

Evaluation of a Wastewater-Based Epidemiological Approach to Estimate the Prevalence of SARS-CoV-2 Infections and the Detection of Viral Variants in Disparate Oregon Communities at City and Neighborhood Scales

Blythe A. Layton,^{1,2} Devrim Kaya,¹ Christine Kelly,¹ Kenneth J. Williamson,² Dana Alegre,³ Silke M. Bachhuber,⁴ Peter G. Banwarth,⁵ Jeffrey W. Bethel,⁶ Katherine Carter,³ Benjamin D. Dalziel,^{4,7} Mark Dasenko,³ Matthew Geniza,³ Andrea George,^{1,2} Anne-Marie Girard,³ Roy Haggerty,⁸ Kathryn A. Higley,⁹ Denise M. Hynes,^{3,10,11} Jane Lubchenco,⁴ Katherine R. McLaughlin,¹² F. Javier Nieto,¹¹ Aslan Noakes,⁸ Matthew Peterson,³ Adriana D. Piemonti,² Justin L. Sanders,¹³ Brett M. Tyler,^{3,14} and Tyler S. Radniecki¹

¹School of Chemical, Biological, and Environmental Engineering, Oregon State University (OSU), Corvallis, Oregon, USA

²Department of Research and Innovation, Clean Water Services, Hillsboro, Oregon, USA

³Center for Quantitative Life Sciences, OSU, Corvallis, Oregon, USA

⁴Department of Integrative Biology, OSU, Corvallis, Oregon, USA

⁵Benton County Health Department, Corvallis, Oregon, USA

⁶School of Biological and Population Health Sciences, OSU, Corvallis, Oregon, USA

⁷Department of Mathematics, OSU, Corvallis, Oregon, USA

⁸College of Science, OSU, Corvallis, Oregon, USA

⁹School of Nuclear Science and Engineering, OSU, Corvallis, Oregon, USA

¹⁰U.S. Department of Veterans Affairs, Portland, Oregon, USA

¹¹College of Public Health and Human Sciences, OSU, Corvallis, Oregon, USA

¹²Department of Statistics, OSU, Corvallis, Oregon, USA

¹³Carlson College of Veterinary Medicine, OSU, Corvallis, Oregon, USA

¹⁴Department of Botany and Plant Pathology, OSU, Corvallis, Oregon, USA

BACKGROUND: Positive correlations have been reported between wastewater SARS-CoV-2 concentrations and a community's burden of infection, disease or both. However, previous studies mostly compared wastewater to clinical case counts or nonrepresentative convenience samples, limiting their quantitative potential.

OBJECTIVES: This study examined whether wastewater SARS-CoV-2 concentrations could provide better estimations for SARS-CoV-2 community prevalence than reported cases of COVID-19. In addition, this study tested whether wastewater-based epidemiology methods could identify neighborhood-level COVID-19 hotspots and SARS-CoV-2 variants.

METHODS: Community SARS-CoV-2 prevalence was estimated from eight randomized door-to-door nasal swab sampling events in six Oregon communities of disparate size, location, and demography over a 10-month period. Simultaneously, wastewater SARS-CoV-2 concentrations were quantified at each community's wastewater treatment plant and from 22 Newport, Oregon, neighborhoods. SARS-CoV-2 RNA was sequenced from all positive wastewater and nasal swab samples. Clinically reported case counts were obtained from the Oregon Health Authority.

RESULTS: Estimated community SARS-CoV-2 prevalence ranged from 8 to 1,687/10,000 persons. Community wastewater SARS-CoV-2 concentrations ranged from 2.9 to 5.1 log₁₀ gene copies per liter. Wastewater SARS-CoV-2 concentrations were more highly correlated (Pearson's $r=0.96$; $R^2=0.91$) with community prevalence than were clinically reported cases of COVID-19 (Pearson's $r=0.85$; $R^2=0.73$). Monte Carlo simulations indicated that wastewater SARS-CoV-2 concentrations were significantly better than clinically reported cases at estimating prevalence ($p<0.05$). In addition, wastewater analyses determined neighborhood-level COVID-19 hot spots and identified SARS-CoV-2 variants (B.1 and B.1.399) at the neighborhood and city scales.

DISCUSSION: The greater reliability of wastewater SARS-CoV-2 concentrations over clinically reported case counts was likely due to systematic biases that affect reported case counts, including variations in access to testing and underreporting of asymptomatic cases. With these advantages, combined with scalability and low costs, wastewater-based epidemiology can be a key component in public health surveillance of COVID-19 and other communicable infections. <https://doi.org/10.1289/EHP10289>

Introduction

Wastewater-based epidemiology (WBE) has emerged as an effective and sensitive approach for monitoring COVID-19 presence in a community through the detection of the novel coronavirus

(SARS-CoV-2) shed by infected individuals into wastewater.^{1–5} Although methods for COVID-19 WBE are still being refined, particularly with respect to optimizing sampling and virus concentration methods,^{6–9} this approach has shown promise, with wastewater SARS-CoV-2 concentration trends mimicking those of clinically reported COVID-19 cases.⁵ In addition, in some cases, SARS-CoV-2 had been detected in wastewater prior to the reporting of clinical COVID-19 cases.^{5,10} With the ability to non-invasively monitor an entire community with a single wastewater sample from a wastewater treatment plant (WWTP), WBE also has clear advantages in terms of cost (compared with traditional surveillance methods) and in areas where clinical testing is limited or residents are hesitant to participate.^{1–5}

What has remained elusive is the quantitative relationship between viral concentrations in wastewater and community infection rates, as well as the representativeness of community viral genotype profiles from wastewater sequencing. These limitations are due to the biological variability of SARS-CoV-2 infections, physical uncertainties of wastewater sampling, and inherent

Address correspondence to Tyler S. Radniecki, School of Chemical, Biological, and Environmental Engineering, Oregon State University, 116 Johnson Hall, 105 SW 26th St., Corvallis, OR 97331 USA. Email: tyler.radniecki@oregonstate.edu

Supplemental Material is available online (<https://doi.org/10.1289/EHP10289>).

The authors declare they have nothing to disclose.

Received 9 September 2021; Revised 24 May 2022; Accepted 31 May 2022; Published 29 June 2022.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehpsubmissions@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

variability in case reporting. Biological variability encompasses latent variation in the magnitude and duration of viral shedding in both symptomatic and asymptomatic infected individuals.^{11–13} For instance, asymptomatic individuals have similar viral loads as symptomatic individuals but tend to have shorter durations of viral shedding into fecal material.^{14,15}

Physical uncertainties include representativeness of the wastewater samples, virus concentration and extraction methodologies, molecular detection methods and polymerase chain reaction (PCR) inhibition as well as RNA persistence and decay rates in sewage conveyance systems.^{6,8,16–18} In addition, it is uncertain how the integrity of the SARS-CoV-2 virome affects PCR and sequencing efforts, with both intact and degraded SARS-CoV-2 envelopes and RNA being present in the wastewater.^{19–21} Finally, there is inherent variability in case reporting results from underreporting of infections owing to limited testing capacity, barriers in access to testing, testing avoidance, self-isolation of individuals with mild symptoms, and widespread asymptomatic transmission of the virus.^{22,23}

In this paper, the SARS-CoV-2 burdens on communities of diverse size, location, climates, and demography were examined through four lines of evidence: wastewater SARS-CoV-2 concentrations, prevalence data estimated via testing nasals swabs of residents participating in random door-to-door sampling events,²⁴ clinically reported COVID-19 cases (which was obtained from the Oregon Health Authority and included cases not identified through the door-to-door sampling events), and sequence data obtained from nasal swabs and wastewater samples. From these data sets, the accuracy of wastewater SARS-CoV-2 concentrations and reported COVID-19 cases were compared with the estimated prevalence of SARS-CoV-2 infection (including both symptomatic and asymptomatic individuals). In addition, the ability of wastewater SARS-CoV-2 concentrations to identify neighborhood-level COVID-19 hotspots, as well as the community SARS-CoV-2 variant profile, was determined.

Methods

Community Nasal Swab Sampling and Prevalence Estimates

The prevalence of SARS-CoV-2 in six Oregon communities (some at multiple time points) was estimated using a two-stage cluster sampling scheme as a part of Oregon State University's Team-based Rapid Assessment of community-level Coronavirus Epidemics (TRACE) project.²⁴ In the first stage, 30 clusters (each comprising one or more census blocks) were randomly selected in each community, with probabilities proportional to the number of housing units (i.e., clusters with a large number of occupied housing units had a higher probability of being selected). Each cluster contained a minimum of 50 occupied housing units. The TRACE project received approval from Oregon State University's institutional review board to conduct this work.

Starting in 2021 with the last two community sampling events (in Redmond and Corvallis), results from community wastewater testing were used to inform the first stage of sampling. These samples represent a refinement of the TRACE methodology based on evidence that estimates of SARS-CoV-2 concentration in wastewater correlated well at the micro-sewershed level to prevalence within the community. For these two samples, clusters were first grouped into approximately five strata based on micro-sewershed boundaries. Then wastewater data collected ~2–3 wk prior to household sampling were used to allocate the 30 sampled clusters to the strata, using optimal allocation to minimize the anticipated standard error of the prevalence estimate. This standard approach in sampling methodology facilitated oversampling from higher prevalence areas in a principled design-based

approach and still allowed for approximately unbiased estimates of prevalence to be obtained. With this approach, a total of 30 clusters were still selected; however, a set number of them were required to be from each stratum based on wastewater data. Within each stratum, the clusters selected for door-to-door sampling were still chosen with probability proportional to the number of occupied housing units in the cluster.

In the second stage, field teams used systematic sampling with a random start (i.e., the housing unit at the northwest corner of the cluster) to identify 12 housing units within each of the 30 clusters. The sampling interval (k) used for this stage was specific to each cluster and was calculated by dividing the total number of housing units in the cluster by 12. Field teams were provided with the sampling interval and starting location in advance of data collection. Teams proceeded through their assigned cluster(s) in a serpentine fashion, adhering to the sampling interval. Teams revisited housing units in which no one was home two additional times before the housing unit was replaced. Housing units in which no one agreed to participate were also replaced.

All residents of the selected households, regardless of age, were invited to participate. After obtaining informed consent, field teams interviewed each participant to collect information including name, date of birth, contact information, symptoms, previous positive test result(s), COVID-19 vaccination status, and demographics. Field teams provided test kits to participants and assisted them in self-collecting a nasal swab specimen. Nasal swab specimens were transported to the laboratory for processing and analysis. Test results were mailed to all participants and available to participants via a secure website if they opted for this method. REDCap software was used to securely collect and track data participation and disseminate results of the COVID-19 tests to the participants.²⁵

In total, the TRACE project collected 4,136 self-administered nasal swab samples from residents in 2,521 randomly selected Oregon households in Bend, Corvallis, Eugene, Hermiston, Newport, and Redmond from 30 May 2020 to 14 March 2021. An average of 517 ± 78 individuals from 315 ± 43 households ($60\% \pm 14\%$ average household participation rate) participated in each random door-to-door nasal swab sampling event. The prevalence of SARS-CoV-2 within the community at the time of sampling was estimated using one of two approaches based on the number of positive cases detected.

When at least one positive case was detected, a design-weighted approach accounting for the corresponding multistage sampling design (with or without wastewater strata) was used. In this approach, the sampling unit is the housing unit, the sampling weight depends on factors of the design, and the response is the observed positivity among participants in that housing unit. For example, in the original approach without incorporating the wastewater information, the sampling weight of housing unit j in cluster i is shown in Equation 1 where M_0 is the total number of housing units in the population; M_i is the total number of housing units in cluster i ; m_i is the number of housing units sampled from cluster i (ideally 12); and h_{ij} is the number of participants from housing unit j in cluster i .

$$w_{ij} = \frac{M_0 M_i}{M_i m_i} h_{ij} = \frac{M_0}{m_i} h_{ij}. \quad (1)$$

These weights are used to estimate prevalence using a standard Horvitz–Thompson estimator.²⁶

When zero positive cases were identified through door-to-door sampling in the whole community, a Bayesian approach was taken that combined the observed data with active case counts within the community at the time of sampling. In particular, a beta-binomial model was used where the count of positive

individuals was assumed to follow a binomial (n, p) distribution, and the prior for p is assumed to be beta (α, β). Values for α and β were obtained via optimization by setting the fifth quantile of the beta distribution equal to the number of cases reported to the Oregon Health Authority in the previous 2 wk divided by the population size, and the 50th quantile equal to the fifth quantile divided by the proportion of cases thought to be symptomatic based on previous data.

A noninformative (uniform [0,1]) prior was also explored, and the approach was not sensitive to the choice of prior. The posterior distribution for p also takes the form of a beta and can be used to estimate prevalence. Imperfect test sensitivity and specificity were accounted for in both approaches. Prevalence estimates were quickly shared with the local public health officials partnering with the TRACE project.

