The wastewater microbiome: a novel insight for COVID-19 surveillance

Cristian Gallardo-Escárate (✉ crisgallardo@udec.cl)
University of Concepción   https://orcid.org/0000-0002-7094-6702

Valentina Valenzuela-Muñoz
University of Concepción

Gustavo Núñez-Acuña
University of Concepción

Diego Valenzuela-Miranda
University of Concepción

Francisco Castellón
University of Concepción

Bárbara Benavente-Cartes
University of Concepción

Constanza Sáez-Vera
University of Concepción

Homero Urrutia
University of Concepción   https://orcid.org/0000-0001-8895-5933

Beatriz Novoa
Instituto de Investigaciones Marinas (IIM - CSIC)

Antonio Figueras
Instituto de Investigaciones Marinas (IIM - CSIC)

Steven Roberts
University Of Washington

Article

Abstract

Wastewater-Based Epidemiology is a tool to face and mitigate COVID-19 outbreaks by evaluating conditions in a specific community. This study aimed to analyze the microbiome profiles using nanopore technology for full-length 16S rRNA sequencing in wastewater samples collected from a penitentiary (P), a residential care home (RCH), and a quarantine or health care facilities (HCF). The HCF microbiome was strongly associated with enteric bacteria previously reported in patients with chronic disease and psychological disorders. During the study, the wastewater samples from the RCH and the P were negative for SARS-CoV-2 based on qPCRs, except during the fourth week when was detected. Unexpectedly, the wastewater microbiome from RCH and P prior to week four was correlated with the samples collected from the HCF, suggesting a core bacterial community is expelled from the digest tract of individuals infected with SARS-CoV-2. We provide novel evidence that the wastewater microbiome associated with gastrointestinal manifestations appears to precede the SARS-CoV-2 detection in sewage. This finding suggests that the wastewaters microbiome can be applied as an indicator of community-wide SARS-CoV-2 surveillance.

Introduction

The SARS-CoV-2 is a new member of the Coronaviridae family and the etiologic of COVID-19, officially declared a pandemic by the World Organization of the Health in January 2020. Patients can present a wide variety of symptoms, and the prognosis varies from mild or moderate illness to severe and death[1,2]. A significant percentage of infected are symptomatic, and between 18 and 40% of cases have evidenced to be asymptomatic, a condition that helps the silent spread of the disease[3,4].

The S protein gives the virus its distinctive crown of spikes and is responsible for binding to angiotensin-converting enzyme receptor 2 (ACE2), which allows the virus to enter the host cell[5,6]. These receptors are present in various human cell types, with particular abundance in respiratory and gastrointestinal epithelial cells[7]. Indeed, an analysis of the ACE2 receptor distribution in human tissues found the highest levels of expression in intestine[8,9]. Although respiratory symptoms are the most frequently described in patients with COVID-19, several studies have shown that the gastrointestinal tract can also be affected by SARS-CoV-2. A meta-analysis found that 15% of patients had gastrointestinal symptoms and that about 10% of patients had gastrointestinal symptoms but no respiratory symptoms[10,11].

In contrast, SARS-CoV-2 RNA has been found in the feces of people without gastrointestinal symptoms[12]. It has been found that more than half (55%) of those tested for fecal viral RNAs were positive and noted that the virus is excreted in the feces for long periods, and in some cases, beyond negative testing or with respiratory symptoms[13]. Wastewater-Based Epidemiology (WBE) is an epidemiological tool improve predictions and assist in mitigating COVID-19 outbreaks by evaluating biomarkers in a specific community[14]. The primary advantage is related to minimizing domino effects such as unnecessarily long stay-at-home policies that stress humans and economies alike[15]. While researchers have evaluated

biomarkers for SARS-CoV-2 there is no information of how the wastewater microbiome is altered in communities infected by coronavirus or if the enteric bacteria can display species-specific signatures detected in sewage either from symptomatic and asymptomatic patients. Notably, cumulative evidence has been associated with dysbiosis of the gastrointestinal microbiota, and the relationship with pathologies such as obesity, diabetes, immunosuppression states, hypertension among others[16-18]. With this in mind, according to the Center for Disease Control and Prevention (CDC), these pathologies increase the risk of severe COVID-19 in the human population. Surprisingly, the WBE tool has not yet been integrated with metagenomic analysis based on sequencing DNA technology and risk factors for viral diseases. This novel approach can help epidemiologists and public health officials improve SARS-CoV-2 surveillance outbreak predictions which could lead to better mitigation strategies.

In this study, we analyze the microbiome signals using nanopore technology for full-length 16S rRNA sequencing in samples collected from sewage produced in a residential care home (RCH), a penitentiary (P), and a healthcare facility (HCF). The HCF corresponds to quarantine accommodations, where COVID-19 positive patients stay-short periods in isolation for health recovery. For each location and sampling time-point, wastewaters were analyzed to detect SARS-CoV-2 by qPCR. The study was conducted from May to August 2020 in Chillan City, southern Chile. Notably, the sewage microbiome collected from HCF was strongly associated with enteric microbiota previously reported in patients with obesity, diabetes mellitus type-2, hypertension, and liver disease. It is important to note that during the study, the residential care home and the penitentiary were negative for SARS-CoV-2 qPCRs, except during the fourth week where the virus was detected. Unexpectedly, the wastewater microbiome from RCH and P prior to week four was correlated with the samples collected from the HCF, suggesting a core bacterial community is expelled from the digest tract of individuals infected with SARS-CoV-2. We provide novel evidence that the wastewater microbiome associated with gastrointestinal disorders appears to precede the SARS-CoV-2 detection in sewage. This finding suggests that integrating metagenomic analyses into WBE provided an early indicator for COVID-19 outbreaks.

Results

SARS-CoV-2 quantification in wastewater surveillance

From May to July 2020 in Chillan, Chile, we monitored wastewater SARS-CoV-2 load from a residential care home (RCH), penitentiary (P), and health care facilities (HCF) selected according to their social isolation level and risk factor. COVID-19 was confirmed by qPCR analysis of ORFab1 and N protein genes. The quantity of virus genome units per liter of sewage samples was calculated by a 10-log standard curve using 10-fold dilutions of a Positive control TaqMan[TM] 2019-nCoV Control Kit v1 (Thermo Fisher Scientific, USA) (Fig. S1). The positive control group, HCF, scored positive for SARS-CoV-2 genes with average C_T values of 32 and 33 for ORFab1 and N protein genes. These C_T values corresponded to an estimated maximum concentration of 1×10^4 Genome Unit/L (Fig. 1). Samples obtained from RCH, with middle-isolation level and high-risk factor, registered an outbreak of COVID-19 from July 10th to July 21st

(corresponding to week 4 of the study) with C_T values between 32 and 33 for both genes to an average virus load of 7×10^3 Genome Unit/L. In addition, penitentiary samples evidenced an outbreak of COVID-19 during July 10th with a 4×10^3 Genome Unit/L, observing C_T values of 35.6 and 33.4 for ORFab1 and N protein genes, respectively. The demography data obtained for each studied location and group of people showed substantial differences in sex proportions, age, and comorbidity. This last characteristic associated with risk factors for COVID-19 indicated less prevalence in youngest individuals and under confinement in prison.

Metagenomic analysis associated with COVID-19 outbreak

Metagenome analyses were conducted in sewage samples collected from the three different sites and four sampling points (Fig. 1) based on Oxford nanopore sequencing. Reads obtained from the DNA sequencing from the different samples collected were annotated through BLASTn against the NCBI RefSeq for 16S ribosomal RNA database for bacteria and archaea, and an equivalent to OTU (Operative Taxonomic Unit) frequency table was built. Rarefaction analysis indicated that all sequenced libraries reached a plateau, indicating that the sequencing effort was enough to capture most of the microbial variability in the samples (Fig. S2). In parallel, NCBI Taxonomy tree from the wastewater microbiome was produced, indicating high bacterial diversity represented by the most abundant (30) taxa (Fig. 2).

