” Stainless Steel Grinding Balls (OPS Diagnostics) were added to each sample which was subsequently homogenized by shaking with a Geno/Grinder 2010 (Spex SamplePrep). Samples were then briefly centrifuged to remove air bubbles introduced during the homogenization process, and then vortexed to re-mix the sample.

RNA was extracted from 10 replicate aliquots per sample. For each replicate, RNA was extracted from 300 µl of solids sample using the Chemagic Viral DNA/RNA 300 Kit H96 for the Perkin Elmer Chemagic 360 followed by PCR Inhibitor Removal with the Zymo OneStep-96 PCR Inhibitor Removal Kit. Extraction negative controls (water) and extraction positive controls were extracted using the same protocol as the solids samples. The positive controls consisted of 500 copies of WT-gRNA in the BCoV-spiked DNA/RNA shield solution described above. Undiluted extract was used for the SARS-CoV-2 assay template and a 1:100 dilution of the extract was used for the BCoV / PMMoV assay template. PMMoV is pepper mild mottle virus and is found in high abundance in feces and wastewater globally and is used here as a fecal strength control and an internal recovery control.

ddRT-PCR

Digital droplet RT-PCR assays for the mutations and N gene targets (see Appendix Table for primer and probe sequences, purchased from IDT) were performed on 20 µl samples from a 22 µl reaction volume, prepared using 5.5 µl template, mixed with 5.5 µl of One-Step RT-ddPCR Advanced Kit for Probes (Bio-Rad 1863021), 2.2 µl Reverse transcription, 1.1 µl DTT and primers and probes at a final concentration of 900 nM and 250 nM respectively. Droplets were generated using the AutoDG Automated Droplet Generator (Bio-Rad). PCR was performed

using Mastercycler Pro with cycling conditions described below. Droplets were analyzed using the QX200 Droplet Reader (Bio-Rad). All liquid transfers were performed using the Agilent Bravo (Agilent Technologies).

PCR was performed using Mastercycler Pro with the following cycling conditions: reverse transcription at 50°C for 60 minutes, enzyme activation at 95°C for 5 minutes, 40 cycles of denaturation at 95°C for 30 seconds and annealing and extension at either 59°C (for single or multiplex SARS-CoV-2 assays that include HV69–70, or do not include any mutation), 61°C (for single or multiplex assays that include Del156–157/R158G) or 56°C (for PMMoV/BCoV duplex assay) for 30 seconds, enzyme deactivation at 98°C for 10 minutes then an indefinite hold at 4°C. The ramp rate for temperature changes were set to 2°C/second and the final hold at 4°C was performed for a minimum of 30 minutes to allow the droplets to stabilize.

For the assay sensitivity and specificity testing, assays were run in single wells, or triplicate wells as specified above; each plate contained positive and negative PCR controls. Each wastewater sample was run in 10 replicate wells, and each plate of wastewater samples included extraction negative controls that were run in 7 wells, and extraction positive controls in 1 well. In addition, PCR positive controls for each target assayed on the plate were run in 1 well, and NTC were run in 7 wells.

Positive controls consisted of BCoV and PMMoV gene block controls (dsDNA purchased from IDT), gRNA of SARS-CoV-2 (strain 2019-nCoV/USA-WA1/2020, ATCC® VR-1986D), gRNA of SARS-CoV-2 Alpha lineage (SARS-CoV-2 variant B.1.1.7, ATCC® VR-3326HK), and gene block controls for Del156–157/R158G. Results from replicate wells were merged for analysis. Thresholding was done using QuantaSoft Analysis Pro Software (Bio-Rad, version 1.0.596). In order for a sample to be recorded as positive, it had to have at least 3 positive droplets. For the wastewater samples, three positive droplets corresponds to a concentration between ~500–1000 cp/g; the range in values is a result of the range in the equivalent mass of dry solids added to the wells.

For the wastewater samples, concentrations of RNA targets were converted to concentrations per dry weight of solids in units of copies/g dry weight using dimensional analysis. The total error is reported as standard deviations and includes the errors associated with the Poisson distribution and the variability among the 10 replicates. The recovery of BCoV was

determined by normalizing the concentration of BCoV by the expected concentration given the value measured in the spiked DNA/RNA shield. BCoV was used solely as a process control; samples were rerun in cases where the recovery of BCoV was less than 10%. PMMoV concentrations at the POTW were consistent within POTWs (Appendix Figure 6) and within the ranges observed previously suggesting consistent recovery of RNA across samples. All wastewater data, along with additional reporting details per the MIQE (5) and EMMI (6) guidelines are available publicly at the Stanford Digital Repository (<https://doi.org/10.25740/zf117dn1545>). The detection limits for the endogenous wastewater RNA targets are as follows: 500–1000 copies/g dry weight for all SARS-CoV-2 RNA gene targets, and 5×10^5 – 10×10^5 copies/g for the PMMoV gene target. The range varies from sample-to-sample and depends on the wet weight of the solids suspended in the DNA/RNA shield.