Clinically Reported COVID-19 Cases

Weekly COVID-19 case data for the participating communities were obtained from ZIP code reported data from the Oregon Health Authority website.²⁷ For weeks when <10 cases occurred in a given ZIP code, the Oregon Health Authority reported the case count as “1–9 cases.” In these instances, to minimize error, a value of 5 cases was substituted in the calculations. The ZIP codes included for each community were as follows: Bend (97701, 97702, 97703, 97707, 97708), Corvallis (97330, 97331, 97333), Eugene (97401, 97402, 97403, 97404, 97405, 97447, 97478), Hermiston (97838), Newport (97365, 97366), and Redmond (97756). Reported test results with location data for the city of Newport were obtained with permission from the Lincoln County Health and Human Services Department, and public health officials anonymized the data by aggregating according to micro-sewershed prior to analysis.

Comparison of Wastewater Concentration Methods

Several wastewater concentration methods were tested to determine which method resulted in the highest viral recovery, as estimated by the maximum mean SARS-CoV-2 signal recovered from replicate composite WWTP samples. The WWTP sample was collected from Newport, Oregon, during the June 2020 sampling event. Two concentration methods, electronegative membrane filtration (EMF) and centrifugal ultrafiltration (CU) were evaluated. In addition, four iterations of EMF with various amendments were tested (EMF Methods A–D). All methods were conducted in triplicate.

For the EMF method, 30 mL of WWTP influent composite sample was mixed gently on a stir plate with a magnetic stir bar. Each of the EMF method iterations differed in the second step, when amendments were made to the sample. In EMF Method A, 0.01 N hydrochloric acid (HCl) was added to titrate the pH below 3.5 and magnesium chloride (MgCl_2) was added to make a final concentration of 25 mM. In EMF Method B, 0.01 N HCl was

used to titrate the pH below 3.5 but no MgCl_2 was added. In EMF Method C, MgCl_2 was added for a final concentration of 25 mM, but the sample was not acidified. In EMF Method D, no additions were made.

All samples were gently mixed for an additional 5 min. After mixing, the entire sample was vacuum filtered through a 0.45- μm pore size, 47-mm diameter, mixed cellulose ester electronegative membrane (Catalog no. 7141-104; Whatman). The filter was stabilized in 1 mL of DNA/RNA Shield (Zymo Research, Irvine, CA) and stored at -20°C until RNA extraction and reverse transcriptase droplet digital PCR (RT-ddPCR) could occur.

In the CU method, 40 mL of wastewater was added to a Centricon Plus-70 centrifugal ultrafiltration device (Catalog no. UFC701008; Merck Millipore) with a 10-kDa cutoff. The ultrafiltration device was centrifuged for 20 min at $3,220 \times g$. The filtrate was discarded, and the concentrate spun out of the ultrafilter back into the concentrate cup at $1,000 \times g$ for 2 min. The concentrate was removed and placed in 1 mL of DNA/RNA Shield.

WWTP Sampling

All WWTP influent samples comprised 24-h time-weighted composites taken prior to primary treatment. The characteristics of each WWTP are given in Table 1. Samples were collected at the time of each random door-to-door nasal swab sampling event, as well as an additional one to three times per week for 6–11 months from April 2020–May 2021.

Twenty-two pump stations serving Newport, Oregon’s Vance Avery WWTP sewershed, were sampled hourly for 24 h from 20–21 June and 11–12 July 2020 (Figure 1). Some pump stations were sampled twice within a single weekend, for a total of 52 pump station samples. The total population of the Newport sewershed was 10,853, with the populations served by each micro-sewershed ranging from 15 to 9,426 persons. The characteristics of each micro-sewershed and the distribution of samples are given in Table 2.

For both WWTPs and pump stations, the 24-h composites consisted of hourly samples and were kept on ice during sampling. Samples collected on or before 31 July 2020 were frozen in 200-mL aliquots and stored at -80°C for up to 33 d prior to concentration (median 6 d). After thawing in cold water baths, samples were concentrated using electronegative filtration, as previously described.²⁸ Briefly, all samples collected on or before 10 July 2020 were acidified to a final pH of 3.5 and MgCl_2 was added to a final concentration of 25 mM. Samples (30–40 mL) were vacuum filtered through a 0.45- μm pore size, 47-mm diameter mixed cellulose ester electronegative filter (HAWP; Millipore).

Influent samples collected after 10 July 2020 were filtered with no amendments given that preliminary data showed improved viral recovery with unamended wastewater (Figure S1). In addition, influent samples collected after 31 July 2020 were neither frozen nor amended prior to filtration. In these samples, filtration occurred within 8 h of sample collection.

Table 1. Oregon city and wastewater treatment plant statistics: the location, population, and average wastewater flow rate of the communities participating in this study.

City	County	Population served (n)	Average daily wastewater flow (MGD)	Sewer system type
Bend	Deschutes	100,421	6.2 ± 0.2	Separated
Corvallis	Benton	58,856	9.6 ± 5.5	Combined ^a
Eugene	Lane	258,910	29.8 ± 11.9	Separated
Hermiston	Umatilla	17,782	1.2 ± 0.0	Separated
Newport	Lincoln	10,853	1.7 ± 0.6	Separated
Redmond	Deschutes	32,421	2.4 ± 0.1	Separated

Note: Collection type for all communities was 24-h composites. Sample matrix for all communities was wastewater influent, which refers to wastewater as it is entering a wastewater treatment plant. MGD, millions of gallons per day.

^a3.5 mi² are combined wastewater and stormwater out of an 8.6-mi² sewershed.

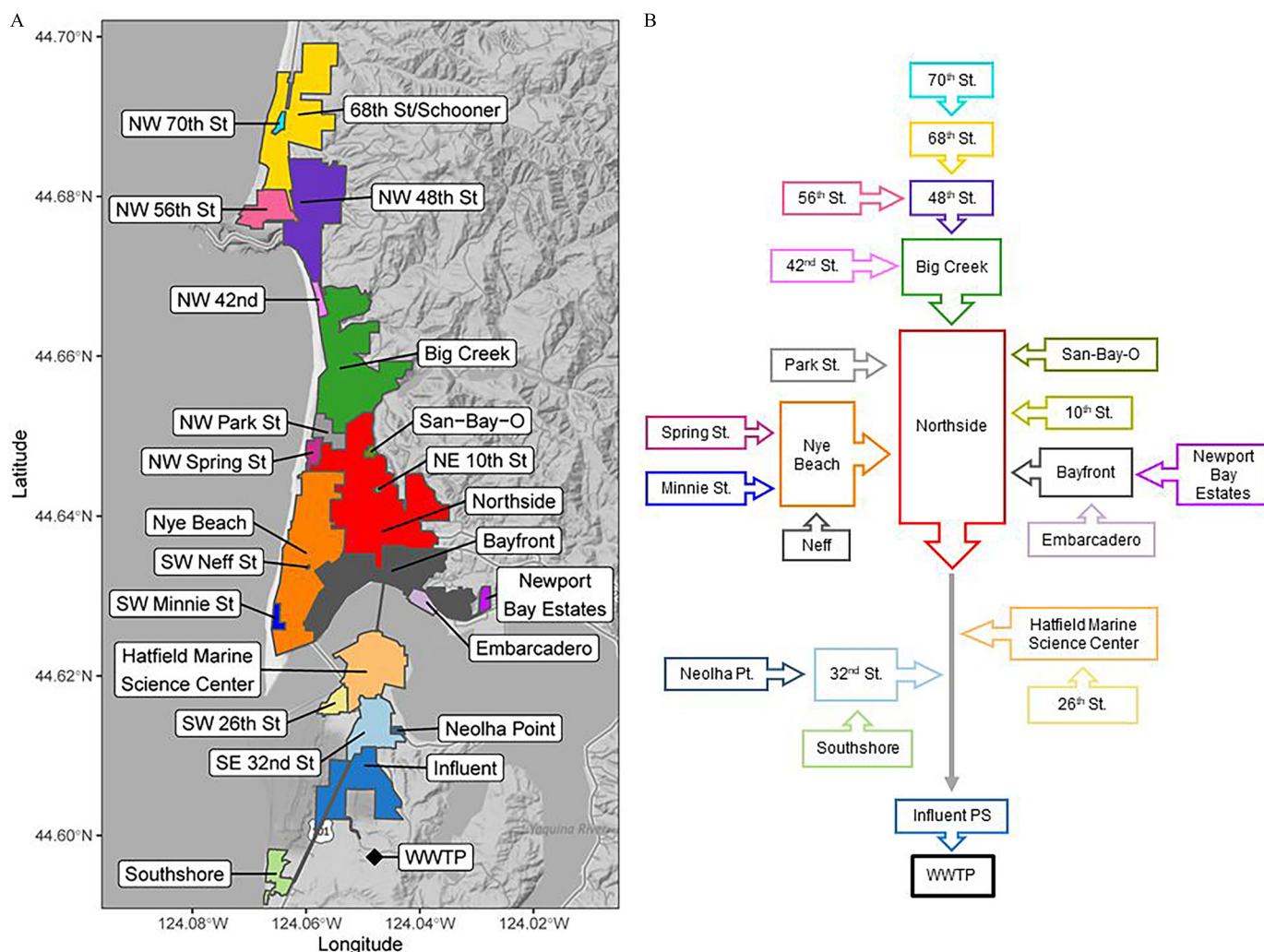


Figure 1. Newport, Oregon, micro-sewersheds. (A) Location and name of the 22 pump stations and their associated micro-sewersheds sampled in Newport. (B) Flow chart depicting the hierarchical relationships between pump stations. The arrow from Northside to Influent PS represents a forced main running under the Yaquina Bay and southward toward the WWTP. Map tiles by Stamen Design, under CC BY 3.0. Basemap data by OpenStreetMap, under ODbL. Note: PS, pump station; WWTP, wastewater treatment plant.

Once filtration was complete, the electronegative membranes were placed into 2-mL tubes containing either 0.7-mm garnet or 0.5-mm glass beads, stabilized in 1 mL of DNA/RNA Shield (Zymo Research), and frozen until analysis. Preliminary experiments showed no difference between these homogenization methods (Figure S2). Field blanks of deionized water were processed with every batch of samples.

Molecular Analysis: Nasal Swabs

Participant nasal swab samples were analyzed at the Oregon Veterinary Diagnostic Laboratory using the TaqPath COVID-19 Combo Kit (Applied Biosystems), in accordance with the manufacturer's instructions for use as required by the Emergency Use Authorization under strict biosafety level 2 (BSL2) conditions. Nucleic acid isolation was performed using the MagMax Viral/Pathogen II Nucleic Acid Isolation Kit (Applied Biosystems/ThermoFisher). Briefly, 200 μ L of transport medium from the swab sample was added to a single well of a KingFisher Deepwell 96-well plate containing 5 μ L of Proteinase K. Each 96-well plate held 94 participant samples and 1 negative control well containing water. The last well was left empty to allow for the positive control to be added during the (RT)-PCR detection

step. After sample addition, nucleic acid magnetic beads were resuspended, 10 μ L was added to 265 μ L of binding solution and then added to the wells. MS2 phage control (5 μ L) was added to all wells as an extraction control and processed on a KingFisher Flex magnetic particle processor. Purified nucleic acids were eluted in 50 μ L of MagMax elution solution. Eluted nucleic acid was stored at -80°C unless RT-PCR was run within 2 h of extraction.

Detection of SARS-CoV-2 viral RNA was performed using the TaqPath RT-PCR COVID-19 Kit on a 7500 Fast Real-Time PCR Instrument (Applied Biosystems). Reactions were run in multiplex with primers and probes specific for three gene sequences specific to SARS-CoV-2: ORF1ab, N Protein, and S Protein. A primer and probe set was included to detect the MS2 phage added during the initial sample processing as an internal control to verify RNA extraction. The thermal protocol for the one-step real-time PCR was as follows: 2 min at 25°C uracil *N*-glycosylase incubation, 10 min at 53°C reverse transcription, 2 min at 95°C activation, followed by 40 cycles of 3 s at 95°C denaturation and 30 s at 60°C anneal/extension/detection. Results were analyzed using SDS Software (version 1.5.1) and interpreted using COVID-19 Interpretive Software (version 1.2; Applied Biosystems).

Table 2. Newport, Oregon, micro-sewershed characteristics: the pump station name, size (area and population) and the percentage of the area that is residential for the 22 Newport micro-sewersheds (neighborhoods) sampled in this study.