Diversity indexes were calculated based on the generated OTU table. These indexes included Shannon, Simpson's, Evenness, and Chao-1 indexes (Fig. 3). Even though no significant variations were perceived on Shannon, Simpson's, and Evenness, a marked variation in the Chao-1 index was evidenced in the different sampled sites and points. Thus, an increase in the diversity value from samples collected during the COVID-19 outbreak was observed using one index. Notably, the highest index for the residence care home and the penitentiary was observed at timepoint S3 (Fig. 3d). Concerning the health care facility (HCF), no differences in taxa diversity during the study were found. Interestingly, a variation in the relative abundance of identified phylum was observed. Specifically, there was a marked reduction in the relative abundance of Proteobacteria and corresponding increase in other phylum at the residential care home and the penitentiary during the outbreak (Fig. 3e).

There was substantial differences in microbiota communities in wastewater collected from all locations (Fig. 4). At the level of species, the penitentiary showed the highest number of exclusive taxa (304); meanwhile, the residence care home and the health care facilities samples contained 135 and 45 species, respectively. Notably, a core microbiome was observed with 333 bacterial species (Fig. 4A and Table 1S). We analyzed samples from all the localities according to the viral quantification with the aim of evaluating differences between exclusives and shared species for wastewaters tested for SARS-CoV-2. Samples collected from locations testing positive or negative for SARS-CoV-2 contained 521 and 198 exclusive taxa, respectively. There were 409 taxa present in both SARS-CoV-2 positive and SARS-CoV-2 negative samples. (Fig. 4B and Table 1S).

The relative OUT abundance was estimated at the genus and species level through normalizing the reads of each unit by the number of reads of the corresponding sample. The top-25 most abundant genera had a significant association with samples according to the SARS-CoV-2 detection in wastewater samples (Fig. 5A). Principal component analyses (PCA) relative to genera abundances as main variables, SARS-CoV-2 quantification (genome Unit/L) as a supplementary quantitative variable, and negative/positive samples characterization as main factors, revealed a differential association of these genera. Specifically, *Simpliscira*, *Flavobacterium*, *Acidovorax,* and *Acinetobacter* genera were associated with negative samples for SARS-CoV-2. In contrast, *Prevotella*, *Bacteroides*, *Aeromonas*, *Sulfurospirillum*, *Arcobacter*, *Tolumonas*, *Citrobacter*, *Zoogloea*, and *Janthinobacteriums* were more associated with positive virus detection in wastewater samples.

Regarding species, the top-25 most abundant species were evaluated in all the samples through heatmaps based on normalized read values, indicating several changes according to the presence of SARS-CoV-2 or the site location where the samples were collected (Fig. 5B). The top species associated with positive samples (particularly samples from HCF in all the sampling times, P and RCH on July 17th) were *Arcobacter suis*, *A. venerupis*, *A. cloacae*, *Prevotella copri*, *A. aquamarines,* and *Bacteroides vulgatus*. Notably, these species were also more abundant in P and RCH sampling points in the sampling time before the first SARS-CoV-2 detection (July 7th) even when the virus was not detected.

A subset of bacteria was used to evaluate the association with COVID-19 outbreaks. These taxonomic units were selected according to their functions and previously reported association with SARS-CoV-2 infections, including commensal symbionts and opportunistic pathogens[30]. Relative abundance was calculated, as previously described. Normalized reads of selected bacteria were correlated with viral load values through Pearson's correlations using all of the data. The commensal symbionts *Eubacterium ventriosum*, and commensal species from the *Roseburia* genus and the Lachnospiraceae family were significantly correlated to the viral load in the samples (Fig. 6A). However, the opportunistic pathogen species *Bacteroides nordii*, and species from the *Rothia* and *Veilonella* genera were positively correlated with SARS-CoV-2 levels in the sewage samples.

In contrast, pathogens from the *Clostridium*, *Actinomyces,* and *Streptococcus* genera were negatively correlated with the viral load, and the same was obtained for the commensal species from the Ruminoccoccaceae family. Evaluating how the abundances of this selected subset vary in the different locations, a heatmap of normalized reads was built (Fig. 6B). In general, the most variable location was the Penitentiary (P), where bacteria from the Lachnospiraceae and Rumnococcaceae families were the most abundant after the COVID-19 outbreak. A similar trend was also obtained for *Alistipes* genus in this location. Most of the selected bacteria were identified in the sampling point before COVID-19 outbreak detection (July 7th), particularly in the P and RCH locations. The location with constant COVID-19 positive detection (HCF) had lower variations of this group.

Discussion

Our data are consistent with the idea that the wastewater microbiome reveals species-specific profiling associated with human communities where SARS-CoV-2 is prevalent, independently the symptomatic/asymptomatic status of those individuals. This finding strongly suggests that the virus alters the microbiome of infected individuals similarly to risk factors of severe COVID-19[19]. Several studies have implicated the microbiome in a range of physiologic processes that are vital to host health, including energy, metabolism, gastrointestinal health, immune response, and neurophysiology[17,18]. The microbial community supports metabolic interactions connecting those with the host and its physiological state. Notably, the enteric microbiome has an active role in humans' well-being, involving complex molecular signaling between the gut-brain axis[19].

Moreover, dysbiosis in the gastrointestinal microbiome is strongly associated with chronic diseases and psychological disorders in humans[20-23]. Our study revealed substantial differences in the microbiota associated with wastewaters collected from three communities with different levels of social isolation and risk factors for COVID-19. However, a core microbiota community was significantly found in sewage laboratory-tested for SARS-Cov-2. This finding suggests that the putative microbiome dysbiosis of infected individuals is the primary factor for the wastewater microbiota variations. However, this study does not support how the risk factors for COVID-19 are involved in the magnitude of the microbiota dysbiosis, and consequently, in the wastewater microbiome associated with SARS-CoV-2 detection. Notably, the demographic data obtained for this study did not indicate a high prevalence of comorbidities related to the Penitentiary population. Unexpectedly, the analysis conducted in wastewaters collected from patients with positive SARS-CoV-2 testing indicated more exclusive bacterial species than samples associated with individuals negatives for coronavirus. Our results indicate that *Simpliscira*, *Flavobacterium*, *Acidovorax*, and *Acinetobacter* genera were negatively associated with SARS-CoV-2.

In contrast, *Prevotella*, *Bacteroides*, *Aeromonas*, *Sulfurospirillum*, *Arcobacter*, *Tolumonas*, *Citrobacter*, *Zoogloea*, and *Janthinobacteriums* showed a clear association with wastewaters positive for SARS-CoV-2. Moreover, the top-25 most abundant species found in sewage samples demonstrate several changes according to the presence of individuals infected with SARS-CoV-2. Specifically, *Arcobacter suis*, *A. venerupis*, *A. cloacae*, *Prevotella copri*, *A. aquamarines*, and *Bacteroides vulgatus* appear to precede the SARS-CoV-2 detection in sewage. Previous studies have suggested strong correlations among these taxonomic units with chronic diseases such as obesity[24], immune-mediated inflammatory diseases[25], and cardiovascular diseases[26].

Studies connecting the microbiota with distal organs in regulatory functions such as gut–lung axis have revealed a key role in the respiratory system[27]. It has been reported that respiratory viral infections alter the gut microbiota, modulating the antiviral immune response through TNF-α and CD8$^+$ cells[28]. For example, gastric disorders during respiratory syncytial virus (RSV) or influenza virus infection in mice resulted in significantly altered gut microbiota diversity, increased Bacteroidetes, and a concomitant decrease in Firmicutes phyla abundance. Viral infection also led to changes in the fecal gut metabolome, with a significant shift in lipid metabolism. Sphingolipids, polyunsaturated fatty acids, and the short-

chain fatty acid valerate were all increased in abundance in the fecal metabolome following RSV infection[29]. Our findings showed that *Eubacterium ventriosum*, and commensal species from the *Roseburia* genus and the Lachnospiraceae family were significantly correlated to the viral load in wastewater samples. Notably, opportunistic pathogens from the *Bacteroides* genera showed high modulation during the study, and some species positively correlated with the SARS-CoV-2 levels. This evidence can be connected with a recent study where the sphingosine gene expression was significantly correlated with increased *Bacteroides* in patients with acute respiratory viral infections[28] and gastrointestinal inflammation disorders[30].

