Microbiome Analysis for Wastewater Surveillance during COVID-19

Kyle D. Brumfield,[a,b] Menu Leddy,[c] Moiz Usmani,[d] Joseph A. Cotruvo,[e] Ching-Tzone Tien,[f] Suzanne Dorsey,[f] Karlis Graubics,[g] Brian Fanelli,[g] Isaac Zhou,[g] Nathaniel Registe,[g] Manoj Dadlani,[g] Malinda Wimalarante,[h] Dilini Jinasena,[h] Rushan Abayagunawardena,[h] Chiran Withanachchi,[h] Anwar Huq,[a] Antarpreet Jutla,[d] Rita R. Colwell[a,b,g]

[a]Maryland Pathogen Research Institute, University of Maryland, College Park, Maryland, USA

[b]University of Maryland Institute for Advanced Computer Studies, University of Maryland, College Park, Maryland, USA

[c]Essential Environmental and Engineering Systems, Huntington Beach, California, USA

[d]Geohealth and Hydrology Laboratory, Department of Environmental Engineering Sciences, University of Florida, Gainesville, Florida, USA

[e]Joseph Cotruvo and Associates LLC, Washington, DC, USA

[f]Maryland Department of Environment, Baltimore, Maryland, USA

[g]CosmosID Inc., Germantown, Maryland, USA

[h]Inspection Experts Inc., Columbia, Maryland, USA

ABSTRACT Wastewater surveillance (WS), when coupled with advanced molecular techniques, offers near real-time monitoring of community-wide transmission of SARS-CoV-2 and allows assessing and mitigating COVID-19 outbreaks, by evaluating the total microbial assemblage in a community. Composite wastewater samples (24 h) were collected weekly from a manhole between December 2020 and November 2021 in Maryland, USA. RT-qPCR results showed concentrations of SARS-CoV-2 RNA recovered from wastewater samples reflected incidence of COVID-19 cases. When a drastic increase in COVID-19 was detected in February 2021, samples were selected for microbiome analysis (DNA metagenomics, RNA metatranscriptomics, and targeted SARS-CoV-2 sequencing). Targeted SARS-CoV-2 sequencing allowed for detection of important genetic mutations, such as spike: K417N, D614G, P681H, T716I, S982A, and D1118H, commonly associated with increased cell entry and reinfection. Microbiome analysis (DNA and RNA) provided important insight with respect to human health-related factors, including detection of pathogens and their virulence/antibiotic resistance genes. Specific microbial species comprising the wastewater microbiome correlated with incidence of SARS-CoV-2 RNA, suggesting potential association with SARS-CoV-2 infection. Climatic conditions, namely, temperature, were related to incidence of COVID-19 and detection of SARS-CoV-2 in wastewater, having been monitored as part of an environmental risk score assessment carried out in this study. In summary, the wastewater microbiome provides useful public health information, and hence, a valuable tool to proactively detect and characterize pathogenic agents circulating in a community. In effect, metagenomics of wastewater can serve as an early warning system for communicable diseases, by providing a larger source of information for health departments and public officials.

IMPORTANCE Traditionally, testing for COVID-19 is done by detecting SARS-CoV-2 in samples collected from nasal swabs and/or saliva. However, SARS-CoV-2 can also be detected in feces of infected individuals. Therefore, wastewater samples can be used to test all individuals of a community contributing to the sewage collection system, i.e., the infrastructure, such as gravity pipes, manholes, tanks, lift stations, control structures, and force mains, that collects used water from residential and commercial sources and conveys the flow to a wastewater treatment plant. Here, we profile community wastewater collected from a manhole, detect presence of SARS-CoV-2, identify genetic mutations of SARS-CoV-2, and perform COVID-19 risk score assessment of the study area. Using metagenomics analysis, we also detect other microorganisms

(bacteria, fungi, protists, and viruses) present in the samples. Results show that by analyzing all microorganisms present in wastewater, pathogens circulating in a community can provide an early warning for contagious diseases.

KEYWORDS COVID-19, SARS-CoV-2, wastewater, microbiome, wastewater-based epidemiology, wastewater monitoring, wastewater surveillance, RT-qPCR, DNA sequencing, RNA sequencing, metagenomics, whole metagenome sequencing, metatranscriptomics, shotgun sequencing, risk assessment, environmental risk

Global change, namely, climate variability, urbanization, rapid long-distance travel, and the projected human population growth, has increased the risk of infectious disease outbreaks (1). Pathogen cross-over from animal reservoirs into human populations increasingly has been reported to occur, both in frequency and diversity, over the last century, namely, Ebola, HIV/AIDS, West Nile Virus, and Middle East respiratory syndrome. The United Nations predicts further emergence of zoonotic diseases as habitats are ravaged by wildlife exploration, unsustainable farming practices, and climate variability (2).

In December 2019, several cases of an unknown pneumonia were reported in the Hebei province of Central China, marking the beginning of the global COVID-19 pandemic caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), and declared by the World Health Organization in March 2020 (3). By January 2022, over 400 million cases and more than 5 million deaths were reported (4). A significant finding of the COVID-19 pandemic to date has established aerosols as the primary transmission route of SARS-CoV-2 (5). However, the role of the ambient and built environment, as well as climate and weather processes, have begun to be considered important factors in promoting transmission, namely, via aerosolized viruses (6).

Progression of the pandemic has been monitored primarily by counting the number of individuals testing positive for presence of SARS-CoV-2, prompted mainly by onset of symptoms of the disease, generally appearing 2 to 4 days after exposure but, in some cases, taking up to 2 weeks to present after infection (7). However, in a large proportion of transmission events, presymptomatic (8) and even asymptomatic (9) transmission has been observed. Clinical testing programs provide information only on a subset of individuals that consent to testing (10). Over the counter self-tests, namely, antigenic tests which are considered less accurate than quantitative reverse transcriptase PCR (RT-qPCR), especially for asymptomatic subjects (11), have become increasingly popular but rely on individuals correctly performing the test and reporting positive results to health departments (12). Hence, the use of clinical data for estimations of COVID-19 prevalence has the potential to be biased, based on factors such as health-seeking behavior, under-testing of asymptomatic cases, inequitable access to testing, and selective testing mandates applying only to certain groups or regions (10).

Clinical diagnosis of COVID-19 is typically determined by detection of acute infection targets in the genome of SARS-CoV-2 present in nasopharyngeal, nasal, and saliva swab samples (13, 14). However, the virus can also be detected in specimens from other sites of infected individuals, notably in feces (15, 16). Proposed as a complement to clinical testing, wastewater surveillance (WS; also known as wastewater monitoring or wastewater-based epidemiology) of samples collected from treatment systems of communities has been used for early detection of community-wide disease prevalence, notably for poliovirus (17–19), noroviruses (20), flu (21), and recently COVID-19 (22). WS has also been employed to assess diverse factors influencing communitywide health, such as monitoring consumption of local diets, alcohol, illicit drugs, and tobacco, and evaluating exposure to hazardous chemicals and pharmaceuticals (23, 24). WS of confined populations, e.g., inhabitants of a single building (25) or university dormitory (26) or small community sewage collection systems (SCS) of a neighborhood (27), can be considered less biased, because the evaluation is pooled contributions of all individuals served by a given catchment area, compared with clinical testing, where only a minority of consenting individuals are tested routinely. Therefore, WS can be a useful public health tool to
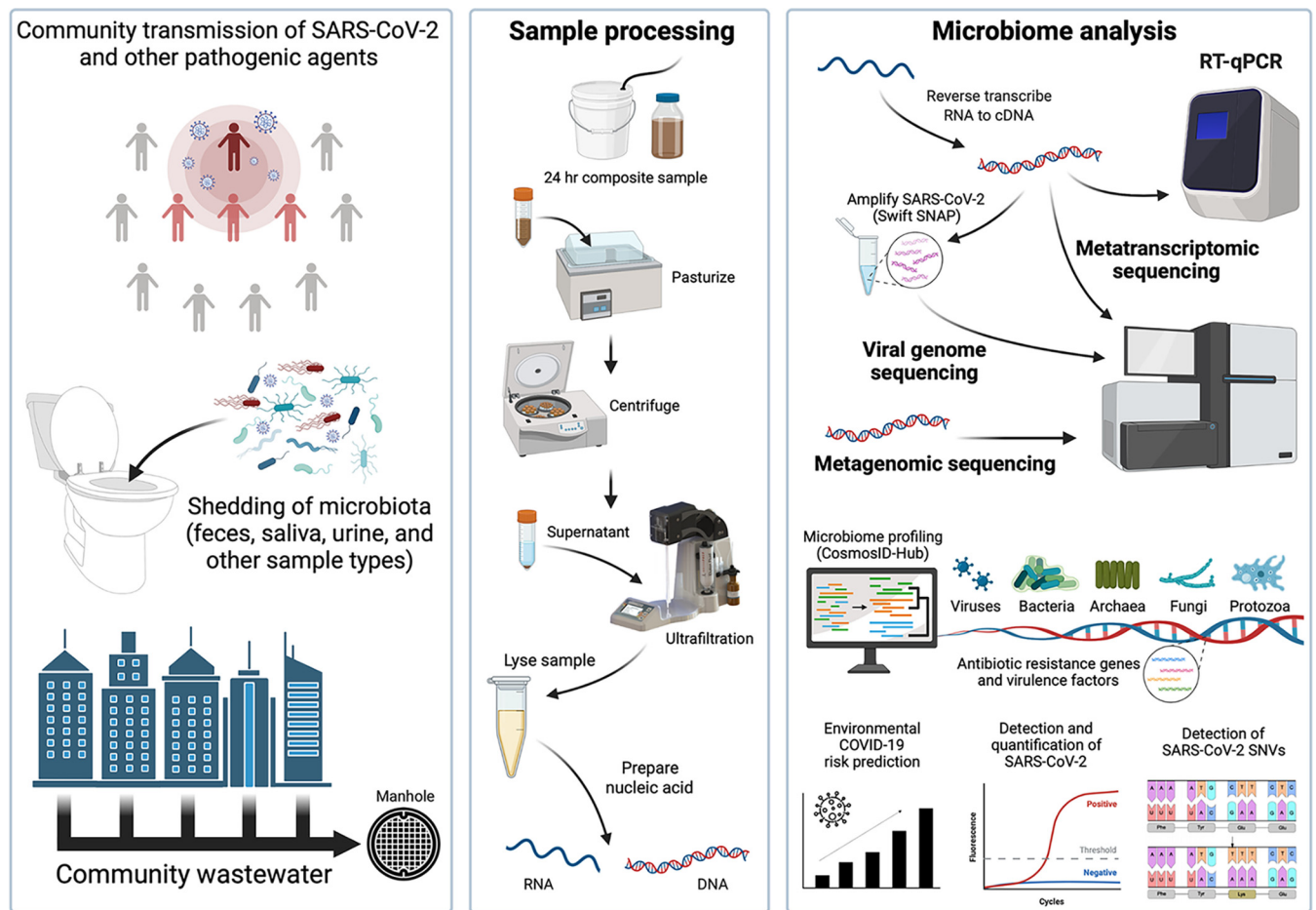
FIG 1 Schematic representation of sample processing. Image created using BioRender.

proactively study emergence and spread of COVID-19 (28). However, it is worth noting that WS, while able to monitor a given community, is most useful in municipalities with SCS infrastructure, and representation of rural towns may be diminished given that WS has the potential to miss those households served by septic systems.

RT-qPCR is widely recognized as the "gold standard" for detection of SARS-CoV-2 RNA in wastewater samples around the world (29–35). However, detection and enumeration only of SARS-CoV-2 markers does not provide comprehensive information, namely, the full range of enteric microorganisms and virulence and antimicrobial resistance (AMR) associated genes present in a community at a given time. Shotgun metagenomic sequencing (SMS) offers an effective WS tool (36–38) that allows bacterial, archaeal, viral, fungal, and protozoan microbiome community members to be identified to subspecies taxonomic level and characterized (39). While both RT-qPCR and SMS are established methods, the interaction of SARS-CoV-2 with other members of the microbial community, both in the patient and environment, remain understudied.

This investigation is the first to employ a complete molecular toolset, i.e., DNA metagenomics, RNA metatranscriptomics, targeted amplicon sequencing, and RT-qPCR, along with advanced bioinformatics, to analysis of community wastewater (Fig. 1). We demonstrated the feasibility and application of assessing trends of SARS-CoV-2 RNA in wastewater samples collected from a SCS in Maryland, USA, for approximately 1 year, from December 2020 to November 2021, and simultaneously profiled the microbiome component (DNA and RNA; bacteria, archaea, fungi, protists, and viruses, and presence of virulence and AMR associated genes) of a subset of samples. We showed significant positive correlation between environmental risk scores and prevalence of COVID-19 cases in the study area. Results suggest remote sensing can be used in near real-time to identify geographic regions with increased
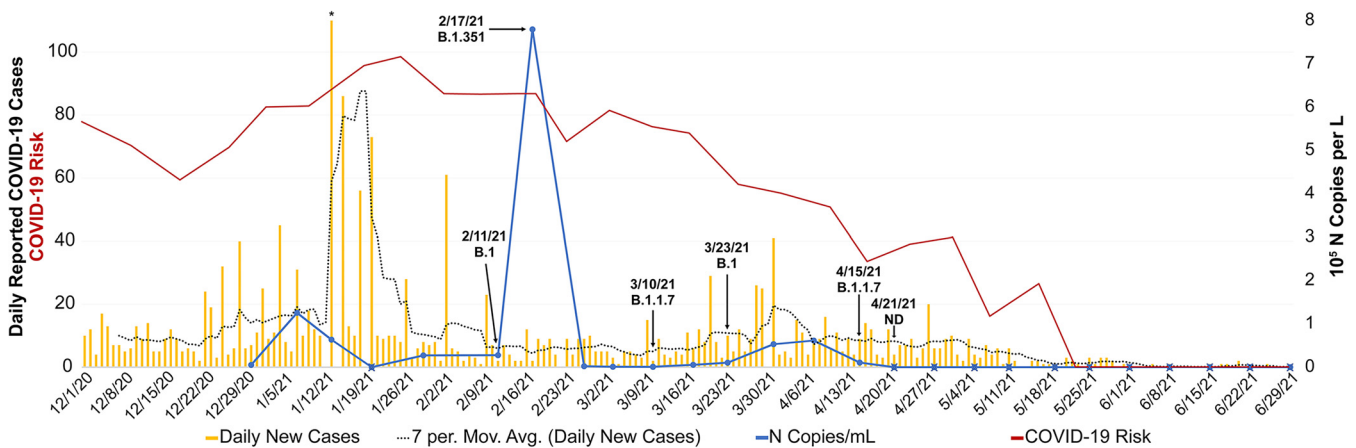
FIG 2 Reported COVID-19 cases and associated risk (left axis) and detection of SARS-CoV-2 N protein via RT-qPCR (right axis). Total daily number of reported COVID-19 cases for the ZIP code where the study took place was retrieved from the Maryland Department of Health (81). Asterisks represent truncated bar plot as 273 cases were reported on January 13, 2021. COVID-19 environmental predictive risk was calculated using ambient air temperature and dew point, as described previously (6). The risk score, a ratio between 0 and 1, is normalized on a scale of 0 to 100, with 100 being the highest risk of transmission. Wastewater samples were collected between December 30, 2020 and June 29, 2021. Concentration of SARS-CoV-2 was determined by RT-qPCR analysis. Blue circle indicates detection of SARS-CoV-2; blue circle with superimposed "X" indicates SARS-CoV-2 was not detected. Arrows indicate date of samples included for advanced molecular analysis (DNA metagenomics, RNA metatranscriptomics, and targeted RNA sequencing). SNVs commonly associated with known variations detected in consensus sequences are shown as SARS-CoV-2 Pango lineages. Gold bars, daily new cases; dotted black line, 7-day moving average of daily new cases; blue line, N copies/L; red line, COVID-19 risk; ND, not detected.

COVID-19 risk scores, and the wastewater microbiome is a useful public health tool to detect and characterize pathogenic agents circulating in communities, hence, an early warning system for communicable diseases.

RESULTS

SARS-CoV-2 RT-qPCR. (i) RT-qPCR quality controls. Calibration model performance parameters, including slope and y-intercept parameters, linearity (R^2), and Amplification efficiency (AE), are provided in the supporting information (Data set S2). Calibration model R^2 values for Nucleocapsid Phosphoprotein (N protein) were $\geq$ 0.98 and E-values ranged from 0.73 to 1.02 (average, 0.91). All no template controls (NTCs) were negative for detection of SARS-CoV-2 RNA. Based on internal control (IC) MS2 bacteriophage testing, amplification inhibition was not observed. Hence, all RNA extracts included in subsequent analysis exhibited negligible matrix interference and passed sample processing control.