Pump station	Drainage area (acres)	Percentage residential (%)	Estimated population (n)	Sample dates
10th St.	1.45	100	15	18 June 2020; 8 July 2020
26th St.	27.02	88	131	19 June 2020; 9 July 2020
32nd St.	157.54	51	391	19 June 2020; 9 July 2020
42nd St.	13.04	100	119	19 June 2020; 8 July 2020
48th St.	593.47	75	1,771	19 June 2020; 8 July 2020
56th St.	68.49	93	611	19 June 2020; 8 July 2020
68th St./ Schooner	297.56	74	7,091	19 June 2020; 8 July 2020
70th St.	6.97	100	24	19 June 2020; 8 July 2020
Bayfront	330.41	67	1,456	19 June 2020; 8 July 2020; 9 July 2020
Big Creek	952.72	66	3,569	19 June 2020; 8 July 2020; 9 July 2020
Embarcadero	34.28	33	470	18 June 2020; 8 July 2020
Hatfield	196.93	12	222	18 June 2020; 9 July 2020
Minnie St.	10.68	100	53	18 June 2020; 8 July 2020
Neff	0.4	100	12	19 June 2020; 9 July 2020
Neolha Point	4.39	100	128	18 June 2020; 9 July 2020
Northside	2163.57	65	9,426	18 June 2020; 8 July 2020; 9 July 2020
Nye Beach	381.78	52	3,023	18 June 2020; 19 June 2020; 8 July 2020; 9 July 2020
Park St.	33.22	86	742	19 June 2020; 8 July 2020; 9 July 2020
Running Springs/Newport Bay Estates	11.46	100	30	18 June 2020; 9 July 2020
San-Bay-O	5.66	28	18	18 June 2020; 8 July 2020; 9 July 2020
Southshore	40.88	100	603	18 June 2020; 9 July 2020
NW Spring St.	18.47	100	402	19 June 2020; 8 July 2020

Note: Collection type for all communities was 24-h composites. Sample matrix for all communities was wastewater conveyance, which refers to wastewater collected in the sewer lines at either pump stations or manhole openings.

Molecular Analysis: Wastewater

Wastewater samples were homogenized with either 0.7-mm garnet beads or 0.5-mm glass in DNA/RNA Shield using either a Qiagen TissueLyser (Qiagen Inc.) or a BioSpec Mini-Beadbeater 16 (BioSpec Products, Inc.) for 2 min. Beads and debris were pelleted by centrifugation at 12,000 rcf for 1 min. The lysate was transferred from each tube to a 96-well plate, and 200–400 μ L was extracted using the MagMAX Viral/Pathogen kit on a KingFisher Flex automated instrument (ThermoFisher Scientific), as described above. Purified RNA was eluted in 50 μ L of MagMAX Elution Solution. Extraction recovery from RNA extraction step was quantified with a commercial standard (Exact Diagnostics), and extraction blanks were included with every run. RNA was stored at -80°C until analysis.

Two SARS-CoV-2 RNA targets (N1/N2) and an internal control (Human RNase P) were measured via RT-ddPCR using a commercial triplex assay (2019-nCoV CDC ddPCR Triplex Probe Assay; Catalog no. 12008202; Bio-Rad) using the One-Step RT-ddPCR Advanced Kit for Probes and run on a QX-200 ddPCR system (Bio-Rad). The primer and probe sequences were published previously.²⁹ RT-ddPCR was chosen over RT-qPCR owing to its superior sensitivity,^{30,31} robustness against inhibitors,³² and its wide application in WBE methodologies.^{17,33,34}

Reactions were partitioned into droplets using an automated droplet generator (ADG). Twenty-two microliter reactions were prepared with 5.5- μ L template RNA, whereas the ADG partitioned only 20 μ L, yielding an effective template volume of 5 μ L. Each reaction had an average of 12,657 droplets [standard deviation (SD) = 1,783]. Commercially prepared RNA standards and negative controls were included on each extraction plate and ddPCR plate (cat. no. COV019 and COV000; Exact Diagnostics). All samples and controls were analyzed in duplicate.

The one-step thermal cycling conditions were as follows: reverse transcription at 50°C for 60 min; enzyme activation at 95°C for 10 min; 40 cycles of denaturation at 94°C for 30 s followed by annealing/extension at 55°C for 60 s; enzyme inactivation at 98°C for 10 min; and, last, a 4°C hold for droplet stabilization, for a minimum of 30 min to a maximum of

overnight. Finally, the amplification in the droplets was determined using the Bio-Rad droplet reader. All assay conditions were performed as specified in the Bio-Rad assay protocol.³⁵

RT-ddPCR Quality Control

The quality controls for the wastewater RT-ddPCR method included field blanks, extraction blanks, negative control reactions [containing only human genomic DNA (gDNA)], positive control reactions (containing synthetic RNA of SARS-CoV-2 assay targets and human gDNA), and no-template controls (NTCs). The results of all quality control (QC) reactions are summarized in Table S1. Only positive detections which passed quality assurance/QC were included in this study.

RT-ddPCR Data Analysis

Wells with <6,000 droplets were omitted. Sample data with positive reactions were accepted only if the corresponding extraction blank and field blanks, as well as the PCR negative and no-template controls, were all negative for the N1/N2 targets. When averaging sample data across replicates, a value of one-half the sample-specific limit of detection was substituted for nondetects. Reactions were regarded as positive if three or more droplets per well amplified in either target. Droplet clusters were manually called for each target using the QuantaSoft Analysis Pro software (version 1.4; Bio-Rad). All other analyses were conducted in R (version 4.1.0; R Development Core Team) with Rstudio Desktop (version 1.4.1717).³⁶ Spatial graphics were created using the *sf*,³⁷ *ggplot2*,³⁸ *gtable*,³⁹ and *ggmap*⁴⁰ packages.

The N1 and N2 markers exhibited generally good agreement in the wastewater samples. The markers were concordant in 91.4% of reactions: 4.1% of reactions ($n=954$) had a positive detection in N1 only (using a threshold of three positive droplets per reaction), and 4.5% of reactions were positive in N2 only. Quantitatively, the markers were also reasonably well aligned according to a simple linear model [slope = 0.94, with N2 as the response variable, adjusted $r^2 = 0.97$, root mean square error (RMSE) = 0.58 copies

per reaction]. Accordingly, the N1 and N2 data were averaged together for all reported concentration values.

For limit of blank (LOB) determination, 104 reactions were run with the triplex assay across three plates using the Exact Diagnostics Negative Control as a template. Due to the nonnormal distribution of the results, the LOB was determined using a nonparametric (rank order) method with a false-positive probability (α) of 0.05.

The limit of detection (LOD) was predicted for each target (N1 and N2) using Equation 2, as follows:

$$LOD_{predicted} = LOB + 2 \times stDev(CopiesPerReaction_{LOB}), \quad (2)$$

where $stDev(CopiesPerReaction_{LOB})$ is the standard deviation (Dev) of the copies per reaction from the LOB assay. The predicted LOD was subsequently tested by running 60 test reactions at concentrations of 4 and 12 copies per reaction of target using the Exact Diagnostics Standard for SARS-CoV-2.

Bovine Coronavirus Process Recovery Control

To determine the viral recovery efficiency of the wastewater processing method used in this study, a spike-in experiment was performed using bovine coronavirus (BCoV) as a surrogate for SARS-CoV-2. BCoV solution was prepared from freeze-dried Calf Guard cattle vaccine (Zoetis). After rehydrating in 3 mL of sterile diluent provided by the manufacturer, the BCoV solution was divided into 100- μ L aliquots and stored at -20°C . To use, the BCoV stock solution aliquot was thawed on ice and vortexed thoroughly; each aliquot was used for a maximum of two freeze-thaw cycles. BCoV stock solution was spiked into wastewater samples at a ratio of 1/1,000 (vol/vol) before the concentration step.

To determine the concentration of the BCoV stock solution, 10 μ L of the stock was spiked into 390 μ L of phosphate-buffered saline (PBS), and 200 μ L of that mixture was extracted following the same protocols used for wastewater sample extractions as described in the main text. One extraction blank, prepared with PBS, was included on each plate as an RNA extraction contamination control. The extracted RNA was then serially diluted (1:10) in nuclease-free water for six dilutions and ran in duplicate using a previously established BCoV assay, following the one-step RT-ddPCR procedure.⁴¹ Stock concentration of BCoV was around 230,000 gene copies (gc)/ μ L.

Process efficiency (i.e., viral recovery) was calculated by dividing the final quantity of BCoV measured in wastewater samples by the quantity of BCoV spiked to each wastewater sample before concentration (Equation 3). Nonspiked wastewater samples were also quantified for BCoV to assess background concentrations.

$$\begin{aligned} \frac{gc}{L} = \frac{copies}{rxn} \times 1 \frac{rxn}{Template Volume (\mu L)} \times Elution Volume (\mu L) \\ \times \frac{1}{Lysate Volume (\mu L)} \times Shield Volume (\mu L) \\ \times \frac{1}{Volume Filtered (mL)} \times 1,000 \frac{mL}{L} \end{aligned} \quad (3)$$

cDNA Library Preparation and SARS-CoV-2 Sequencing

For cDNA synthesis, 11 μ L of lysate from positive wastewater or participant samples, were used for single-strand cDNA synthesis, using the Thermo Superscript IV kit with the following modifications: No host gDNA/RNA removal steps were performed, no RNase H step was performed, and the reverse transcriptase incubation step at 50°C was increased from 10 min to 30 min.

The cDNA was used for amplification and sequencing, using the Swift Amplicon SARS-CoV-2 Panel (AL-COV48) together with Swift Amplicon Combinatorial Dual Indexed Adapters (AL-S1A96, AL-S1BA96, AL-S2A96, AL-S2B96). The Swift Biosciences protocol was followed, except that, after optimization experiments, the volume of the G1 reagent was reduced to 25% of recommended level. This resulted in an increase in read coverage for wastewater samples of 2- to 8-fold for the experiments described in this study (with the release of version 2 of the Swift primer set, the volume of this reagent was no longer reduced). Individual libraries (30–96 samples) were quantified by fluorescence, normalized, and pooled for 2×150 bp sequencing on a lane of an Illumina HiSeq3000 sequencer.

Except for some initial sequencing experiments, libraries from individuals were prepared on different days and sequenced in different lanes than samples from wastewater to reduce the possibility of contamination of sequences from low titer wastewater samples with sequences from high titer individual samples. Pooled libraries from wastewater samples were often run on two to three lanes of the Hi-Seq 3,000 to increase read depth.

Wastewater sequencing on the HiSeq3000 produced a median percentage of reads mapped of 0.86% compared with a median of 21.7% for nasal swab samples. In general, wastewater samples with an RNA concentration of $\log_{10} > 4.0$ gene copies of N1/N2 per liter of wastewater (gc/L) reliably produced usable amounts of sequence. Samples with lower RNA concentrations produced variable amounts of sequence, and those $< \log_{10} 3.0$ gc/L were routinely unsuccessful.

Multi-Locus Sequence Typing

After demultiplexing, Illumina primer sequences were trimmed using BBduk (v.38.96), sequences were aligned to the reference sequence (NC_045512.2; Wuhan-Hu-1) using BWA-MEM (v.0.7.17-r1188), and the SARS-Cov-2 primers were removed using the Swift Biosciences *Primerclip* package. Genome Analysis Toolkit software (GATK) (v.4.2.0.0) was used to identify variants compared with the reference sequence and to count the numbers of reference and variants reads at each single nucleotide polymorphism (SNP) site. The ploidy was set to 4 and the down-sampling function was not employed. By not using the down-sampling function, all sequence reads that passed standard QC parameters were used in making variant calls. When required, an Integrated Genomics Viewer (IGV) was used to manually inspect sequence alignments and variant calls.

To identify viral genotypes represented in wastewater RNA by multi-locus sequencing typing, sequences from individual samples from Oregon were used to define multi-locus genotypes as a set of polymorphic sites that were unique to each variant and were not shared with any other variant known at the time (Table S2).

To estimate the fraction of each variant present in the wastewater RNA sequences, the number of reads supporting each variant were summed across all variant-specific SNP sites and divided by the total number of reads spanning those SNP sites. At SNP sites where the total number of reads was > 100 , both the variant read number and total read number were scaled down to a total read number of 100 prior to the estimation calculation. At SNP sites with > 100 total reads, the actual number of reads was used, thus decreasing the weight of those counts proportionally to the read coverage.

After this estimation was conducted for each variant, the fractions were summed for all identified variants. If the sum was > 0.7 , it was assumed that the identified variants comprised all of the RNA molecules present and each fraction was divided by the total to normalize the total to 1.0. If the initial sum was < 0.7 , it was assumed that the difference was comprised of RNA molecules

from unidentified variants, and the fraction of RNA attributable to the unidentified variants was set to the difference between the total and 1.0. If the sum was >1.5 , the SNP data were manually reviewed to identify and remove artifacts (miscalled 1-bp indels) and remove common SNPs that were shared between the variants present to prevent double-counting. The custom software to conduct these calculations was implemented in R and packaged into a *SnakeMake* pipeline.⁴² All individual sequences were deposited in the data sharing platform Global Initiative on Sharing All Influenza Data (GISAID; see Table S3 for accession numbers) and all wastewater sequences were deposited in NCBI's short read archive, under BioProject PRJNA719837.^{43,44}

Monte Carlo Simulations

To compare the accuracy of SARS-CoV-2 wastewater concentrations and clinically reported COVID-19 cases as estimators of community prevalence, Monte Carlo simulations were performed to account for the uncertainty in the point estimates for each sampling event. Clinically reported COVID-19 cases included cases identified through the TRACE random door-to-door sampling events as well as cases reported through standard health surveillance efforts by the Oregon Health Authority. The Monte Carlo simulations accounted for the uncertainty in the point estimates for each sampling event with the underlying assumption that the community prevalence based off the TRACE random door-to-door sampling events were the closest to the ground truth.