Alterations in the gut microbiota of patients with COVID-19 during hospitalization have recently been studied[31]. Herein, the taxa reported with an association to SARS-CoV-2 infection were analyzed in our study to evaluate their abundance in wastewater samples. Interestingly, the Lachnospiraceae and Rumnococcaceae families, and the *Alistipes* genus were the most abundant after the COVID-19 outbreak in this study. However, a clear trend was seen in all the selected bacteria, where the abundance increased in the sampling point just before COVID-19 detection in sewage.

The modern world is interconnected, and globalization arrives in our society with unpredictable circumstances and risks for human health. Today we have created an extensive network and hubs for quickly traveling by air, ground, and sea. This superposition of human trajectories has never been considered the leading risk factor for human society and its survival. Despite the planning efforts carried out by each country, the current COVID-19 pandemic illustrates how even the most extensive works may be inadequate if the system is not designed to adopt novel strategies that quickly respond to the changing and unpredictable conditions involved outbreak infection diseases[32]. The COVID-19 pandemic has revealed several gaps to face virus disease[33]. Among them, SARS-CoV-2 surveillance is one of the most critical issues for preparedness between nations. The individual testing of SARS-CoV-2 by qPCR or alternative analytical methods is not enough to capture how coronavirus is spread among the human communities and how SARS-CoV-2 infects asymptomatic individuals can be detected as super-spreaders for COVID-19[34]. Recently, patients from Korea with COVID-19 in quarantine showed that the shedding of virus into feces persisted until day 50 after diagnosis even in patients with asymptomatic or mild disease, or with no gastrointestinal symptoms[35]. This information is pivotal to establish risk factors for COVID-19, preventing the severity and socio-economic impacts of COVID-19 outbreaks. Herein, we propose the integration of the wastewater microbiome as an indicator of community-wide SARS-CoV-2 surveillance. This approach, based on nanopore technology for full-16S rRNA sequencing, will substantially improve the Wastewater-Based Epidemiology to predict and mitigate COVID-19 outbreaks.

Methods

Sample collection. Untreated wastewater (sewage) samples were obtained from May to July 2020 in Chillan (~200,000 people), a city in southern Chile. Wastewater operators collected a 24h flow-dependent composite sample of 10 L using a submersible in-situ high-frequency autosampler. Samples were

transported on ice to the laboratory and stored at 4 °C until further analysis. For this study, three sewage sampled points were selected according to their social isolation level, risk factor, and the confirmation for COVID-19 by qPCR. Herein, for high-isolation level and low-risk factor, a penitentiary (P) comprising a total population of 818 people (convicts: 574 and correctional officers/administrative staff: 244) was evaluated. For a middle-isolation level and high-risk factor, a residential care home (RCH) composed of 42 people (28 residents and 14 assistants) was also analyzed. As a positive control group, sewage samples from patients who were positive for COVID-19, and isolated in health care facilities (HCF) for a short-stay until their recovery was selected. The HCF was comprised of 38-100 patients during the study and corresponded to a hotel supported by the Chilean Government as part of the strategy to control the coronavirus outbreak. All localizations were selected based on the epidemiological analysis provided by the Ministry of Health (MINSAL) and the Ministry of Science, Technology, Knowledge, and Innovation (MINCIENCIA), Chile.

Ethical approval. Ethical approval to conduct the analysis was granted by the institutional review board of the University of Concepción as the responsible ethics committee for all participating institutions. Written informed consent is not mandatory for wastewater samples. MINSAL, Chile granted the authorization to conduct the study.

Demographic characteristics. Age, sex, and comorbidities were documented for individuals where the data was available and given by the MINSAL authorities.

SARS-CoV-2 molecular detection. A subsample of 100 ml from the collected composite wastewaters was centrifugated at 4500xg for 30 min, followed by filtration of supernatant using Millipore Membrane Filter, 0.22 μm pore size. The filtrates were concentrated in Amicon®-50 kDa (Millipore) by centrifugation at 5000xg for 15 min. Finally, 200 μL of concentrate wastewater was used for RNA isolation using the SV Total RNA Isolation System (Promega, USA). The concentration and purity of the RNA were determined by Nanodrop™ One/One^C (Thermo Fisher Scientific, USA). The integrity of total RNA was evaluated by TapeStation 2200 system (Agilent Technologies, USA). For SARS-CoV-2 detection, RNAs were analyzed with TaqMan™ 2019nCoV Assay Kit v1 (Thermo Fisher Scientific, USA), by the detection of ORF1ab, N gene, and a positive control assay RNAaseP. RT-PCR reactions were performed using TaqPath^{TM} 1-Step RT-qPCR Master Mix (Applied Biosystems, USA) in a 10 μL reaction mixture with 3 μL of RNA. Positive control TaqMan^{TM} 2019-nCoV Control Kit v1 (Thermo Fisher Scientific, USA) was used. RT-PCR experiment was performed in QuantStudio^{TM} 3 Real-Time PCR system (Applied Biosystems, USA) under the following condition: UNG incubation 2 min at 25ºC, Reverse transcription 15 min at 50ºC, activation 2 min at 95ºC, 40 cycles of 3 sec at 95ºC and 30 sec at 60ºC. Results were interpreted as TaqMan™ 2019nCoV Assay Kit v1 (Applied Biosystems) instructions. Finally, for SARS-CoV-2 quantification, a standard curve was performed by serial dilution of TaqManTM 2019-nCoV Control Kit v1 and amplification of ORF1ab, N gene (Fig. 1S).

Microbiome sequencing. Wastewater samples from the three sample points and times were centrifuged at 4500xg for 30 min, and then the supernatant was removed, and the organic pellet was used for DNA

isolation using the QIAamp Fast DNA Stool Mini Kit (Qiagen, USA) following the manufacturer's instructions. DNA integrity was verified by electrophoresis as above mentioned, and concentration measured by fluorescence in a Qubit 4 (ThermoScientific, USA) using the Qubit dsDNA BR Assay Kit (ThermoScientific, USA) following the manufacturer's instructions. Genomic DNA samples were diluted to 10 ng/µL, and amplification of full-16S rRNA gene was performed by PCR in 25 µL with the 27 F 5'-AGAGTTTGATCCTGGCTCAG-3' y 1492 R 5' GGTTACCTTGTTACGACTT-3' primers and Taq DNA polymerase LongAmp (NewEngland Biolabs, USA). PCR conditions were the followings: initial denaturing step at 95 C for 1 min, followed by 25 cycles of 95 °C for 20 sec, 56 °C for 30 sec and 65 °C for 2 min, and an extension at 65 °C for 5 min. PCR products were evaluated by electrophoresis, as previously described.

After confirming the presence of full-16S rRNA amplicon, PCR products were purified with Agencourt AMPure XP beads (Beckman Coulter, USA) in a 1:2 sample-beads ratio. The mixture was incubated for 5 min at room temperature, and the sample placed in the magnetic stand for ethanol washing (freshly prepared at 70%). The washed mixture was resuspended in ultrapure sterilized water, and the purified product was separated from beads by placing it in a magnetic stand. The purified amplicons were used as a template for library synthesis using the 16S Barcoding Kit (SQK-RAB204, Oxford Nanopore Technologies, UK) following the manufacturer's instructions for 1D sequencing strategy. Purified 16S amplicon was mixed with barcodes accordingly, and using the LongAmp Taq Polymerase (New England Biolabs, USA). PCR was performed as described in the previous step, with a final reaction volume of 50 µL. The PCR product was purified again with Agencourt AMPure XP beads and incubated in HulaMixer for 5 min. After the magnetic beads washing step, the purified product was eluted in 10 µL of elution buffer (10 mM Tris-HCl pH8.0 with 50 mM NaCl).