(ii) Detection of SARS-CoV-2 RNA and association with epidemiological data. Detection of genetic markers against SARS-CoV-2 N protein was determined from wastewater concentrates, representing temporal sampling between December 30, 2020 and November 16, 2021. Of the 48 measurements, 24 (50%) were negative for detection of N protein, 23 were within the range of quantification (ROQ), and one (11/2/2021; 6.53×10^6 N copies/L) was above the ROQ. Excluding the measurement above ROQ, N protein was detected at the highest concentration during the week of February 17, 2021 (7.80×10^5 N copies/L), followed by October 5, 2021 (1.65×10^5 N copies/L) and at varying concentrations throughout the study period (Data set S2).

For the location where the study was conducted, the number of reported COVID-19 cases was higher during the winter compared with warmer months, with a spike in number of COVID-19 cases reported during the second week of January 2021. Generally, the amount of SARS-CoV-2 RNA detected in the community wastewater showed a similar pattern to the number of reported COVID-19 cases (Fig. 2). However, because the population served by this SCS represents *ca.* 3% of the total population included in the epidemiological reports, and viral shedding in feces of infected individuals is extremely variable, no statistical comparisons between the number of COVID-19 cases and concentrations of SARS-CoV-2 RNA were determined.

(iii) COVID-19 environmental predictive risk. Previous work showed a tolerable ambient temperature range of 17°C to 24°C corresponded to a decrease in number of

COVID-19 cases in the human population (6). A schematic representation of this generalizable hypothesis governing the dynamics of COVID-19 disease in a human population, where AT and DPT play a role in potential aerosolization of SARS-CoV-2 via particles in ambient and built environments (indoors) is shown in the supporting information (Fig. S1). During this investigation, the change in COVID-19 cases and detected concentration of SARS-CoV-2 RNA in wastewater followed a similar pattern to the change in the environmental risk scores (Fig. 2). Using parametric (Pearson) and nonparametric (Kendall tau) measures to determine the association between predicative risk scores and reported cases yielded statistically significant ($P < 0.01$) positive correlation of 0.58 and 0.54, respectively. On a weekly scale between December 2020 and November 2021 (47 weeks), when the risk score was greater than 0.5, SARS-CoV-2 was detected in wastewater samples 88% of the time (15/18 weeks), whereas when the risk score was below 0.5, SARS-CoV-2 was detected at low concentration 31% of the time (9/29 weeks). The greatest risk scores were detected during the end of January and the beginning of February, which coincided with highest number of COVID-19 cases. During the week of April 19, COVID-19 risk scores dropped below 0.5 and remained low throughout the summer, with slight increases in risk occurring in July and August. During this period of time only a few COVID-19 cases were reported and SARS-CoV-2 RNA was not detected by RT-qPCR in the wastewater samples. Subsequently, environmental risk scores increased during the fall, reaching above 0.5 in November 2021, which was mirrored by both increases in the number of COVID-19 cases and detection of SARS-CoV-2 at higher concentrations in wastewater samples. However, it is worth noting that various social and demographic factors, e.g., asymptomatic carriers, unreported home tests, human behavior, etc., impact the number of reported cases and could not be controlled for in this study. Therefore, we present calculated environmental COVID-19 risk scores as a ratio (Data set S3).

(iv) Characteristics of SARS-CoV-2 variants. Using a targeted sequencing approach, presence of SARS-CoV-2 was confirmed, and consensus sequences were characterized (Fig. 3). Presence of SARS-CoV-2 was confirmed in five of the six samples processed for targeted sequencing. The sample collected on April 21, 2021 was negative for presence of SARS-CoV-2 RNA, in agreement with results of the RT-qPCR analysis for these samples. An increase in the concentration of SARS-CoV-2 RNA in wastewater was observed on February 17, and genetic variations detected in the consensus sequence of this sample were similar to those associated with B.1.351 (Beta, V2). Prior to the February increase, various mutations commonly described in basal lineages were detected, and genetic variations associated with the B.1.1.7 (Alpha, V1) lineage were detected in consensus sequences recovered in April. Characteristics of SARS-CoV-2 variants recovered from wastewater, including the percent of reads that contained each mutation, are shown in Fig. 3. Across consensus sequences recovered from wastewater, between 20 (February 11, 2021) and 32 (April 15, 2021) single nucleotide variants (SNVs) were detected. All samples encoded a pyrimidine nucleotide variation at position 241 in the 5'-untranslated region of the virus genome (C241T). The synonymous C3037T (ORF1a: F924F) mutation was also detected at all time points. Two mutations were observed at genome position 21,801, including S:D80A (February 17, 2021) and S:D80G (March 10, 2021). Nonsynonymous mutations were detected frequently in the structural protein coding regions of spike (D614G, P681H, T716I, and S982A) and nucleocapsid (R203K, G204R, and S235F). However, SNVs in the envelope (E) and membrane (M) proteins were less common. The E:V57F mutation was observed in the sample collected on April 15, 2021, but no other E protein SNVs were detected. Detected membrane SNVs include I81T detected in the sample collected on February 11, 2021 and A97S detected on April 15, 2021. A genomic deletion was observed on March 10, 2021 (ORF1ab: GTCTGGTTTT11287G; penetrance = 100%), and three stop codons with 100% penetrance (ORF1ab:Q4618*, ORF8:Q26*, and OFR8:K67*) were detected in the sample collected on April 15, 2021. Other notable mutations detected in more than one sample include two synonymous mutations in ORF1ab (S216S and L4715L), four nonsynonymous mutations
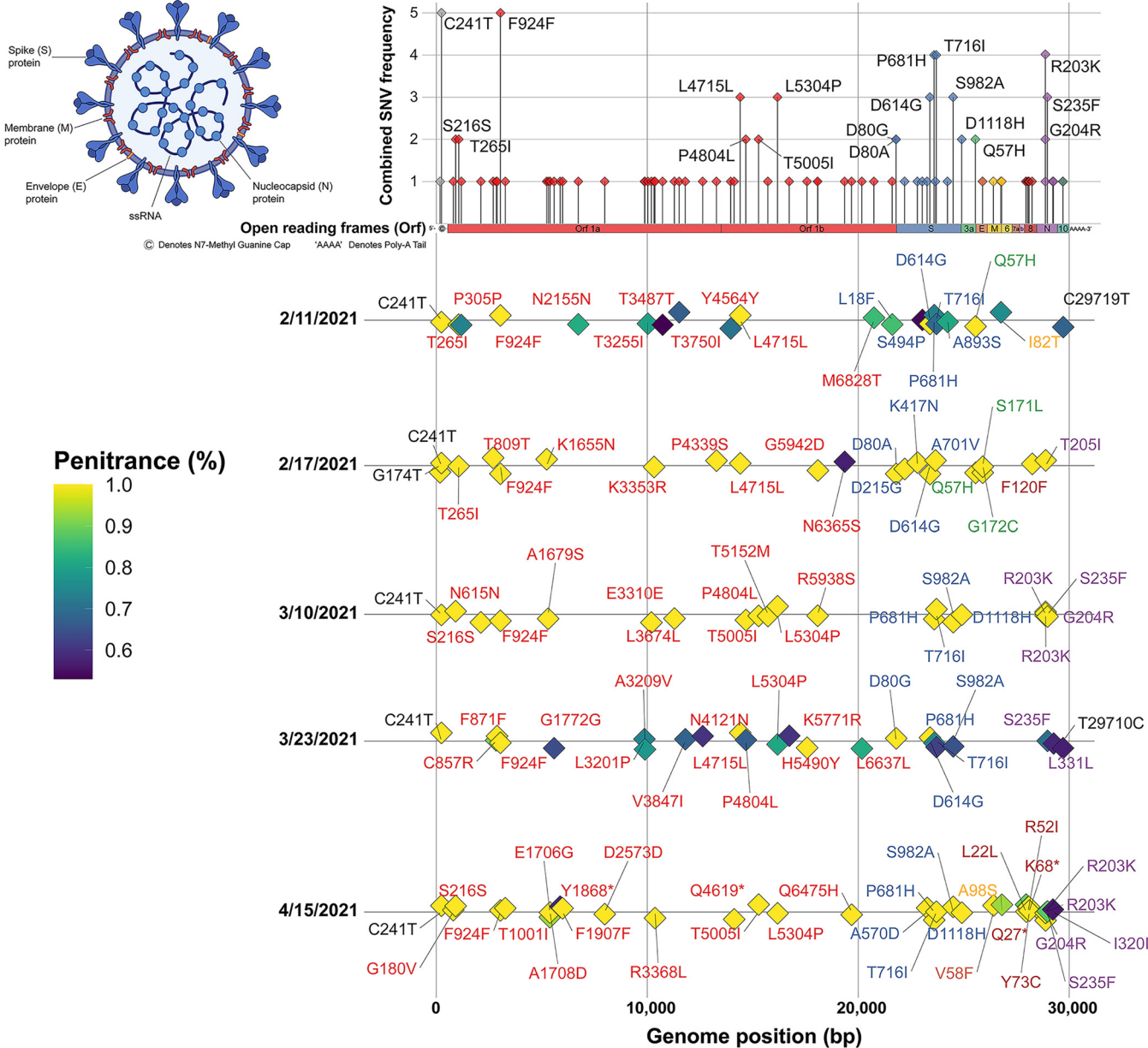
FIG 3 Penetrance of SARS-CoV-2 variants recovered from wastewater. Bar plot (top panel) showing the combined number of samples for which each genetic variant was detected. Heatmap (bottom panel) showing the penetrance (%) of reads associated with each genetic variant detected in SARS-CoV-2 consensus sequences recovered from wastewater relative to the total number of reads mapping to a given position. Inset shows SARS-CoV-2 structure.

detected in ORF1ab (T265I, 4804L, T5005I, L5304P) and one nonsynonymous SNV in ORF3a (Q57H).

Metagenomic data analysis. (i) Community microbiome. Shotgun metagenomics and metatranscriptomics using DNA and RNA prepared from the wastewater samples generated approximately 3.92×10^8 and 5.53×10^8 reads across the raw sequence libraries, with a mean of 65.3 and 92.1 million unique reads for metagenomic and metatranscriptomic samples, respectively (Data set S4). Alpha diversity was calculated using CHAO1 index and ranged from 1,059 to 1,358 (average = 1,180.5) for total bacteria and from 29 to 51 (average = 43.3) for total RNA viruses (Table S1). Bacterial communities in wastewater samples were analyzed by three-dimensional PCoA using Bray-Curtis dissimilarity index (Fig. S2A), where distance between points indicates degree of difference in bacterial DNA sequence composition. Each sample contained a relatively distinct bacterial composition.
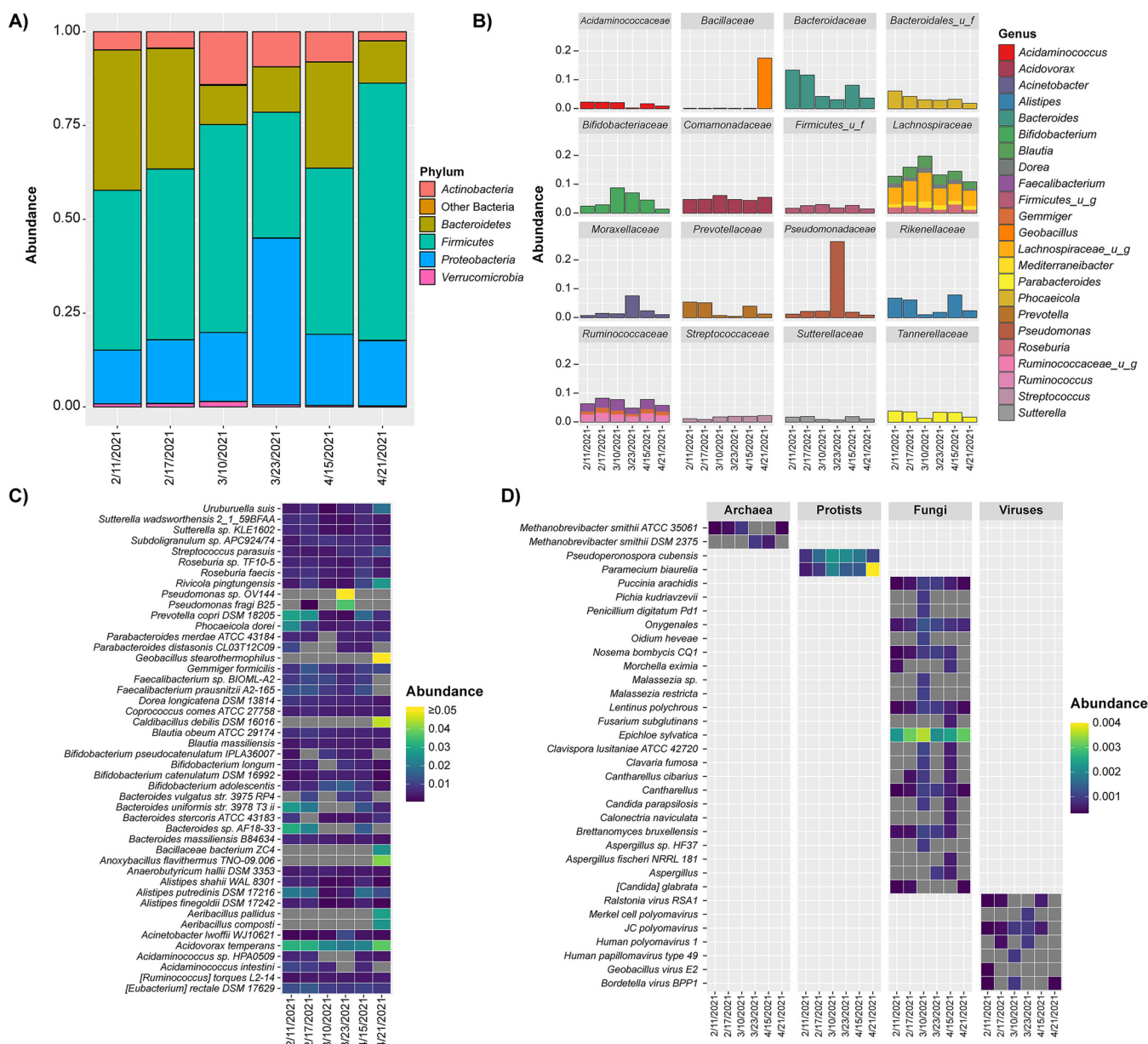
FIG 4 Microbiome profiles employing DNA metagenomic sequencing. (A) Stacked bar plot showing relative abundance of detected bacterial phyla. Top five most abundant phyla are shown, and all other phyla are labeled "other bacteria." (B) Stacked bar plots showing relative abundance of detected bacterial genera. Shown are the 20 most abundant genera grouped by family. (C) Heatmap showing universal kingdom relative abundance of most abundant species. Shown are most abundant taxa, representing >0.2% relative abundance. (D) Heatmap showing universal kingdom relative abundance of detected archaea, protists, fungi, and viruses.

However, samples collected during February 2021, clustered more closely compared with those collected during March and early April. Interestingly, the April 21, 2021 sample that tested negative for SARS-CoV-2 RNA by RT-qPCR did not cluster with any of the other samples. Bacteria, archaea, fungi, protozoa, and viruses identified by DNA metagenomics are shown in Fig. 4, representing relative abundance (RA) of microbial species.