For each simulation, a new wastewater concentration ($\log_{10}$ gc/L) was redrawn for each community sample from a Gaussian distribution with mean equal to the point estimate and standard deviation equal to the standard error of the point estimate. Similarly, a new prevalence was redrawn for each community from a method-specific distribution: a truncated Gaussian distribution with mean equal to the point estimate, standard deviation equal to the standard error of the point estimate, and lower truncation bound equal to zero when the design-based estimator was used; or a Beta distribution with shape and scale parameters estimated using optimization to fit the posterior mean and 95% credible interval when the Bayesian estimator was used.⁴⁵ The reported COVID-19 case numbers were not redrawn because these were the publicly available standard. For each simulation, a simple linear regression model was fit using the new wastewater concentration draw to estimate the $\log_{10}$ of the new community prevalence draw, and a separate model was fit using the $\log_{10}$ of the observed COVID-19 case count to estimate the $\log_{10}$ of the new community prevalence draw. For each model, the slope, intercept, and R^2 value were recorded. In total, 10,000 simulations were performed.

Results

LOB, LOD, and BCoV Process Recovery

The LOB for N1 and N2 was 2.0 and 4.2 copies per reaction, respectively. Both of these values were below the three-droplet threshold; only reactions with one or two droplets yielded copy numbers at that level, which provided confidence in the chosen threshold for calling positive reactions. All LOB reactions were below the positive threshold for N1, and only 4/104 nontarget reactions had three or more droplets in N2, which was a false-positive rate of 4%.

The predicted LOD based on the LOB results were 4 and 12 copies per reaction for N1 and N2, respectively. For an LOD estimate to be valid, $>95\%$ of test reactions at the predicted LOD value need to amplify above the LOB. The LOD of N2 was confirmed to be 12 copies per reaction, as 58/60 (97%) test reactions

at that concentration had copy numbers above the N2 LOB. The N1 LOD was somewhere between 4 and 12 copies per reaction: at 12 copies per reaction, all 60 reactions amplified above the N1 LOB, but at 4 copies per reaction, 13/60 (22%) reactions amplified below the N1 LOB. Using a parametric method (which is an imperfect estimate in this case because the test reaction data were not normally distributed), the N1 LOD was estimated to be 8 copies per reaction. The process efficiency based on BCoV recoveries was $57 \pm 4\%$. BCoV was not detected in nonspiked wastewater samples.

Community SARS-CoV-2 Prevalence

Over the course of 10 months, from May 2020 through March 2021, eight sampling events were conducted in six Oregon communities by the TRACE project to determine the COVID-19 prevalence within those communities. The selected communities represented a diverse cross-section of Oregon, ranging from a small coastal commercial fishing community (Newport), to mid-sized and large university communities in the temperate Willamette Valley (Corvallis and Eugene), mid-sized and large arid high desert communities (Redmond and Bend), and a small agricultural community in eastern Oregon (Hermiston).

Community prevalence results are summarized in Table 3. The response rates ranged from 38% to 71%, with an average response rate of $60 \pm 14\%$. The estimated SARS-CoV-2 prevalence, including both symptomatic and asymptomatic infections, ranged from 8/10,000 to 1,687/10,000 persons. The SARS-CoV-2 concentrations in the influent wastewater corresponding to the door-to-door nasal swab samples ranged from 2.92 to 5.13 $\log_{10}$ gene copies per liter of wastewater (gc/L).

The majority of the prevalence sampling events took place in the absence of significant precipitation (Table S4). The only exception was in Newport on 20 June 2020. During this 24-h period, 3.56 mm of rain fell, which increased the wastewater flow rate by 5% compared with 11 July 2020, a 24-h period in which no rain fell.

The SARS-CoV-2 concentrations in the influent wastewater of these same cities over a 6- to 11-month period ranged from non-detect to 5.58 $\log_{10}$ gc/L (Figure S3). The wastewater SARS-CoV-2 concentrations showed a moderate positive correlation (Pearson's $r = 0.71$) when compared with the log to the base 10 of reported cases per 10,000 persons (to normalize for differences in population size). The accuracy of estimating reported COVID-19 cases using wastewater SARS-CoV-2 concentrations was also moderate [root mean square logarithmic error (RMSLE) = 0.14 and mean absolute percentage error (MAPE) = 0.29], with wastewater concentrations differing by up to 62-fold representing the same number of reported COVID-19 cases per 10,000 persons (Figure 2A; Tables S6 and S8). Similar correlation strengths and levels of accuracy were observed with each city when analyzed individually (Figure S4).

When the wastewater SARS-CoV-2 concentrations were compared with the TRACE-informed prevalence estimates that include both symptomatic and asymptomatic cases (Figure 2B, Table 3), the positive correlation was much stronger (Pearson's $r = 0.96$), and the accuracy was much higher, as demonstrated by the lower RMSLE (0.03) and MAPE (0.17) values. This correlation also suggests that the detection limit of our WBE method was 3 infections/10,000 persons. Compared with wastewater SARS-CoV-2 concentrations, reported COVID-19 cases were more weakly correlated with estimated prevalence (Pearson's $r = 0.85$) and had a lower accuracy (RMSLE = 0.05 and MAPE = 0.31) than wastewater SARS-CoV-2 concentrations (Figure 2C; Tables S6 and S8).

Table 3. Oregon SARS-CoV-2 prevalence sampling event statistics, 30 May 2020–14 March 2021: summary of the results of the eight random door-to-door sampling events and its corresponding wastewater analyses.

City in Oregon	Date	Participating households (n)	Household participation rate (%)	Samples (n)	Estimated prevalence per 10,000 ^a	Lower bounds per 10,000 ^b	Upper bounds per 10,000 ^c	Wastewater SARS-CoV-2 concentration (log ₁₀ gc/L ± SE)
Bend	30–31 May 2020	342	68	615	8 ^d	1	38	2.9 ± 0.2
Newport	20–21 June 2020	336	71	569	335	109	562	4.9 ± 0.1
Newport ^e	11–12 July 2020	338	70	550	60	0	150	4.1 ± 0.1
Hermiston	25–26 July 2020	249	44	469	1,687	714	2,660	5.1 ± 0.1
Corvallis	26–27 September 2020	354	71	580	30	0	89	3.8 ± 0.2
Eugene	7–8 November 2020	304	49	463	50 ^d	10	140	4.3 ± 0.1
Redmond	29–31 January 2021	251	38	376	320	0	760	4.8 ± 0.0
Corvallis	12–14 March 2021	347	67	514	131	0	310	4.2 ± 0.1

Note: gc, gene copies; SE, standard error.

^aPrevalence estimates were calculated using design-weighted estimators appropriate to the respective community sampling design when positive cases were identified or using a beta-binomial Bayesian model when zero positive cases were found.

^bRepresents the lower bounds of 95% intervals.

^cRepresents the upper bounds of 95% intervals.

^dZero positive cases were observed and the prevalence estimate used additional data on active infections in the community.

^eSecond sampling event was conducted at the request of the local county public health officials.

In addition, Monte Carlo simulations demonstrated that wastewater SARS-CoV-2 concentrations were significantly better than reported COVID-19 cases counts at estimating COVID-19 prevalence, even after accounting for uncertainty inherent in the wastewater and prevalence estimates (Figure 3; Tables S6 and S8). The median R^2 for the wastewater SARS-CoV-2 concentration model was 0.82 compared with 0.71 for the reported-cases model, based on the Monte Carlo simulations. A Wilcoxon rank-sum test, which determines whether the two distributions for R^2 have a location shift of 0 vs. the alternative that the wastewater SARS-CoV-2 concentration R^2 distribution has a positive shift, gave a p -value of <0.0001. Thus, even after accounting for the uncertainty inherent in the wastewater SARS-CoV-2 concentration and community prevalence estimates, the difference between the wastewater SARS-CoV-2 concentration and reported-cases median R^2 values was significant, and the wastewater SARS-CoV-2 concentration had a larger median R^2 .

For each simulation, the two nonnested regression models were directly compared using Vuong's test, which is a likelihood-ratio-based test for model selection using the Kullback–Leibler information criterion.^{46,47} Vuong's test revealed that the wastewater SARS-CoV-2 concentration model fit significantly better than the reported-cases model ($p < 0.05$) 53% of the time. The reported-cases model fit significantly better only 2% of the time; the models could not be distinguished on the remaining occasions. Although Vuong's test cannot determine whether the preferred model is the true model, the high R^2 value (0.91 using the observed data) and low RMSLE (0.03) and MAPE (0.17) are indicative of a good fit.

Localized Micro-Sewershed Surveillance (Newport, Oregon)

To identify COVID-19 “hotspots” at the neighborhood-scale in Newport, the wastewater SARS-CoV-2 concentrations were quantified from wastewater samples collected at 22 pump stations located throughout the small community of 10,853 persons (Figure 1). These pump stations divided the community into neighborhood-scale micro-sewersheds with estimated populations ranging from 15 persons at the 10th St. pump station up to 9,426 persons at the Northside pump station (Table 2). These wastewater samples were collected during the same time period as the two Newport random nasal swab prevalence sampling events.

During the peak of a Newport outbreak in mid-June, 2020, the random nasal swab sampling event estimated a community COVID-19 prevalence of 3.4% (Table 3). During this time, SARS-CoV-2 was

detected in wastewater samples from 13/22 micro-sewersheds, with concentrations ranging from 3.16 to 5.25 log₁₀ gc/L (Figure 4; Table S5). Three weeks later, in July 2020, the estimated COVID-19 prevalence in Newport had decreased to 0.6%. The number of micro-sewersheds with detectable levels of SARS-CoV-2 in their wastewater also decreased to 5, with concentrations ranging from 3.36 to 4.71 log₁₀ gc/L. In addition to decreased micro-sewershed wastewater SARS-CoV-2 concentrations, the decrease in community prevalence also corresponded with marked decreases in nasal swab positivity rates and clinically reported COVID-19 cases (Figure 4; Tables S5 and S7).

Wastewater SARS-CoV-2 concentrations collected from Newport's 22 micro-sewersheds were significantly correlated with the positivity rate of random swab samples [Pearson's product-moment correlation = 0.64; 95% confidence interval (CI): 0.32, 0.83, $df = 22$; $p < 0.001$] and reported cases (Pearson's product-moment correlation = 0.61; 95% CI: 0.39, 0.76; $df = 44$; $p < 0.001$). Similarly, the presence of SARS-CoV-2 in wastewater was significantly correlated with reported cases within each micro-sewershed [Fisher's exact test; odds ratio (OR) = 24.3; 95% CI: 3.1, 341; $p < 0.001$, $n = 33$]. In addition, the presence/absence of SARS-CoV-2 was also significantly correlated between wastewater and random nasal swabs within each micro-sewershed (Fisher's exact test; OR could not be calculated; 95% CI: 2.16, ∞; $p = 0.003$; $n = 22$).

Newport Community SARS-CoV-2 Genotype Profiling

Using Multi-Locus Sequence Typing,⁴⁸ two distinct viral variants, a B.1.399 lineage variant designated NA and a B.1 lineage variant designated NB, were detected both in the wastewater and among the individuals who tested positive during the June 2020 Newport sampling event (Table S2). The B.1.399/NA consensus most closely matched sequences found in Europe and then in California in March and April 2020, respectively. The B.1/NB consensus sequence most closely matched that from Yakima County, Washington, on 29 April 2020, and sequences found in Europe in March 2020.

