Environmental Research 276 (2025) 121509

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

From sewage to genomes: Expanding our understanding of the urban and semi-urban wastewater RNA virome

Sergio Guajardo-Leiva [a,b,c,*], Beatriz Díez [d,e,f], Cecilia Rojas-Fuentes [g,h], Jonás Chnaiderman [g], Eduardo Castro-Nallar [b,i], Valentina Catril [g], Manuel Ampuero [g], Aldo Gaggero [g,**]

[a] Dirección de Investigación, Vicerrectoría Académica, Universidad de Talca, Talca, Chile
[b] Centro de Ecología Integrativa, Universidad de Talca, Talca, Chile
[c] Facultad de Ciencias Agrarias, Universidad de Talca, Talca, Chile
[d] GEMA Center for Genomics, Ecology & Environment, Universidad Mayor, Santiago, Chile
[e] Center for Climate and Resilience Research (CR)2, Chile
[f] Millennium Institute Center for Genome Regulation (CGR), Chile
[g] Programa de Virología, ICBM, Facultad de Medicina, Universidad de Chile, Chile
[h] Millennium Institute on Immunology and Immunotherapy, Santiago, Chile
[i] Departamento de Microbiología, Facultad de Ciencias de la Salud, Universidad de Talca, Talca, Chile

ARTICLE INFO

Keywords:
RNA viruses
Wastewater RNA virosphere
Viral metagenomics
Uncultured RNA viral genomes

ABSTRACT

Wastewater is a hotspot for viral diversity, harboring various microbial, plant, and animal viruses, including those that infect humans. However, the dynamics, resilience, and ecological roles of viral communities during treatment are largely unknown.

In this study, we explored RNA virus ecogenomics using metagenomics from influent and effluent samples across three wastewater catchment areas in Chile, with a population of 7.05 million equivalent inhabitants. We identified 14,212 RNA-dependent RNA polymerase (RdRP)-coding sequences from the Orthornavirae kingdom, clustering into 4989 viral species. Using extensive databases of 14,150 family-level representative sequences, we classified 90 % of our sequences at the family level. Our analysis revealed that treatment reduced viral richness and evenness (Shannon index), but phylogenetic diversity remained unchanged. Effluents showed lower richness and evenness than influents with similar phylogenetic diversity. Species turnover, influenced by catchment area and treatment, accounted for 54 % of sample dissimilarities (Weighted Unifrac). Biomarker analysis indicated that families like Astroviridae and Fiersviridae were more abundant in influents, while Reoviridae and Virgaviridae dominated effluents. This suggests that viral resistance to treatment varies and cannot be solely attributed to genome type, size, or morphology. We traced viral genomes through time and space, identifying sequences like the Pepper Mild Mottle Virus (PMMoV) from the Virgaviridae family over large distances and periods, highlighting its wastewater marker potential. High concentrations of human pathogens, such as Rotavirus (Reoviridae) and Human Astrovirus (Astroviridae), were found in both influents and effluents, stressing the need for continuous monitoring, especially for treated wastewater reuse.

1. Introduction

RNA viruses can be important pathogens affecting humans and agricultural organisms, often transmitted through the fecal-oral route (Rothman et al., 2021). These viruses mutate and evolve more rapidly than DNA viruses, which allows them to adapt quickly to changing environments, hosts, or selective pressures (Sanjuán and Domingo-Calap, 2016). This rapid evolution results in untapped diversity and emerging variants that must be investigated to enhance our understanding of viral diversity and evolution.

Wastewater serves as a significant reservoir for various RNA viruses, including those that infect humans, plants, animals, fungi, and bacteria

* Corresponding author. Dirección de Investigación, Vicerrectoría Académica, Universidad de Talca, Talca, Chile.
** Corresponding author.
E-mail addresses: sergio.guajardo@utalca.cl (S. Guajardo-Leiva), agaggero@med.uchile.cl (A. Gaggero).

https://doi.org/10.1016/j.envres.2025.121509
Received 26 December 2024; Received in revised form 28 March 2025; Accepted 29 March 2025
Available online 2 April 2025

(Adriaenssens et al., 2021; Guajardo-Leiva et al., 2020; Martínez-Puchol et al., 2020; Zhang et al., 2024). Monitoring these viruses in wastewater has proven valuable for tracking viral diseases within populations, from the seminal 1932 study in Philadelphia, which tested the hypothesis of polio transmission via the fecal-oral route, to its recent application as an early warning system during the SARS-CoV-2 pandemic (Li et al., 2022; Paul et al., 1940; Rothman et al., 2021; Zhang et al., 2024). However, viral pathogens in wastewater present risks that extend beyond public health, affecting agriculture, water ecosystems, and overall environmental health (Al-Hazmi et al., 2023). Contamination of water sources designated for drinking or irrigation by wastewater treatment plants (WWTPs) can lead to gastrointestinal outbreaks in humans and viral infections in crops and livestock, particularly in areas lacking adequate sanitation or that rely on reclaimed water (Al-Hazmi et al., 2023; Farkas et al., 2020; Plaza-Garrido et al., 2022; Smith et al., 2024). This issue is particularly pertinent in regions experiencing freshwater shortages, where wastewater reuse may cause ecological disruptions (Al-Hazmi et al., 2023; Fernandes et al., 2023). Therefore, understanding the ecological dynamics of RNA viruses in wastewater is essential for mitigating health risks and enhancing treatment practices.

Wastewater treatment is essential for public health, particularly since effluents are frequently reused for irrigation, recreation, or discharged into rivers. The treatment process typically consists of preliminary, primary, secondary, and sometimes tertiary stages, depending on regulatory requirements (Al-Hazmi al., 2022; Mousazadeh et al., 2022; Plaza-Garrido et al., 2022; Qiu et al., 2015). Preliminary treatment encompasses the mechanical removal of coarse solids, sand, and oils. In primary treatment, high-density solids are settled by gravity, forming the primary sludge. Secondary treatment utilizes microorganisms to decompose organic matter into simpler compounds such as carbon dioxide, water, and methane, resulting in secondary sludge comprised of dead microorganisms and organic residues. The tertiary treatment further enhances effluent quality by removing nutrients and inert organic matter or by disinfection of pathogens as necessary (Al-Hazmi et al., 2022; Plaza-Garrido et al., 2022). However, viral inactivation during treatment relies on various physical and chemical factors, including temperature, light, pH, metal oxides, organic matter, colloids, ionic strength, salinity, ultraviolet (UV) exposure, and microbial interactions (Al-Hazmi et al., 2022; Lofrano and Brown, 2010; Mousazadeh et al., 2022). Membrane bioreactors (MBRs) have demonstrated higher efficiency in eliminating a wide range of viruses compared to conventional methods, including adenoviruses, noroviruses, enteroviruses, sapoviruses, rotaviruses, coliphages, and SARS-CoV- 2 (O'Brien and Xagoraraki, 2020). Despite the advantages of MBR technologies, chlorination remains the most commonly employed disinfection method due to its simplicity, high efficacy, and low cost. Chlorination exhibits virus-selective properties, with free chlorine being effective against enteroviruses and coxsackie virus, while monochloramine is more effective against adenoviruses, necessitating adjustments in dosage and contact time. Sodium hypochlorite has also effectively removed SARS-CoV-2 from wastewater (Plaza-Garrido et al., 2022; Polanco et al., 2023). These findings underscore the complexity of virus removal from wastewater, which depends on various factors and presents significant challenges to public health and safe wastewater reuse in irrigation. Furthermore, the specific impacts of these treatments on viral communities and their broader ecological implications remain poorly understood.

To better understand viral dynamics during wastewater treatment, we performed high-throughput sequencing (HTS) to comprehensively analyze the RNA virosphere in influent and effluent samples from three wastewater treatment plants in Chile. Our results revealed shifts in viral abundance, community composition, and diversity between influent and effluent samples, providing insights into how treatment processes impact viral populations in urban and semi-urban aquatic environments. These results emphasize the need for targeted strategies to enhance viral removal and ensure the safe reuse of treated wastewater, with broader implications for public health and environmental management.