For all wastewater samples, significant differences were observed in the microbiota (Fig. 4, 5A). While SARS-CoV-2 RNA was detected in five of six samples by RT-qPCR, the virus was not detected by RNA metatranscriptomics in any of the samples. However, there appeared to be a relation between the number of shared and exclusive bacteria, i.e., unique taxa detected in a set of samples and not detected in other samples, and the lineage of SARS-CoV-2 that was detected. A Venn diagram representing bacteria profiles, with respect to detection of genetic mutations commonly associated with
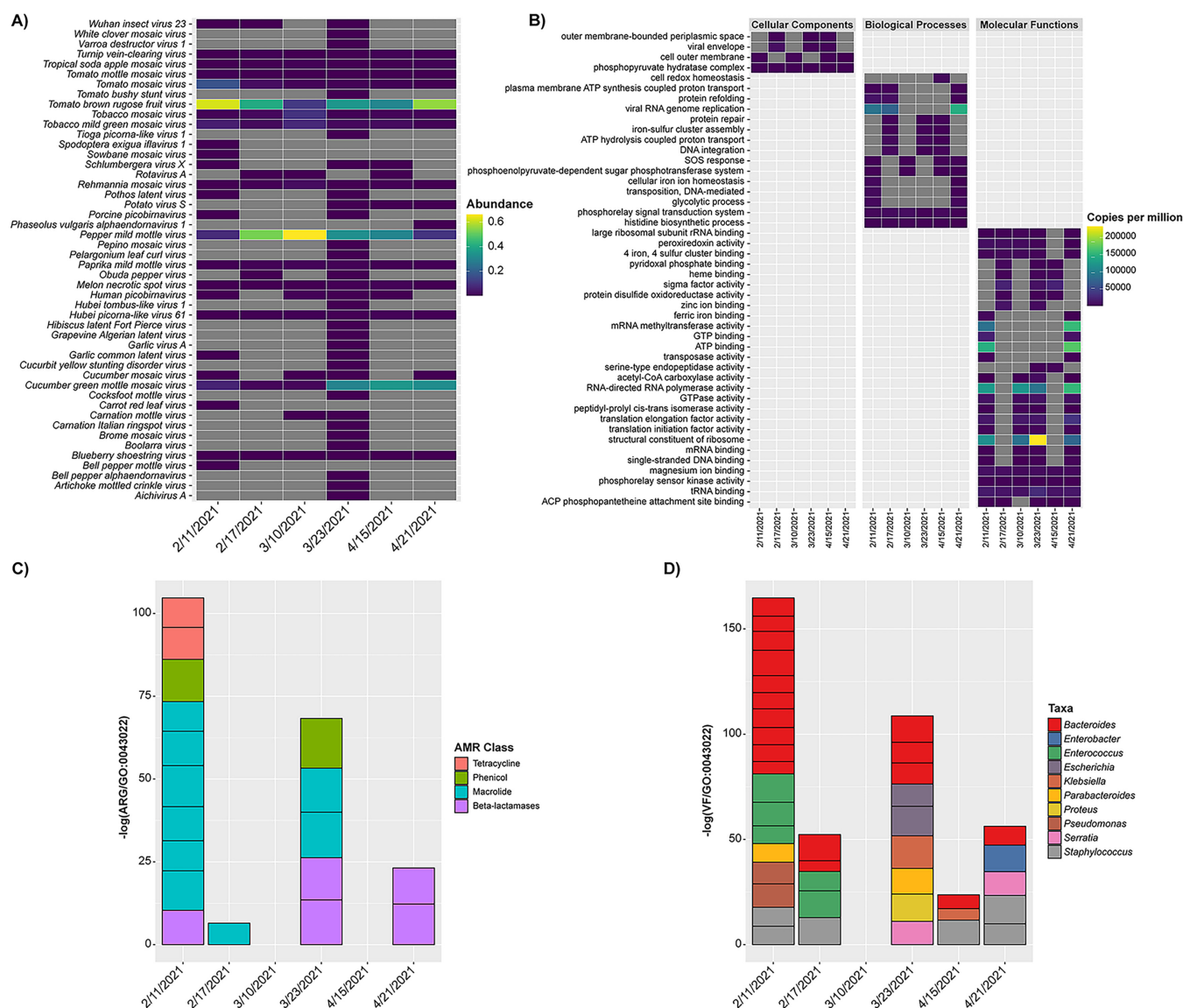
FIG 5 Microbiome profiles employing RNA metatranscriptomic sequencing. (A) Heatmap showing relative abundance of detected RNA viruses. (B) Stacked bar plot showing expression of detected AMR genes. Shown are individual AMR genes colored by class. Qualitative expression is shown relative to GO:0043022. (C) Stacked bar plot showing expression of detected virulence factors. Shown are individual virulence-associated genes colored by taxa. Qualitative expression is shown relative to GO:0043022. (D) Heatmap showing relative expression of detected gene ontology terms. Shown are relative expression (copies per million reads) of the most abundant gene ontology terms grouped by category.

SARS-CoV-2 variants of concern, is included in the supporting information (Fig. S2B). An increased number of exclusive bacteria was observed in samples corresponding to detection of Alpha variants (486 spp.) compared with samples where Beta variants were detected or samples where SARS-CoV-2 was not detected, which contained 137 and 127 exclusive species, respectively. The sample negative for SARS-CoV-2, collected on April 21, 2021, contained fewer shared bacteria with samples where Beta variants were detected (26 spp.) compared with samples where Alpha variants were detected (127 spp.). A total of 167 shared species was observed between samples with Alpha variants and samples with Beta variants. Analysis of the RNA viral community showed the sample collected on March 23 contained the greatest number of exclusive taxa (18), while the other samples contained relatively few exclusive taxa (≤5).

A core microbiome, i.e., taxa detected in all samples, was observed for all of the wastewater samples, with 544 bacterial species detected by DNA metagenomics and 14 RNA viruses detected by metatranscriptomics. RNA metatranscriptomics also allows

for detection of active members of the microbiome. Overall, a similar core microbiome was detected between RNA and DNA sequencing. Microbiome profiles employing RNA metatranscriptomics, including the most abundant genera detected by both sequencing methods is included in the supporting information (Fig. S3). Generally, core bacteria detected were assigned to phyla *Firmicutes*, *Proteobacteria*, *Bacteroidetes*, or *Actinobacteria*. At the family level, *Lachnospiraceae*, *Bacteroidaceae*, and *Ruminoccoccaceae* were the most abundant core bacteria detected. However, there was a substantial increase in *Pseudomonadaceae*, namely, members of the genus *Pseudomonas*, detected in the March 23, 2021 sample (>20% RA), and *Bacillaceae*, primarily members of the genus *Geobacillus*, detected in the sample negative for SARS-CoV-2 (April 21; >15% RA). Across all samples, the most abundant core microbial species were *Acidovorax temperans* (bacteria; *ca.* 3 to 4% RA) and *Epichloe sylvatica* (fungi; *ca.* 0.02 to 0.04% RA). Pepper mild mottle virus and tomato brown rugose fruit virus were the most abundant RNA viruses detected in the core wastewater microbiome. Notable exclusive microbial species detected in the sample negative for SARS-CoV-2 (April 21) include *Geobacillus stearothermophilus*, *Caldibacillus debilis*, and *Anoxybacillus flavithermus*. With respect to samples positive for detection of SARS-CoV-2 RNA, the results showed a minimal number of exclusive microbial species detected in samples at high SARS-CoV-2 concentrations compared with lower concentrations.

(ii) Pathogenic wastewater-associated microorganisms. Incorporating an *in situ* control for DNA metagenomics allowed for absolute cell number quantification of detected microbiota following metagenomic profiling (Fig. S4). The sample collected on February 17, 2021 contained the highest concentration of microbial cells/L (1.34×10^{10}), followed by April 21 (1.00×10^{10}) and February 11 (8.42×10^9). The lowest concentration was detected on March 10, 2021 (1.23×10^9 cell/L). Most microbial cells were profiled as bacteria (min = 1.21×10^9, March 10; max = 1.33×10^{10}, February 17; average = 6.30×10^9), followed by fungi (min = 5.39×10^6, March 23; max = 5.97×10^7, February 17; average = 2.64×10^7) and protists (min = 5.39×10^6, March 23; max = 5.47×10^7, April 21; average = 1.89×10^7). DNA viruses were detected in all samples at low abundance ($<2 \times 10^5$ cells/L), and archaea were rarely detected ($\leq 1 \times 10^3$ cells/L). Table 1 shows the number of multiple pathogenic microorganisms detected in wastewater samples. Members of the *Enterobacteriaceae* family, namely, *Escherichia coli* and *Klebsiella pneumoniae*, and were detected in all samples. *Aeromonas hydrophilia* was also detected in all samples, but at lower cell number. *Vibrio cholerae* was detected at a concentration of 1.16×10^6 cells/L in the February 11 sample, and at a lower cell number on February 17 and March 23. Human polyomavirus 2 was the most common DNA virus, detected in all samples except the sample collected on April 21.

(iii) Functional analysis of wastewater microbiome. RNA metatranscriptomics allows biological insight into gene expression and functional analysis. Fig. 5B shows the most abundant gene ontology (GO) terms profiled in each sample. Genes associated with cellular components (GO:0000015, phosphopyruvate hydratase complex), biological processes (GO:0000160, phosphorelay signal transduction system, and GO:0070180, large ribosomal subunit rRNA binding), and molecular functions (GO:0000287, magnesium ion binding, GO:0000155, phosphorelay sensor kinase activity, and GO:0000049, tRNA binding) were detected in all samples. Genes associated with viral RNA genome replication (GO:0019079) were detected in three wastewater samples (February 11 and 17, and April 21).

(iv) Community resistome and virulome. AMR genes detected in the wastewater samples are shown in Fig. 5C. The largest number of AMR genes was detected in the February 11, 2021 sample (*n* = 10), followed by March 23, 2021 (*n* = 5). AMR genes associated with macrolide and beta-lactamase antibiotic classes were dominant in these samples. During the time when an increase in SARS-CoV-2 RNA was detected (February 17, 2021), macrolide was the only antibiotic class detected, at low abundance, while beta-lactamase was the only AMR class detected in the sample negative for SARS-CoV-2 RNA (April 21, 2021). A similar pattern was observed for total number of virulence associated genes (Fig. 5D). Highest

TABLE 1 Quantification of microorganisms employing DNA metagenomics[a]

Genus	Species	Illness/description	2/11/21	2/17/21	3/10/21	3/23/21	4/15/21	4/21/21
Bacteria								
Aeromonas	A. hydrophilia	Gastroenteritis; septicemia	4.22×10^5	6.93×10^5	2.5×10^4	3.3×10^4	5.2×10^4	3.68×10^5
Bacteroides	B. thetaiotaomicron	FIB	5.21×10^7	6.33×10^7	2.04×10^6	2.54×10^6	1.44×10^7	5.88×10^6
Bifidobacterium	B. adolescentis	FIB	5.29×10^7	1.02×10^8	2.89×10^7	4.83×10^7	4.33×10^7	3.6×10^7
	B. dentium	FIB	1.16×10^6	6.93×10^5	2.24×10^5	7.91×10^5	1.19×10^6	1.23×10^5
Burkholderia	B. vietnamiensis	Respiratory illness	1.05×10^5	-	-	-	5.2×10^4	-
Campylobacter	Campylobacter spp.[b]	Campylobacteriosis	1.05×10^5	1.73×10^5	-	-	-	-
Escherichia	E. coli	FIB	1.72×10^9	1.01×10^9	1.66×10^6	2.18×10^6	2.79×10^6	8.09×10^6
Enterococcus	E. avium	FIB	2.64×10^6	3.47×10^5	-	-	-	1.23×10^5
	E. casseliflavus	FIB	2.11×10^5	-	-	6.6×10^4	-	-
	E. faecalis	FIB	9.49×10^5	-	1.74×10^5	5.6×10^5	3.1×10^5	1.23×10^5
	E. faecium	FIB	1.37×10^6	6.93×10^5	2.48×10^5	4.81×10^6	4.31×10^6	4.44×10^6
Klebsiella	K. pneumoniae	Pneumonia; UTI	6.38×10^6	9.27×10^6	1.78×10^6	7.35×10^6	-	-
Legionella	L. septentrionalis	Respiratory illness	-	-	-	6.6×10^4	5.2×10^4	-
Methanobrevibacter	M. Smithii	FIB	7.38×10^5	2.08×10^6	1.74×10^5	8.24×10^5	1.40×10^6	1.35×10^6
Mycobacterium	Nontuberculosis Mycobacterium[c]	Respiratory illness; skin infections	-	-	-	-	-	1.23×10^5
Salmonella	S. enterica	Gastroenteritis	5.27×10^5	3.47×10^5	-	-	5.2×10^4	1.23×10^5
Shigella	S. dysenteriae	Dysentery	1.05×10^5	-	-	-	-	-
Vibrio	V. cholerae	Cholera	1.16×10^6	1.73×10^5	-	3.3×10^4	-	-
Viruses								
Papillomaviridae	Human papillomavirus	HPV	-	-	4×10^3	-	-	-
Polyomaviridae	Human polyomavirus 2[d]	Progressive multifocal leukoencephalopathy	1.06×10^5	1.34×10^5	2.1×10^4	2×10^4	5.2×10^4	-
	Merkel cell polyomavirus	Merkel cell carcinoma	-	-	-	3×10^3	-	-
Fungi								
Malassenzia	M. restricta	Skin infection	-	-	2.34×10^5	-	-	-
Candida	Candida spp.[e]	Candidiasis	1.97×10^5	2.48×10^5	4.85×10^5	-	1.40×10^5	2.01×10^5
Aspergillus	Aspergillus spp.[f]	Lung infection	-	-	4.6×10^4	-	1.25×10^5	-

[a]Number of each taxon was normalized to cell number of an *in situ* control (ZymoBIOMICS High Microbial Load Spike-in Control I; Zymo Research, Irvine, CA, USA) comprised of *Imtechella halotolerans* (Gram-negative) and *Allobacillus halotolerans* (Gram-positive). Quantification of microbiota is shown as cells/L. UTI, Urinary tract infection; FIB, fecal indicator bacterium; HPV, Human papillomavirus; -, not detected.
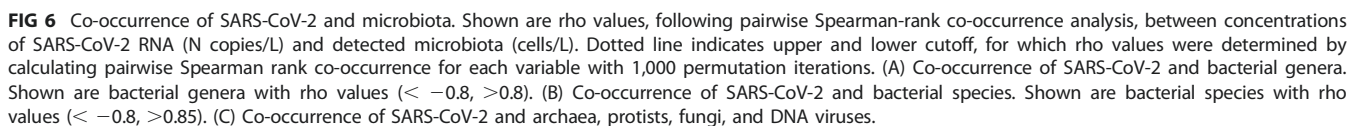[b]*Campylobacter gracilis, Campylobacter upsaliensis, Campylobacter ureolyticus.*
[c]*Mycobacterium avium, Mycobacterium pseudokansasii.*
[d]JC polyomavirus.
[e]*Clavispora lusitaniae, Candida parapsilosis, Candida glabrata.*
[f]*Aspergillus fischeri, Aspergillus spp. HF37.*

FIG 6 Co-occurrence of SARS-CoV-2 and microbiota. Shown are rho values, following pairwise Spearman-rank co-occurrence analysis, between concentrations of SARS-CoV-2 RNA (N copies/L) and detected microbiota (cells/L). Dotted line indicates upper and lower cutoff, for which rho values were determined by calculating pairwise Spearman rank co-occurrence for each variable with 1,000 permutation iterations. (A) Co-occurrence of SARS-CoV-2 and bacterial genera. Shown are bacterial genera with rho values (< -0.8, >0.8). (B) Co-occurrence of SARS-CoV-2 and bacterial species. Shown are bacterial species with rho values (< -0.8, >0.85). (C) Co-occurrence of SARS-CoV-2 and archaea, protists, fungi, and DNA viruses.

frequencies of occurrence of virulence factors were detected in samples collected on February 11, 2021 ($n = 18$) and March 23, 2021 ($n = 9$). Virulence factors (VF) associated with *Bacteroides* were dominant in these samples. VF associated with *Staphylococcus* were detected in four of the six samples. No VF or AMR associated genes were detected in the March 10, 2021 sample, collected the week following increase in detected SARS-CoV-2 RNA on February 17, 2021.

(v) Co-occurrence of SARS-CoV-2 and detected microbiota. Normalizing DNA sequencing reads to the *in situ* control allowed successful conversion of sequencing relative abundance to cells per L, a comparable unit of conversion to concentration of SARS-CoV-2 RNA detected in wastewater following RT-qPCR. After calculating pairwise Spearman rank co-occurrence for each variable with 1,000 permutation iterations, upper and lower quantiles for permuted rho values between concentration of SARS-CoV-2 (N copies per L) and detected microbiota (cells per L) were calculated, including bacterial genera (-0.87, 0.88), species (-0.86, 0.88), and combined kingdom (-0.86, 0.88). Microbiota with highest co-occurrence correlation is shown in Fig. 6. Both positive and negative correlations were observed. However, no microbiota associated with negative correlation were significant. Of the 345 genera detected across samples, 14 were associated with positive correlation of SARS-CoV-2 RNA and greatest correlation was associated with *Bifidobacterium*, *Leclercia*, *Pyramidobacter*, *Tannerella*, *Massilimaliae*, and *Erythrobacter*. A total of 39 bacterial species were associated with significant positive correlations, most notably, *Bordetella bronchiseptica*, *Enterobacter cloacae*, and *Leclercia adecarboxylata*. However, various *Bifidobacterium spp.*, *Pseudomonas spp.*, *Bacteroides spp.*, and *Prevotella spp.*, also showed significant positive correlations. As a result of combined kingdom analysis, JC polyomavirus was the only taxa to show correlation with SARS-CoV-2, and correlations with archaea, protists, fungi, and other DNA viruses were not determined to be significant.