The final concentration of the library was quantified by fluorometry using Qubit 4 (Thermo Fisher, USA) and evaluated by TapeStation Bioanalyzer 2200 system (Agilent Technologies, USA) using DNA ScreenTape (Agilent Technologies, USA) following the manufacturer's instructions. A single library was synthesized from wastewater¢s DNAs for three sample points at three different times. Then libraries were pooled in multiplex mode following the protocol of Oxford Nanopore Technologies, and the flowcell MK1 Spot-ON FLO-MIN107-R9 was accordingly loaded. For each sample point, three replicates were sequenced. A total of 36 libraries were run in three flowcells for the MinION platform (Oxford Nanopore Technologies, UK). Sequencing efficiency was monitored through the software MinKNOW 2.0 (Oxford Nanopore Technologies, UK). Finally, as internal control, the DNA of a microbial mock community (ZymoBiomics Microbial Community Standard, USA) was extracted and sequenced following the same described procedures to evaluate the sequencing run accuracy through the observed/expected taxa abundances.

Bioinformatic analyses. After the sequencing runs, the yielded fast5 files were base-called using Guppy (version3.2.2, Oxford Nanopore Technologies, UK), and a filter step was applied to retain only sequences with Q-score ≥7 (quality filter). Sequences were then evaluated using the Fastq 16S v3.2.1 (Oxford Nanopore Technologies, UK) pipeline, which included demultiplexing, primers, and barcode trimming. Sequences with ambiguous identification were excluded from subsequent analysis. BLASTN aligned

cleaned sequences against NCBI RefSeq for 16S ribosomal RNA database for bacteria and archaea within the EPI2ME software package (Oxford Nanopore Technologies), and only the sequences with identity ≥75% were retained (confidence level filter) as mentioned by Edwards, *et al*.[36]. The occurrence of each NCBI taxonomy identifier was evaluated for each sample using R, and the relative abundance of each NCBI taxa id was calculated per sample to normalize. Taxonomical identifiers were ordered, and to minimize singletons, doubletons product of sequencing errors taxon with <0.01% of relative abundance in at least one sample were eliminated. For each sample, an equivalent to OTU (Operative Taxonomic Unit) frequency table was built using R, and rarefaction analysis was performed to assess sequencing depth performance. The alpha diversity of microbiota communities associated with wastewater was compared among sample points by calculating the Shannon and Simpson 1-D diversity indices, the Chao-1 richness index and community evenness using "Vegan" package in R. The difference among sample points was evaluated with One-Way ANOVA followed by the posthoc multiple comparison Tukey-HSD test when applicable, considering a P-value < 0.05. Microbial community structure was analyzed based on taxon relative abundance data, and principal coordinate analysis (PCoA) were performed as a multivariate unsupervised data exploration. Variations of the microbial community along sample time points were explored by plotting relative bacteria abundance, which corresponded to normalized reads of each taxon by the total of reads of the corresponding sample. Also, multivariate exploratory analyses were performed through PCoA from the microbiome collected in positive and negative sewage samples for SARS-Cov2. Distance between communities of different sample and time points was evaluated at the species level by hierarchical clustering using the UPGMA method. A group of selected species and genera were also evaluated due to previous findings related to COVID-19 outbreaks[30]. Pearson's correlations between the relative abundance of these bacteria and the SARS-CoV-2 viral load in the samples were calculated using the R software by the Corrplot package[37].

Declarations

Data availability

Sequence data that support the findings of this study have been deposited in GenBank with the accession codes SRX8936604 - 12 (https://www.ncbi.nlm.nih.gov/sra/PRJNA656810).

Author contributions

C.G., V.V., H. U designed the experiment. F.C., B.B., and C.S. quantified the viral load in sewage samples. G.N., V.V., and D.V. conducted DNA sequencing and metagenomic analyses. C.G., V.V., G.N., D.V., B. N., A. F.,

and S. R., led the manuscript writing. All the authors revised and agreed with the final version of the manuscript.

Competing interest

The authors declare no competing interest.

References

1   Guan, W. *et al.* Clinical Characteristics of Coronavirus Disease 2019 in China. *New Engl J Med* **382**, 1708-1720, doi:10.1056/NEJMoa2002032 (2020).

2   He, F., Deng, Y. & Li, W. N. Coronavirus disease 2019: What we know? *J Med Virol* **92**, 719-725, doi:10.1002/jmv.25766 (2020).

3   Jiang, F. *et al.* Review of the Clinical Characteristics of Coronavirus Disease 2019 (COVID-19). *Journal of General Internal Medicine* **35**, 1545-1549, doi:10.1007/s11606-020-05762-w (2020).

4   Hadi, A. G., Kadhom, M., Hairunisa, N., Yousif, E. & Mohammed, S. A. A Review on COVID-19: Origin, Spread, Symptoms, Treatment, and Prevention. *Biointerface Research in Applied Chemistry* **10**, 7234-7242, doi:10.33263/briac106.7234.7242 (2020).

5   Ou, X. Y. *et al.* Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Commun* **11**, doi:10.1038/s41467-020-15562-9 (2020).

6   Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-+, doi:10.1038/s41586-020-2012-7 (2020).

7   Li, Y. W., Zhou, W., Yang, L. & You, R. Physiological and pathological regulation of ACE2, the SARS-CoV-2 receptor. *Pharmacological Research* **157**, doi:10.1016/j.phrs.2020.104833 (2020).

8   Vuille-dit-Bille, R. N., Liechty, K. W., Verrey, F. & Guglielmetti, L. C. SARS-CoV-2 receptor ACE2 gene expression in small intestine correlates with age. *Amino Acids*, doi:10.1007/s00726-020-02870-z.

9   Zhang, H. *et al.* Specific ACE2 expression in small intestinal enterocytes may cause gastrointestinal symptoms and injury after 2019-nCoV infection. *Int J Infect Dis* **96**, 19-24, doi:10.1016/j.ijid.2020.04.027 (2020).

10   Liang, W. C. *et al.* Diarrhoea may be underestimated: a missing link in 2019 novel coronavirus. *Gut* **69**, 1141-+, doi:10.1136/gutjnl-2020-320832 (2020).

11   Cheung, K. S. *et al.* Gastrointestinal Manifestations of SARS-CoV-2 Infection and Virus Load in Fecal Samples From a Hong Kong Cohort: Systematic Review and Meta-analysis. *Gastroenterology* **159**, 81-95, doi:10.1053/j.gastro.2020.03.065 (2020).

12   Xu, Y. *et al.* Characteristics of pediatric SARS-CoV-2 infection and potential evidence for persistent fecal viral shedding. *Nat Med*, doi:10.1038/s41591-020-0817-4.

13   Effenberger, M. *et al.* Faecal calprotectin indicates intestinal inflammation in COVID-19. *Gut* **69**, 1543-1544, doi:10.1136/gutjnl-2020-321388 (2020).

14   Daughton, C. G. Wastewater surveillance for population -wide Covid-19: The present and future. *Sci Total Environ* **736**, doi:10.1016/j.scitotenv.2020.139631 (2020).

15   de Lima, C. V. C. *et al.* Effects of quarantine on mental health of populations affected by Covid-19. *Journal of Affective Disorders* **275**, 253-254, doi:10.1016/j.jad.2020.06.063 (2020).

16   Lee, C. J., Sears, C. L. & Maruthur, N. Gut microbiome and its role in obesity and insulin resistance. *Annals of the New York Academy of Sciences* **1461**, 37-52, doi:10.1111/nyas.14107 (2020).

17   Cani, P. D. Human gut microbiome: hopes, threats and promises. *Gut* **67**, 1716-1725, doi:10.1136/gutjnl-2018-316723 (2018).