2. Material and methods

2.1. Study site and sample preparation

El Trebal (33.541064 S, 70.835618 W) and La Farfana (33.472653 S, 70.788820 W) are urban municipal WWTPs in Santiago-Chile. San Pedro (36.975278 S, 73.174167 W) is a semi-urban municipal WWTP in Coronel, Chile (Fig. 1A). The El Trebal and La Farfana catchment areas serve 6.9 million (3.2M and 3.7M, respectively), while the San Pedro catchment area serves 0.15 million equivalent inhabitants. The three WWTPs employ activated sludge, biological tertiary treatment (anammox), and chlorine gas disinfection.

Thirteen composite samples, equivalent to 24 h of wastewater influent and effluent (1 L each), were obtained over several months from 2018 to 2019. Supplementary Table S1 provides detailed information about the samples.

Samples were processed by centrifugation at 5000*g* for 30 min, followed by sequential filtration through 8, 3, and 0.8 μm pore-size polycarbonate filters (Isopore, 47 mm diameter, Millipore, Milford, MA, USA) using a Swinex filter holder (Millipore, Milford, MA, USA) and a 0.22 μm pore-size PES Sterivex (Millipore, Milford, MA, USA). The 0.22 μm filtrate particles were concentrated by skim milk flocculation, as described by Gonzales-Gustavson et al. (2017), with in-house modifications. Briefly, the conductivity and pH of the filtered wastewater samples were adjusted to 1.5 mS and pH 3.5, respectively. Samples were flocculated with 0.01 vol of a 2 % pH 3.5 skim milk solution (Difco, Becton Dickinson, Wokingham, UK) with agitation at room temperature overnight or centrifugated. Flocs were resuspended in 4 mL of glycine buffer (glycine 0.25M, Tris 1M, EDTA 0.13M, pH 9.5) and amended with 4 mL of 2X PBS buffer. The resuspended flocs solution was centrifuged at 3,000g for 20 min to remove residual skim milk. Viral particles in the supernatant were concentrated by ultracentrifugation at 100,000 g for 60 min, and the pellet was resuspended in 1 mL of PBS buffer. The sample preparation process included a negative control (Milli-Q water).

2.2. RNA extraction and high-throughput sequencing

The resuspended viral particles (500 μL) and negative control were treated with 600U of DNAse-I and incubated for 12 h at 4 °C to remove the remaining free DNA from the cellular fraction, followed by inactivation of DNAse with EDTA to a final concentration of 100 mM. Viral RNA was extracted with TRIzol^{TM}LS following the manufacturer's instructions. The aqueous phase was mixed with one volume of 100 % ethanol, and the RNA was purified using the RNA Clean & Concentrator-5 kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's instructions. RNA was quantified using the Qubit RNA HS Assay and visualized in a 1 % Agarose gel. The RNA concentration in the negative control was below the detection threshold (<0.2 ng/μL).

Bacterial DNA contamination was assessed by 16S rRNA gene PCR amplification using a universal primer set, 515F (Parada) 5′-GTGYCAGCMGCCGCGGTAA-3' - 806R (Apprill) 5′-GGACTACNVGGGTWTCTAAT-3 (Thompson et al., 2017). No PCR product was visible on a 1 % agarose gel. The purified viral RNAs were sequenced by Illumina HiSeq technology (Roy J. Carver Biotechnology Center, Illinois, USA). Briefly, ribosomal RNA contamination was removed with the Ribo-Zero Plus kit (Illumina, San Diego, California, USA), and the RNAseq libraries were prepared with the Illumina TruSeq Stranded Total RNA Sample Prep kit (Illumina, San Diego, California, USA). The thirteen libraries were quantified by qPCR and sequenced in one lane for 151 cycles paired-end on a NovaSeq 6000 (Flow Cell Type SP). FASTQ files were generated and demultiplexed using the bcl2fastq v2.2 Conversion Software (Illumina, San Diego, California, USA).

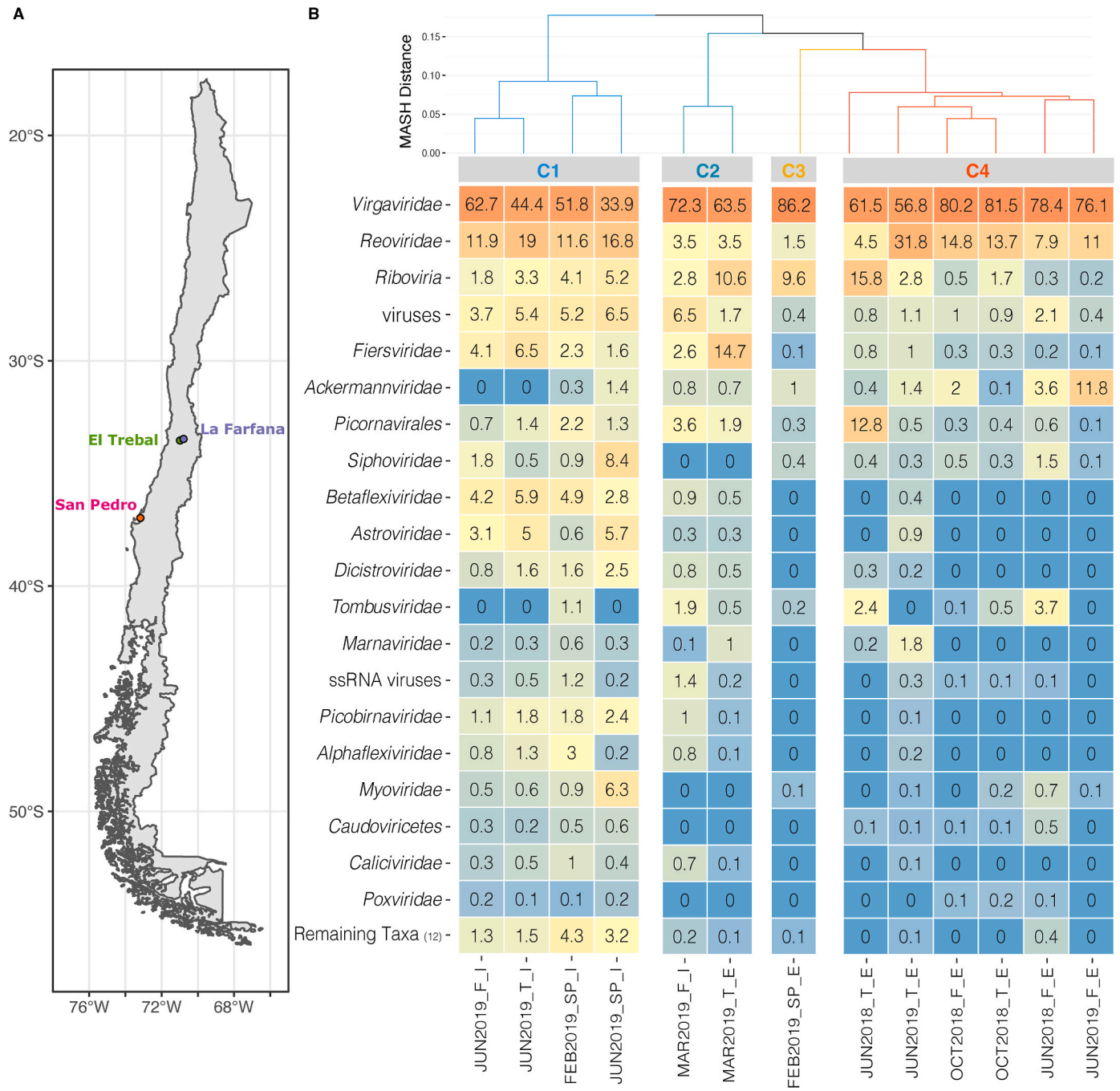


Fig. 1. A) Geographical location of El Trebal, La Farfana and San Pedro wastewater treatment plants in Chile. **B)** Hierarchical clustering and taxonomic composition of wastewater viral communities. The dendrogram was constructed using Ward's minimum variance method based on a matrix of MASH distances from metagenomic reads. Taxonomic classification of reads was performed with the LCA algorithm through local alignment to the NCBI nr database. Heatmap colors represent relative abundances on a logarithmic scale, normalized by library size. Samples were coded as follows: F_E/F_I (Farfana Effluent/Influent), T_E/T_I (Trebal Effluent/Influent), and SP_E/SP_I (San Pedro Effluent/Influent). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

2.3. RNA viral metagenome processing and assembly

Raw metagenomic reads were filtered by quality using FASTX-Toolkit (RRID:SCR_005534) using hard-clipping (fastx_trimmer) of the ten paired-bases (pb) in the 5′ end and five pb in the 3′ end. The fastx_artifacts_filter script filtered sequencing artifacts, and sequencing adaptors were removed using the fastx-clipper script. The remaining unpaired sequences were removed using Pairfq (https://github.com/sestaton/Pairfq), and the paired sequences were retained (makepairs). Ribosomal RNA sequences were removed using SortMeRNA software (Kopylova et al., 2012) and the SILVA database Release 138.1 (Quast et al., 2013). Detailed information on the number of sequences obtained is shown in Supplementary Table S1.