Effect of Storage on RNA

RNA samples from San Jose (SJ) were stored between 0 and 7 days before they were analyzed for HV69–70 and the N gene in a multiplex assay. SJ and Sacramento RNA samples were stored between 15 and 190 days before analysis for Del156–157/R158G and the N gene; Sacramento RNA samples were stored between 15 and 300 days before being analyzed for HV69–70 and the N gene. The N gene was run a second time in all samples as a check for RNA degradation during storage.

Additional Details on Incident COVID-19 Cases and COVID-19 Case Isolate Sequences

Sewershed service area boundary shapefiles were provided by each POTW and compared to postal code boundaries (2020 U.S. Census TIGER/Line shapefiles) in ArcGIS Pro version 2.7.3 (ESRI, Redlands CA). All postal codes with >50% of land area within sewershed boundaries were considered within the POTW service area and used for further analysis.

Supplementary Results

Recovery of BCoV was higher than 10% in all samples suggesting acceptable viral RNA recovery, and absence of substantial inhibition. PMMoV concentrations ranged from 108 to 109 copies per gram at both plants (Appendix Figure 6) on the sampling dates included in this study, within expected ranges previously measured at the POTWs (7) and indicating consistent

recovery of viral RNA from the wastewater solids. Negative controls including extraction and PCR negative controls were negative on all plates and positive controls all amplified. N genes concentrations measured contemporaneously with the mutation targets were similar to those measured immediately after wastewater sample collection (average ratio of the measurements ~ 0.9) indicating RNA stability during -80°C storage.