18   Cryan, J. F. *et al.* THE MICROBIOTA-GUT-BRAIN AXIS. *Physiol Rev* **99**, 1877-2013, doi:10.1152/physrev.00018.2018 (2019).

19   Zhou, F. *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* **395**, 1054-1062, doi:10.1016/s0140-6736(20)30566-3 (2020).

20   Jie, Z. Y. *et al.* The gut microbiome in atherosclerotic cardiovascular disease. *Nat Commun* **8**, doi:10.1038/s41467-017-00900-1 (2017).

21   Villette, R. *et al.* Unraveling Host-Gut Microbiota Dialogue and Its Impact on Cholesterol Levels. *Frontiers in Pharmacology* **11**, doi:10.3389/fphar.2020.00278 (2020).

22   Rieder, R., Wisniewski, P. J., Alderman, B. L. & Campbell, S. C. Microbes and mental health: A review. *Brain Behavior and Immunity* **66**, 9-17, doi:10.1016/j.bbi.2017.01.016 (2017).

23   Taylor, A. M. & Holscher, H. D. A review of dietary and microbial connections to depression, anxiety, and stress. *Nutritional Neuroscience* **23**, 237-250, doi:10.1080/1028415x.2018.1493808 (2020).

24   Crovesy, L., Masterson, D. & Rosado, E. L. Profile of the gut microbiota of adults with obesity: a systematic review. *European Journal of Clinical Nutrition*, doi:10.1038/s41430-020-0607-6.

25   Geva-Zatorsky, N. *et al.* Mining the Human Gut Microbiota for Immunomodulatory Organisms. *Cell* **168**, 928-+, doi:10.1016/j.cell.2017.01.022 (2017).

26   Tang, W. H. W., Kitai, T. & Hazen, S. L. Gut Microbiota in Cardiovascular Health and Disease. *Circ Res* **120**, 1183-1196, doi:10.1161/circresaha.117.309715 (2017).

27   Groves, H. T. *et al.* Respiratory Disease following Viral Lung Infection Alters the Murine Gut Microbiota. *Frontiers in Immunology* **9**, doi:10.3389/fimmu.2018.00182 (2018).

28   Groves, H. T., Higham, S. L., Moffatt, M. F., Cox, M. J. & Tregoning, J. S. Respiratory Viral Infection Alters the Gut Microbiota by Inducing Inappetence. *Mbio* **11**, doi:10.1128/mBio.03236-19 (2020).

29   Hasegawa, K. *et al.* Sphingolipid metabolism potential in fecal microbiome and bronchiolitis in infants: a case–control study. *BMC Research Notes* **10**, 325, doi:10.1186/s13104-017-2659-9 (2017).

30   Brown, E. M. *et al.* Bacteroides-Derived Sphingolipids Are Critical for Maintaining Intestinal Homeostasis and Symbiosis. *Cell Host & Microbe* **25**, 668-+, doi:10.1016/j.chom.2019.04.002 (2019).

31   Zuo, T. *et al.* Alterations in Gut Microbiota of Patients With COVID-19 During Time of Hospitalization. *Gastroenterology*, S0016-5085(0020)34701-34706, doi:10.1053/j.gastro.2020.05.048 (2020).

32   Cheng, Z. K. J. & Shan, J. 2019 Novel coronavirus: where we are and what we know. *Infection* **48**, 155-163, doi:10.1007/s15010-020-01401-y (2020).

33   Stawicki, S. P. *et al.* The 2019-2020 Novel Coronavirus (Severe Acute Respiratory Syndrome Coronavirus 2) Pandemic: A Joint American College of Academic International Medicine-World Academic Council of Emergency Medicine Multidisciplinary COVID-19 Working Group Consensus Paper. *J Glob Infect Dis* **12**, 47-93, doi:10.4103/jgid.jgid_86_20 (2020).

34   Casella, F. Can the COVID-19 Epidemic Be Controlled on the Basis of Daily Test Reports? *Ieee Control Systems Letters* **5**, 1079-1084, doi:10.1109/lcsys.2020.3009912 (2021).

35   Park, S. K. *et al.* Detection of SARS-CoV-2 in Fecal Samples From Patients With Asymptomatic and Mild COVID-19 in Korea. *Clin Gastroenterol Hepatol*, doi:10.1016/j.cgh.2020.06.005 (2020).

36   Edwards, A. *et al.* In-field metagenome and 16S rRNA gene amplicon nanopore sequencing robustly characterize glacier microbiota. *BioRxiv*, 073965 (2019).

37   Wei, T. *et al.* Package 'corrplot'. *Statistician* **56**, 316-324 (2017).
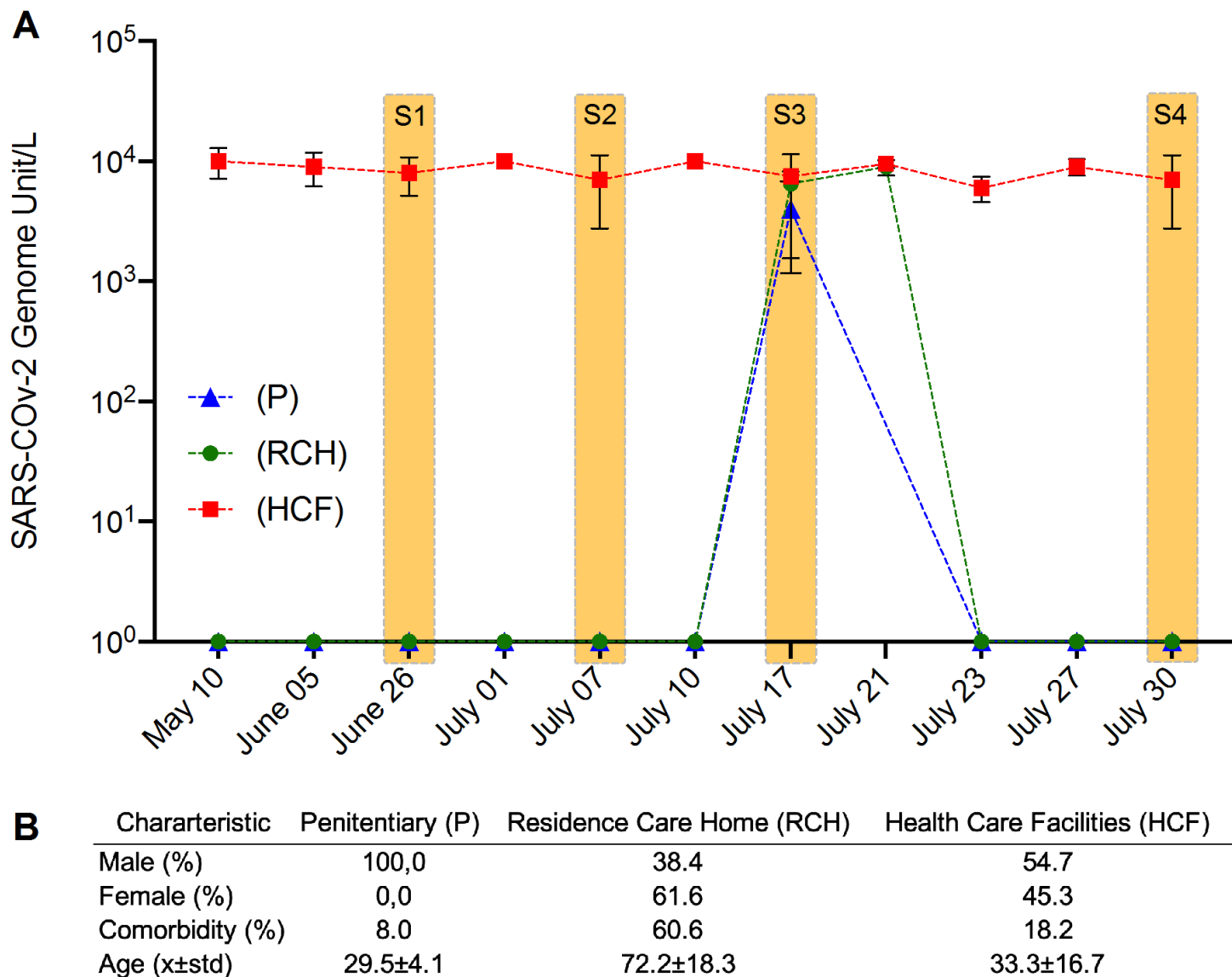
Figures