During the June sampling event (from 569 samples), B.1.399/NA was detected in 77% of the individuals who tested positive (10/13), whereas B.1/NB was detected in 23% of the individuals who tested positive (3/13) (Figure 5, Table 4). In the July sampling event, two individuals tested positive (from 550 samples). Although both samples both yielded low coverage sequence data, they were identified as B.1.1.291 by the international database,

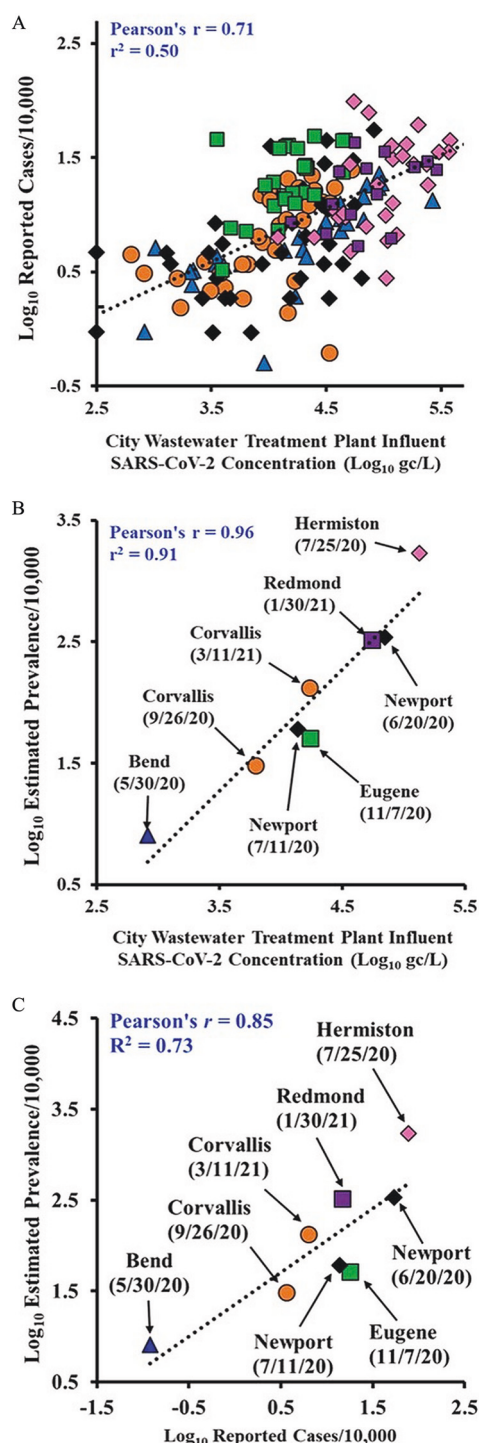


Figure 2. Wastewater concentrations vs. reported COVID-19 cases or estimated prevalence. (A) $\log_{10}$ of wastewater SARS-CoV-2 concentrations (gene copies per liter of wastewater) vs. the $\log_{10}$ of weekly reported COVID-19 cases (reported by ZIP code), (B) $\log_{10}$ of wastewater SARS-CoV-2 concentrations vs. the $\log_{10}$ of estimated prevalence, and (C) $\log_{10}$ of weekly reported COVID-19 cases (reported by ZIP code) vs. the $\log_{10}$ of estimated prevalence for Bend (blue triangles), Corvallis (orange circles), Eugene (green squares), Hermiston (pink diamonds), Newport (black diamonds), and Redmond (purple squares), Oregon. See Table 3 for corresponding numeric values and upper and lower bounds on 95% intervals for estimated prevalence per 10,000 values. See Table S6 for corresponding wastewater SARS-CoV-2 concentration numeric values and standard errors. See Table S8 for corresponding case count numeric data. Prevalence estimates were calculated using design-weighted estimators appropriate to the respective community sampling design when positive cases were identified or using a beta-binomial Bayesian model when zero positive cases were found. Note: gc, gene copies.

GISAID and, thus, were different from the individuals from the June sampling event.

In the wastewater samples, the fraction of viral RNA accounted for by each variant was estimated from the total fraction of variant reads summed across all SNP positions specific to that variant. During the June sampling event, the viral variant distribution in the WWTP influent was dominated by B.1.399/NA, accounting for 70% of the reads, whereas B.1/NB accounted for a minority of the reads at 4% (Figure 5, Table 4). This mirrors the observations made through the random nasal swab assays (77% B.1.399/NA and 23% B.1/NB).

B.1.399/NA was also dominant in 11/12 positive micro-sewershed wastewater samples collected. During the June sampling event, B.1.399/NA accounted for 30–98% of the total viral sequence reads across the micro-sewersheds, whereas during the July sampling event it accounted for 48–92% of the viral sequence reads across the micro-sewersheds (Table 4). In addition, although the abundance of B.1/NB was always in the minority, ranging from 0% to 52% of total virus sequence reads during the June sampling event and 0% to 11% of total virus sequence reads during the July sampling event, it was always detected in the wastewater of micro-sewersheds, when B.1/NB was detected among individuals via random nasal swabs (Figure 5, Table 4). The only exception to this relationship was the Bayfront micro-sewershed. In Bayfront, B.1/NB was not detected in the wastewater, and the single positive individual discovered in Bayfront carried B.1/NB.

Finally, from the weeks of 8 June to 30 November 2020, five SARS-CoV-2 variants were detected consistently in wastewater samples from the Newport WWTP (Figure 6A), representing at least 5% of the sequence reads in samples from at least 3 wk. Variant B.1.399/NA was the dominant variant during the weeks of 8 June to 6 July, with lesser amounts of B.1/NB detected. During the weeks of 20 July and 27 July, a B.1.369 subvariant, designated EE (Table S2B), was most abundant, and it was also detected during the weeks of 29 June and 3 August. From 3 August onward, a B.1.2 subvariant, designated FF, was the most abundant variant detected. The detection of these variants in Newport broadly mirrored trends across the state of Oregon observed in individuals whose positive samples were sequenced and deposited in GISAID (Figure 6B; Figure S5).

Discussion

Estimating COVID-19 Prevalence

WBE has been widely used to quantify wastewater SARS-CoV-2 concentrations during the COVID-19 pandemic to provide a relative sense of viral burden in a community and track its trend over time.^{2,49,50} However, the moderate correlation between wastewater SARS-CoV-2 concentrations and reported COVID-19 cases, similar to those observed in this study, has resulted in the perception that WBE is limited in its ability to estimate community COVID-19 infections.⁵¹ It has been hypothesized that this moderate correlation can be attributed to a variety of factors related to the wastewater sample, including virus decay in sewage conveyance systems, variability in sampling and concentrating techniques, and variability in the magnitude and duration of shedding by infected individuals.^{6,11,16} It should be noted that rainfall events were largely absent during the wastewater collection periods of this study. Thus, dilution of wastewater virus concentrations by rainfall events was not an important source of the observed variability in wastewater SARS-CoV-2 concentrations.

Although the correlation between wastewater SARS-CoV-2 concentrations and COVID-19 cases (as reported by health departments) has been moderate,^{51,52} this study demonstrated a very

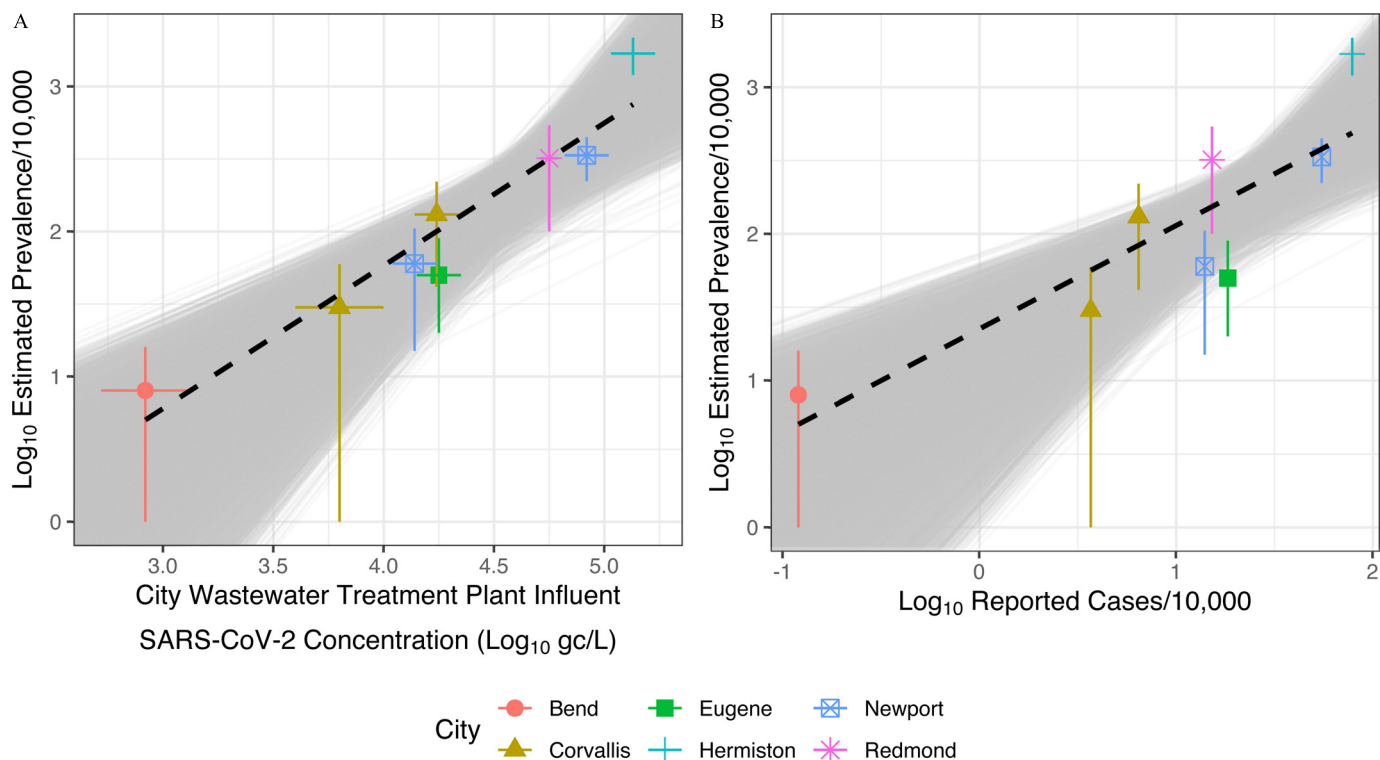


Figure 3. Estimated prevalence vs. wastewater concentration and reported cases with uncertainty from Monte Carlo Simulations. (A) Log₁₀ of wastewater SARS-CoV-2 concentrations (gene copies per liter of wastewater) vs. the log₁₀ of estimated prevalence, where the regression line from observed data is shown as a dashed black line. Horizontal and vertical segments indicate 1 SE or a 68% credible interval. The gray band is made up of individual regression lines from Monte Carlo simulations. (B) Log₁₀ of wastewater SARS-CoV-2 concentrations vs. the log₁₀ of reported cases, where the regression line from observed data is shown as a dashed black line. Vertical segments indicate 1 SE or a 68% credible interval. The gray band is made up of individual regression lines from Monte Carlo simulations. Prevalence estimates were calculated using design-weighted estimators appropriate to the respective community sampling design when positive cases were identified or using a beta-binomial Bayesian model when zero positive cases were found. See Table 3 for corresponding numeric values and upper and lower bounds on 95% intervals for estimated prevalence per 10,000 values. See Table S6 for corresponding wastewater SARS-CoV-2 concentration numeric values and SEs. See Table S8 for corresponding case count numeric data. Note: SE, standard error.

strong correlation between wastewater SARS-CoV-2 concentrations and estimated COVID-19 prevalence. With an average response rate of $60 \pm 14\%$, the response rate of this study was in the upper end of the range reported by other studies doing a door-to-door health survey (20–50%), lending strength to the correlation analyses.^{53–55} In addition, Monte Carlo simulations indicated that wastewater SARS-CoV-2 concentrations were significantly better ($p < 0.05$) than reported COVID-19 cases at estimating COVID-19 prevalence in a community. The COVID-19 prevalence estimates minimized many of the uncertainties inherent to reported COVID-19 cases, including asymptomatic individuals and limited access to testing. This suggested that the weaker correlations between wastewater SARS-CoV-2 concentrations and reported COVID-19 cases had less to do with the inherent variability within the wastewater measurements and more to do with the inherent variability in the reporting of COVID-19 cases. In addition, the inherent variability in reported COVID-19 cases is expected to become even more complex with the anticipated unreliability of self-reporting positive results from at-home testing kits.

Thus, when compared with a metric with much smaller uncertainties, such as the prevalence estimates presented in this study, wastewater data were quantitative and could accurately estimate COVID-19 prevalence. This improved correlation may also explain why wastewater has shown significantly better estimation capabilities when compared with COVID-19 hospitalizations and deaths.^{2,5,49} These metrics also have lower uncertainties than reported COVID-19 cases, which are likely to be more strongly impacted by unequal distribution of testing availability, asymptomatic individuals, and individuals

who do not seek testing for various reasons, including being vaccinated.

Identifying Neighborhood COVID-19 Hotspots and SARS-CoV-2 Variants

In addition to estimating community-wide prevalence with high precision, this study demonstrated that quantifying wastewater SARS-CoV-2 was a powerful method for detecting infection “hot spots” at the micro-sewershed (i.e., neighborhood) level. It should be noted that owing to the hierarchical flows between pump stations, the Northside micro-sewershed was receiving flow from the Bayfront micro-sewershed, which contains the Samaritan Pacific Communities Hospital (Figure 1). Thus, the Northside micro-sewershed wastewater results may have captured individuals who would not be linked to the Northside micro-sewershed by case reporting or random household surveillance (e.g., infected individuals at the Samaritan Pacific Communities Hospital).