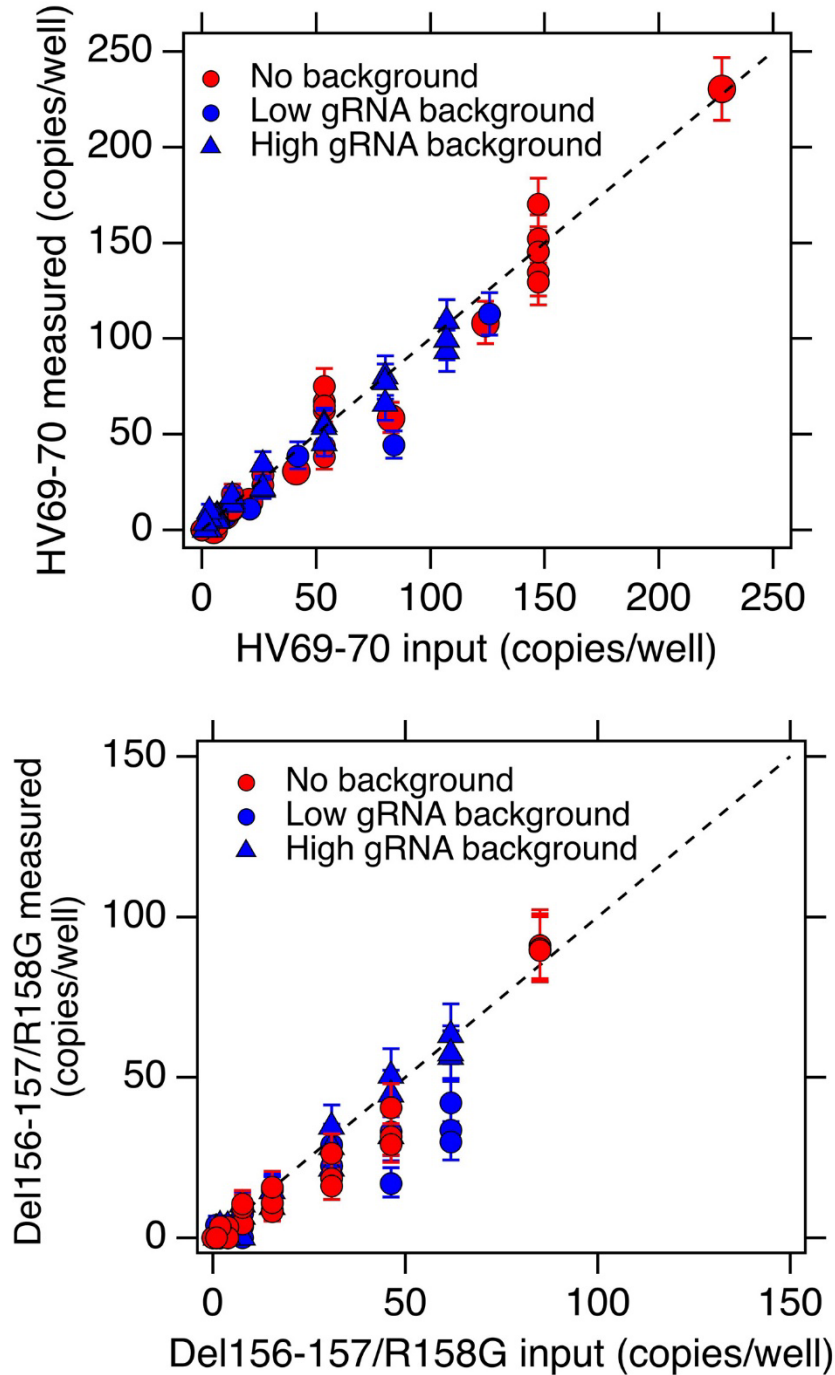
References

1. Haramoto E, Kitajima M, Kishida N, Konno Y, Katayama H, Asami M, et al. Occurrence of pepper mild mottle virus in drinking water sources in Japan. *Appl Environ Microbiol.* 2013;79:7413–8. [PubMed https://doi.org/10.1128/AEM.02354-13](https://doi.org/10.1128/AEM.02354-13)
2. Zhang T, Breitbart M, Lee WH, Run J-Q, Wei CL, Soh SWL, et al. RNA viral community in human feces: prevalence of plant pathogenic viruses. *PLoS Biol.* 2006;4:e3. [PubMed https://doi.org/10.1371/journal.pbio.0040003](https://doi.org/10.1371/journal.pbio.0040003)
3. Decaro N, Elia G, Campolo M, Desario C, Mari V, Radogna A, et al. Detection of bovine coronavirus using a TaqMan-based real-time RT-PCR assay. *J Virol Methods.* 2008;151:167–71. [PubMed https://doi.org/10.1016/j.jviromet.2008.05.016](https://doi.org/10.1016/j.jviromet.2008.05.016)
4. Wolfe MK, Topol A, Knudson A, Simpson A, White B, Vugia DJ, et al. High-frequency, high-throughput quantification of SARS-CoV-2 RNA in wastewater settled solids at eight publicly owned treatment works in northern California shows strong association with COVID-19 incidence. *mSystems.* 2021;6:e00829–21. [PubMed https://doi.org/10.1128/mSystems.00829-21](https://doi.org/10.1128/mSystems.00829-21)
5. Huggett JF; dMIQE Group. The Digital MIQE guidelines update: minimum information for publication of quantitative digital PCR experiments for 2020. *Clin Chem.* 2020;66:1012–29. [PubMed https://doi.org/10.1093/clinchem/hvaa125](https://doi.org/10.1093/clinchem/hvaa125)
6. Borchardt MA, Boehm AB, Salit M, Spencer SK, Wigginton KR, Noble RT. The environmental microbiology minimum information (EMMI) guidelines: qPCR and dPCR quality and reporting for environmental microbiology. *Environ Sci Technol.* 2021;55:10210–23. [PubMed https://doi.org/10.1021/acs.est.1c01767](https://doi.org/10.1021/acs.est.1c01767)
7. Wolfe MK, Archana A, Catoe D, Coffman MM, Dorevich S, Graham KE, et al. Scaling of SARS-CoV-2 RNA in settled solids from multiple wastewater treatment plants to compare incidence rates of laboratory-confirmed COVID-19 in their sewersheds. *Environ Sci Technol Lett.* 2021;8(5):393–404. <https://doi.org/10.1021/acs.estlett.1c00184>