Viral RNA metagenomes were assembled using De Bruijn graphs implemented in the SPAdes assembler (Bankevich et al., 2012) in metagenomic mode (metaSPAdes). Only contigs sequences >200 pb were further analyzed. The remaining contigs from each metagenome (>200 pb) were *in silico* decontaminated using BLASTN (Camacho et al., 2009) (-evalue 0.00001, query coverage of ≥5 %) against Bacteria, Archaea, Eukarya, and phiX-174 sequences in the NCBI Nucleotide database as

described in reference (Guajardo-Leiva et al., 2021).

2.4. Comparison and taxonomic assignment of viral reads

Reads from each metagenome were mapped to each assembly (clean contigs) using Bowtie2 (Langmead and Salzberg, 2012) parameters (–end-to-end –very-sensitive -N 1). We used the aligned read pairs (–al-conc) to generate FASTQ paired-end files to obtain a putative set of viral reads.

Viral reads were used to determine pairwise mutation distances from the samples using the MinHash dimensionality-reduction technique implemented in MASH (Ondov et al., 2016) parameters (-k 21 and -s 1000000). The genetic distance between samples (MASH distance) was visualized by hierarchical clustering (hclust function in R) using the minimal increase of the sum of squares method (Ward.d2 function in R). The most optimal number of clusters was determined by a silhouette plot analysis using the silhouette function in R.

Viral reads were aligned against viral sequences in the NCBI nr database using DIAMOND (Buchfink et al., 2015) (blastx, –evalue 0.00001, –taxonlist 10239) and parsed using the lowest common ancestor algorithm through MEGAN 6 (Huson et al., 2016) (LCA score = 50) using the NCBI taxonomy tree obtaining the taxonomic annotation of each viral read according to NCBI taxonomy. We constructed heatmaps of the most abundant taxa (>1 %) agglomerated at the genus or the best taxonomic hit levels to assess the taxonomic composition of viral communities using the Ampvis2 v2.7.6 package (Andersen et al., 2018).

2.5. RNA-dependent RNA polymerase (RdRP) identification and taxonomic classification

We used the clean contig dataset (3618612 contigs) to identify RdRP sequences using Position Specific Scoring Matrices (PSSM) of the polymerase-palmprint (pp) motif implemented in the Palmscan software (Babaian and Edgar, 2022) that searches for a defined segment of the palm sub-domain delimited by conserved catalytic motifs, as described in reference (Edgar et al., 2022).

To generate a non-redundant polymerase-palmprint contig dataset (pp-contigs), pp positive contigs (14212 contigs) were clustered at 90 % of nucleotide identity across the 80 % of the shortest sequence as described in reference (Zayed et al., 2022) using cd-hit-est (Li and Godzik, 2006) parameters (-c 0.90 -n 8,9 -G 0 -aS 0.8 -g 1) resulting in 4989 pp-contigs.

To further classify the pp-contigs according to their pp-motifs, we constructed a unified RdRP database. As the first step, we identified pp-motifs in the IMG/VR v3.0 (Roux et al., 2020) and the RVMT (Neri et al., 2022) databases to subsequently cluster these sequences with those obtained from the PALMdb database (Babaian and Edgar, 2022; Edgar et al., 2022) at 40 % of aminoacidic identity using usearch (Edgar, 2010) parameters (-cluster_fast -id 0.4) defined in (Edgar et al., 2022). The resulting unified RdRP database comprises 951540 pp-motifs clustered into 14150 centroid sequences corresponding to family-level representative sequences.

To determine the taxonomic classification of each sequence in the pp-contig dataset, we clustered the 4989 wastewater pp-motifs with 14150 pp-motifs of reference. Clustering was done at the viral family threshold (≥40 % of the pp-motifs amino acid identity) as in (Edgar et al., 2022) using cd-hit (Li and Godzik, 2006) parameters (-c 0.4 -n 2 -G 1 -g 1). Clustering-based classification allowed us to assign taxonomy to 4523 pp-contigs sequences, following the taxonomic classification of the reference sequences with which they cluster.

2.6. RNA-dependent RNA polymerase (RdRP) diversity

Palmprint motifs nucleotide sequences from pp-contigs were quantified in each viral metagenome using bowtie2 (Langmead and Salzberg, 2012) parameters (–end-to-end –very-sensitive -N 1). Relative abundances of pp-motifs were normalized by gene length and library size of each viral metagenome. The normalized pp-motifs abundances were used to calculate alpha diversity (Chao1 index, Pielou index, Effective Number of Species [ENS] from Shannon H index, and Faith's Phylogenetic Diversity [PD]) and beta diversity (Bray-Curtis and Weighted-Unifrac) with the Phyloseq package in R (McMurdie and Holmes, 2013). The normality of the alpha diversity distribution was assessed using a Shapiro-Wilk test via the shapiro.test function in R. Heteroscedasticity of alpha diversity across predictors was evaluated with Levene's test implemented in the leveneTest function from de car package in R. To test the statistical significance of alpha diversity variation across predictor variables (i.e., wastewater type, catchment area, city, year, and month), we employed a type II Analysis of Variance (ANOVA) implemented in R (Anova function from the car package).

We fitted a generalized linear model (the glm function from the stats package in R, with a Gaussian distribution) to model the alpha diversity variation between influent and effluent wastewater samples. We visualized the generalized linear model using the Visreg package in R. the model's statistical significance was assessed using a Type II ANOVA implemented in the Anova function from the car package in R.

To identify drivers of beta diversity, we first tested whether the dispersion among groups was homogeneous using the function betadisper implemented in the Vegan v2.5.7 package (Oksanen, 2020). Then, we assessed the statistical significance of the beta diversity among wastewater type, catchment area, and their interaction using a Permutational Multivariate Analysis Of Variance (PERMANOVA) and pairwise comparisons with pairwise-PERMANOVA implemented in the adonis and pairwise.adonis functions in the Vegan v2.5.7 (Oksanen, 2020) and pairwiseAdonis packages, respectively.

We performed a constrained ordination analysis to extract and summarize the sample variation that the explanatory variables can explain. To decide whether to apply the linear or unimodal ordination method, we first calculated a Detrended Correspondence Analysis (DCA) using the decorana function in the Vegan v2.5.7 package (Oksanen, 2020). Then, we carried out a redundancy analysis (RDAs) using the Hellinger transformed Bray-Curtis and Weighted-Unifrac distances based on the pp-motifs abundance matrix using Ampvis2 (Andersen et al., 2018). To avoid overfitting, we reduced the number of explanatory variables using a forward variable selection procedure implemented in the ordi2step function in the Vegan v2.5.7 package(Oksanen, 2020). The statistical significance of the selected variables was quantified by an ANOVA test with Holm correction for multiple testing. The statistical significance of the RDAs was tested using 9999 permutations.

2.7. Core virome and biomarkers analyses

To visualize unique and shared pp-motifs between influent and effluent wastewater samples, we used Venn diagram analysis implemented in the R package "microeco v0.6.0" (Liu et al., 2021). To differentiate viral markers in each wastewater type (influent and effluent), we fitted a Negative Binomial model for each pp-motif and performed a differential abundance testing (Wald test) implemented in DESeq2 (Love et al., 2014) on taxa that were present in the 100 % of the samples. Viral markers (13 pp-motifs) were selected based on their statistical significance (p-value ≤0.05).