Similar to the observations at the community level, wastewater SARS-CoV-2 concentrations taken from Newport micro-sewersheds correlated slightly more strongly with nasal swab positivity rates (taken during the prevalence sampling event) than with reported COVID-19 cases at the neighborhood level. Thus, wastewater SARS-CoV-2 concentrations were generally more accurate at identifying hot spots of COVID-19 infections within a community than reported clinical cases. Other studies have used similar approaches to demonstrate the utility of wastewater monitoring at the building level on college and university campuses.^{56,57}

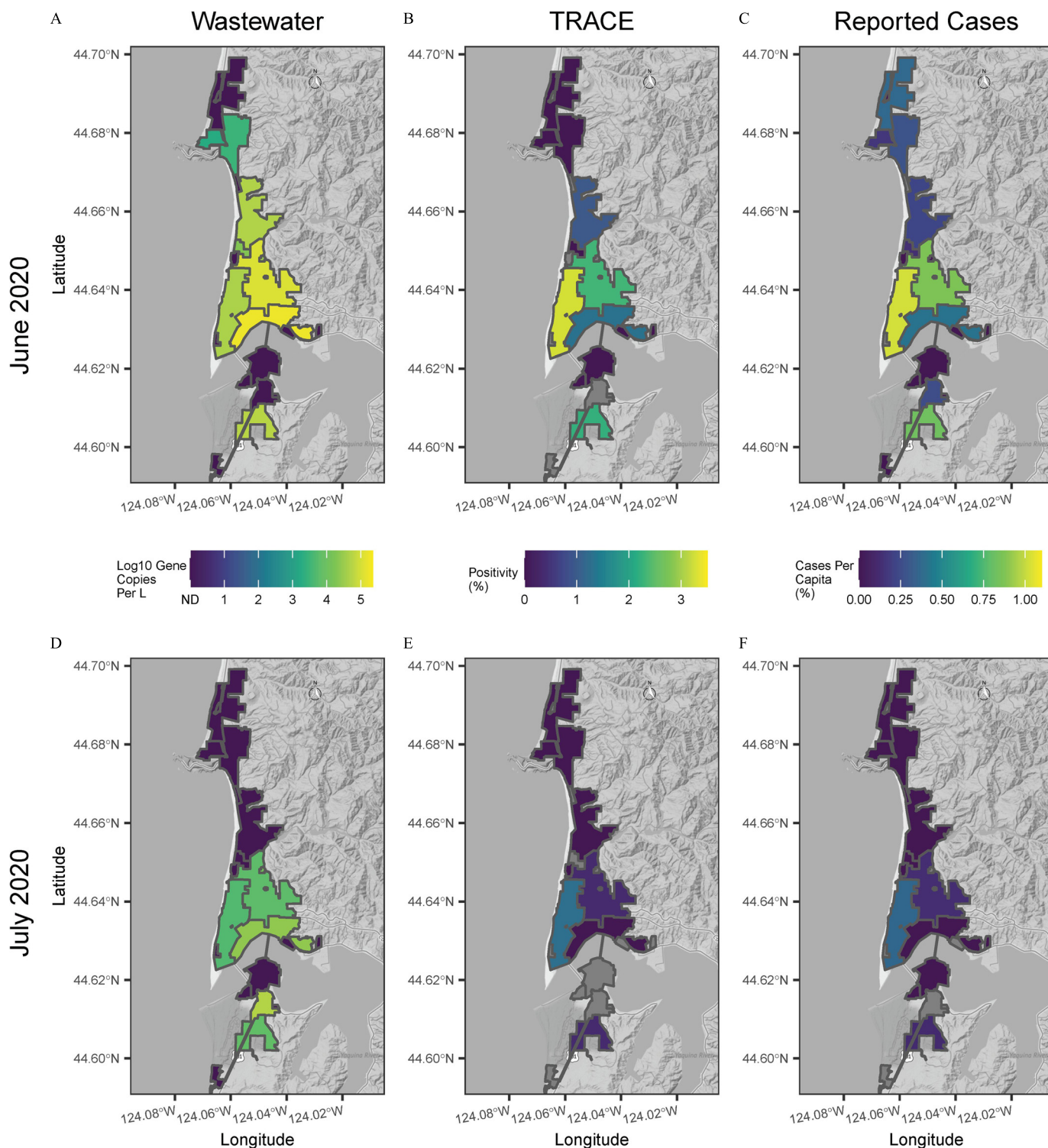


Figure 4. COVID-19 burden heat maps. (A,D) Wastewater SARS-CoV-2 concentrations, (B,E) percentage positivity from random door-to-door nasal swab sampling events, and (C,F) percentage reported cases per capita of the 22 micro-sewersheds sampled in Newport, Oregon. Wastewater and nasal swab sampling events were conducted during (A–C) 18–19 June 2020 and (D–F) 8–9 July 2020; reported case windows were the 10 d prior to and including the sampling periods. See Table S5 for corresponding wastewater SARS-CoV-2 concentration (WBE) numeric values and 95% CIs. See Table S7 for corresponding positivity (TRACE) and reported cases per capita (Cases) percentages. Map tiles by Stamen Design, under CC BY 3.0. Basemap data by OpenStreetMap, under ODbL. Note: CI, confidence interval; TRACE, Team-based Rapid Assessment of community-level Coronavirus Epidemics (project); WBE, wastewater-based epidemiology.

WBE was also successful at identifying the dominant and minor SARS-CoV-2 variants (B.1.399/NA and B.1/NB, respectively) at both the micro-sewershed and city levels in Newport. At the city WWTP, as well as at three of the four micro-

sewersheds that contained positive nasal swab samples during the random door-to-door sampling events (Big Creek, Nye Beach, and Northside), the SARS-CoV-2 variants identified in the wastewater were the same as those identified via nasal swab analyses

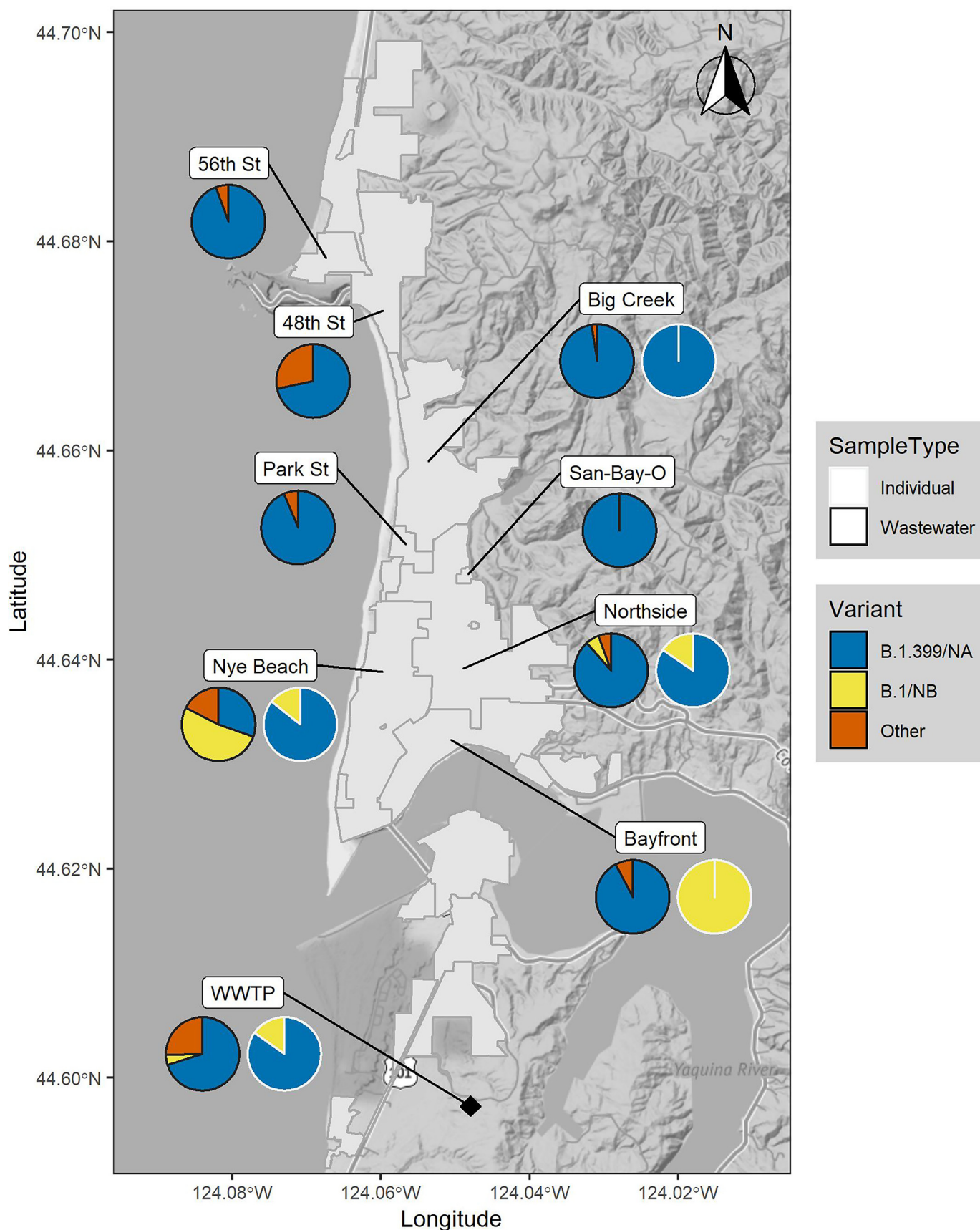


Figure 5. Spatial distribution of SARS-CoV-2 variants. The percentage sequence reads of variants (A) B.1.399/NA and (B) B.1/NB located within the various Newport, Oregon, micro-sewershed boundaries during the 18–19 June 2020 prevalence sampling event. Sequences were obtained from both micro-sewershed wastewater, as well as random door-to-door nasal swab sampling events. See [Table 4](#) for corresponding numeric data. Map tiles by Stamen Design, under CC BY 3.0. Basemap data by OpenStreetMap, under ODbL. Note: WWTP, wastewater treatment plant.

Table 4. SARS-CoV-2 variant relative abundances in Newport, Oregon: relative abundances of variants detected in samples from individuals and wastewater across micro-sewersheds (neighborhoods) during the 20–21 June 2020 prevalence sampling event in Newport.

Pump station	Individual numbers ^a		Individual fraction		Wastewater fraction $\pm$ SE ^b		Reps ^c
	NA ^d	NB ^e	NA ^d	NB ^e	NA ^d	NB ^e	
48th St.	0	0	0	0	0.72	0	1
56th St.	0	0	0	0	0.9	0	1
Bayfront	0	1	0	1	0.92 $\pm$ 0.01	0 $\pm$ 0	2
Big Creek	1	0	1	0	0.98 $\pm$ 0.01	0 $\pm$ 0	2
Northside	11	2	0.85	0.15	0.88 $\pm$ 0.04	0.06 $\pm$ 0.06	3
Nye Beach	6	1	0.86	0.14	0.3 $\pm$ 0.16	0.52 $\pm$ 0.08	3
Park St.	0	0	0	0	0.94	0	1
San-Bay-O	0	0	0	0	1	0	1
Influent	11	2	0.85	0.15	0.7 $\pm$ 0	0.04 $\pm$ 0.04	3

Note: Rep, replicate samples; SE, standard error.

^aThe number of individuals counted for each micro-sewershed include not only those residing within the micro-sewershed itself but also those residing within upstream micro-sewersheds flowing into the micro-sewershed.

^bSE was calculated on untransformed fractions.

^cMeasurements were from a) a single assay, b) assays of two independent water samples, c) two assays of RNA from one sample and one assay of a second independent sample.

^dNA = B.1.399/NA.

^eNB = B.1/NB.

and at similar relative proportions. The lone exception was the Bayfront micro-sewershed whose wastewater sample was dominated by the B.1.399/NA variant, whereas the nasal swab samples from that area were dominated by the B.1/NB variant. However, this micro-sewershed was home to the Samaritan Pacific

Communities Hospital, which may have skewed the micro-sewershed variant information with inputs from hospitalized individuals who would not be normally associated with that micro-sewershed.

WBE data was also able to monitor the change in SARS-CoV-2 variants present in the Newport community over time. The five major SARS-CoV-2 variants identified in the WBE data from 8 June to 30 November 2020 mirrored the primary SARS-CoV-2 variants identified through clinical samples in the state of Oregon during that same time period. Interestingly, spikes observed in the WWTP influent SARS-CoV-2 concentrations often corresponded to the appearance of a new dominant variant in the wastewater sequences. This suggests that rises in viral RNA concentrations in wastewater samples may signal the appearance of a new dominant variant.