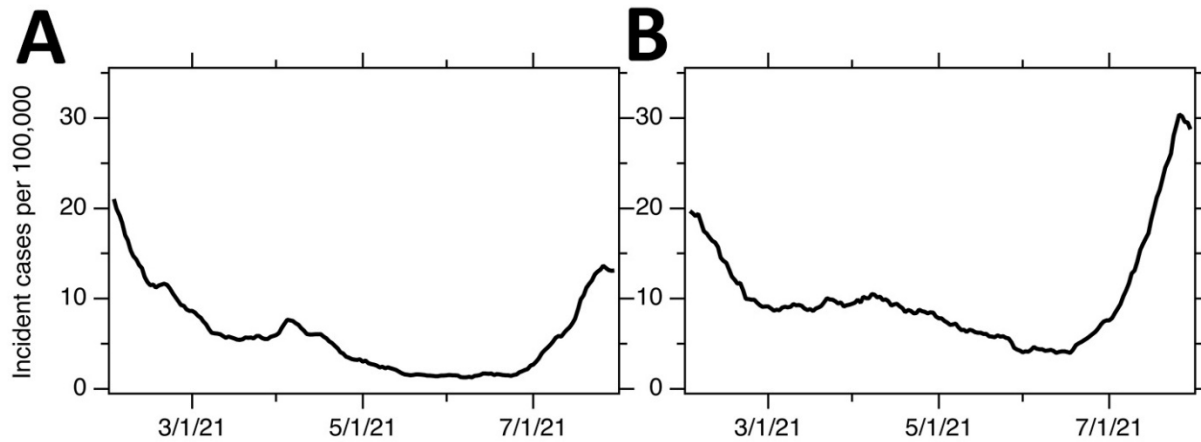
Appendix Table. Primer and probe sequences*

Target	Primer/Probe	Sequence
N Gene	Forward	CATTACGTTTGGTGGACCCT
	Reverse	CCTTGCCATGTTGAGTGAGA
	Probe	CGCGATCAAAACAACGTCGG (5' FAM/ZEN/3' IBFQ)
BCoV	Forward	CTGGAAGTTGGTGGAGTT
	Reverse	ATTATCGGCCTAACATACATC
	Probe	CCTTCATATCTATACACATCAAGTTGTT (5' FAM/ZEN/3' IBFQ)
PMMoV	Forward	GAGTGGTTTGACCTAACGTTTGA
	Reverse	TTGTCGGTTGCAATGCAAGT
	Probe	CCTACCGAAGCAAATG (5' HEX/ZEN/3' IBFQ)
HV69–70 del (Alpha)	Forward	ACTCAGGACTTGTCTTACCT
	Reverse	TGGTAGGACAGGGTTATCAAAC
	Probe	ATGCTATCTCTGGGACCAAT (5' FAM or HEX/ZEN/3' IBFQ)
Del156–157/R158G (Delta)	Forward	ATTCGAAGACCCAGTCCCTA
	Reverse	AGGTCCATAAGAAAAGGCTGA
	Probe	TGGATGGAAAGTGGAGTTTATTCTAG (5' FAM or HEX/ZEN/3' IBFQ)

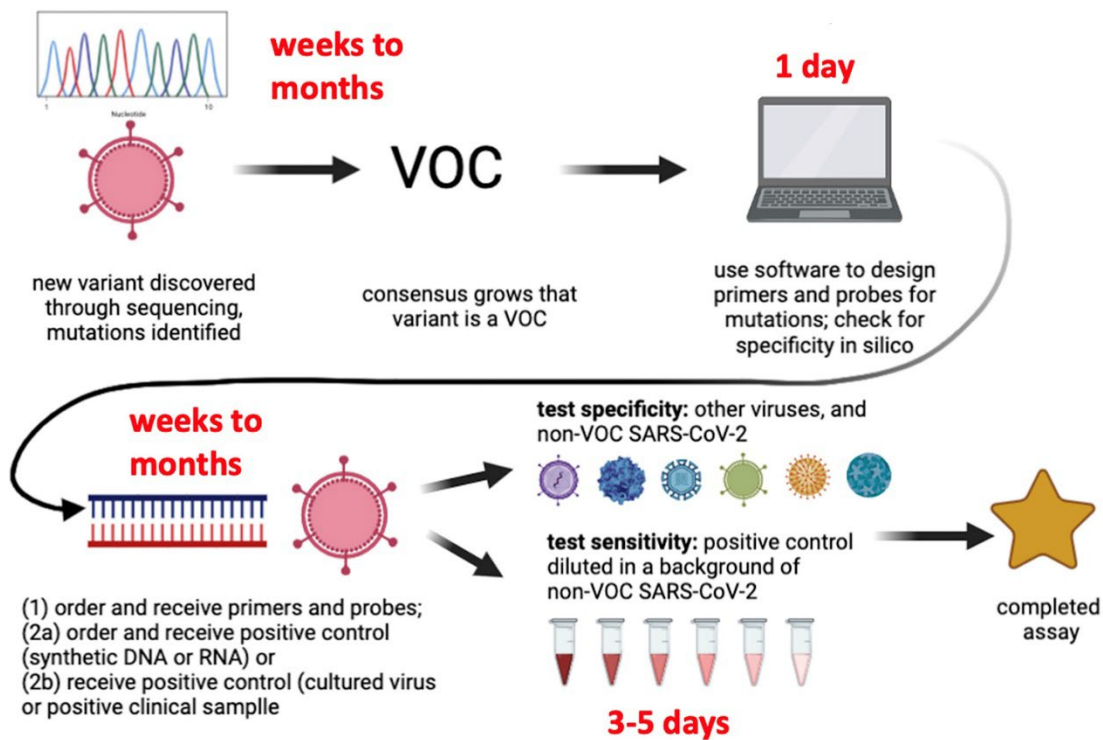
*The N gene assay is described in detail in Huisman et al. (J.S. Huisman et al., unpub. data, <https://doi.org/10.1101/2021.04.29.21255961>). The PMMoV assay is from Haramoto et al. (5) and Breitbart et al. (6). The BCoV assay is from Decaro et al. (7).