2.8. Identification of viral operational taxonomic units

To generate a set of RNA viral genome representatives (RNA vOTUs), we filtered the pp-contig dataset by length, keeping contigs ≥1 kb, obtaining 3718 contigs as described in (Neri et al., 2022; Zayed et al., 2022). We used CheckV (Nayfach et al., 2021) to estimate the genome completeness for the 3718 contigs as described in (Zayed et al., 2022) with an in-house updated database (2,086,109 sequences) derived from a manually curated subset of NCBI nt viral sequences (taxid 10239), downloaded on May 19, 2023. We kept 3588 contigs with a quality

indicator from low to complete, discarding all the undetermined quality contigs (130 sequences). From now on, we will refer to this set of 3588 contigs as the vOTUs dataset.

To explore the novelty of our vOTUs dataset in the viral genomic space, we generated a protein monopartite network implemented in vContact2 (Jang et al., 2019). Briefly predicted proteins from the vOTUs dataset (3588 contigs) and Orthornavirae (all RNA viruses encoding an RdRP) reference sequences (81461 contigs) from IMG/VR (Roux et al., 2020) and RVMT (Neri et al., 2022) were compared by DIAMOND (Buchfink et al., 2015) in an all-versus-all pairwise comparison (-evalue 0.00001, bitscore 50). Protein clusters were subsequently identified using the ClusterONE algorithm (Nepusz et al., 2012) based on DIAMOND E values, building protein cluster profiles for each genome and generating a similarity network. For network visualization, we used an edge-weighted spring-embedded model implemented in Cytoscape 3.7.1 (Shannon et al., 2003).

The relative abundance of vOTUs was obtained in each viral metagenome by read-mapping using Bowtie2 (Langmead and Salzberg, 2012) parameters (–end-to-end –very-sensitive -N 1). Relative abundances of vOTUs were normalized by sequence length and sequencing depth using the trimmed mean of the M-values algorithm, implemented in the edgeR package (Robinson et al., 2010) as described in (Zayed et al., 2022). Only vOTUs with 30 % of the length or 1 kb horizontally covered by reads at 95 % identity were considered present in the sample.

To identify vOTUs (relative abundances above 0.5 % across all samples) homologous sequences in the NCBI nt and nr databases, we performed BLAST (Camacho et al., 2009) analyses (BLASTn or BLASTx if no hit was obtained by BLASTn) using the following cutoffs (evalue 0.00001, pident 90 %).

3. Results

3.1. Overview of wastewater RNA viral communities

RNA viral metagenomes from wastewater in three catchment areas were analyzed to assess and compare the genetic diversity of viral communities in urban and semi-urban environments in Chile. Samples were collected from three wastewater treatment plants: El Trebal and La Farfana in Santiago and San Pedro in Coronel (Fig. 1A). Hierarchical clustering (Fig. 1B) identified four distinct viral clusters (C1–C4), which differentiated influents from effluents and suggested that wastewater treatment processes influence the genetic structuring of the samples. No evidence indicated that samples were structured based on the viral enrichment procedure.

Taxonomic classification identified 136.43 million reads across all samples, with the most abundant viral groups being plant viruses from the *Virgaviridae* (33.9–86.2 %) and *Betaflexiviridae* (0–5.9 %) families, followed by animal viruses from the *Reoviridae* (1.5–31.8 %) and *Astroviridae* (0–5.7 %) families, and bacteriophages from the *Fiersviridae* (0.1–14.7 %) and *Ackermannviridae* (0–11.8 %) families. Additionally, 0.4 %–6.5 % of the sequences were identified as having viral origins (viruses), but LCA did not provide a more precise taxonomic resolution.

The analysis indicated a dominance of viruses targeting plant, animal, and bacterial hosts.

3.2. Diversity of RNA viruses in wastewater influents and effluents

To assess the diversity of RNA viruses in our wastewater dataset, we identified 14,212 contigs containing RNA-dependent RNA polymerase (RdRPs) by detecting a polymerase palm print motif (pp-motif) in the polymerase catalytic core. These contigs were refined into a non-redundant dataset ($\geq$90 % nucleotide identity at $\geq$ 80 % sequence length), resulting in a set of 4989 "pp-contigs."

To determine the taxonomic classification of the pp-contig sequences, we conducted clustering with 14,227 reference sequences at the family-level threshold (40 % identity in the palmprint amino acid sequence). This process resulted in the classification of 4523 sequences (Fig. 2A), with the most prevalent sequences belonging to the *Fiersviridae* (1376 sequences), *Picobirnaviridae* (445 sequences), *Dicistroviridae* (240 sequences), *Marnaviridae* (236 sequences), *Duinviridae* (209 sequences), *Atkinsviridae* (142 sequences), *Solspiviridae* (138 sequences), *Nodaviridae* (119 sequences), and *Partitiviridae* (104 sequences) families. A total of 466 contigs remained unclassified by this methodology, likely representing viral dark matter.

Because the samples were collected before the SARS-CoV-2 pandemic, we identified only 20 sequences corresponding to Coronavirus type 1 in El Trebal effluents collected in March 2019, with a low relative abundance of 0.0002 % in the MAR2019_T_E sample.

Subsequently, we determine whether statistically significant differences exist in species diversity (considered as pp-motif diversity) among various factors, including wastewater type (effluent and influent), catchment area, geographic location (city), and sampling date (year or month). Despite the unbalanced design (yet homogenous variance), our findings indicate statistically significant differences between wastewater types (ANOVA type II, p-value $\leq$0.05). In contrast, no significant differences were observed among catchment areas, geographic locations, or sampling months or years. Next, we modeled viral diversity (Shannon and PD), richness (Chao 1), and evenness (Pielou) across influent and effluent samples (Fig. 2B). We identified significant differences in richness, diversity (Shannon), and evenness between influent and effluent samples (analysis of deviance, p-value $\leq$0.05). Our models underscore that influent wastewater harbors higher diversity (Shannon), richness (Chao1), and evenness (Pielou) among viral communities compared to effluent wastewater. However, no significant differences were observed in phylogenetic diversity (Faith PD) between influent and effluent samples (analysis of deviance, p-value $>$0.05).

3.3. Drivers of the viral community structure

We detected statistically significant differences (PERMANOVA, p-value $\leq$0.05) in viral communities (Bray-Curtis and weighted Unifrac) among wastewater types (Fig. 3A) and across different catchment areas. However, we found no significant differences between cities or aggregated temporal factors (years or months). Notably, pairwise comparisons within the three catchment areas revealed significant variations in beta diversity between La Farfana and San Pedro (pairwise PERMANOVA test, p-value $\leq$0.05). These results indicate a high turnover of viral species between wastewater influent and effluent and between the La Farfana and San Pedro catchment areas.

We also measured the variation in viral composition based on wastewater type and catchment area. These analyses revealed that the catchment area accounted for the highest variance in viral community composition, representing 28 % and 32 % for Bray-Curtis and Weighted Unifrac, respectively. Wastewater treatment contributed 22 % to the variance in viral species for both Bray-Curtis and Weighted Unifrac. The remaining variance, ranging from 50 % to 46 %, can be attributed to unaccounted variables.

To further investigate the community structure within our dataset, we conducted a constrained ordination analysis (RDA) that summarizes the maximum variation in viral composition explained by statistically significant variables. As a result, the multivariate space was constrained (Fig. 3B) solely by wastewater type, as other variables such as catchment area, geographic location, or temporal factors did not improve the RDA model. The RDA analysis indicated that the constrained space accounted for 23 % of the total variance in viral communities using both Bray-Curtis and Weighted UniFrac. The first two axes represented approximately 57 % of the constrained variance within the total constrained space for either Bray-Curtis or Weighted UniFrac. This highlights the importance of wastewater treatment as a driver of dissimilarities in viral community turnover.