DISCUSSION

Wastewater-based epidemiology and COVID-19. An outbreak traditionally is monitored by testing individuals presenting symptoms of the disease. However, disease incidence thus determined can be delayed or underreported because epidemiological reports rely on infected people seeking medical aid and medical practitioners reporting test results. The Centers for Disease Control and Prevention (CDC) reported *ca.* 41% of U.S. adults delayed or avoided medical care because of concerns about COVID-19 (98). Most testing methods for COVID-19 focused on RT-qPCR, an effective method for

detecting very small amounts of SARS-CoV-2 RNA. However, such molecular tests typically are performed by centralized high-complexity laboratories with special equipment and scaling to meet demands of public health testing has been a major challenge globally (99). During the COVID-19 pandemic, the testing capacity of many public health systems was overwhelmed and a global shortage of testing reagents and supplies occurred (100). These major bottlenecks in testing led to turnaround times exceeding 5 to 10 days in some regions of the U.S., making such tests unreliable in preventing disease transmission (99). Over-the-counter diagnostic tests, such as rapid antigen detection assays, are intrinsically quicker and less laborious than RT-qPCR. However, antigenic assays rely on individuals correctly performing self-tests and reporting positive results and are considered less accurate than RT-qPCR, especially for asymptomatic subjects (11, 12). These issues led to underreported cases, along with increases in false positive and false negative results (101).

Shortly after SARS-CoV-2 was reported to be present in human waste (102), monitoring the virus in wastewater gained attention as an effective epidemiological tool for tracking spatial and temporal dynamics of COVID-19 at the community level (22, 29–35). A few studies employing WS detected community transmission of SARS-CoV-2 prior to onset of new COVID-19 cases (31, 103, 104), suggesting longitudinal wastewater analysis could be used to detect viral shedding of infected (symptomatic and asymptomatic) individuals and identify trends sooner than clinical case reporting. However, it has been suggested that WS serves as a true early warning system only when community incidence of COVID-19 is low and clinical testing in the region of interest is scarce or deficient (105). Furthermore, due to the nature of analyzing pooled samples, e.g., samples collected from a SCS, individual contributions to a positive sample cannot be determined, and WS is not as effectively employed for contact tracing as clinical testing of individuals. Hence, WS is best suited as a complement to clinical testing, serving as an independent indicator of disease prevalence (10).

High-throughput molecular methods for microbiome analysis, e.g., DNA metagenomics and RNA metatranscriptomics, and SNV analysis of pathogens in wastewater allows microbiome comparisons between regions and detection of sources of infection and transmission dynamics (36–38, 106, 40). However, complete microbiome analysis of wastewater for community surveillance is very much in its infancy. With continued emergence of new SARS-CoV-2 variants, it is increasingly evident that COVID-19 will not be eradicated globally. While human-to-human interaction is the major route of transmission, and wastewater testing can elucidate community epidemiology, understanding the possible role of climatic and weather processes in accelerating such interactions is also an important component. Elucidating the role of environmental parameters and how they interact with the microbiome can be very helpful in developing a climate-informed understanding, translatable to predictive models for estimating risk of the infection in humans.

Detection and characterization of SARS-CoV-2. In the study reported here, longitudinal wastewater analysis was employed for community surveillance (Fig. 1), with a manhole (Fig. 7) selected for sample collection because it offers distinct advantages over testing at a wastewater treatment plant (WWTP). SCSs are designed as tree networks, with the community (homes, schools, businesses, industries, etc.) flowing raw sewage and excess water into the network, which eventually feeds downstream into a local WWTP. Sampling at an individual manhole provides better understanding of how microbial communities are associated, distributed, and vary temporally within a geographically defined area, without being diluted with sewage from other geographic areas or other kinds of wastes, e.g., industrial. Generally, the amount of SARS-CoV-2 RNA detected in the community wastewater showed a pattern similar to the number of reported COVID-19 cases in the study area (Fig. 2). Previous studies have reported detection of an increased RNA load in wastewater prior to manifestation of positive cases, *ca.* 2 and 8 days (22), and prior to other pandemic indicators, e.g., COVID-19 hospital and intensive care unit administrations, up to 9 days (107). However, shedding of
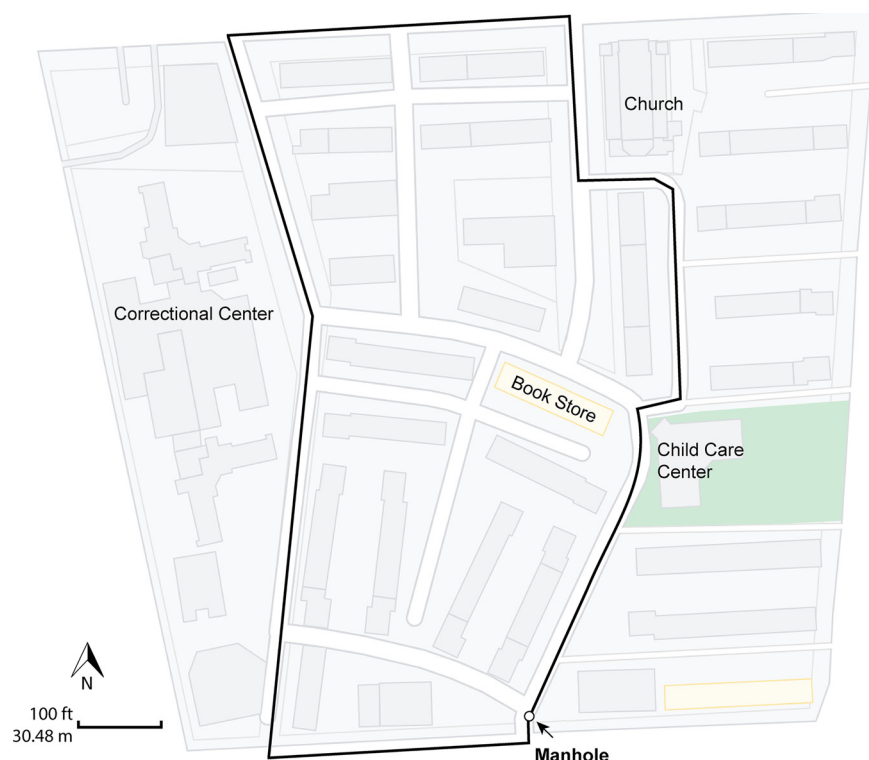
FIG 7 Map of sewage collection system. Map shows location of manhole where wastewater samples were collected. Black line indicates boundary of the area serviced by the sewer collection system. Arrow indicates location of manhole where samples were collected. The manhole is downstream of the area serviced, and wastewater flows from north to southeast. Scale bar corresponds to 100 ft (30.48 m).

SARS-CoV-2 in feces is variable both in viral load and time course (108). Furthermore, the population served by this SCS represents *ca.* 3% of the total population covered by the epidemiological reports. It is also worth noting that the rate of clinical COVID-19 testing in MD varied in frequency throughout the study (81) and, despite the likelihood of increased number of reported COVID-19 cases reflecting active spread of the virus, the potential for a confounding influence of variations in testing frequency must be considered (109), which complicates calculating correlations between RNA concentration and number of infected individuals. Therefore, calculation of number of COVID-19 cases versus SARS-CoV-2 RNA concentrations was not done. However, RT-qPCR data are useful to assess trends in abundance of SARS-CoV-2 RNA detected in the SCS. Excluding the measurement above ROQ, the highest concentration of SARS-CoV-2 RNA detected over the course of the 48-week study (7.80×10^5 N copies/L) was similar to other studies in the United States. employing RT-qPCR to detect SARS-CoV-2 RNA in raw wastewater influent, with *ca.* 10^5 or 10^7 N copies/L reported (22, 35).

Targeted amplicon sequencing for SARS-CoV-2 for variant calling among population-based samples, such as wastewater, compared with SNV analysis of single isolates recovered from patients and the environment, results in a consensus sequence that will identify the most prevalent genetic mutations within a sample. However, the resulting consensus has the potential to mask minor variants or merge mutations from multiple variants into a single sequence. Nonetheless, population-based variant calling in wastewater has recently been used to track viral evolution occurring in large WWTPs (110, 111). Yet, only a few studies have used it for analysis at the small community level, e.g., single building (25) or university dormitory (26). Those studies and others reported viral concentrations of *ca.* 10^6 genome copies/L to be sufficient for sequencing SARS-CoV-2. Using methods reported here, we successfully employed targeted amplicon sequencing for detection and characterization of SARS-CoV-2 in wastewater

samples with viral concentrations as low as 10^3 N copies/L. However, calculations across studies are variable and depend on sample collection and preparation methods, as well as the sequencing method used, and, therefore, are not directly comparable. We were not able to detect SARS-CoV-2 using shotgun RNA metatranscriptomics, even when viral concentrations were relatively high (7.80×10^5 N copies/L). That is, SARS-CoV-2 detection was successful when more sensitive methods were used, i.e., targeted amplicon sequencing and RT-qPCR. Spurbeck et al. (25) employed RT-qPCR, targeted sequencing, and untargeted metatranscriptomics to examine the local dynamics of SARS-CoV-2 strains and identify other pathogens circulating in the community, and concluded that alternative sequencing approaches may be required to consistently detect SARS-CoV-2 for biosurveillance, namely, due to low abundance of viral RNA compared to total RNA present in wastewater. Hence, multiple methods should be considered for wastewater surveillance.

On January 12, 2021, the first observation of the alpha variant (B.1.1.7) was reported in Maryland (112). Interestingly, this is around the time that a dramatic increase in the number of cases was observed in the study area. Targeted amplicon sequencing (Fig. 3) confirmed the presence of genetic mutations associated with this variant and others. An increase in SARS-CoV-2 RNA was detected in the wastewater samples in early February 2021, and SNVs commonly associated with the B.1.351 (Beta) variant were also detected. In April 2021, SNVs associated with B.1.1.7 (Alpha) were detected along with genetic mutations associated other SARS-CoV-2 lineages prior to the February increase and on March 23, 2021, indicating that multiple variants may have been circulating in the region. Similar observations have been made during genomic surveillance of SARS-CoV-2 in Delaware, USA (113), showing how quickly Alpha variants spread throughout a community. These data, coupled with what is known about increased transmissibility of B.1.617.2 (Delta) and B.1.1.529 (Omicron), highlight the importance of this type of proactive genomic surveillance.

Although most mutations may be deleterious or neutral, a few have the potential to affect viral function, potentially altering infectivity and/or disease severity (76). The SARS-CoV-2 spike protein recognizes host cells and is the main target of the human immune response, and some of the protein mutations have been linked to heightened infectivity and the ability to evade infection-blocking antibodies (114). During the onset of COVID-19, two major locally transmitted outbreaks occurred in China, one in Wuhan from December 2019 to April 2020, another in Beijing-Xinfadi in June 2020 (115). Interestingly, SARS-CoV-2 sequences recovered from Beijing-Xinfadi can be distinguished by a single nucleotide at position 241, a conserved pyrimidine nucleotide located in the loop region of stem-loop 5B of the 5'-untranslated region (116), compared with those collected in Wuhan. Viral isolates from patients and the environment generally encode uridine, whereas the viruses isolated in Wuhan encode a cytidine nucleotide at the corresponding position (115). In this study, the C241T variation was observed at all time points. The S protein mutation D614G, detected in samples, collected on February 11 and 17, 2021 and March 23, 2021, is an important genetic mutation associated with increased SARS-CoV-2 infectivity (117). Interestingly, detection of S:D614G coincided with the detection of higher SARS-CoV-2 RNA concentrations (Data set S2). The P681H mutation, detected at four time points in this study—February 11, 2021, March 10 and 23, 2021, and April 15, 2021—is common to B.1.1.7 isolates and has been characterized as part of the spike S1/S2 cleavage site, with potential to enhance viral cell entry (118). Other spike SNVs detected in this study, namely, T716I, S982A, and D1118H, have also been associated with natural mutations observed in B.1.1.7 isolates that show greater transmissibility than wild-type SARS-CoV-2 isolates (119). However, it is worth noting that the spike mutation K417N was detected in the sample collected on February 17, 2021, which corresponds to the time point with highest concentration of SARS-CoV-2 RNA detected via RT-qPCR (Fig. 2), is commonly associated with B.1.351 isolates of high concern because they have potential to compromise neutralization generated by previous infection or vaccination (120, 121).

Environmental COVID-19 risk score analysis and wastewater surveillance. Usmani et al. (6) reported transmission of SARS-CoV-2 in the human population increased when the ambient air temperature (AT) is out of the comfort zone (17°C to 24°C). When temperatures are uncomfortable, i.e., outside the comfort zone, human behavior in ambient and built environments is related to increased COVID-19 transmission (Fig. S1). In addition to AT, increased transmission of SARS-CoV-2 was reported in cold and dry environments. In this study, the association of AT and DPT with incidence of COVID-19 in the human population was related to detection of SARS-CoV-2 in wastewater (Fig. 2). Results showed a statistically significant ($P < 0.1$) positive correlation between environmental risk scores and reported cases. When the environmental risk was greater than 0.5, more COVID-19 cases were reported, and SARS-CoV-2 was detected 88% of the time (via RT-qPCR) in wastewater. *Per contra*, when the risk score dropped below 0.5, only a few COVID-19 cases were reported, and SARS-CoV-2 RNA was rarely (31% of the time) detected in wastewater and was at low abundance. Hence, incidence of SARS-CoV-2 in the human population examined here appears associated with AT and DPT, and hence, WS as a COVID-19 risk score can be a useful public health tool. Because the environmental predictive risk model has a spatial resolution of 4 km $\times$ 4 km, the proposed model is best suited to identify regions with heightened environmental risk. In turn, WS can be used to provide ground truth for presence of SARS-CoV-2, along with other pathogens, and determine potential for disease transmission at a local level, e.g., community or neighborhood. Employing such an approach can reduce human labor and identify targeted regions where there may be a need for WS to capture finer spatial resolution of disease transmission, thereby serving as an effective tool to mitigate disease.

Wastewater microbiome. (i) Microbial diversity. DNA metagenomics can be used to detect and identify the bacteria, virus, fungi, and protists comprising the microbial community of environmental samples (39), with viability or infectious potential of the detected microorganisms requiring additional analyses of metabolic activity. RNA metatranscriptomics is useful for detecting RNA viruses, but can also give insight into the functional profile of the microbiome (122). Both DNA metagenomics and RNA metatranscriptomics have been used to explore the microbiome of wastewater (36–38, 123).

Bacteria dominated the microbiome of wastewater samples in this study (Fig. S4), consistent with prior observations of microbiomes of samples collected at WWTPs (36, 38, 124). Similarly, the bacterial phyla (Fig. 4A) were those commonly detected in WWTPs (124), namely, *Firmicutes*, *Proteobacteria*, and *Bacteroidetes*, as the most abundant. Similarly, *Ruminococcaceae*, a microbial family detected in most WWTP samples (124), was present in all samples. However, *Bacteroidaceae*, *Bifidobacteriaceae*, *Comamonadaceae*, and *Lachnospiraceae*, were also dominant in the wastewater microbiome. Genera (Fig. 4B) included *Bacteroides*, *Bifidobacterium*, *Acidovorax*, *Acinetobacter*, *Pseudomonas*, *Alistipes*, and *Prevotella*, similar to reports of other investigators (36, 38, 124). Previous studies suggest *Trichococcus* spp. are dominant members of the microbial community in urban sewers (125), but species of this genus were present at lower RA in this study. During a metagenomic survey of wastewater in the United Kingdom, *Arcobacter* and *Aeromonas* were identified as predominant wastewater indicators (126) and those genera were detected in this study.

Several studies investigating the wastewater microbiome have been published. However, most of those studies use partial 16S rRNA gene sequencing, which usually does not allow microbial profiling to lowest taxonomic resolution, i.e., subspecies (39). Few studies detailed microbiome profiles of wastewater to species. Here, we employed SMS to profile the microbial composition of wastewater to subspecies (Fig. 4C). In addition, archaea, protists, fungi, and viruses were detected and identified (Fig. 4D). Protozoan species common to wastewater, such as *Pseudoperonospora cubensis* and *Paramecium biaurelia*, are environmental protists associated with infection in plants (127) and frequent symbionts of green algae in the aquatic environment (128), respectively. Fungi are ubiquitous in the environment, coexisting and interacting with other microorganisms and regulating a range of ecological functions. Fungi also are an important group of microorganisms found in

wastewater, aiding organic decomposition (129). However, some fungal genera commonly detected in WWTPs are opportunistic pathogens for humans and plants (130), including *Olipidium*, *Aspergillus*, *Candida*, and *Penicillium*, and these were detected in this study. Interestingly, *Epichloe sylvatica*, found previously to be associated with biofilm in WWTPs (36), was the most frequent fungal species detected in all samples examined in this study.