The agreement between wastewater sequence data with clinical sequence data at both the micro-sewershed and city levels supports the reliability of wastewater sequence data to accurately represent the SARS-CoV-2 variant distribution in a city or neighborhood. To our knowledge, this is the first study of variant abundance in wastewater at the neighborhood level, although others have sequenced SARS-CoV-2 in wastewater at the community level.^{58–60}

Limitations of WBE Methodologies

Although results of this study have demonstrated the power of WBE methodologies, care should be taken when comparing quantitative results across studies. A large interlaboratory study found that the quantified SARS-CoV-2 concentrations in a single wastewater sample can vary by several orders of magnitude, depending on the concentration methods (e.g., polyethylene glycol precipitation, ultrafiltration, direct extraction, or HA membrane filtration), pretreatments (e.g., removal of solids and/or pasteurization of the wastewater samples), and PCR platforms (e.g., digital or quantitative) employed.¹⁸

However, this same study also reported very low variability in replicate samples from each participating laboratory, as well as between participating labs that used the exact same methodologies, ($SD < 0.2 \log_{10} \text{ gc/L}$). This finding suggests that when the same methodologies were employed, consistent SARS-CoV-2 concentrations both within and across laboratories was achievable. Further research to optimize and standardize the wastewater surveillance process is still needed to unify the results from national and global wastewater surveillance efforts.

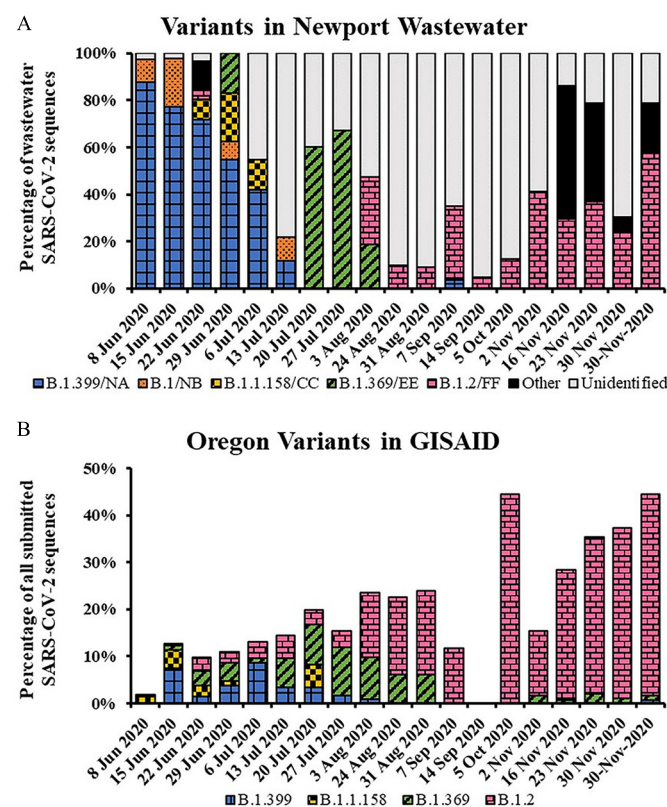


Figure 6. SARS-CoV-2 variant temporal distribution. (A) The average estimated percentage viral sequence reads of the indicated SARS-CoV-2 variant RNAs and the $\log_{10}$ SARS-CoV-2 concentrations quantified in the Vance Avery Wastewater Treatment Plant influent (Newport, Oregon) from 10 June to 2 December 2020. Each data point represents the mean of measurements collected during the week beginning on the date shown. Sequence data from some dates derived from a single measurement, and standard errors are expected to be comparable to those shown in Table 4. (B) Percentage of all variants of the indicated lineages among SARS-CoV-2 sequences from samples collected in Oregon during the week beginning on the indicated date and deposited in GISAID. Note: GISAID, Global Initiative on Sharing All Influenza Data.

Conclusion

This study demonstrated that wastewater SARS-CoV-2 concentrations were more accurate than clinically reported COVID-19 cases at estimating COVID-19 prevalence in communities that are diverse in terms of size, location, access to testing, and pandemic stage. This was likely due to the representativeness and noninvasive nature of the WBE methodologies used to quantify wastewater SARS-CoV-2 concentrations. The sampling of wastewater was not subject to many of the major biases present in clinical COVID-19 case reporting, including limited testing capacity, barriers in access to testing, testing avoidance, self-isolation of individuals with mild symptoms, and the presence of asymptomatic transmissions. The avoidance of these biases appeared to be more important for estimating COVID-19 prevalence than avoiding the biases inherent to WBE methodologies, including variabilities in shedding rates over time and between individuals, in addition to viral signal decay in the wastewater.

In addition, this study demonstrated that WBE methodologies can identify neighborhood COVID-19 hotspots and accurately profile the SARS-CoV-2 variant diversity present in a city or neighborhood. Considering that prevalence, hotspot identification, and variant compositions can be determined from a single to relatively few wastewater samples, compared with the hundreds or thousands of clinical samples required for traditional surveillance techniques, WBE has clear advantages in terms of cost. In addition, wastewater SARS-CoV-2 concentrations can be reported within 24–48 h and variant compositions for an entire community determined within 5–7 d. Thus, with relatively low costs, rapid turnaround times and proven reliability, WBE should be considered an essential tool for long-term public health surveillance strategies.

Acknowledgments

We are grateful to the Public Works Departments of the Oregon cities of Bend, Corvallis, Eugene, Hermiston, Newport, and Redmond, the Benton County Health Department, the Deschutes County Health Services, the Lane County Public Health Department, the Lincoln County Health and Human Services Department, the Umatilla County Public Health Department, and the Oregon Health Authority for their support in sample collection and public health data acquisition. We also thank the staff of Clean Water Services (especially S. Mansell, L. Barker, and K. Bailey), Oregon Veterinary Diagnostic Laboratory (A. Hunkapiller, N. Lawler, D. Mulrooney, B. Nagamine, and A. Al-Fotis), and M. Harry from OSU's School of Chemical, Biological and Environmental Engineering. Funding for this work was provided by a National Science Foundation Rapid Response Research grant (2027679 to T.S.R.), the David and Lucille Packard Foundation (to B.D.D.), PacificSource Health Plans (to J.W.B.), the Oregon State University Foundation (to B.D.D.), and Clean Water Services with funding under the Clean Water Services Coronavirus Aid, Relief, and Economic Security Act through the Washington County Cities and Special District Assistance program (to K.J.W.).

All individual sequences were deposited in the data sharing platform Global Initiative on Sharing All Influenza Data (GISAID; see Table S3 for accession numbers) and all wastewater sequences were deposited in the National Center for Biotechnology Information's short read archive, under BioProject PRJNA719837. Data from community prevalence sampling events is given in Table 3. Wastewater reverse transcriptase droplet digital polymerase chain reaction data supporting the findings of this study are located within the paper and its supplementary information files. All R scripts used in this analysis are available on GitHub at <https://github.com/laytonba/LaytonEHP22>.

References

1. Bivins A, North D, Ahmad A, Ahmed W, Alm E, Been F, et al. 2020. Wastewater-based epidemiology: global collaborative to maximize contributions in the fight against COVID-19. *Environ Sci Technol* 54(13):7754–7757, PMID: 32530639, <https://doi.org/10.1021/acs.est.0c02388>.
2. Medema G, Heijnen L, Elsinga G, Italiaander R, Brouwer A. 2020. Presence of SARS-Coronavirus-2 RNA in sewage and correlation with reported COVID-19 prevalence in the early stage of the epidemic in the Netherlands. *Environ Sci Technol Lett* 7(7):511–516, <https://doi.org/10.1021/acs.estlett.0c00357>.
3. Ahmed W, Angel N, Edson J, Bibby K, Bivins A, O'Brien JW, et al. 2020. First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: a proof of concept for the wastewater surveillance of COVID-19 in the community. *Sci Total Environ* 728:138764, PMID: 32387778, <https://doi.org/10.1016/j.scitotenv.2020.138764>.
4. La Rosa G, Iaconelli M, Mancini P, Bonanno Ferraro G, Veneri C, Bonadonna L, et al. 2020. First detection of SARS-CoV-2 in untreated wastewaters in Italy. *Sci Total Environ* 736:139652, PMID: 32464333, <https://doi.org/10.1016/j.scitotenv.2020.139652>.
5. Wurtzer S, Marechal V, Mouchel JM, Maday Y, Teyssou R, Richard E, et al. 2020. Evaluation of lockdown effect on SARS-CoV-2 dynamics through viral genome quantification in waste water, Greater Paris, France, 5 March to 23 April 2020. *Euro Surveill* 25(50):2000776, PMID: 33334397, <https://doi.org/10.2807/1560-7917.ES.2020.25.50.2000776>.
6. Curtis K, Keeling D, Yetka K, Larson A, Gonzalez R. 2020. Wastewater SARS-CoV-2 concentration and loading variability from grab and 24-hour composite samples. *medRxiv*. Preprint posted online 11 July 2020, <https://doi.org/10.1101/2020.07.10.20150607>.
7. Ahmed W, Bertsch PM, Bivins A, Bibby K, Farkas K, Gathercole A, et al. 2020. Comparison of virus concentration methods for the RT-qPCR-based recovery of murine hepatitis virus, a surrogate for SARS-CoV-2 from untreated wastewater. *Sci Total Environ* 739:139960, PMID: 32758945, <https://doi.org/10.1016/j.scitotenv.2020.139960>.
8. LaTurner ZW, Zong DM, Kalvapalle P, Gamas KR, Terwilliger A, Crosby T, et al. 2021. Evaluating recovery, cost, and throughput of different concentration methods for SARS-CoV-2 wastewater-based epidemiology. *Water Res* 197:117043, PMID: 33784608, <https://doi.org/10.1016/j.watres.2021.117043>.
9. Whitney ON, Kennedy LC, Fan V, Hinkle A, Kantor R, Greenwald H, et al. 2020. Sewage, Salt, Silica and SARS-CoV-2 (4S): an economical kit-free method for direct capture of SARS-CoV-2 RNA from wastewater. *medRxiv*. Preprint posted online 3 December 2020, <https://doi.org/10.1101/2020.12.01.20242131>.
10. Randazzo W, Truchado P, Cuevas-Ferrando E, Simón P, Allende A, Sánchez G. 2020. SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. *Water Res* 181:115942, PMID: 32425251, <https://doi.org/10.1016/j.watres.2020.115942>.
11. Jones DL, Baluja MQ, Graham DW, Corbishley A, McDonald JE, Malham SK, et al. 2020. Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19. *Sci Total Environ* 749:141364, PMID: 32836117, <https://doi.org/10.1016/j.scitotenv.2020.141364>.
12. Cheung KS, Hung IFN, Chan PPY, Lung KC, Tso E, Liu R, et al. 2020. Gastrointestinal manifestations of SARS-CoV-2 infection and virus load in fecal samples from a Hong Kong cohort: systematic review and meta-analysis. *Gastroenterology* 159(1):81–95, PMID: 32251668, <https://doi.org/10.1053/j.gastro.2020.03.065>.
13. Chen Y, Chen L, Deng Q, Zhang G, Wu K, Ni L, et al. 2020. The presence of SARS-CoV-2 RNA in the feces of COVID-19 patients. *J Med Virol* 92(7):833–840, PMID: 32243607, <https://doi.org/10.1002/jmv.25825>.
14. Zhang Y, Cen M, Hu M, Du L, Hu W, Kim JJ, et al. 2021. Prevalence and persistent shedding of fecal SARS-CoV-2 RNA in patients with COVID-19 infection: a systematic review and meta-analysis. *Clin Transl Gastroenterol* 12(4):e00343, PMID: 33835096, <https://doi.org/10.14309/ctg.0000000000000343>.
15. Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. 2021. 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* 2(1):e13–e22, PMID: 33521734, [https://doi.org/10.1016/S2666-5247\(20\)30172-5](https://doi.org/10.1016/S2666-5247(20)30172-5).
16. Bivins A, Greaves J, Fischer R, Yinda KC, Ahmed W, Kitajima M, et al. 2020. Persistence of SARS-CoV-2 in water and wastewater. *Environ Sci Technol Lett* 7(12):937–942, <https://doi.org/10.1021/acs.estlett.0c00730>.
17. Bivins A, North D, Wu Z, Shaffer M, Ahmed W, Bibby K. 2021. Within-day variability of SARS-CoV-2 RNA in municipal wastewater influent during periods of varying COVID-19 prevalence and positivity. *medRxiv*. Preprint posted online 24 March 2021, <https://doi.org/10.1101/2021.03.16.21253652>.
18. Pecson BM, Darby E, Haas CN, Amha YM, Bartolo M, Danielson R, et al. 2021. Reproducibility and sensitivity of 36 methods to quantify the SARS-CoV-2 genetic signal in raw wastewater: findings from an interlaboratory methods