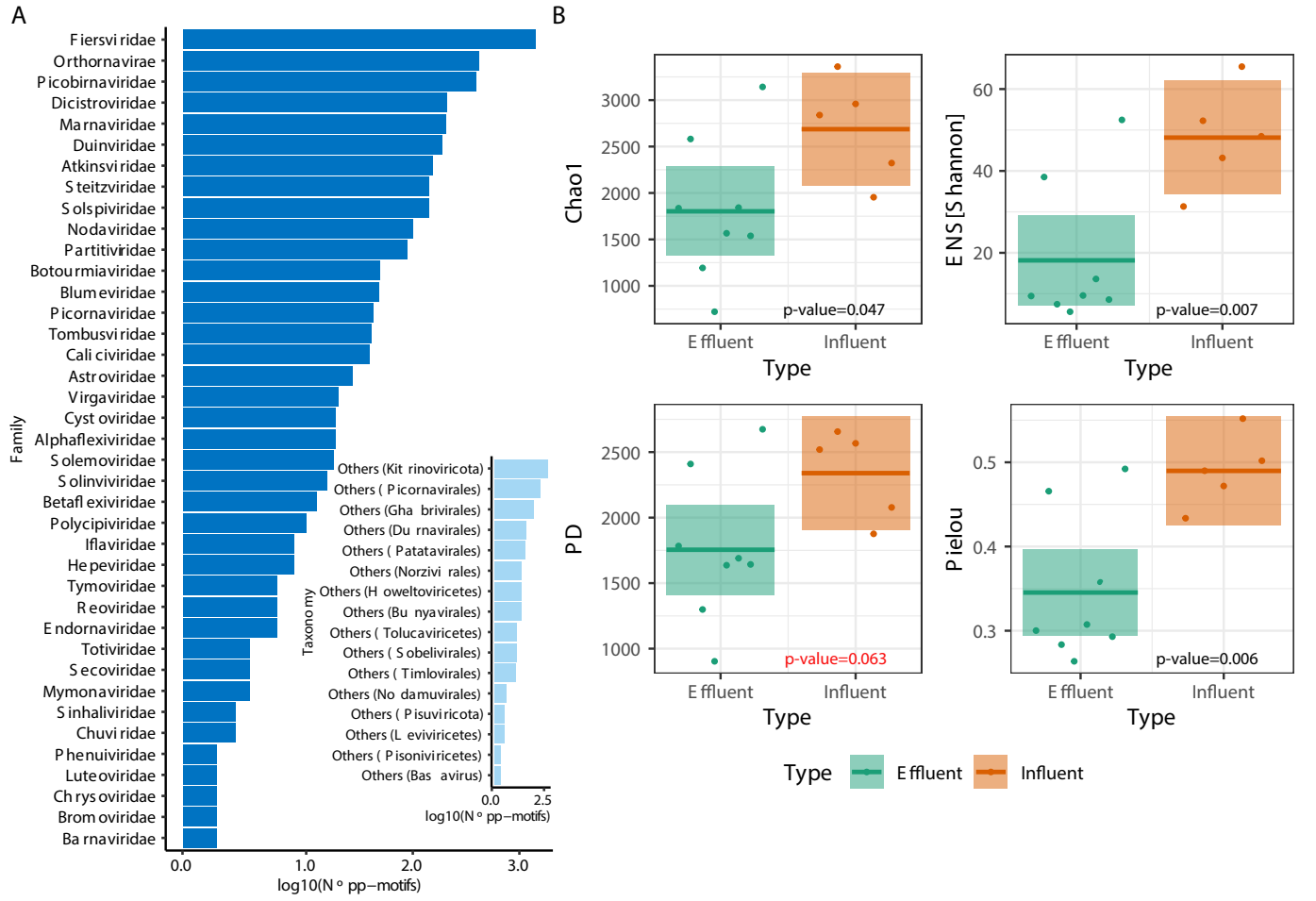


Fig. 2. Alpha diversity analysis for RNA viruses. **A)** The taxonomic composition was determined based on the clustering of contigs encoding RNA-dependent RNA polymerase. Classification was performed at the family level or assigned as "Others" based on the best-hit **B)** Generalized Linear Models (GLMs) were used to analyze alpha diversity metrics (Chao1, Effective Number of Species [Shannon], Phylogenetic Diversity, and Pielou's Evenness) in influent and effluent wastewater samples. Confidence bands are represented by colored boxes, dots indicate partial residuals, and lines depict predicted values inferred from the models. The statistical significance of the predictor fit (sample type) is indicated by the analysis of deviance *p*-value at the bottom of each plot.

3.4. Core virome and biomarkers in wastewater influents and effluents

The core virome, consisting of shared viral sequences between influent and effluent wastewater, comprised 3463 sequences representing 65 viral families, accounting for 99 % of the total abundance (Fig. 4A). Additionally, exclusive sequences were identified in both influent (614 sequences) and effluent (446 sequences) samples.

The most abundant core taxa were *Fiersviridae* (1075 sequences), followed by *Picobirnaviridae* (352 sequences), *Marnaviridae* (196 sequences), *Dicistroviridae* (188 sequences), and *Duinviridae* (160 sequences). In contrast, only three families were found exclusively in wastewater effluents, represented by five sequences from *Phenuiviridae*, *Leishbuviridae*, and the Bunya-like Bridouvirus.

Biomarker analysis (Fig. 4B) identified specific viral families that were more abundant in influent samples, such as *Astroviridae*, *Duinviridae*, *Picobirnaviridae*, *Alfaflexiviridae*, *Betaflexiviridae*, *Fiersviridae*, and *Partitiviridae*. In contrast, families like *Virgaviridae*, *Reoviridae*, *Picornavirales*, *Marnaviridae*, and *Solspiviridae* were more abundant in effluent samples.

These findings indicate that the core virome comprises the most prevalent viral taxa, but distinct biomarkers differentiate treated and untreated wastewater.

3.5. RNA operational taxonomic units in wastewater effluents and influents

We conducted a genomic analysis of viral communities in wastewater, generating 3588 viral operational taxonomic units (vOTUs) categorized into three quality tiers: 381 high-quality, 1631 medium-quality, and 1576 low-quality.

To explore the genomic landscape of RNA viruses in wastewater, we constructed a gene-sharing network of 81,461 genomes from IMG/VR, RVMT, and NCBI RefSeq. This network included 17,400 sequences grouped into 2353 viral clusters (VCs), with sizes ranging from two to 148 sequences.

Of the genomes, 5.78 % (1005 sequences) came from wastewater vOTUs, forming 321 viral clusters (Fig. 5A). These clusters are composed of 3210 database genomes representing various viral families. The largest cluster of wastewater vOTUs, with 145 sequences, was linked to *Steitzviridae*, while 72 clusters contained only two sequences each from diverse viral families. Many of these VCs belonged to viral families in the *Leviviricetes* class, known to infect prokaryotes, including *Fiersviridae*, *Atkinsviridae*, and *Steitzviridae*. Additionally, we found clusters of plant-infecting viruses from *Virgaviridae*, *Tombusviridae*, and *Alphaflexiviridae*. Smaller clusters included viruses infecting various hosts, such as *Hepeviridae* and *Astroviridae* for vertebrates, *Dicistroviridae* for insects, and *Marnaviridae* for phytoplankton.