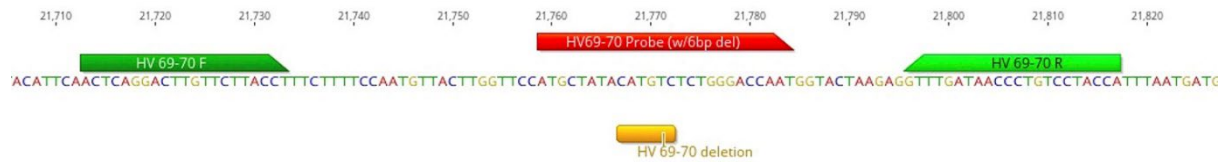
Appendix Figure 1. Measured copies of mutation targets versus input copies of mutation targets in no, low, (100 copies/well) and high (10,000 copies/well) background of WT-gRNA from 2019-nCoV/USA-WA1/2020 which does not contain the mutations. Error bars represent standard deviations of the measurements. No background is shown in red; and measurements where background gRNA was included are shown in blue. We were able to detect concentrations as low as 5 copies/well for the mutation assays when there was no, low, and high background of WT-gRNA.



Appendix Figure 2. Incident cases of COVID-19 (seven day smoothed) in the sewersheds.



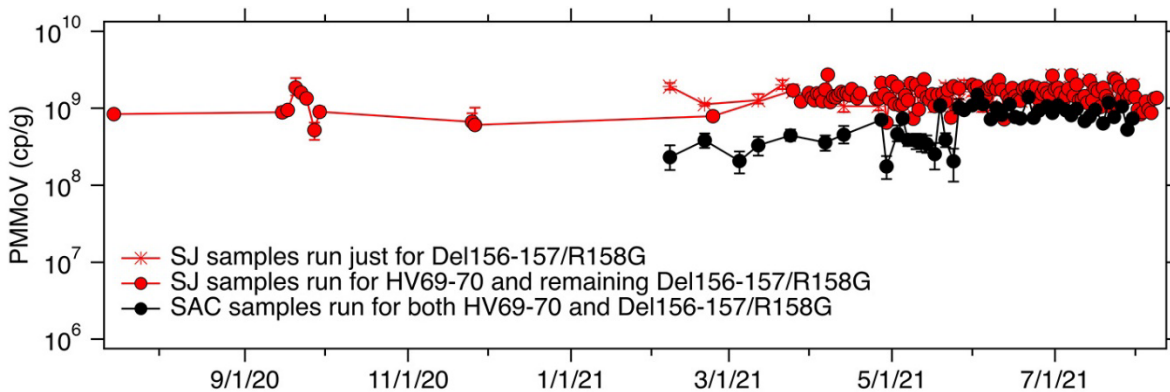
Appendix Figure 3. Timeline to develop completed variant mutation assay. The time for variant discovery to consensus that variant is a VOC and time for receiving reagents are obstacles for the rapid development of mutation assays. Figure was created using Biorender.com.



Appendix Figure 4. Location of HV69–70 primers and probes within the SARS-CoV-2 genome. Probe was designed with the deletion in the center (image shows the location, but the annotated probe has a 6bp deletion at amino acids 69 and 70).



Appendix Figure 5. Location of Del156–157/R158G primers and probes within the SARS-CoV-2 genome. Probe was designed with the deletion in the center.



Appendix Figure 6. PMMoV RNA target concentrations in wastewater solids used in this study in units of copies per gram dry weight. PMMoV is used as an endogenous control as we expect it in high concentrations and low concentrations would suggest failure in extraction efficiency.