(ii) Virome. Traditionally, microbial communities have been defined using culture dependent methods to detect and enumerate microorganisms. However, it is estimated that the vast majority of prokaryotic genospecies remain uncultured. Accordingly, metagenomics and metatranscriptomics obviate the need to isolate and culture microorganisms. Furthermore, molecular methods are used to profile viral communities. In the current study, a combination of molecular methods, i.e., RT-qPCR, targeted amplicon sequencing, DNA metagenomics, and RNA metatranscriptomics, were used to detect and identify the wastewater virome. DNA viruses were detected in all samples, with *Papillomaviridae* and *Polyomaviridae*, human viruses, detected in samples positive for SARS-CoV-2 (Fig. 4D). The most frequently detected RNA viruses included tomato brown rugose fruit virus, pepper mild mottle virus, cucumber green mottle mosaic virus, tomato mosaic virus, tobacco mild green mosaic virus, tropical soda apple mosaic virus, tomato mottle mosaic virus, and melon necrotic spot virus (Fig. 5A). This viral composition was reported in a study of a WWTP in southern California during the COVID-19 pandemic (123). Results reported here show RNA metatranscriptomics is effective in detecting the dominant viruses in circulation, in agreement with prior research (25).

(iii) Resistome and virulome. Metagenomic sequencing has been used to explore AMR trends in anthropogenically impacted environments (41) and to detect wastewater associated enteric pathogens (36, 95). However, viability or infectious potential of detected microorganisms requires additional analysis of metabolic activity. In addition to detecting RNA viruses (25, 43), high-throughput sequencing of cDNA for RNA analysis, i.e., RNA sequencing, provides insight into gene expression (44) and has been used for metatranscriptomic analysis of wastewater (45). In the current study, untargeted RNA metatranscriptomic sequencing was employed successfully to analyze the active microbial community structure of wastewater samples, including functional activities, active pathways, and AMR and VF associated gene expression (Fig. 5). RNA viral genes are replicated, expressed, and assembled in association with living host cells (46). It has been suggested that SARS-CoV-2 can remain viable in sewage up to 4.3 days (47) and other coronaviruses may remain viable in aqueous matrices for over 1 year (48). In the current study, evidence of viral RNA genome replication was detected in 50% of the wastewater samples examined, suggesting infectious potential of RNA viruses in wastewater.

AMR is a growing global threat, claiming 700,000 deaths per year, further complicated by the COVID-19 pandemic, (49). Antibiotics have been remarkably successful in treating cobacterial infections associated with COVID-19, but these drugs do not work on viruses, including coronaviruses or influenza (50). In this study, the greatest number of AMR associated genes, namely those of the Macrolides, were detected in the sample collected on February 17, 2021, which was the week prior to the observed increase in SARS-CoV-2 RNA in wastewater. Interestingly, no AMR associated genes were detected in the sample collected the week following the increase in detected SARS-CoV-2 RNA, February 17, 2021. A similar pattern was observed with genes coding for VF, where an increase in VF was observed the week prior to the observed increase in SARS-CoV-2 RNA. VF associated with *Bacteroides* was dominant in these samples. *Bacteroides* are important members of the gut microbiome and have been linked to dysbiosis and the fecal-viral load during COVID-19 infection (51). While no attempts were made to associate clinical data to the resistome or virulome of wastewater detected and identified in this study, the results support previous observations that antibiotic use may be associated with progression of COVID-19 and bacterial coinfection (52), and may shed light on changes to the human fecal microbiome during COVID-19 infection (53), with respect to abundance of opportunistic pathogens and the VF carried.

Quantitative comparative metagenomics. Detection of wastewater associated pathogens by employing molecular techniques has been done at WWTPs (36, 38, 124) and used for source tracking of fecal pollution (126, 95, 54). However, a major hurdle for WS is communication of the results and their interpretation to public health officials. Metagenomic surveys are typically presented as the RA of sequencing reads, not a meaningful parameter for microbiota quantification and can be biased, depending on the experimental protocol employed (55). To circumvent these issues, we included an *in situ* control to the DNA metagenomics, allowing absolute cell number quantification, a recent technique increasingly cited in literature (56, 58). Presenting metagenomic sequencing results as cell numbers provides more precise microbial load estimates and permits normalization of studies. To present the utility of this application, the cell number of microorganisms detected in wastewater samples, e.g., pathogens and bacteria commonly employed as indicators of human fecal contamination (95, 59), is presented in Table 1, and can be used as a comparative baseline for metagenomic investigations employing this technique.

Wastewater microbiome and SARS-CoV-2 surveillance. While COVID-19 is generally considered to be primarily a respiratory illness, gastrointestinal symptoms, e.g., diarrhea, vomiting, nausea, or abdominal pain, are commonly associated with this disease (60). Therefore, gut microbiota have been linked to a variety of COVID-19 risk factors for infection and the gut microbiome is a potential therapeutic target for many other diseases (61). Emerging evidence suggests SARS-CoV-2 can infect the gastrointestinal tract directly (63).

Using DNA metagenomics, we aimed to evaluate if changes in the wastewater microbiota could be associated with prevalence of SARS-CoV-2 RNA in a specific community. The inclusion of an *in situ* control allowed correlations to be made between quantified microbial load (cells/L) and concentration of SARS-CoV-2 RNA (cells/L) detected in wastewater samples (Fig. 6). A similar study used Nanopore 16S rRNA sequencing to profile the microbiome of wastewater samples during a COVID-19 outbreak in Chile and showed correlation between RA, namely percentage, of both bacteria and SARS-CoV-2 (64). Strikingly, members of families *Lachnospiraceae* and *Rumnococcaceae* and genus *Alistipes* were detected at increased abundance following a COVID-19 outbreak, an observation supported by findings presented here. Gallardo-Escárate et al. (64) reported strong association with *Prevotella*, *Bacteroides*, *Aeromonas*, *Sulfurospirillum*, *Arcobacter*, *Tolumonas*, *Citrobacter*, *Zoologea*, and *Janthinobacterium*. In the current study, a similar correlation was observed for members of these taxa, but below our threshold of confidence. Statistical differences for these taxa may result from the sequencing method used, i.e., 16S rather than SMS (39). Because the wastewater microbiome is dominated by human gut bacteria, similar investigations done in different geographical regions, including areas where the diet differs, e.g., western versus eastern diets, which can influence the gut microbiome, are required to confirm these observations.

Here, bacterial genera with co-occurrence correlated with SARS-CoV-2 include *Bifidobacterium*, *Leclercia*, *Pyramidobacter*, *Tannerella*, *Massilimaliae*, and *Erythrobacter* (Fig. 6A). Most of the species determined to have significant co-occurrence correlation with SARS-CoV-2 were bacteria commonly found in the human gut (Fig. 6B). These include *Paraprevotella xylaniphila*, *Phascolarctobacterium succinatutens*, and *Slackia piriformis*, and also various *Bifidobacterium* spp. (*B. adolescents*, *B. catenulatum*, *B. pseudocatenulatum*, and *B. pseudolongum*), *Bacteroides* spp. (*B. caecimuris*, *B. nordii*, and *B. rodentium*), and *Prevotella* spp. (*P. lascolaii*, P. *marseillensis* and *P. pectinovora*), supporting the hypothesis that the resident microbial communities of the gastrointestinal and respiratory tracts can act as modulators of local and systemic inflammatory activity, e.g., the gut-lung axis (65). Remarkably, several opportunistic species shown to have significant correlation with SARS-CoV-2 RNA in wastewater, including *Bordetella bronchiseptica* (66), *Enterobacter cloacae* (67), *Leclercia adecarboxylata* (68), and *Pseudomonas monteilii* (69). These have been identified clinically as coinfecting microorganisms during COVID-19. Co-occurrence of archaea, protozoa, fungi, and DNA viruses was analyzed

(Fig. 6C), but results of combined kingdom analysis were inconclusive. Additional investigation is required to elucidate these complex interactions.

Challenges and future directions. Observations in this study were characterized using samples limited in number, yet illustrative of an outbreak in a single location and representing temporal shifts in wastewater microbial communities during an increase in COVID-19 cases. The data presented in this study were not normalized to flow rate or fecal and urine load, which have shown to be important in comparing microbial loads in wastewater (70–72). Additional studies are needed to characterize fully the core wastewater microbiome and establish a comparative baseline to verify microbial shifts between outbreak and nonoutbreak conditions, and establish public health significance. The wastewater microbiome of samples collected from a manhole may differ from wastewater samples collected from other SCSs, e.g., hospitals, prisons, nursing homes, WWTPs, etc. In additional to more samples and other locations, different seasons of the year need to be studied. In particular, such investigations will shed light on significant shifts in SMS and qRT-PCR results relative to COVID-19 outbreak size and duration, as well as incidence of other communicable diseases, notably influenza.

With respect to methods used in this study, because of the complexities of detecting SARS-CoV-2 in the environment, the decision was to centrifuge samples after collection to remove large debris and then employ ultrafiltration. It is possible that centrifugation removed larger microorganisms and those associated with sediment thereby influencing the microbiome profile. Hence, comparison of sample processing, sequencing, and bioinformatics is in progress.

Implementation of SMS for WS requires access to laboratories, trained personnel, bioinformatics support, and infrastructure for data storage. SMS can take longer to perform and currently is more expensive than traditional microbiological testing. The information provided, however, will prove valuable for public health.

Many short-read sequencing technologies rely on clonal PCR for signal detection. However, new technologies with long reads and single-molecular sequencing, such as those of Pacific Biosciences and Oxford Nanopore, are proving useful (42). Recently it has been shown that rapid variant identification of SARS-CoV-2 can be achieved in near real-time using a miniature-sized and field-deployable Oxford Nanopore MinION sequencing device (73) and also for microbiome analysis (74). Similarly, single cell sequencing has emerged as a useful molecular tool to identify genetically diverse viral genomes within single infectious units (75); its application in WS would help navigate variant calling in samples of a pooled population.

It is important to note that municipalities across the world now track SARS-CoV-2 in wastewater, and the addition of SMS will provide detailed characterization of wastewater microbiome and detection of infectious agents circulating within local populations. This is a crucial step in preparing for future epidemics and pandemics. This approach can also be used to monitor an array of public health indicators, e.g., obesity (77), consumption of alcohol, illicit drugs and tobacco, exposure to hazardous chemicals and pharmaceuticals (23, 24), and detection of sexually transmitted disease etiological agents (78, 79). In the meantime, SMS can be usefully applied to clinical and WS practices directed at both community and global public health.

Conclusion. Wastewater surveillance (WS) is accomplished by WWTPs in many municipalities globally to track SARS-CoV-2. In the study reported here, detection was achieved using RT-qPCR but was also combined with advanced molecular sequencing to characterize SARS-CoV-2 genetic mutations (targeted amplicon sequencing) and profile the complete microbiome (DNA metagenomics and RNA metatranscriptomics) of wastewater samples collected from a sewage collection system (a manhole). The microbial communities of the wastewater were associated, distributed, and varied temporally within a defined geographical area. Trends observed in this study are for a single location, but when environmental COVID-19 risk prediction was combined with the WS results, it was observed that after introduction of SARS-CoV-2 into a community, climatic conditions, namely, temperature, is associated with viral transmission. WS serves as a useful indicator of disease prevalence, but results show early detection, characterization of SARS-CoV-2,

and co-infections in a community can be determined using microbiome profiling of wastewater. Ongoing research will integrate clinical case reports of gastrointestinal and other communicable diseases prevalent in a community, in addition to COVID-19, to assess whether changes in the wastewater microbiome are indicative of community health and if reliable biomarkers can be detected that are applicable to public health.

MATERIALS AND METHODS

Site description. WS was carried out by monitoring a component of a SCS in Maryland, USA (Fig. 7). The exact location of the study has been masked to protect the privacy of persons living in the area. Total daily number of reported COVID-19 cases among Maryland residents within the ZIP code where the study was conducted, which has an estimated population of *ca*. 22,000 (80), was retrieved from the Maryland Department of Health (81). The cumulative number of positive COVID-19 cases is presented in the supporting information (Data set S1). At the beginning of the study period, the SCS serviced 45 buildings with 669 housing units, equating to a population density of roughly 760 residents. The average daily wastewater flow (*ca*. 17,610 L per day) was estimated from the average daily accumulative water flow of all buildings serviced by the SCS between October 26 and November 25, 2021, using a conversion factor of 80%. Justification of the flow projection based on water consumption in the serviced community is supported by calculations detailed elsewhere (82).

Sample collection. Composite samples were collected weekly from a manhole downstream of the SCS between December 20, 2020 and November 16, 2021 (n = 48). During each sampling event, an automated composite sampling unit (Teledyne ISCO, Lincoln, NE, USA) was prepared in the morning with ice surrounding the collection jar. Samples (60 mL) were collected at 15-min intervals for 24 h, totaling 96 individual sampling attempts and a composite volume of 5.76 L. Temperature of the composite samples was recorded at time of retrieval and is provided in the supporting information (Data set S2). Samples were homogenized manually, and an aliquot (110 mL) was pasteurized in a water bath (Polyscience, Niles, IL, USA) at 60°C for 30 min, transported to the laboratory on ice, and processed the same day. To remove larger debris, 45 mL were transferred to a sterile polypropylene 50 mL conical Falcon tube (Corning, Glendale, AZ, USA) and centrifuged at 7,500 RCF for 10 min at 4°C. The resulting supernatant was transferred to a clean conical tube and concentrated using InnovaPrep Concentrating Pipette Select with Ultrafilitration PS Hollow Fiber pipette tips (nominal molecular weight limit of *ca*.100 to 120 kDa; InnovaPrep, Drexel, MO, USA), following the manufacturer's recommendations for "Wastewater Application Note, Revision B" (83). InnovaPrep Wet Foam Elution was stored in DNA/RNA shield (Zymo Research, Irvine, CA, USA), following the manufacturer's specifications, at −80°C until nucleic acid was prepared (<48 h). Detection of SARS-CoV-2 RNA was done on all the samples that were collected and six samples were selected for microbiome analysis.

COVID-19 risk assessment. Usmani et al. (6) hypothesized transmission of SARS-CoV-2 occurs in the human population when the ambient AT rises above or falls below the "comfort zone" (17°C to 24°C) and transmission of the virus thereby increases (Fig. S1). This phenomenon is the result of changes in human behavior in the ambient and built environment, namely, when temperatures become uncomfortable. That is, when AT is within the comfort zone, there will be a decrease in COVID-19 transmission. Furthermore, an increase in aerosolization of SARS-CoV-2 occurs in cold and dry environments. For example, an increased number of COVID-19 cases were reported in the northern part of the U.S. during winter months, compared with other seasons. In the current study, we sought to evaluate the feasibility of employing this hypothesis as a predictive COVID-19 risk score model and demonstrate its potential use with WS.

To examine this model, details of which are described elsewhere (6), environmental COVID-19 risk scores were predicted along with epidemiological and experimental data obtained throughout the study. The difference between AT and dew point temperature (DPT) was used as an indication of moisture availability in the atmosphere, an observation described previously to impact viral survival as a seasonal factor in influenza and poliomyelitis (84, 85). Briefly, DPT, defined as the temperature at which water vapor condenses to form liquid droplets large enough to settle quickly on Earth's surface due to gravity, is directly proportional to moisture content in the air. That is, low DPT indicates lower moisture in the air and high DPT suggests high moisture content. Hence, DPT must be less than or equal to AT. The difference between AT and DPT indicates moisture saturation of the ambient air, and the difference is inversely proportional to air moisture saturation. If the difference between AT and DPT is zero, then the air is considered fully saturated. As the difference between AT and DPT increases, the air becomes less saturated. In summary, DPT indicates specific moisture in the air, while the difference between AT and DPT can be used to determine moisture saturation in the ambient air.

Daily AT and DPT were obtained from the Oregon State University Parameter-elevation Regressions on Independent Slopes Model (PRISM) climate group products at a resolution of 4 km × 4 km (86). Prediction of environmental COVID-19 risk was computed weekly using AT and DPT as model input and quantified according to the hypothesis, where deviation of AT outside the comfort zone was calculated for the previous 2 weeks. The number of days with a negative DPT and when the difference between AT and DPT was greater than 5°C was also determined for the same time interval. Variables were determined using a scale of 0 to 4, with 4 being the maximum value. Environmental risk of COVID-19 was determined from the normalized sum of quantified variables with equal weight. That is, the risk model output varies between 0 and 1, with 1 being the highest risk of transmission. Following environmental

risk assessment, weekly risk score values were compared with RT-qPCR observations of SARS-CoV-2 in wastewater samples.