To assess vOTUs prevalence in wastewater catchment areas defined

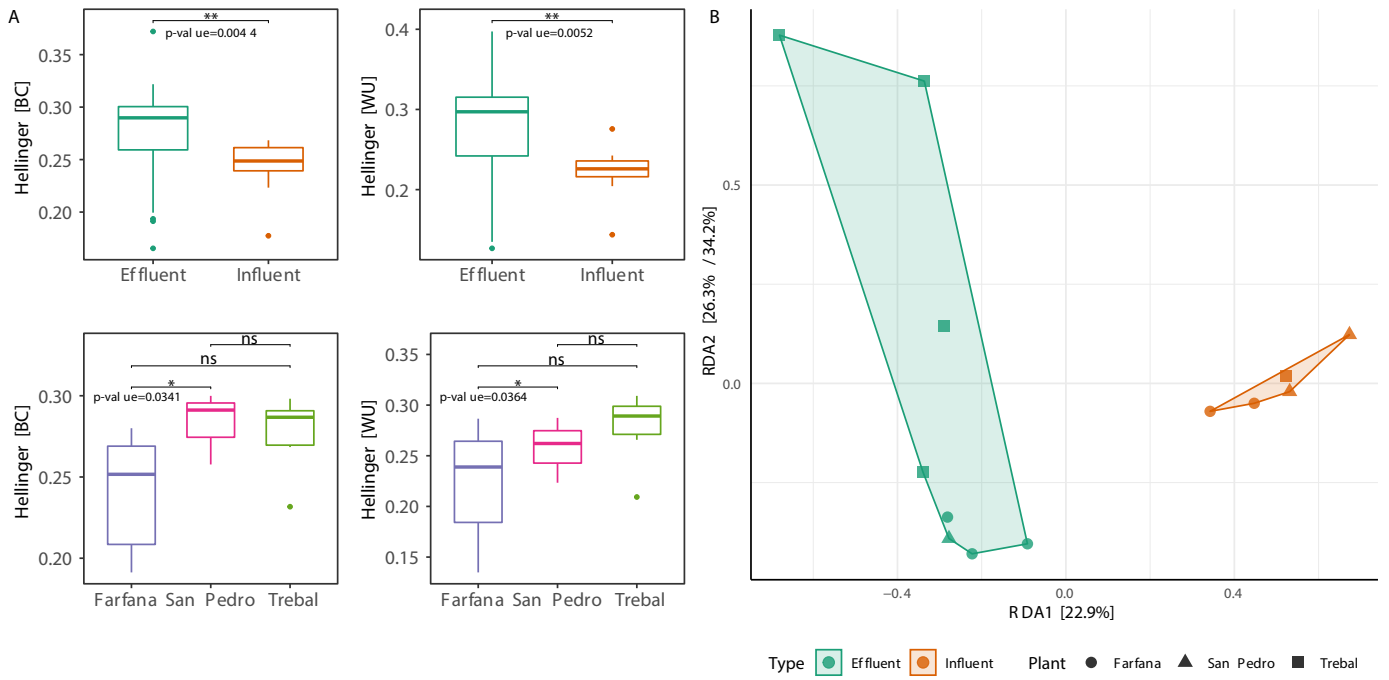


Fig. 3. Beta diversity analysis for RNA viruses. **A)** Permutational analysis of variance (PERMANOVA) was performed using Hellinger-transformed Bray-Curtis (BC) and Weighted UniFrac (WU) distances to assess the influence of sample type (Effluent vs. Influent) and treatment plant (Farfana, San Pedro, Trebal) on community composition. Boxplots represent the interquartile range of BC or WU distances, with dots indicating outliers. Statistically significant differences (*p*-values) between species turnover predictors are highlighted with brackets. **B)** Redundancy Analysis (RDA) of Hellinger-transformed Bray-Curtis distances was selected to describe microbial community structure in a supervised framework. Each axis indicates the percentage of variance explained, reflecting both unsupervised and supervised analyses.

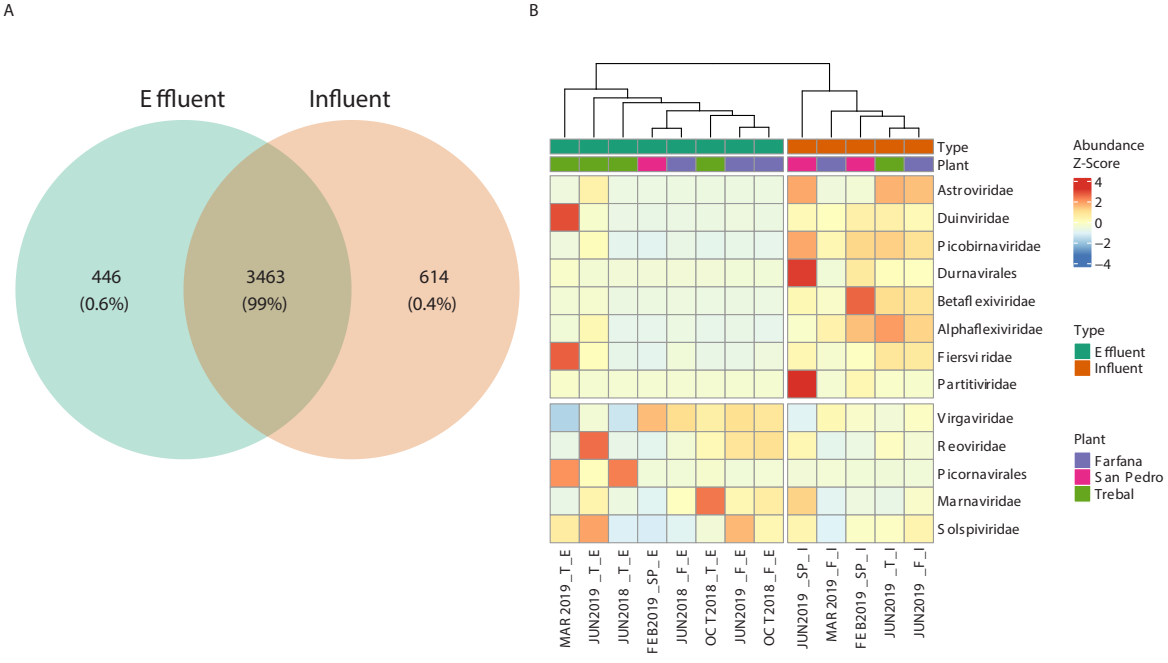
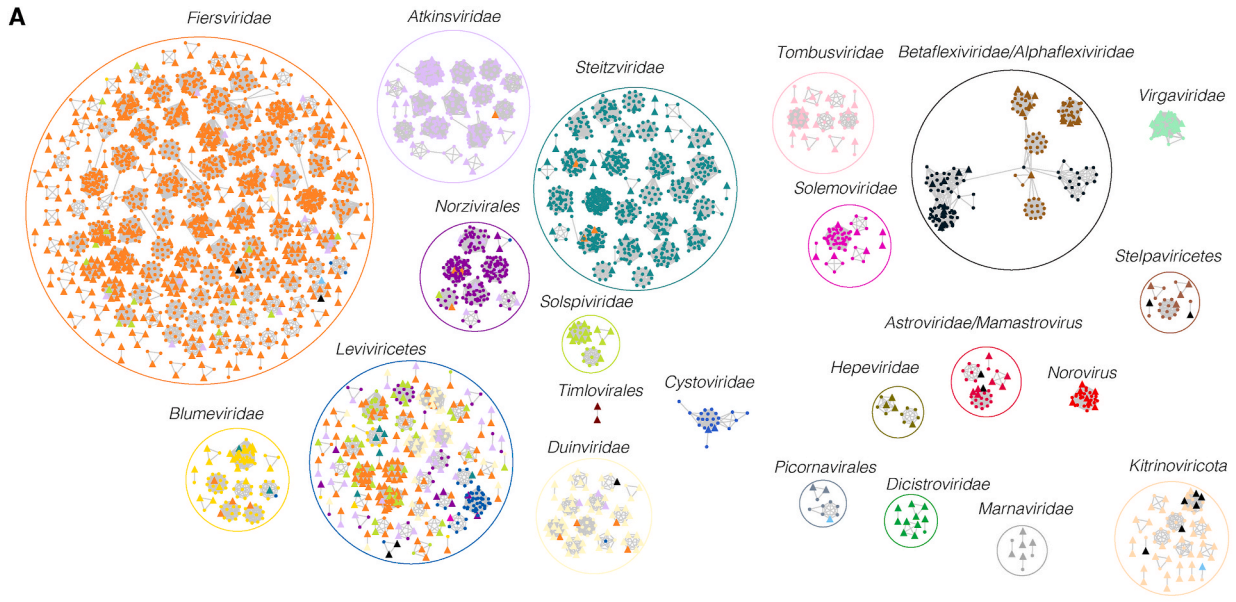


Fig. 4. Core virome analysis and taxonomic markers of wastewater effluent and influent samples from la Farfana, San Pedro and el Trebal treatment plants. **A)** Venn diagrams illustrate the number of shared and exclusive pp-contigs between effluent and influent samples from the La Farfana, San Pedro, and El Trebal treatment plants. Integers represent the absolute count of pp-contigs, while percentages (in parentheses) indicate the relative abundance of core virome contigs (intersection) and those exclusive to effluents or influents. **B)** Heatmap colors display the normalized and centered taxa abundance (Z-Score) for each sample type (effluent vs. influent). Samples on the X-axis are arranged by hierarchical clustering, with color stripes below the dendrogram denoting sample type and treatment plant. Samples are coded as follows: F_E/F_I (Farfana Effluent/Influent), T_E/T_I (Trebal Effluent/Influent), and SP_E/SP_I (San Pedro Effluent/Influent). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)