Figure 1

Experimental design and sampling points/times for microbiome sequencing. (A) SARS-CoV-2 quantification. The full-16S rRNA gene was sequenced using MinION platform (Oxford Nanopore Technology). The sequencing runs are indicated as orange rectangles (S1-4). The virus load was estimated by qPCR in untreated wastewaters from three locations: (P) Penitentiary, (RCH) Residence Care Home and (HCF) Health Care Facilities. The study was conducted from May to July 2020. (B) Demography data for P, RCH and HCF.
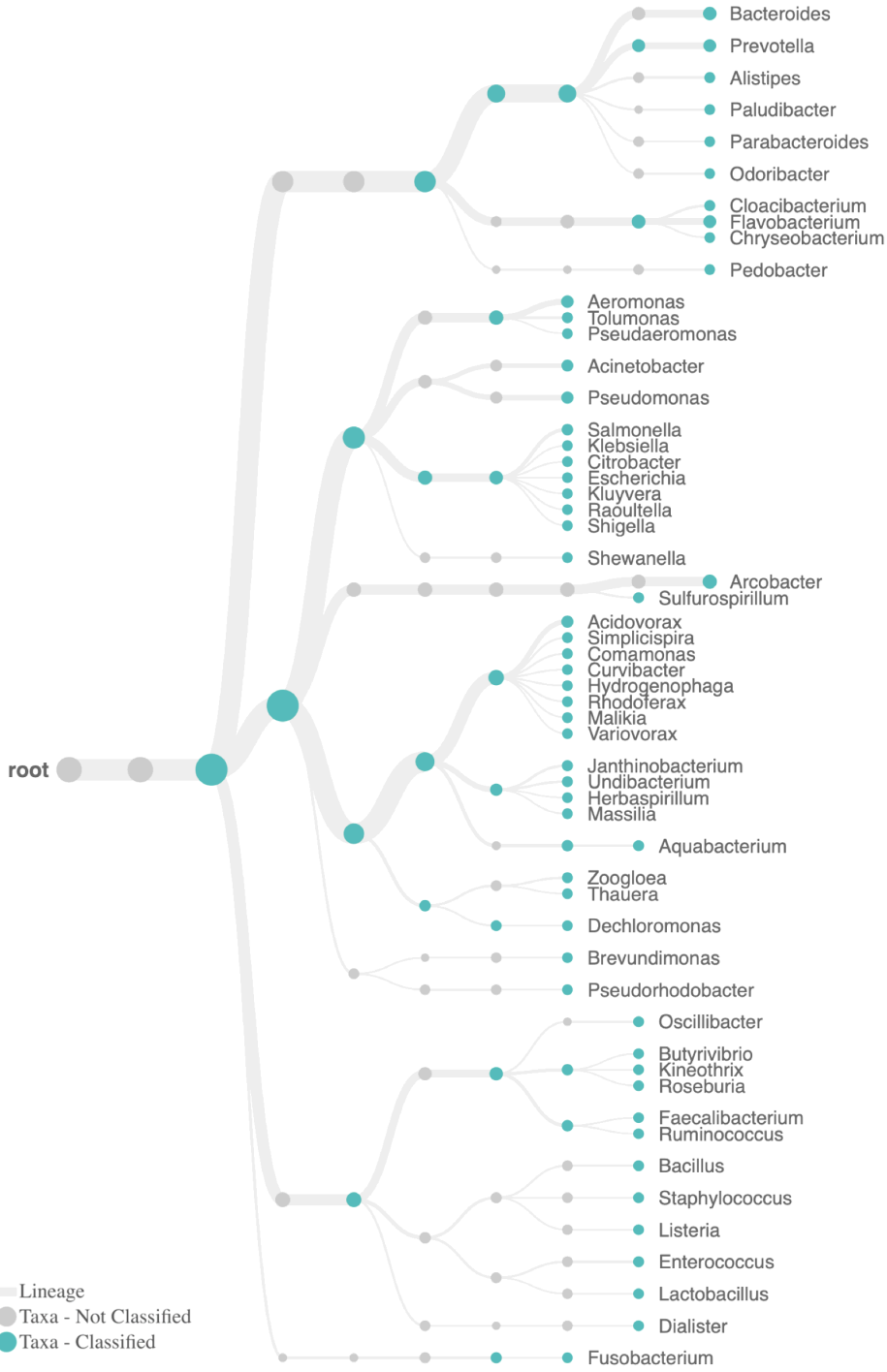
Figure 2

NCBI Taxonomy tree from the wastewater microbiome showing the top 30 taxa. Minimum abundance cutoff was 0.5%.
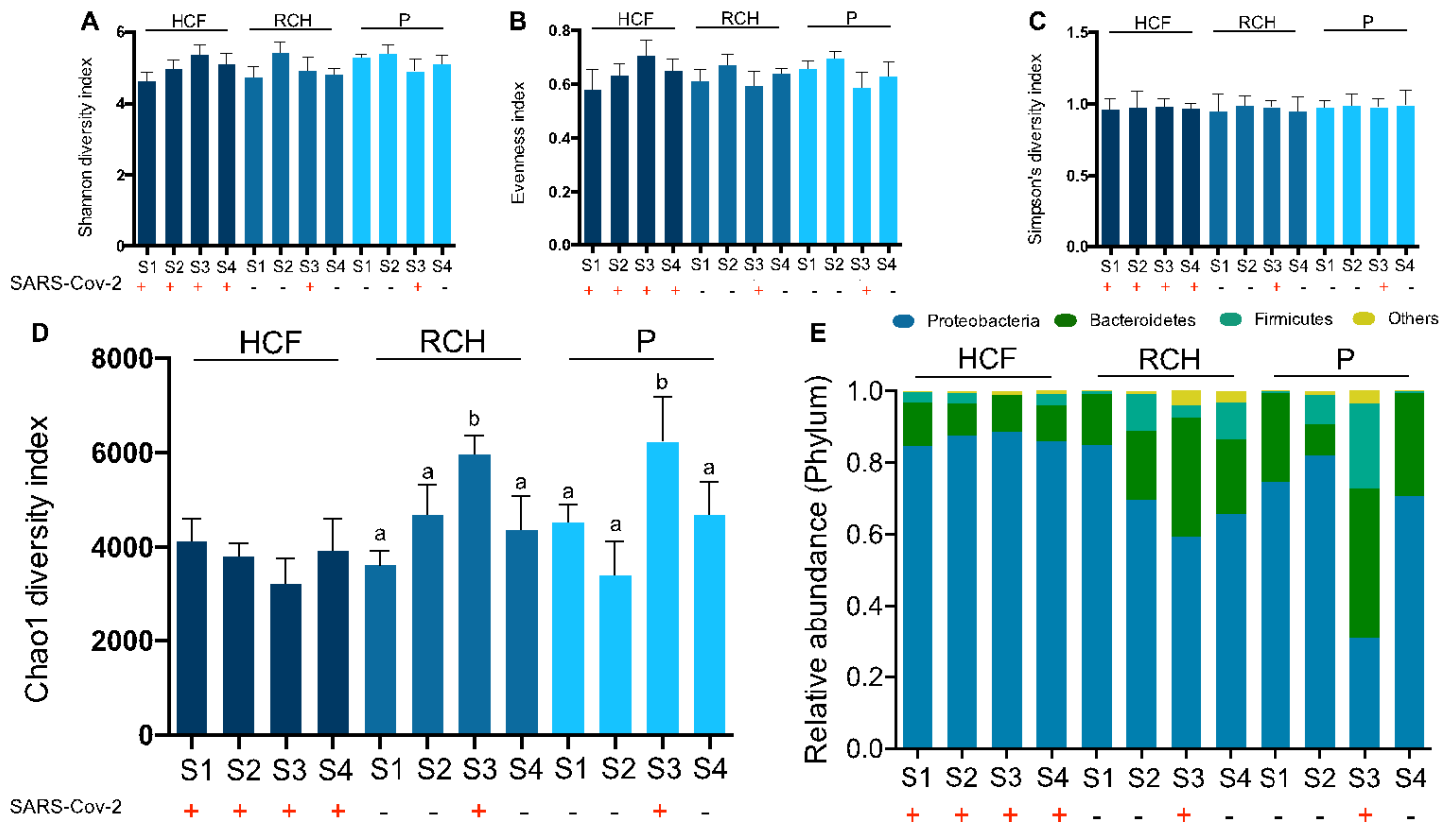
Figure 3