Quantitative reverse transcriptase PCR for detection of SARS-CoV-2 RNA. (i) RNA purification. To serve as internal control (IC) for the RT-qPCR assay, 25 μL of MS2 Bacteriophage, target titer of 1.0×10^3 PFU/mL, (ZeptoMetrix, Buffalo, NY, USA), was added to 115 μL InnovaPrep Wet Foam Elution. Total RNA was prepared from wastewater concentrates containing MS2 Bacteriophage, employing the QIAamp Viral RNA minikit (Qiagen, Germantown, MD, USA), following the manufacturer's instructions for use on the automated QIAcube Connect platform. RNA extracts were stored in LoBind microcentrifuge tubes (Eppendorf, Hamburg, Gernamny) at $-80°$C prior to RT-qPCR amplification ($<$48 h).

(ii) Reference RNA materials. The SARS-CoV-2 Research Grade Test Material (RGTM 10169; National Institute of Standards and Technology, Gaithersburg, MD, USA), consisting of two synthetic RNA fragments from the SARS-CoV-2 genome (including SARS-CoV-2 sequences 25949 to 29698 and 12409 to 15962 of isolate USA-WA1/2020) in a background of 5 ng/μL human Jurkat RNA, was diluted over six serial log dilutions with nuclease-free water for use as calibration standards. RNA template of SARS-CoV-2 Nucleocapsid Phosphoprotein (N protein) encapsulated in the MS2 bacteriophage construct (PerkinElmer, Waltham, MA, USA) was prepared following the manufacturer's instructions for use as positive control (PC). Single-use aliquots of reference RNA were stored in LoBind microcentrifuge tubes (Eppendorf, Hamburg, Gernamny) at $-80°$C.

(iii) RT-qPCR amplification. The New Coronavirus Nucleic Acid Detection Kit v.7.0 (PerkinElmer, Waltham, MA, USA) was employed using the QuantStudio 3 System (Thermo Fisher, Waltham, MA, USA), for multiplexed detection of SARS-CoV-2 N protein and IC, following the manufacturer's specifications. The New Coronavirus Nucleic Acid Detection Kit also allows for detection of SARS-CoV-2 Open Reading Frame 1ab (ORF1ab). However, under the methods detailed below, controlled spike-in experiments have suggested that N protein provides a more accurate estimation of SARS-CoV-2 abundance in wastewater. Internal lab results, including multiple temporal samples from different locations, have shown that wastewater samples presenting detection of ORF1ab alone (ORF1ab positive and N negative) were extremely rare. Hence, N protein was used as the representative SARS-CoV-2 marker for subsequent analysis. Amplification efficiency (*AE*) for each reference DNA calibration model was calculated as follows: $AE = 10^{-1/slope} - 1$. The lower limit of quantification (LLOQ) was determined by signal-to-noise ratio of 10:1 and approximated using residual standard deviation (SD) of regression, defined by: $LLOQ = 10\frac{SD}{-slope}$. The upper ROQ was defined as the highest calibration standard included in the standard curve (5×10^6 N copies per reaction). Calibration model performance of each RT-qPCR run is provided in the supporting information (Data set S2). Following each run, log florescence thresholds were set manually, and melt curve analysis was done to identify spurious amplicons that could confound data interpretation (no spurious amplicons detected). Quantification cycle (Cq) values were exported to Excel (Microsoft, Redmond, WA, USA) for further analysis.

(iv) Quality control. Method extraction controls (positive, PC; negative, NC; and internal, IC) were prepared using the New Coronavirus Nucleic Acid Detection Kit (PerkinElmer, Waltham, MA, USA) to monitor sample processing. Synthetic SARS-CoV-2 N protein RNA template encapsulated in MS2 bacteriophage (PerkinElmer, Waltham, MA, USA) was included as PC. For IC, bacteriophage MS2 was used to monitor the process from nucleic acid extraction to fluorescence detection for amplification inhibition. Failure to detect PC resulted in an invalid run for all samples. Samples negative for IC, suggestive of amplification inhibition, or positive for NC, suggestive of potential contamination, were discarded from further analysis.

Complete microbiome analysis employing next generation sequencing. A total of six samples were selected for complete microbiome analysis, including shotgun DNA metagenomic and RNA metatranscriptomic sequencing, and targeted amplicon sequencing for identification and characterization of full SARS-CoV-2 genome.

(i) Nucleic acid preparation. Total RNA was prepared from InnovaPrep Wet Foam Elution using QIAamp Viral RNA minikit, as mentioned previously. To serve as an *in situ* control for DNA sequencing, the ZymoBIOMICS High Microbial Load Spike-in Control I (Zymo Research, Irvine, CA, USA) was added to each sample, per manufacturer's specifications. Total DNA was prepared from 250 μL of InnovaPrep Wet Foam Elution using the DNeasy PowerSoil Pro Kit (Qiagen, Germantown, MD, USA), following the manufacturer's instructions for use on the automated QIAcube Connect platform. Nucleic acid concentrations were determined using Qubit Broad Range Assay Kits (Thermo Fisher Scientific, Waltham, MA, USA) on an Invitrogen Qubit 4.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) for RNA and double-stranded DNA, respectively.

(ii) Targeted sequencing for identification of full SARS-CoV-2 genome. Following detection and quantification of SARS-CoV-2 by RT-qPCR, excess cDNA of selected samples and sample processing controls (method extraction blank and synthetic SARS-CoV-2 RNA) was used as input to establish mutation profiles. Sequencing libraries were prepared using the Swift Normalase Amplicon SARS-CoV-2 Panel kit (Swift Biosciences, Ann Arbor, MI, USA), following the manufacturer's instructions. The Swift Normalase Amplicon SARS-CoV-2 Panel involves multiplex PCR technology that utilizes multiple overlapping amplicons. Using tiled primer pairs targeting the entire SARS-CoV-2 genome enables library construction from cDNA, even from low input quantities. The resulting libraries were quantified using Qubit dsDNA High Sensitivity assay kit (Thermo Fisher Scientific, Waltham, MA, USA) on an Invitrogen Qubit 4.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), and sequenced 2 $\times$ 150 bp on an Illumina HiSeq4000 Instrument (Illumina Inc., San Diego, CA, USA).

General sequencing statistics for all samples and mean sequence quality distribution were measured by FastQC v.0.11.6 (87). Base-calling error probabilities (P) were evaluated using Phred Quality Score (Q),

defined by: $Q = -10log_{10}(P)$. Using a previously defined read quality threshold (88), adapter and quality trimming of raw sequencing reads were done using Trimmomatic v0.38 (89) to ensure a Phred Quality Score of at least 17 for at least 80% of the read lengths, i.e., probability of correct base call was at least 98%. Primers used during panel amplification were trimmed using Primerclip v.0.3.8 (Swift Biosciences, Ann Arbor, MI, USA). Processed sequencing reads were aligned to the Wuhan-Hu-1 genome (NCBI accession no. NC_045512.2) using BBmap v38.70 (90), and variants were called using bcftools v.1.14 (91).

(iii) DNA metagenomics and RNA metatranscriptomics for complete microbiome analysis. rRNA (rRNA) of abundant species (human, mouse, and rat) was reduced using Ribo-Zero Plus rRNA Depletion Kit (Illumina Inc., San Diego, CA, USA). Second strand synthesis of cDNA and preparation of RNA libraries was achieved using NEBNext Ultra II Nondirectional RNA Library Prep Kit (New England BioLabs, Ipswich, MA, USA), following the manufacturer's instructions for use with Illumina chemistry. DNA libraries were prepared using Nextera XT DNA Library Prep Kit (Illumina Inc., San Diego, CA, USA). All sequencing libraries were quantified and sequenced using Illumina chemistry, as mentioned previously. A no template control (NTC), consisting of nuclease free water, and sequencing standard, i.e., ZymoBIOMICS™ Microbial Community Standard (Zymo Research, Irvine, CA, USA), were included for quality control.

(iv) Microbiome taxonomic profiling. For shotgun DNA metagenomics and RNA metatranscriptomics, adapter and quality trimming of raw sequencing reads was done as described previously. Unassembled metagenomic and metatranscriptomic sequencing reads were analyzed, as reported elsewhere (62, 92–95), using CosmosID-HUB Microbiome Platform (CosmosID Inc., Germantown, MD, USA) to achieve multikingdom microbiome analysis. Virulence and AMR associated gene profiling was done using RNA metatranscriptomic data through the CosmosID-HUB. Qualitative gene expression of detected Virulence and AMR associated genes is presented as the RA compared to ribosome binding molecular function (GO:0043022), which was detected at roughly equal abundance across all samples when normalized to copies per million reads. Additional information on the bioinformatics pipeline employed for taxonomic classification of sequencing reads can be found in the supporting information (Text S1).

(v) Statistical analysis. Microbiota detected in the NTC (*Pseudoxanthomonas* spp.) were subtracted from taxonomic profiling of DNA metagenomics; no microbiota were detected in the RNA metatranscriptomic NTC. For DNA metagenomics, microbial cell number of each taxon was normalized to the cell number of *Imtechella halotolerans* (Gram-negative) and *Allobacillus halotolerans* (Gram-positive), following the manufacturer's specifications of the ZymoBIOMICS High Microbial Load Spike-in Control I (Zymo Research, Irvine, CA, USA). Quantification of the organism RA was defined as the proportion of unique organism-specific k-mers annotated by each database relative to the total number of unique sequencing reads generated for that sample. Analysis of community virulome and resistome were achieved by identifying virulence and AMR associated genes based on percent coverage as a function of gene-specific k-mer frequency in each sample.

The alpha diversity of bacterial communities associated with wastewater was compared among samples by calculating the Shannon and Simpson 1-D diversity indices and the Chao-1 richness index using Vegan v.2.5.7 (96). RA of bacterial taxa in each sample was used for principal coordinate analysis (PCoA) employing Bray-Curtis dissimilarity index, using CosmosID-HUB Microbiome Platform v.2.0 (CosmosID Inc., Germantown, MD, USA). Pair-wise Spearman-rank co-occurrence between concentrations of SARS-CoV-2 RNA (N copies/L) and microbiota (cells/L) was determined using phylosmith v.1.0.6 (57), which is based on methods described elsewhere (97). Upper and lower cutoffs for which rho values are not likely to have occurred by random chance were determined by calculating the pairwise Spearman rank co-occurrence for each variable with 1,000 permutation iterations. Significance for co-occurrence of each taxa and SARS-CoV-2 was determined using rho value cutoffs and P value ($\leq$0.02).

Data availability. Sequencing data generated for all samples included in this study are deposited in the NCBI Sequence Read Archive database under BioProject ID PRJNA812772. Accession numbers for individual sample sequencing read libraries (SAMN2649443 to SAMN2649459) are provided in the supplementary information (Data set S4).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
DATA SET S1, XLSX file, 0.02 MB.
DATA SET S2, XLSX file, 0.01 MB.
DATA SET S3, XLSX file, 0.01 MB.
DATA SET S4, XLSX file, 0.01 MB.
TEXT S1, DOCX file, 0.03 MB.
FIG. S1, DOCX file, 0.4 MB.
FIG. S2, DOCX file, 0.4 MB.
FIG. S3, DOCX file, 0.2 MB.
FIG. S4, DOCX file, 0.2 MB.
TABLE S1, DOCX file, 0.02 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Baker RE, Mahmud AS, Miller IF, Rajeev M, Rasambainarivo F, Rice BL, Takahashi S, Tatem AJ, Wagner CE, Wang L-F, Wesolowski A, Metcalf CJE. 2022. Infectious disease in an era of global change. Nat Rev Microbiol 20: 193–113. https://doi.org/10.1038/s41579-021-00639-z.

2. United Nations Environment Program and International Livestock Research Institute. 2020. Preventing the next pandemic - Zoonotic diseases and hot to break the chain of transmission. United Nations. https://www.unep.org/resources/report/preventing-future-zoonotic-disease-outbreaks-protecting-environment-animals-and.

3. WHO. 2020. WHO Director-General's remarks at the media briefing on 2019-nCoV on 11 February 2020. https://www.who.int/director-general/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020.

4. Dong E, Du H, Gardner L. 2020. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 20:533–534. https://doi.org/10.1016/S1473-3099(20)30120-1.

5. WHO. 2020. Modes of transmission of virus causing COVID-19: implications for IPC precaution recommendations: scientific brief, 27 March 2020. World Health Organization. https://apps.who.int/iris/handle/10665/331601.

6. Usmani M, Yusuf J, Gangwar M, Magers B, Chaves-Gonzalez J, Wu C-Y, Colwell RR, Jutla AS. 2022. Asymmetric relationship between ambient air temperature and incidence of COVID-19 in the human population. Am J Trop Med Hyg 106:1–9. https://doi.org/10.4269/ajtmh.21-0328.

7. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, Azman AS, Reich NG, Lessler J. 2020. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann Intern Med 172:577–582. https://doi.org/10.7326/M20-0504.

8. Tindale LC, Stockdale JE, Coombe M, Garlock ES, Lau WYV, Saraswat M, Zhang L, Chen D, Wallinga J, Colijn C. 2020. Evidence for transmission of COVID-19 prior to symptom onset. Elife 9:e57149. https://doi.org/10.7554/eLife.57149.

9. Moghadas SM, Fitzpatrick MC, Sah P, Pandey A, Shoukat A, Singer BH, Galvani AP. 2020. The implications of silent transmission for the control of COVID-19 outbreaks. Proc Natl Acad Sci U S A 117:17513–17515. https://doi.org/10.1073/pnas.2008373117.

10. Safford HR, Shapiro K, Bischel HN. 2022. Opinion: wastewater analysis can be a powerful public health tool—if it's done sensibly. Proc Natl Acad Sci U S A 119:e2119600119. https://doi.org/10.1073/pnas.2119600119.

11. Ferté T, Ramel V, Cazanave C, Lafon M-E, Bébéar C, Malvy D, Georges-Walryck A, Dehail P. 2021. Accuracy of COVID-19 rapid antigenic tests compared to RT-PCR in a student population: The StudyCov study. J Clin Virol 141:104878. https://doi.org/10.1016/j.jcv.2021.104878.

12. Syal K. 2021. Guidelines on newly identified limitations of diagnostic tools for COVID-19 and consequences. J Med Virol 93:1837–1842. https://doi.org/10.1002/jmv.26673.

13. Bustin SA, Nolan T. 2020. RT-qPCR Testing of SARS-CoV-2: A Primer. Int J Mol Sci 21:3004. https://doi.org/10.3390/ijms21083004.

14. Tsujimoto Y, Terada J, Kimura M, Moriya A, Motohashi A, Izumi S, Kawajiri K, Hakkaku K, Morishita M, Saito S, Takumida H, Watanabe H, Tsukada A, Morita C, Yamaguchi Y, Katsuno T, Kusaba Y, Sakamoto K, Hashimoto M,

Suzuki M, Takasaki J, Hojo M, Miyoshi-Akiyama T, Sugiyama H. 2021. Diagnostic accuracy of nasopharyngeal swab, nasal swab and saliva swab samples for the detection of SARS-CoV-2 using RT-PCR. Infect Dis (Lond) 53:581–589. https://doi.org/10.1080/23744235.2021.1903550.

15. Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, Tan W. 2020. Detection of SARS-CoV-2 in different types of clinical specimens. JAMA 323:1843–1844. https://doi.org/10.1001/jama.2020.3786.

16. Chen Y, Chen L, Deng Q, Zhang G, Wu K, Ni L, Yang Y, Liu B, Wang W, Wei C, Yang J, Ye G, Cheng Z. 2020. The presence of SARS-CoV-2 RNA in the feces of COVID-19 patients. J Med Virol 92:833–840. https://doi.org/10.1002/jmv.25825.

17. Berchenko Y, Manor Y, Freedman LS, Kaliner E, Grotto I, Mendelson E, Huppert A. 2017. Estimation of polio infection prevalence from environmental surveillance data. Sci Transl Med 9:280–281. https://doi.org/10.1126/scitranslmed.aaf6786.

18. Hovi T, Shulman LM, Van Der Avoort H, Deshpande J, Roivainen M, De Gourville EM. 2012. Role of environmental poliovirus surveillance in global polio eradication and beyond. Epidemiol Infect 140:1–13. https://doi.org/10.1017/S095026881000316X.

19. Nakamura T, Hamasaki M, Yoshitomi H, Ishibashi T, Yoshiyama C, Maeda E, Sera N, Yoshida H. 2015. Environmental surveillance of poliovirus in sewage water around the introduction period for inactivated polio vaccine in Japan. Appl Environ Microbiol 81:1859–1864. https://doi.org/10.1128/AEM.03575-14.

20. La Rosa G, Iaconelli M, Pourshaban M, Fratini M, Muscillo M. 2010. Molecular detection and genetic diversity of norovirus genogroup IV: a year-long monitoring of sewage throughout Italy. Arch Virol 155:589–593. https://doi.org/10.1007/s00705-010-0619-y.

21. Heijnen L, Medema G. 2011. Surveillance of influenza A and the pandemic influenza A (H1N1) 2009 in sewage and surface water in the Netherlands. J Water Health 9:434–442. https://doi.org/10.2166/wh.2011.019.