B

	Farfana					San Pedro			Trebal				
WW_contig_278; Virgaviridae	34.7	32.8	25.1	22.9	31.9	15.2	21.2	23.9	8.9	17	21.8	8.7	34.7
WW_contig_115; Virgaviridae	33.3	40.8	22.4	21.2	29.4	12.3	14.4	5.9	6.3	11.5	19.3	5.2	24.4
WW_contig_4797; Picornavirales	0	0.1	0.1	0.1	1.5	0.1	0	0.1	22.9	6.3	0	26.1	6.6
WW_contig_2527; Picornavirales	0	0	0	0	0	52.6	0	0	0	0	0	0	0
WW_contig_2554; Virgaviridae	3.8	0.4	0.3	3	1.9	5.5	8.9	0.1	0.3	0.2	0.4	1.4	1.6
WW_contig_3504; Dicistroviridae	1.5	0.7	2.5	3	1.5	0.2	6.6	2.8	0.5	1.1	3.4	0.6	1.2
WW_contig_2985; Dicistroviridae	1.2	0.6	1.9	3	1.4	0.3	5.7	2.4	0.4	0.9	2.7	0.6	1
WW_contig_1865; Picornavirales	0	0	0	0	0	0	0	0	18.6	0.2	0	1.5	0.1
WW_contig_3724; Reoviridae	1.1	3.3	0.9	0.3	4.1	0.1	0.5	1.5	0.2	3.4	1.4	0.3	2.7
WW_contig_3716; Dicistroviridae	1.6	0.4	1.9	4.2	1.2	0.1	1.8	0.9	0.4	1.2	2.8	0.8	0.9
WW_contig_1841; Kitrinoviricota	0	2.7	1	0	5	0	0	0	0	7.1	0.3	0.1	1.8
WW_contig_2064; Botourmiaviridae	0	0	0	0	0.1	0	0	0	10.5	1.3	0	5	0.3
WW_contig_145; Botourmiaviridae	0	0	0	0	0.1	0	0	0	10.7	1.3	0	4.7	0.3
WW_contig_1835; Unknown	0.1	0.1	0	0	0.9	0	0	0	0	10.6	0	0.1	1.6
WW_contig_200; Marnaviridae	0.9	1.9	0.7	0	2.4	0	0	0	0	0.7	0.2	0	5.5
WW_contig_2975; Dicistroviridae	0.2	0.2	1.5	3.1	0.1	0	1.8	1.7	0	0.1	2.4	0.5	0.2
WW_contig_99; Kitrinoviricota	0	1.5	0.6	0	3.5	0	0	0	0	3.2	0.2	0.1	1.2
WW_contig_2973; Picornavirales	0	0	0	0	0	8.8	0	0	0	0	0	0	0
WW_contig_4485; Astroviridae	0	0	2	0.1	0	0	0.1	2.2	0	0.9	2.8	0.1	0
WW_contig_1897; Kitrinoviricota	0	0	0	0	0	0	0	0	0	0	0	6.3	0
WW_contig_2966; Picornavirales	0	0	0	0	0	0.1	0	0	4.1	0.1	0	1.6	0.1
WW_contig_2544; Astroviridae	0	0	1	0.1	0	0	0.1	3.4	0	0.3	1	0.1	0
Remaining taxa (3696)	21.5	14.4	38.3	39.1	14.8	4.6	38.7	55	16	32.5	41.3	36.3	15.8
	JUN2018_F_E	JUN2019_F_E	JUN2019_F_I	MAR2019_F_I	OCT2018_F_E	FEB2019_SP_E	FEB2019_SP_I	JUN2019_SP_I	JUN2018_T_E	JUN2019_T_E	JUN2019_T_I	MAR2019_T_E	OCT2018_T_E

(caption on next page)

Fig. 5. Wastewater RNA genome analyses. **A)** Gene-sharing network of wastewater RNA vOTUs. Nodes represent vOTUs (triangles) or reference sequences (circles) from the IMG/VR and RVMT databases. Edges indicate statistically significant relationships between the protein profiles of viral genomes. Modules within the network consist of groups of similar sequences. **B)** Relative abundance profile of wastewater vOTUs by catchment area. We took only the most abundant taxa (abundance ≥0.5 % in the total samples) at their best-hit taxonomic classification Heatmap colors represent the relative abundance on a logarithmic scale. Abundances were normalized by library and genome size. Samples were codified as following, F_E/F_I (Farfana Effluent/Influent), T_E/T_I (Trebal Effluent/Influent) and SP_E/SP_I (San Pedro Effluent/Influent). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

by WWTP, we quantified the relative abundance of each across samples (Fig. 5B). Only 22 vOTUs exceeded the 0.5 % abundance threshold. Three vOTUs from the Virgaviridae family dominated, comprising 43 % of mapped reads in all areas. Blastn analysis (Table 1) revealed two had high identity with Pepper mild mottle virus and a third matched *Tomato mosaic virus*. Five abundant vOTUs from the *Picornavirales* order were restricted to specific areas. Blastx analyses indicated these vOTUs are homologous to sequences from *Picornavirales* sp. sampled from the Havel River, Germany (Table 1). Additionally, four widely dispersed *Dicistroviridae* vOTUs were abundant. Three of these were homologous to sequences from Red Panda feces and tissues, while one matched *Laverivirus* sequences from San Francisco wastewater. A prominent Reoviridae vOTU was prevalent in all areas, with Blastn analysis showing homology to G3P[8] Human rotavirus A. Two *Kritinoviricota* vOTUs and one from the *Marnaviridae* family were prevalent in metropolitan areas, homologous to Zhaovirus-like and picorna-like sequences, respectively (Table 1). Two Botourmiaviridae vOTUs were abundant in El Trebal, homologous to *Fiersviridae* from sewage (Table 1). An unclassified vOTU in metropolitan effluents showed homology to a Tombus-like virus (Table 1). Lastly, two *Astroviridae* vOTUs were prevalent in influent samples across all catchment areas, showing high homology to known Astrovirus isolates (Table 1).

These results demonstrate that a genomic approach via metaviromics allows tracking specific viruses in the environment over time and space.

4. Discussion

RNA viruses in wastewater are of significant interest due to their potential risks to public health and the silvo-agricultural industry. Understanding their presence and dynamics in wastewater ecosystems is crucial for several reasons. First, RNA viruses can serve as indicators of human and animal health, revealing infection patterns and potential outbreaks within communities. Second, their interactions with wastewater treatment processes can inform the effectiveness of current treatment methods and identify areas needing improvement. Despite the importance, there is a notable gap in comprehensive data on the RNA virosphere in wastewater, particularly regarding how treatment processes influence viral diversity and composition.

Using High Throughput Sequencing (HTS), our study comprehensively analyzes RNA viral communities in influent and effluent wastewater from three catchment areas. This approach identifies the taxonomy of RNA viruses present and monitors changes in their abundance and diversity throughout the treatment process. Our findings may enhance strategies for detecting and managing viral pathogens in wastewater, thereby supporting public health and providing scientific evidence to policymakers regarding reclaimed water use in agriculture. Additionally, identifying viral biomarkers may improve our understanding of wastewater viral ecology and the impact of treatment processes on viral communities.

4.1. Plant viruses dominate the wastewater RNA virosphere

The observed segregation of viral communities between influent and effluent samples suggests that wastewater treatment processes significantly shape viral taxonomic composition, consistent with findings from smaller-scale studies (Adriaenssens et al., 2021). This underscores the critical role of treatment procedures as drivers of wastewater viral community structure. The dominant presence of *Pisuviricota* viruses, which includes both + ssRNA and dsRNA viruses, aligns with findings from previous wastewater studies (Adriaenssens et al., 2018, 2021; Cantalupo et al., 2011; Fernandez-Cassi et al., 2018; Guajardo-Leiva et al., 2020; Martínez-Puchol et al., 2020; Muniesa et al., 2011), highlighting the widespread occurrence of specific viral taxa and the potential role of treatment in shaping these communities.