Diversity analysis of wastewater microbiome community associated with COVID-19 outbreak. Shannon (A), evenness index (B), Simpson (C) and richness Chao1 estimator (D). The relative abundance (Phylum) was estimated for Protobacteria, Bacteroidetes, Firmicutes and others. The letters S1-4 represent the nanopore sequencing conducted in sewage from (P) Penitentiary, (RCH) Residence Care Home and (HCF) Health Care Facility. Symbols (+) and (-) indicate the testing conducted for SARS-CoV-2.

A

B
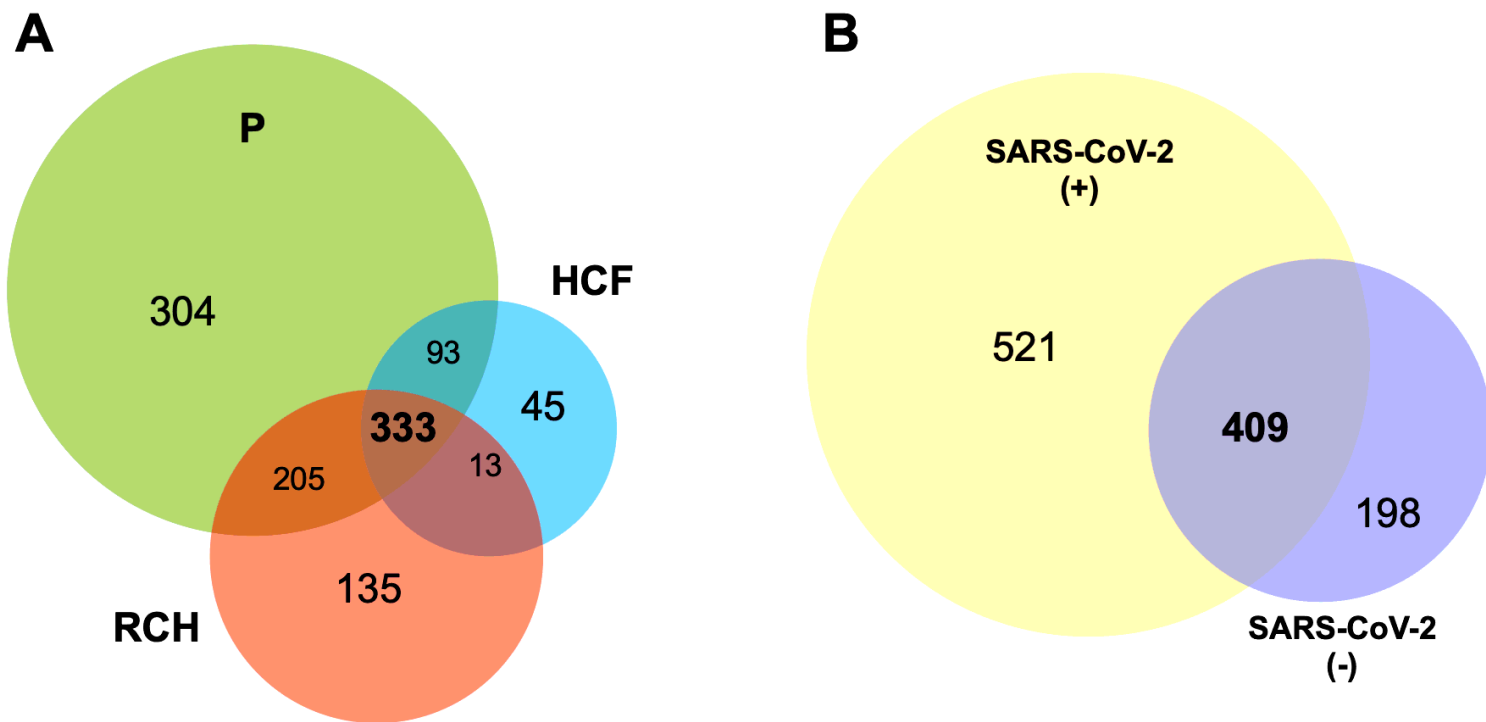
Figure 4

Venn diagram representing the taxonomical features identified in wastewater analyzed in this study. (A) Exclusives and shared bacteria among the Penitentiary (P), RCH) Residence Care Home and (HCF) Health Care Facility sewages. (B) Comparison between wastewaters samples collected from localities tested as positive and negative for SARS-Cov-2.
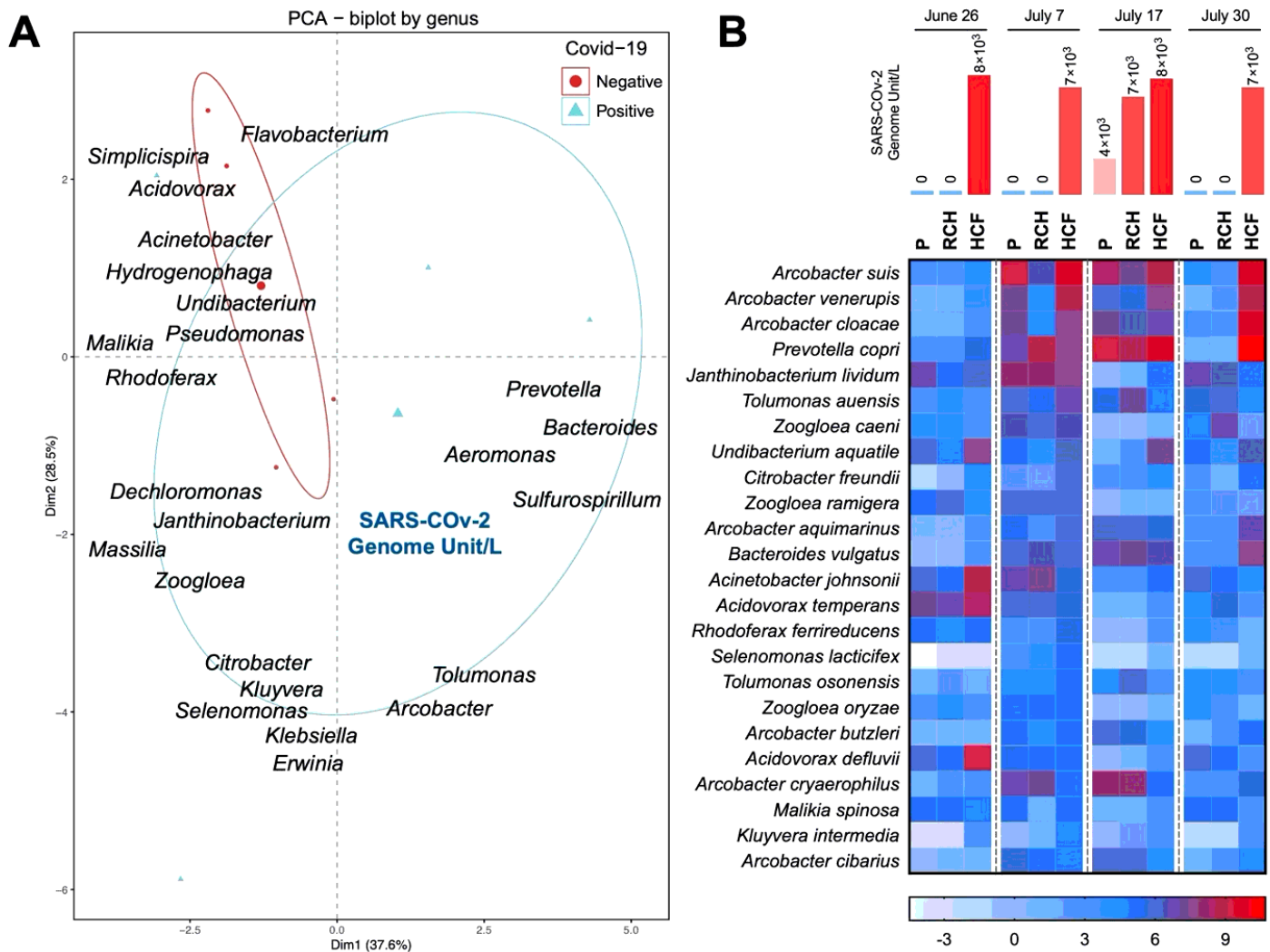
Figure 5