22. Peccia J, Zulli A, Brackney DE, Grubaugh ND, Kaplan EH, Casanovas-Massana A, Ko AI, Malik AA, Wang D, Wang M, Warren JL, Weinberger DM, Arnold W, Omer SB. 2020. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. Nat Biotechnol 38: 1164–1167. https://doi.org/10.1038/s41587-020-0684-z.

23. Adhikari S, Halden RU. 2022. Opportunities and limits of wastewater-based epidemiology for tracking global health and attainment of UN sustainable development goals. Environ Int 163:107217. https://doi.org/10.1016/j.envint.2022.107217.

24. Gushgari AJ, Venkatesan AK, Chen J, Steele JC, Halden RU. 2019. Long-term tracking of opioid consumption in two United States cities using wastewater-based epidemiology approach. Water Res 161:171–180. https://doi.org/10.1016/j.watres.2019.06.003.

25. 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. https://doi.org/10.1016/j.scitotenv.2021.147829.

26. Vo V, Tillett RL, Chang C-L, Gerrity D, Betancourt WQ, Oh EC. 2022. SARS-CoV-2 variant detection at a university dormitory using wastewater

genomic tools. Sci Total Environ 805:149930. https://doi.org/10.1016/j.scitotenv.2021.149930.

27. Larson RC, Berman O, Nourinejad M. 2020. Sampling manholes to home in on SARS-CoV-2 infections. PLoS One 15:e0240007. https://doi.org/10.1371/journal.pone.0240007.

28. Ahmed W, Tscharke B, Bertsch PM, Bibby K, Bivins A, Choi P, Clarke L, Dwyer J, Edson J, Nguyen TMH, O'Brien JW, Simpson SL, Sherman P, Thomas KV, Verhagen R, Zaugg J, Mueller JF. 2021. SARS-CoV-2 RNA monitoring in wastewater as a potential early warning system for COVID-19 transmission in the community: A temporal case study. Sci Total Environ 761:144216. https://doi.org/10.1016/j.scitotenv.2020.144216.

29. Ahmed W, Angel N, Edson J, Bibby K, Bivins A, O'Brien JW, Choi PM, Kitajima M, Simpson SL, Li J, Tscharke B, Verhagen R, Smith WJM, Zaugg J, Dierens L, Hugenholtz P, Thomas KV, Mueller JF. 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. https://doi.org/10.1016/j.scitotenv.2020.138764.

30. Kumar M, Patel AK, Shah AV, Raval J, Rajpara N, Joshi M, Joshi CG. 2020. First proof of the capability of wastewater surveillance for COVID-19 in India through detection of genetic material of SARS-CoV-2. Sci Total Environ 746:141326. https://doi.org/10.1016/j.scitotenv.2020.141326.

31. 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 7:511–516. https://doi.org/10.1021/acs.estlett.0c00357.

32. 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. https://doi.org/10.1016/j.watres.2020.115942.

33. Sherchan SP, Shahin S, Ward LM, Tandukar S, Aw TG, Schmitz B, Ahmed W, Kitajima M. 2020. First detection of SARS-CoV-2 RNA in wastewater in North America: a study in Louisiana, USA. Sci Total Environ 743:140621. https://doi.org/10.1016/j.scitotenv.2020.140621.

34. Wu F, Zhang J, Xiao A, Gu X, Lee WL, Armas F, Kauffman K, Hanage W, Matus M, Ghaeli N, Endo N, Duvallet C, Poyet M, Moniz K, Washburne AD, Erickson TB, Chai PR, Thompson J, Alm EJ. 2020. SARS-CoV-2 titers in wastewater are higher than expected from clinically confirmed cases. mSystems 5:e00614-20. https://doi.org/10.1128/mSystems.00614-20.

35. Gonzalez R, Curtis K, Bivins A, Bibby K, Weir MH, Yetka K, Thompson H, Keeling D, Mitchell J, Gonzalez D. 2020. COVID-19 surveillance in Southeastern Virginia using wastewater-based epidemiology. Water Res 186:116296. https://doi.org/10.1016/j.watres.2020.116296.

36. Stamps BW, Leddy MB, Plumlee MH, Hasan NA, Colwell RR, Spear JR. 2018. Characterization of the microbiome at the world's largest potable water reuse facility. 9:2435. https://doi.org/10.3389/fmicb.2018.02435.

37. Karkman A, Berglund F, Flach C-F, Kristiansson E, Larsson DGJ. 2020. Predicting clinical resistance prevalence using sewage metagenomic data. Commun Biol 3:711. https://doi.org/10.1038/s42003-020-01439-6.

38. Leddy MB, Hasan NA, Subramanian P, Heberling C, Cotruvo J, Colwell RR. 2017. Characterization of microbial signatures from advanced treated wastewater biofilms. jawwa 109:E503–E512. https://doi.org/10.5942/jawwa.2017.109.0116.

39. Brumfield KD, Huq A, Colwell RR, Olds JL, Leddy MB. 2020. Microbial resolution of whole genome shotgun and 16S amplicon metagenomic sequencing using publicly available NEON data. PLoS One 15:e0228899. https://doi.org/10.1371/journal.pone.0228899.

40. Rockett RJ, Arnott A, Lam C, Sadsad R, Timms V, Gray K-A, Eden J-S, Chang S, Gall M, Draper J, Sim EM, Bachmann NL, Carter I, Basile K, Byun R, O'Sullivan MV, Chen SC-A, Maddocks S, Sorrell TC, Dwyer DE, Holmes EC, Kok J, Prokopenko M, Sintchenko V. 2020. Revealing COVID-19 transmission in Australia by SARS-CoV-2 genome sequencing and agent-based modeling. Nat Med 26:1398–1404. https://doi.org/10.1038/s41591-020-1000-7.

41. Karkman A, Pärnänen K, Larsson DGJ. 2019. Fecal pollution can explain antibiotic resistance gene abundances in anthropogenically impacted environments. Nat Commun 10:80. https://doi.org/10.1038/s41467-018-07992-3.

42. Zhong Y, Xu F, Wu J, Schubert J, Li MM. 2021. Application of next generation sequencing in laboratory medicine. Ann Lab Med 41:25–43. https://doi.org/10.3343/alm.2021.41.1.25.

43. 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 10:e00280-21. https://doi.org/10.1128/MRA.00280-21.

44. Kukurba KR, Montgomery SB. 2015. RNA sequencing and analysis. Cold Spring Harb Protoc 2015:951–969. https://doi.org/10.1101/pdb.top084970.

45. Ekwanzala MD, Budeli P, Unuofin JO. 2021. Application of metatranscriptomics in wastewater treatment processes, p 187–204. In Shah MP, Sarkar A, Mandal S (ed), Wastewater treatment. Elsevier. https://doi.org/10.1016/B978-0-12-821881-5.00008-8. (https://www.sciencedirect.com/science/article/pii/B9780128218815000088).

46. Rampersad S, Tennant P. 2018. Replication and expression strategies of viruses: virology, p 55–82. In Tennant P, Fermin G, Foster JE (ed), Viruses, Academic Press. https://doi.org/10.1016/B978-0-12-811257-1.00003-6.

47. Bivins A, Greaves J, Fischer R, Yinda KC, Ahmed W, Kitajima M, Munster VJ, Bibby K. 2020. Persistence of SARS-CoV-2 in water and wastewater. Environ Sci Technol Lett 7:937–942. https://doi.org/10.1021/acs.estlett.0c00730.

48. Giacobbo A, Rodrigues MAS, Zoppas Ferreira J, Bernardes AM, de Pinho MN. 2021. A critical review on SARS-CoV-2 infectivity in water and wastewater. What do we know? Sci Total Environ 774:145721. https://doi.org/10.1016/j.scitotenv.2021.145721.

49. Rezasoltani S, Yadegar A, Hatami B, Aghdaei HA, Zali MR. 2020. Antimicrobial resistance as a hidden menace lurking behind the COVID-19 outbreak: the global impacts of too much hygiene on AMR. Front Microbiol 11:3097.

50. CDC. 2021. Be antibiotics aware: smart use, best care. https://www.cdc.gov/patientsafety/features/be-antibiotics-aware.html.

51. Zuo T, Zhang F, Lui GCY, Yeoh YK, Li AYL, Zhan H, Wan Y, Chung ACK, Cheung CP, Chen N, Lai CKC, Chen Z, Tso EYK, Fung KSC, Chan V, Ling L, Joynt G, Hui DSC, Chan FKL, Chan PKS, Ng SC. 2020. Alterations in gut microbiota of patients with COVID-19 during time of hospitalization. Gastroenterology 159:944–955.e8. https://doi.org/10.1053/j.gastro.2020.05.048.

52. Nori P, Cowman K, Chen V, Bartash R, Szymczak W, Madaline T, Punjabi Katiyar C, Jain R, Aldrich M, Weston G, Gialanella P, Corpuz M, Gendlina I, Guo Y. 2021. Bacterial and fungal coinfections in COVID-19 patients hospitalized during the New York City pandemic surge. Infect Control Hosp Epidemiol 42:84–88. https://doi.org/10.1017/ice.2020.368.

53. Yamamoto S, Saito M, Tamura A, Prawisuda D, Mizutani T, Yotsuyanagi H. 2021. The human microbiome and COVID-19: a systematic review. PLoS One 16:e0253293. https://doi.org/10.1371/journal.pone.0253293.

54. Stachler E, Bibby K. 2014. Metagenomic evaluation of the highly abundant human gut bacteriophage CrAssphage for source tracking of human fecal pollution. Environ Sci Technol Lett 1:405–409. https://doi.org/10.1021/ez500266s.

55. Nayfach S, Pollard KS. 2016. Toward accurate and quantitative comparative metagenomics. Cell 166:1103–1116. https://doi.org/10.1016/j.cell.2016.08.007.

56. Jonas K, Pamela F, U MC, Katja H, Annette H, Janina M, Philipp H, Hannes H, Sabine P, Dorothee V, Michael Z, Kirke FS, Christoph H, Peer B, Stephan G, Thea VR. 2022. Meconium microbiome of very preterm infants across Germany. mSphere 7:e00808-21. https://doi.org/10.1128/msphere.00808-21.

57. Smith SD. 2019. phylosmith: an R-package for reproducible and efficient microbiome analysis with phyloseq-objects. JOSS 4:1442. https://doi.org/10.21105/joss.01442.

58. Risely A, Wilhelm K, Clutton-Brock T, Manser MB, Sommer S. 2021. Diurnal oscillations in gut bacterial load and composition eclipse seasonal and lifetime dynamics in wild meerkats. Nat Commun 12:6017. https://doi.org/10.1038/s41467-021-26298-5.

59. Harwood V, Shanks O, Koraijkic A, Verbyla M, Ahmed W, Iriate M. 2017. General and host-associated bacterial indicators of faecal pollution p 3–41. In Rose JB, Jiménez-Cisneros B (eds), Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management (Global Water Pathogen Project). Michigan State University, East Lansing, MI. https://www.waterpathogens.org/book/bacterial-indicators.

60. Brooks EF, Bhatt AS. 2021. The gut microbiome: a missing link in understanding the gastrointestinal manifestations of COVID-19? Cold Spring Harb Mol Case Stud 7:a006031. https://doi.org/10.1101/mcs.a006031.

61. Segal JP, Mak JWY, Mullish BH, Alexander JL, Ng SC, Marchesi JR. 2020. The gut microbiome: an under-recognised contributor to the COVID-19 pandemic? Therap Adv Gastroenterol 13:175628482097491–175628482097414. https://doi.org/10.1177/1756284820974914.

62. Roy MA, Arnaud JM, Jasmin PM, Hamner S, Hasan NA, Colwell RR, Ford TE. 2018. A metagenomic approach to evaluating surface water quality in Haiti. IJERPH 15:2211. https://doi.org/10.3390/ijerph15102211.

63. Xiao F, Tang M, Zheng X, Liu Y, Li X, Shan H. 2020. Evidence for gastrointestinal infection of SARS-CoV-2. Gastroenterology 158:1831–1833.e3. https://doi.org/10.1053/j.gastro.2020.02.055.

64. Gallardo-Escárate C, Valenzuela-Muñoz V, Núñez-Acuña G, Valenzuela-Miranda D, Benaventel BP, Sáez-Vera C, Urrutia H, Novoa B, Figueras A, Roberts S, Assmann P, Bravo M. 2021. The wastewater microbiome: a novel insight for COVID-19 surveillance. Sci Total Environ 764:142867. https://doi.org/10.1016/j.scitotenv.2020.142867.

65. Hussain I, Cher GLY, Abid MA, Abid MB. 2021. Role of gut microbiome in COVID-19: an insight into pathogenesis and therapeutic potential. Front Immunol 12:765965. https://doi.org/10.3389/fimmu.2021.765965.

66. Faqihi F, Alharthy A, Pirompanich P, Noor A, Shahzad A, Nasim N, Balhamar A, Memish ZA, Karakitsos D. 2020. Co-infection of SARS-CoV-2 and Bordetella bronchiseptica in a young man with idiopathic non-cystic bronchiectasis and vitamin D(3) deficiency. Respir Med Case Rep 31:101203. https://doi.org/10.1016/j.rmcr.2020.101203.

67. Fontana C, Favaro M, Minelli S, Bossa MC, Altieri A. 2021. Co-infections observed in SARS-CoV-2 positive patients using a rapid diagnostic test. Sci Rep 11:16355. https://doi.org/10.1038/s41598-021-95772-3.

68. Gamon E, Tammena D, Wattenberg M, Augenstein T. 2022. Rare superinfection in a COVID-19 patient: a chronology. Anaesthesist 71:38–49. https://doi.org/10.1007/s00101-021-01018-2.

69. Thuy-Boun PS, Mehta S, Gruening B, McGowan T, Nguyen A, Rajczewski AT, Johnson JE, Griffin TJ, Wolan DW, Jagtap PD. 2021. Metaproteomics analysis of SARS-CoV-2-infected patient samples reveals presence of potential coinfecting microorganisms. J Proteome Res 20:1451–1454. https://doi.org/10.1021/acs.jproteome.0c00822.

70. Holm RH, Nagarkar M, Yeager RA, Talley D, Chaney AC, Rai JP, Mukherjee A, Rai SN, Bhatnagar A, Smith T. 2022. Surveillance of RNase P, PMMoV, and CrAssphage in wastewater as indicators of human fecal concentration across urban sewer neighborhoods, Kentucky. FEMS Microbes 3: xtac003. https://doi.org/10.1093/femsmc/xtac003.

71. Feng S, Roguet A, McClary-Gutierrez JS, Newton RJ, Kloczko N, Meiman JG, McLellan SL. 2021. Evaluation of sampling, analysis, and normalization methods for SARS-CoV-2 concentrations in wastewater to assess COVID-19 burdens in Wisconsin communities. ACS Est Water 1:1955–1965. https://doi.org/10.1021/acsestwater.1c00160.

72. Daughton CG. 2018. Monitoring wastewater for assessing community health: Sewage Chemical-Information Mining (SCIM). Sci Total Environ 619–620:748–764. https://doi.org/10.1016/j.scitotenv.2017.11.102.

73. Brinkmann A, Ulm S-L, Uddin S, Förster S, Seifert D, Oehme R, Corty M, Schaade L, Michel J, Nitsche A. 2021. AmpliCoV: rapid whole-genome sequencing using multiplex PCR amplification and real-time Oxford Nanopore MinION sequencing enables rapid variant identification of SARS-CoV-2. Front Microbiol 12:1703. https://doi.org/10.3389/fmicb.2021.651151.

74. Kerkhof LJ. 2021. Is Oxford Nanopore sequencing ready for analyzing complex microbiomes? FEMS Microbiol Ecol 97:fiab001. https://doi.org/10.1093/femsec/fiab001.

75. Combe M, Garijo R, Geller R, Cuevas JM, Sanjuán R. 2015. Single-cell analysis of RNA virus infection identifies multiple genetically diverse viral genomes within single infectious units. Cell Host Microbe 18:424–432. https://doi.org/10.1016/j.chom.2015.09.009.

76. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, Ludden C, Reeve R, Rambaut A, Peacock SJ, Robertson DL, Consortium C-19 GUK. 2021. SARS-CoV-2 variants, spike mutations and immune escape. Nat Rev Microbiol 19:409–424. https://doi.org/10.1038/s41579-021-00573-0.

77. Newton RJ, McLellan SL, Dila DK, Vineis JH, Morrison HG, Eren AM, Sogin ML. 2015. Sewage reflects the microbiomes of human populations. mBio 6:e02574-14. https://doi.org/10.1128/mBio.02574-14.