- evaluation in the U.S. *Environ Sci (Camb)* 7(3):504–520, PMID: [34017594](https://doi.org/10.1039/d0ew00946f), <https://doi.org/10.1039/d0ew00946f>.
19. Xiao F, Sun J, Xu Y, Li F, Huang X, Li H, et al. 2020. Infectious SARS-CoV-2 in feces of patient with severe COVID-19. *Emerg Infect Dis* 26(8):1920–1922, PMID: [32421494](https://doi.org/10.3201/eid2608.200681), <https://doi.org/10.3201/eid2608.200681>.
20. Kantor RS, Nelson KL, Greenwald HD, Kennedy LC. 2021. Challenges in measuring the recovery of SARS-CoV-2 from wastewater. *Environ Sci Technol* 55(6):3514–3519, PMID: [33656856](https://doi.org/10.1021/acs.est.0c08210), <https://doi.org/10.1021/acs.est.0c08210>.
21. Canh VD, Torii S, Yasui M, Kyuwa S, Katayama H. 2021. Capsid integrity RT-qPCR for the selective detection of intact SARS-CoV-2 in wastewater. *Sci Total Environ* 791:148342, PMID: [34139497](https://doi.org/10.1016/j.scitotenv.2021.148342), <https://doi.org/10.1016/j.scitotenv.2021.148342>.
22. Gandhi M, Yokoe DS, Havlir DV. 2020. Asymptomatic transmission, the Achilles' heel of current strategies to control COVID-19. *N Engl J Med* 382(22):2158–2160, PMID: [32329972](https://doi.org/10.1056/NEJMe2009758), <https://doi.org/10.1056/NEJMe2009758>.
23. Arons MM, Hatfield KM, Reddy SC, Kimball A, James A, Jacobs JR, et al. 2020. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. *N Engl J Med* 382(22):2081–2090, PMID: [32329971](https://doi.org/10.1056/NEJMoa2008457), <https://doi.org/10.1056/NEJMoa2008457>.
24. Oregon State University. 2021. TRACE COVID-19. <https://trace.oregonstate.edu> [accessed 6 June 2021].
25. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. 2009. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 42(2):377–381, PMID: [18929686](https://doi.org/10.1016/j.jbi.2008.08.010), <https://doi.org/10.1016/j.jbi.2008.08.010>.
26. Thompson SK. 2012. *Sampling*. 3rd ed. Hoboken, New Jersey: John Wiley & Sons, Inc.
27. Oregon Health Authority. 2020. *COVID-19 Weekly Report: Oregon's Weekly Surveillance Summary of the Nonrel Coronavirus (COVID-19)*. 5 August 2020. <https://www.oregon.gov/oha/PH/DISEASES/CONDITIONS/DISEASESAZ/EmergingRespiratoryInfections/COVID-19-Weekly-Report-2020-08-05-FINAL.pdf> [accessed 6 June 2021].
28. Steele JA, Blackwood AD, Griffith JF, Noble RT, Schiff KC. 2018. Quantification of pathogens and markers of fecal contamination during storm events along popular surfing beaches in San Diego, California. *Water Res* 136:137–149, PMID: [29501758](https://doi.org/10.1016/j.watres.2018.01.056), <https://doi.org/10.1016/j.watres.2018.01.056>.
29. Lu X, Wang L, Sakthivel SK, Whitaker B, Murray J, Kamili S, et al. 2020. US CDC real-time reverse transcription PCR panel for detection of severe acute respiratory syndrome coronavirus 2. *Emerg Infect Dis* 26(8):1654–1665, PMID: [32396505](https://doi.org/10.3201/eid2608.201246), <https://doi.org/10.3201/eid2608.201246>.
30. Suo T, Liu X, Feng J, Guo M, Hu W, Guo D, et al. 2020. ddPCR: a more accurate tool for SARS-CoV-2 detection in low viral load specimens. *Emerg Microbes Infect* 9(1):1259–1268, PMID: [32438868](https://doi.org/10.1080/22221751.2020.1772678), <https://doi.org/10.1080/22221751.2020.1772678>.
31. Dong L, Zhou J, Niu C, Wang Q, Pan Y, Sheng S, et al. 2021. Highly accurate and sensitive diagnostic detection of SARS-CoV-2 by digital PCR. *Talanta* 224:121726, PMID: [33379001](https://doi.org/10.1016/j.talanta.2020.121726), <https://doi.org/10.1016/j.talanta.2020.121726>.
32. Cao Y, Raith MR, Griffith JF. 2015. Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. *Water Res* 70:337–349, PMID: [25543243](https://doi.org/10.1016/j.watres.2014.12.008), <https://doi.org/10.1016/j.watres.2014.12.008>.
33. Gonzalez R, Curtis K, Bivins A, Bibby K, Weir MH, Yetka K, et al. 2020. COVID-19 surveillance in southeastern Virginia using wastewater-based epidemiology. *Water Res* 186:116296, PMID: [32841929](https://doi.org/10.1016/j.watres.2020.116296), <https://doi.org/10.1016/j.watres.2020.116296>.
34. D'Aoust PM, Mercier E, Montpetit D, Jia JJ, Alexandrov I, Neault N, et al. 2021. Quantitative analysis of SARS-CoV-2 RNA from wastewater solids in communities with low COVID-19 incidence and prevalence. *Water Res* 188:116560, PMID: [33137526](https://doi.org/10.1016/j.watres.2020.116560), <https://doi.org/10.1016/j.watres.2020.116560>.
35. Bio-Rad. 2020. Instructions for Use: SARS-CoV-2 ddPCR Kit—Part Number 12013743. <https://www.bio-rad.com/webroot/web/pdf/lsr/literature/10000130776.pdf> [accessed 1 May 2020].
36. Rstudio Team. 2020. RStudio: Integrated Development Environment for R. Boston, MA: RStudio, PBC.
37. Pebesma E. 2018. Simple features for R: standardized support for spatial vector data. *R J* 10(1):439–446, <https://doi.org/10.32614/RJ-2018-009>.
38. Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag.
39. Wickham H, Pedersen TL. 2019. gtable: arrange “grobs” in tables. Version 0.3.0 from CRAN. <https://rdr.io/cran/gtable/> (accessed 31 January, 2022).
40. Kahle D, Wickham H. 2013. ggmap: spatial visualization with ggplot2. *R J* 5(1):144–161, <https://doi.org/10.32614/rj-2013-014>.
41. Decaro N, Elia G, Campolo M, Desario C, Mari V, Radogna A, et al. 2008. Detection of bovine coronavirus using a TaqMan-based real-time RT-PCR assay. *J Virol Methods* 151(2):167–171, PMID: [18579223](https://doi.org/10.1016/j.jviromet.2008.05.016), <https://doi.org/10.1016/j.jviromet.2008.05.016>.
42. Mölder F, Jablonski KP, Letcher B, Hall MB, Tomkins-Tinch CH, Sochat V, et al. 2021. Sustainable data analysis with Snakemake. *F1000Res* 10:33, PMID: [34035898](https://doi.org/10.12688/f1000research.29032.2), <https://doi.org/10.12688/f1000research.29032.2>.
43. NCBI Short Read Archive. 2020. Severe acute respiratory syndrome coronavirus 2 genome sequencing and assembly. BioProject PRJNA719837. ID: 719837. <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA719837> [accessed 4 May 2021].
44. GISAID (Global Initiative on Sharing All Influenza Data). 2021. GISAID database. <https://www.gisaid.org/> [accessed 4 June 2021].
45. Devleeschauwer B, Torgerson P, Charlier J, Levecke B, Praet N, Roelandt S, et al. 2022. prevalence: tools for prevalence assessment studies. R package version 0.4.0. <http://cran.r-project.org/package=prevalence> [accessed 5 July 2021].
46. Vuong QH. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57(2):307–333, <https://doi.org/10.2307/1912557>.
47. Merkle E, You D. 2016. nonnest2: tests of non-nested models. R package version 0.5-5. <https://mran.microsoft.com/snapshot/2016-04-25/web/packages/nonnest2/nonnest2.pdf> [accessed 5 July 2021].
48. Ibarz Pavón AB, Maiden MCJ. 2009. Multilocus sequence typing. *Methods Mol Biol* 551:129–140, PMID: [19521872](https://doi.org/10.1007/978-1-60327-999-4_11), https://doi.org/10.1007/978-1-60327-999-4_11.
49. Peccia J, Zulli A, Brackney DE, Grubaugh ND, Kaplan EH, Casanovas-Massana A, et al. 2020. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nat Biotechnol* 38(10):1164–1167, PMID: [32948856](https://doi.org/10.1038/s41587-020-0684-z), <https://doi.org/10.1038/s41587-020-0684-z>.
50. Graham KE, Loeb SK, Wolfe MK, Catoe D, Sinnott-Armstrong N, Kim S, et al. 2021. SARS-CoV-2 RNA in wastewater settled solids is associated with COVID-19 cases in a large urban watershed. *Environ Sci Technol* 55(1):488–498, PMID: [33283515](https://doi.org/10.1021/acs.est.0c06191), <https://doi.org/10.1021/acs.est.0c06191>.
51. Cao Y, Francis R. 2021. On forecasting the community-level COVID-19 cases from the concentration of SARS-CoV-2 in wastewater. *Sci Total Environ* 786:147451, PMID: [33971608](https://doi.org/10.1016/j.scitotenv.2021.147451), <https://doi.org/10.1016/j.scitotenv.2021.147451>.
52. Weidhaas J, Aanderud ZT, Roper DK, VanDerslice J, Gaddis EB, Ostermiller J, et al. 2021. Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds. *Sci Total Environ* 775:145790, PMID: [33618308](https://doi.org/10.1016/j.scitotenv.2021.145790), <https://doi.org/10.1016/j.scitotenv.2021.145790>.
53. Bodewes AJ, Kunst AE. 2016. Involving hard-to-reach ethnic minorities in low-budget health research: lessons from a health survey among Moluccans in the Netherlands. *BMC Res. Notes* 9(1):319, PMID: [27328767](https://doi.org/10.1186/S13104-016-2124-1), <https://doi.org/10.1186/S13104-016-2124-1>.
54. Flynn A, Tremblay PF, Rehm J, Wells S. 2013. A modified random walk door-to-door recruitment strategy for collecting social and biological data relating to mental health, substance use/addictions and violence problems in a Canadian community. *Int J Alcohol Drug Res* 2(2):7–16, PMID: [27279929](https://doi.org/10.7895/ijadr.v2i2.143), <https://doi.org/10.7895/ijadr.v2i2.143>.
55. Saramunee K, Mackridge A, Phillips-Howard P, Richards J, Suttajit S, Krksa J. 2016. Methodological and economic evaluations of seven survey modes applied to health service research. *J Pharm Health Serv Res* 7(1):43–52, <https://doi.org/10.1111/jphs.12121>.
56. Spurbeck RR, Minard-Smith A, Catlin L. 2021. Feasibility of neighborhood and building scale wastewater-based genomic epidemiology for pathogen surveillance. *Sci Total Environ* 789:147829, PMID: [34051492](https://doi.org/10.1016/j.scitotenv.2021.147829), <https://doi.org/10.1016/j.scitotenv.2021.147829>.
57. Harris-Lovett S, Nelson KL, Beamer P, Bischel HN, Bivins A, Bruder A, et al. 2021. Wastewater surveillance for SARS-CoV-2 on college campuses: initial efforts, lessons learned, and research needs. *Int J Environ Res Public Health* 18(9):4455, PMID: [33922263](https://doi.org/10.3390/ijerph18094455), <https://doi.org/10.3390/ijerph18094455>.
58. Crits-Christoph A, Kantor RS, Olm MR, Whitney ON, Al-Shayeb B, Lou YC, et al. 2021. Genome sequencing of sewage detects regionally prevalent SARS-CoV-2 variants. *mBio* 12(1):e02703-20, PMID: [33468686](https://doi.org/10.1128/mBio.02703-20), <https://doi.org/10.1128/mBio.02703-20>.
59. Izquierdo-Lara R, Elsinga G, Heijnen L, Munnink BBO, Schapendonk CME, Nieuwenhuijse D, et al. 2021. Monitoring SARS-CoV-2 circulation and diversity through community wastewater sequencing, the Netherlands and Belgium. *Emerg Infect Dis* 27(5):1405–1415, PMID: [33900177](https://doi.org/10.3201/eid2705.204410), <https://doi.org/10.3201/eid2705.204410>.
60. Agrawal S, Orschler L, Lackner S. 2021. Metatranscriptomic analysis reveals SARS-CoV-2 mutations in wastewater of the Frankfurt metropolitan area in southern Germany. *Microbiol Resour Announc* (15):e00280-21, PMID: [33858934](https://doi.org/10.1128/MRA.00280-21), <https://doi.org/10.1128/MRA.00280-21>.