Microbiome relative abundance in wastewater associated with SARS-CoV-2 detection. Relative abundance was measured as the number of reads of top-25 genera/species normalized by the total of reads of a given sample. (A) Principal component analysis of top-25 most abundant genera using normalized reads as quantitative variables, the SARS-CoV-2 viral load as supplementary quantitative variable, and the qualitative negative/positive test for the virus as grouping factor. Confidence intervals are colored as these factors: blue for positive samples and red for negative. (B) Heatmap based on normalized relative abundance of top-25 species. Red color represents higher abundance, blue average abundance, and gray lower abundances. On top, the viral load per sampling point from (P) Penitentiary, (RCH) Residence Care Home and (HCF) Health Care Facility.
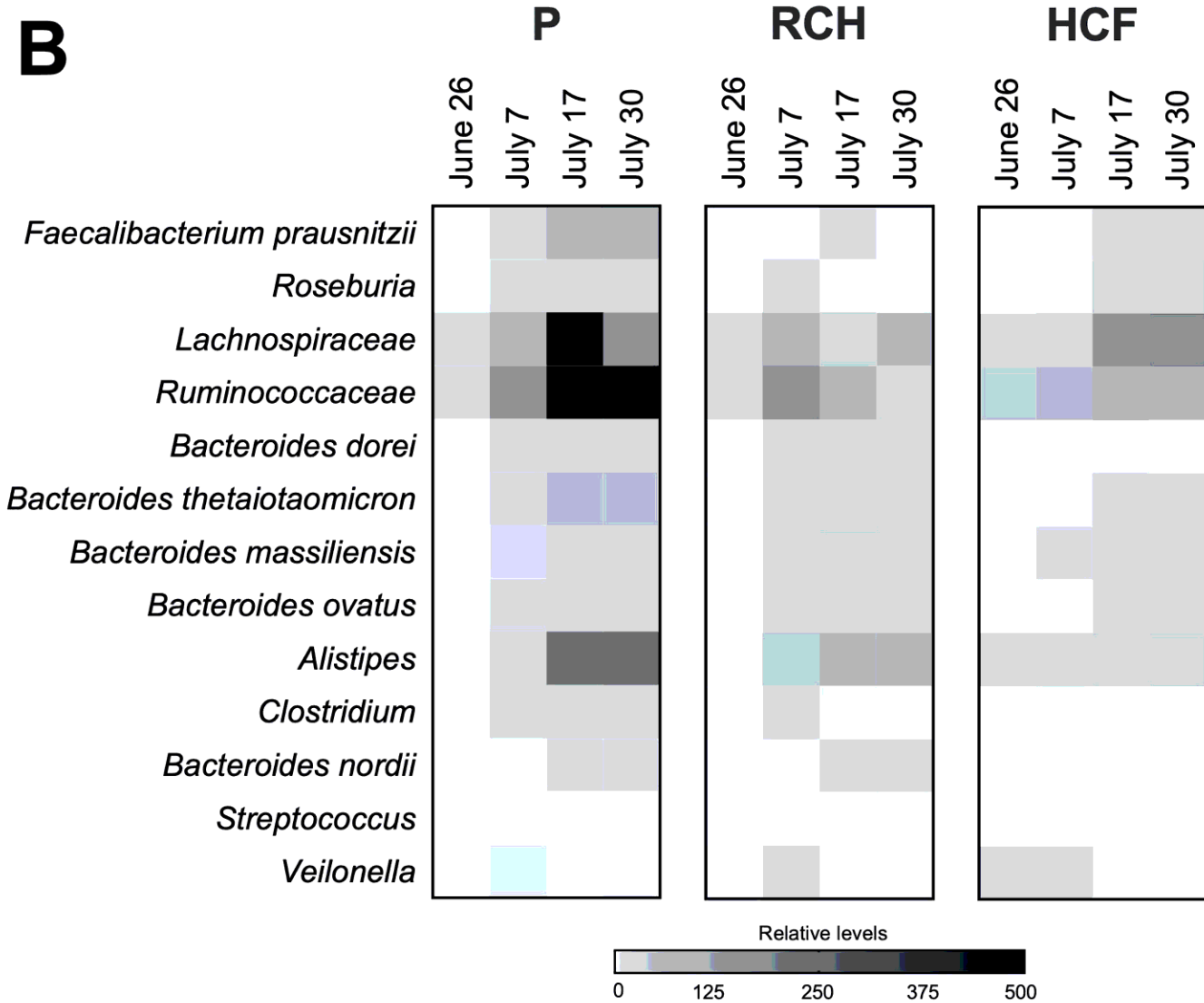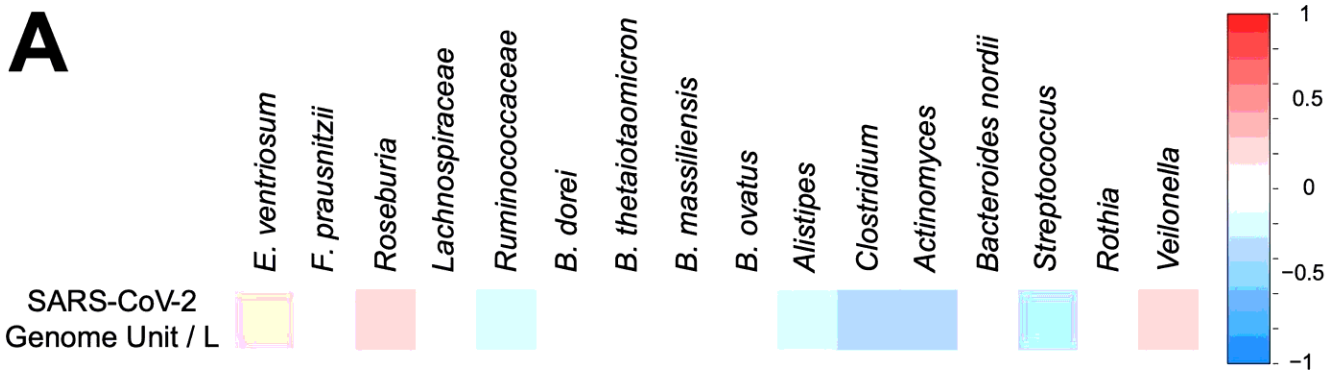
Figure 6

Relative abundances of selected variables and correlation with SARS-CoV-2 viral load. (A) Pearson's correlation analyses between normalized abundance of selected bacteria with viral load expressed as Genome Unit / L. In the correlation plot red color represent positive correlation between variables and blue inverse correlation. (B) Heatmap of relative abundance of selected bacteria from sampling location including (P) Penitentiary, (RCH) Residence Care Home and (HCF) Health Care Facility.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigures.docx
- SupplementaryTable1.xlsx