78. Gerardi MH, Zimmerman MC. 2004. p 3–6. Wastewater pathogens. John Wiley & Sons. https://books.google.com/books?hl=en&lr=&id=FyIrqPeBf6YC&oi=fnd&pg=PR7&dq=Gerardi+MH,+Zimmerman+MC.+2004.+Wastewater+pathogens.+John+Wiley+%26+Sons.&ots=C9ez59SsMY&sig=qs7LMHBKM4vmKQcDCoG0qawktio#v=onepage&q=Gerardi%20MH%2C%20Zimmerman%20MC.%202004.%20Wastewater%20pathogens.%20John%20Wiley%20%26%20Sons&f=false.

79. Andersson P, Tong SYC, Lilliebridge RA, Brenner NC, Martin LM, Spencer E, Delima J, Singh G, McCann F, Hudson C, Johns T, Giffard PM. 2014. Multisite direct determination of the potential for environmental

80. contamination of urine samples used for diagnosis of sexually transmitted infections. J Pediatric Infect Dis Soc 3:189–196. https://doi.org/10.1093/jpids/pit085.

80. United States Census Bureau. 2020. Annual estimates of the resident population: April 1, 2010 to July 1, 2019. http://www.census.gov/.

81. Maryland Department of Health. 2021. Coronavirus Disease 2019 (COVID-19) Outbreak. https://coronavirus.maryland.gov.

82. Environmental Health Administration. 1978. Design Guidelines for Sewerage Facilities ,M-DHMH-EHA-S-001. Department of Health and Mental Hygiene, Maryland. https://books.google.com/books?id=JO6ctgAACAAJ.

83. INNOVAPrep. 2020. Concentrating pipette select wastewater application note-Revision B. https://uploads-ssl.webflow.com/57aa3257c3e841c509f276e2/5f888d1b3bddf35ae661965c_CONCENTRATINGPIPETTESELECT%20WASTEWATER%20APPLICATION%20NOTE%201.17.03%20PM-compressed.pdf.

84. Hemmes JH, Winkler K, Kool SM. 1960. Virus survival as a seasonal factor in influenza and poliomyelitis. Nature 188:430–431. https://doi.org/10.1038/188430a0.

85. Hemmes JH, Winkler KC, Kool SM. 1962. Virus survival as a seasonal factor in influenza and poliomyelitis. Antonie Van Leeuwenhoek 28:221–233. https://doi.org/10.1007/BF02538737.

86. PRISM Climate Group. 2022. PRISM climate data. https://www.prism.oregonstate.edu.

87. Andrews SC. 2019. FastQC. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

88. Brumfield KD, Hasan NA, Leddy MB, Cotruvo J, Rashed SM, Colwell RR, Huq A. 2020. A comparative analysis of drinking water employing metagenomics. PLoS One 15:e0231210. https://doi.org/10.1371/journal.pone.0231210.

89. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

90. Joint Genome Institute. 2020. BBMap short read aligner, and other bioinformatics tools. https://jgi.doe.gov/data-and-tools/software-tools/bbtools/.

91. Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–2993. https://doi.org/10.1093/bioinformatics/btr509.

92. Lax S, Smith DP, Hampton-Marcell J, Owens SM, Handley KM, Scott NM, Gibbons SM, Larsen P, Shogan BD, Weiss S, Metcalf JL, Ursell LK, Vázquez-Baeza Y, Van Treuren W, Hasan NA, Gibson MK, Colwell R, Dantas G, Knight R, Gilbert JA. 2014. Longitudinal analysis of microbial interaction between humans and the indoor environment. Science 345: 1048–1052. https://doi.org/10.1126/science.1254529.

93. Ponnusamy D, Kozlova EV, Sha J, Erova TE, Azar SR, Fitts EC, Kirtley ML, Tiner BL, Andersson JA, Grim CJ, Isom RP, Hasan NA, Colwell RR, Chopra AK. 2016. Cross-talk among flesh-eating Aeromonas hydrophila strains in mixed infection leading to necrotizing fasciitis. Proc Natl Acad Sci U S A 113:722–727. https://doi.org/10.1073/pnas.1523817113.

94. Connelly S, Fanelli B, Hasan NA, Kaleko M, Colwell RR. 2019. Oral metallo-Beta-Lactamase protects the gut microbiome from carbapenem-mediated damage and reduces propagation of antibiotic resistance in pigs. Front Microbiol 10:1–12. https://doi.org/10.3389/fmicb.2019.00101.

95. Brumfield KD, Cotruvo JA, Shanks OC, Sivaganesan M, Hey J, Hasan NA, Huq A, Colwell RR, Leddy MB. 2021. Metagenomic sequencing and quantitative real-time PCR for fecal pollution assessment in an urban watershed. Front Water 3:626849. https://doi.org/10.3389/frwa.2021.626849.

96. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Henry M, Stevens H, Szoecs E, Wagner H. 2020. Vegan: Community Ecology Package. https://cran.r-project.org/web/packages/vegan/vegan.pdf.

97. Williams RJ, Howe A, Hofmockel KS. 2014. Demonstrating microbial co-occurrence pattern analyses within and between ecosystems. Front Microbiol 5:358. https://doi.org/10.3389/fmicb.2014.00358.

98. Czeisler MÉ, Marynak K, Clarke KEN, Salah Z, Shakya I, Thierry JM, Ali N, McMillan H, Wiley JF, Weaver MD, Czeisler CA, Rajaratnam SMW, Howard ME. 2020. Delay or avoidance of medical care because of COVID-19–related concerns—United States, June 2020. MMWR Morb Mortal Wkly Rep 69:1250–1257. https://doi.org/10.15585/mmwr.mm6936a4.

99. Mina MJ, Andersen KG. 2021. COVID-19 testing: one size does not fit all. Science 371:126–127. https://doi.org/10.1126/science.abe9187.

100. Tangcharoensathien V, Bassett MT, Meng Q, Mills A. 2021. Are overwhelmed health systems an inevitable consequence of covid-19? Experiences from

China, Thailand, and New York State. BMJ 372:n83. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8896039/.

101. European Centre for Disease Prevention and Control. 2021. Considerations on the use of self-tests for COVID-19 in the EU/EEA. Author, Stockholm. https://www.ecdc.europa.eu/sites/default/files/documents/Considerations-for-the-use-of-self-tests-for-COVID-19-in-the-EU-EEA.pdf.

102. Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, Spitters C, Ericson K, Wilkerson S, Tural A, Diaz G, Cohn A, Fox L, Patel A, Gerber SI, Kim L, Tong S, Lu X, Lindstrom S, Pallansch MA, Weldon WC, Biggs HM, Uyeki TM, Pillai SK, Washington State 2019-nCoV Case Investigation Team. 2020. First case of 2019 novel coronavirus in the United States. N Engl J Med 382:929–936. https://doi.org/10.1056/NEJMoa2001191.

103. La Rosa G, Iaconelli M, Mancini P, Ferraro GB, Veneri C, Bonadonna L, Lucentini L, Suffredini E. 2020. First detection of SARS-CoV-2 in untreated wastewaters in Italy. Sci Total Environ 736:139652. https://doi.org/10.1016/j.scitotenv.2020.139652.

104. Wu F, Xiao A, Zhang J, Moniz K, Endo N, Armas F, Bonneau R, Brown MA, Bushman M, Chai PR, Duvallet C, Erickson TB, Foppe K, Ghaeli N, Gu X, Hanage WP, Huang KH, Lee WL, Matus M, McElroy KA, Nagler J, Rhode SF, Santillana M, Tucker JA, Wuertz S, Zhao S, Thompson J, Alm EJ. 2022. SARS-CoV-2 RNA concentrations in wastewater foreshadow dynamics and clinical presentation of new COVID-19 cases. Sci Total Environ 805:150121. https://doi.org/10.1016/j.scitotenv.2021.150121.

105. Olesen SW, Imakaev M, Duvallet C. 2021. Making waves: Defining the lead time of wastewater-based epidemiology for COVID-19. Water Res 202:117433. https://doi.org/10.1016/j.watres.2021.117433.

106. Lindner BG, Suttner B, Zhu KJ, Conrad RE, Rodriguez-R LM, Hatt JK, Brown J, Konstantinidis KT. 2022. Toward shotgun metagenomic approaches for microbial source tracking sewage spills based on laboratory mesocosms. Water Res 210:117993. https://doi.org/10.1016/j.watres.2021.117993.

107. Galani A, Aalizadeh R, Kostakis M, Markou A, Alygizakis N, Lytras T, Adamopoulos PG, Peccia J, Thompson DC, Kontou A, Karagiannidis A, Lianidou ES, Avgeris M, Paraskevis D, Tsiodras S, Scorilas A, Vasiliou V, Dimopoulos M-A, Thomaidis NS. 2022. SARS-CoV-2 wastewater surveillance data can predict hospitalizations and ICU admissions. Sci Total Environ 804:150151. https://doi.org/10.1016/j.scitotenv.2021.150151.

108. Miura F, Kitajima M, Omori R. 2021. Duration of SARS-CoV-2 viral shedding in faeces as a parameter for wastewater-based epidemiology: Re-analysis of patient data using a shedding dynamics model. Sci Total Environ 769:144549. https://doi.org/10.1016/j.scitotenv.2020.144549.

109. Murakami M, Hata A, Honda R, Watanabe T. 2020. Letter to the editor: wastewater-based epidemiology can overcome representativeness and stigma issues related to COVID-19. Environ Sci Technol 54:5311. https://doi.org/10.1021/acs.est.0c02172.

110. Izquierdo-Lara R, Elsinga G, Heijnen L, Munnink BBO, Schapendonk CME, Nieuwenhuijse D, Kon M, Lu L, Aarestrup FM, Lycett S, Medema G, Koopmans MPG, de Graaf M. 2021. Monitoring SARS-CoV-2 circulation and diversity through community wastewater sequencing, the Netherlands and Belgium. Emerg Infect Dis 27:1405–1415. https://doi.org/10.3201/eid2705.204410.

111. Crits-Christoph A, Kantor RS, Olm MR, Whitney ON, Al-Shayeb B, Lou YC, Flamholz A, Kennedy LC, Greenwald H, Hinkle A, Hetzel J, Spitzer S, Koble J, Tan A, Hyde F, Schroth G, Kuersten S, Banfield JF, Nelson KL. 2021. Genome sequencing of sewage detects regionally prevalent SARS-CoV-2 variants. mBio 12:e02703-20. https://doi.org/10.1128/mBio.02703-20.

112. CDC. 2022. COVID Data Tracker. https://covid.cdc.gov/covid-data-tracker/#datatracker-home.

113. Franke KR, Isett R, Robbins A, Paquette-Straub C, Shapiro CA, Lee MM, Crowgey EL. 2022. Genomic surveillance of SARS-CoV-2 in the state of Delaware reveals tremendous genomic diversity. PLoS One 17:e0262573. https://doi.org/10.1371/journal.pone.0262573.

114. Noh JY, Jeong HW, Shin E-C. 2021. SARS-CoV-2 mutations, vaccines, and immunity: implication of variants of concern. Signal Transduct Target Ther 6:203. https://doi.org/10.1038/s41392-021-00623-2.

115. Pan Y-Q, Bahoussi AN, Guo F, Xing L. 2021. A single nucleotide distinguishes the SARS-CoV-2 in the Wuhan outbreak in December 2019 from that in Beijing-Xinfadi in June 2020, China. New Microbes New Infect 39:100835. https://doi.org/10.1016/j.nmni.2020.100835.

116. Rangan R, Zheludev IN, Hagey RJ, Pham EA, Wayment-Steele HK, Glenn JS, Das R. 2020. RNA genome conservation and secondary structure in SARS-CoV-2 and SARS-related viruses: a first look. RNA 26:937–959. https://doi.org/10.1261/rna.076141.120.

117. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, Hengartner N, Giorgi EE, Bhattacharya T, Foley B, Hastie KM, Parker MD, Partridge DG, Evans CM, Freeman TM, de Silva TI, Group S-1G, McDanal C, Perez LG, Tang H, Moon-Walker A, Whelan SP, LaBranche CC, Saphire EO, Montefiori DC, Sheffield COVID-19 Genomics Group. 2020. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell 182:812–827.e19. https://doi.org/10.1016/j.cell.2020.06.043.

118. Lubinski B, Fernandes MV, Frazier L, Tang T, Daniel S, Diel DG, Jaimes JA, Whittaker GR. 2022. Functional evaluation of the P681H mutation on the proteolytic activation of the SARS-CoV-2 variant B.1.1.7 (Alpha) spike. iScience 25:103589. https://doi.org/10.1016/j.isci.2021.103589.

119. Xia S, Wen Z, Wang L, Lan Q, Jiao F, Tai L, Wang Q, Sun F, Jiang S, Lu L, Zhu Y. 2021. Structure-based evidence for the enhanced transmissibility of the dominant SARS-CoV-2 B.1.1.7 variant (Alpha). Cell Discov 7:109. https://doi.org/10.1038/s41421-021-00349-z.

120. Planas D, Bruel T, Grzelak L, Guivel-Benhassine F, Staropoli I, Porrot F, Planchais C, Buchrieser J, Rajah MM, Bishop E, Albert M, Donati F, Prot M, Behillil S, Enouf V, Maquart M, Smati-Lafarge M, Varon E, Schortgen F, Yahyaoui L, Gonzalez M, De Sèze J, Péré H, Veyer D, Sève A, Simon-Lorière E, Fafi-Kremer S, Stefic K, Mouquet H, Hocqueloux L, van der Werf S, Prazuck T, Schwartz O. 2021. Sensitivity of infectious SARS-CoV-2 B.1.1.7 and B.1.351 variants to neutralizing antibodies. Nat Med 27:917–924. https://doi.org/10.1038/s41591-021-01318-5.

121. Wibmer CK, Ayres F, Hermanus T, Madzivhandila M, Kgagudi P, Oosthuysen B, Lambson BE, de Oliveira T, Vermeulen M, van der Berg K, Rossouw T, Boswell M, Ueckermann V, Meiring S, von Gottberg A, Cohen C, Morris L, Bhiman JN, Moore PL. 2021. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. Nat Med 27:622–625. https://doi.org/10.1038/s41591-021-01285-x.

122. Shakya M, Lo C-C, Chain PSG. 2019. Advances and challenges in meta-transcriptomic analysis. Front Genet 10:904. https://doi.org/10.3389/fgene.2019.00904.

123. Rothman JA, Loveless TB, Kapcia J, Adams EE, Steele JA, Zimmer-Faust AG, Langlois K, Hideaki N, Wanless D, Griffith M, Mao L, Chokry J, Griffith JF, Whiteson KL. 2021. RNA viromics of Southern California wastewater and detection of SARS-CoV-2 single-nucleotide variants. Appl Environ Microbiol 87:e01448-21. https://doi.org/10.1128/AEM.01448-21.

124. Palanisamy V, Gajendiran V, Mani K. 2021. Meta-analysis to identify the core microbiome in diverse wastewater. Int J Environ Sci Technol 19:5079–5096. https://doi.org/10.1007/s13762-021-03349-4.

125. VandeWalle JL, Goetz GW, Huse SM, Morrison HG, Sogin ML, Hoffmann RG, Yan K, McLellan SL. 2012. Acinetobacter, Aeromonas and Trichococcus populations dominate the microbial community within urban sewer infrastructure. Environ Microbiol 14:2538–2552. https://doi.org/10.1111/j.1462-2920.2012.02757.x.

126. Acharya K, Blackburn A, Mohammed J, Haile AT, Hiruy AM, Werner D. 2020. Metagenomic water quality monitoring with a portable laboratory. Water Res 184:116112. https://doi.org/10.1016/j.watres.2020.116112.

127. Miao J, Dong X, Chi Y, Lin D, Chen F, Du Y, Liu P, Liu X. 2018. Pseudoperonospora cubensis in China: Its sensitivity to and control by oxathiapiprolin. Pestic Biochem Physiol 147:96–101. https://doi.org/10.1016/j.pestbp.2018.03.009.

128. Spanner C, Darienko T, Biehler T, Sonntag B, Pröschold T. 2020. Endosymbiotic green algae in Paramecium bursaria: a new isolation method and a simple diagnostic PCR approach for the identification. Diversity 12:240. https://doi.org/10.3390/d12060240.

129. Zhang H, Feng J, Chen S, Li B, Sekar R, Zhao Z, Jia J, Wang Y, Kang P. 2018. Disentangling the drivers of diversity and distribution of fungal community composition in wastewater treatment plants across spatial scales. Front Microbiol 9:1291. https://doi.org/10.3389/fmicb.2018.01291.

130. Assress HA, Selvarajan R, Nyoni H, Ntushelo K, Mamba BB, Msagati TAM. 2019. Diversity, co-occurrence and implications of fungal communities in wastewater treatment plants. Sci Rep 9:14056. https://doi.org/10.1038/s41598-019-50624-z.