Notably, plant viruses, particularly those in the *Tobamovirus* genus, dominated the wastewater virome, consistent with global findings (Nieuwenhuijse et al., 2020). Consequently, viruses like *Pepper mild mottle virus* (PMMV) have been proposed as potential indicators of

Table 1
Summary of Blast best-hit results for the 22 most abundant (≥0.5 %) vOTUs.

vOTU	Best hit	Blast type	Query coverage (%)	Identity (%)	E-value
WW_contig_278	Pepper mild mottle virus isolate PRO54348	blastn	99	99.53	0
WW_contig_115	Pepper mild mottle virus isolate M2-PMMoV	blastn	98	98.94	0
WW_contig_4797	Picornavirales sp.	blastx	53	35.33	0
WW_contig_2527	Picornavirales sp.	blastx	51	36.06	0
WW_contig_2554	Tomato mosaic virus strain ToMV1-2	blastn	99	99.08	0
WW_contig_3504	Red panda dicistro-like virus	blastx	46	48.68	3.00E-121
WW_contig_2985	Red panda dicistro-like virus	blastx	67	42.04	3.00E-173
WW_contig_1865	Picornavirales sp.	blastx	75	43.22	2.00E-162
WW_contig_3724	Human rotavirus A strain RVA/Human-wt/CHN/Z1557/2011/G3P[8]	blastn	98	99.27	0
WW_contig_3716	Red panda dicistro-like virus	blastx	59	42.43	3.00E-170
WW_contig_1841	Beihai zhaovirus-like virus 4	blastx	24	34.65	4.00E-67
WW_contig_2064	Leviviridae sp./Fiersviridae	blastx	86	49.71	0
WW_contig_145	Leviviridae sp./Fiersviridae	blastx	89	51.72	0
WW_contig_1835	Changjiang tombus-like virus 20	blastx	12	27.5	1.00E-20
WW_contig_200	Beihai picorna-like virus 57	blastx	65	26.35	1.00E-91
WW_contig_2975	Laverivirus UC1	blastx	51	36.39	0.00E+00
WW_contig_99	Beihai zhaovirus-like virus 4	blastx	24	35.23	3.00E-68
WW_contig_2973	Picornavirales sp.	blastx	51	36.99	0.00E+00
WW_contig_4485	Human astrovirus 4 isolate CA-RGDS-1016	blastn	99	99.78	0
WW_contig_1897	Hubei tombus-like virus 36	blastx	36	37.3	1.00E-87
WW_contig_2966	Picornavirales sp. isolate HPLV-16	blastn	99	94.57	0
WW_contig_2544	Human astrovirus 1 strain MAstV-1/Hu/BRA/TO-208/2015	blastn	100	95.14	0

human fecal pollution and wastewater in aquatic environments (Farkas et al., 2020; Kitajima et al., 2014, 2018; Ochar et al., 2023). Furthermore, the frequent detection of animal viruses, such as Rotavirus, Astrovirus, and Norovirus, underscores the role of wastewater as a reservoir for viruses that cause non-bacterial enteritis worldwide (Adriaenssens et al., 2018, 2021; Calgua et al., 2013; Farkas et al., 2018; Fumian et al., 2010; Guajardo-Leiva et al., 2020; Martínez-Puchol et al., 2020; Nieuwenhuijse et al., 2020; Symonds et al., 2009). Seasonal fluctuations in their abundance suggest connections to human health and environmental conditions (Kitajima et al., 2014).

Bacteriophages were also prevalent, particularly those in the Fiersviridae family and the *Caudoviricetes* class, likely originating from human gut-associated bacteria and propagated in wastewater treatment plants (Farkas et al., 2020; Runa et al., 2021). However, their reduced abundance in the effluent suggests that bacteriophages are removed along with their microbial hosts during treatment (Runa et al., 2021 and references therein).

Lastly, environmental viruses from the *Picornavirales* order and *Dicistroviridae* family have contributed to the high diversity of wastewater viromes. These viruses, which infect plants and invertebrates, often show substantial divergence from sequences in known databases (Adriaenssens et al., 2018, 2021; Cantalupo et al., 2011; Fernandez--Cassi et al., 2018; Ng et al., 2012; Nieuwenhuijse et al., 2020). Although their origins remain largely unknown, they are thought to involve pathways such as food consumption and insect activity in drainage systems, highlighting the complex sources that contribute to the wastewater virome (Adriaenssens et al., 2018, 2021; Cantalupo et al., 2011; Ng et al., 2012).

4.2. The richness of RNA viruses in wastewater

All known RNA viruses in the Orthornavirae kingdom share the RNA-dependent RNA polymerase (RdRP) gene, essential for viral genome replication and mRNA transcription (Edgar et al., 2022). This gene contains three conserved catalytic motifs, A, B, and C, located within the palm sub-domain, forming a unique "palmprint" (Babaian and Edgar, 2022; Velthuiste, 2014). Consequently, our study focuses on the Orthornavirae kingdom, where we discovered significant taxonomic diversity, particularly among bacteriophages and eukaryotic viruses, by analyzing palmprints

A key finding was the richness of ssRNA bacteriophages within the *Leviviricetes* class (formerly the *Leviviridae* family), which includes families such as *Fiersviridae*, *Atkinsviridae*, and *Solspiviridae* in the *Norzivirales* order, as well as *Steitzviridae* and *Blumeviridae* in the *Timlovirales* order. In this aspect, our study contributes 1865 new species-representative sequences to the *Norzivirales* order, revealing previously uncharacterized viral taxa that expand the known sequence space for this order. These findings align with recent advancements in metatranscriptomics, which have significantly increased the known diversity within these groups (Callanan et al., 2020, 2021; Krishnamurthy et al., 2016; Neri et al., 2022; Shi et al., 2016; Starr et al., 2019; Zayed et al., 2022).

The *Picornavirales* order also exhibited significant diversity in our analysis, encompassing viruses that infect a wide range of hosts, including vertebrates, arthropods, and plants (Sanfaçon et al., 2009; Smertina et al., 2021; Valles et al., 2017a, 2017b; Vlok et al., 2019; Zell, 2018). These viruses were detected in all three catchment areas, highlighting their prevalence in diverse ecosystems, including invertebrate and mammalian microbiomes, as well as marine and freshwater environments (Lang et al., 2009; Shi et al., 2016; Wolf et al., 2020; Yinda et al., 2017; Zell et al., 2022). This accumulation in wastewater suggests a widespread presence, although there is no evidence of active infections.

Furthermore, considerable species richness was observed in the Nodaviridae family and two families within the *Durnavirales* order: *Picobirnaviridae* and *Partitiviridae*. *Nodaviruses*, such as those from the *Betanodavirus* genus, are known to cause severe pathologies in aquaculture, highlighting the ecological and economic significance of these findings (Hameed et al., 2018; Rosario et al., 2009a). *Picobirnaviruses* exhibit a broad geographic distribution and host diversity, being commonly detected in mammals, birds, reptiles, and invertebrates (Delmas et al., 2018; Woo et al., 2019). However, they remain non-cultivable, with the greatest known diversity identified in feces and raw wastewater (Delmas et al., 2018; Ghosh and Malik, 2021). Genomic evidence suggests that these viruses might be prokaryotic or fungal viruses rather than eukaryotic viruses (Adriaenssens et al., 2018; Boros et al., 2018; Ghosh and Malik, 2021; Guajardo-Leiva et al., 2020; Krishnamurthy and Wang, 2018).

4.3. Wastewater viral ecology

The multi-stage wastewater treatment process utilizes mechanical, biological, and chemical methods to remove inorganic compounds, bacteria, viruses, and parasites. The treated effluent is usually discharged into freshwater sources or reclaimed for other purposes (Qiu et al., 2015; Wang et al., 2018). However, the impact of these wastewater treatments on viral diversity remains under investigation. Current evidence indicates that tertiary treatment processes decrease viral diversity in effluents compared to influents.

Previous studies have primarily focused on estimating the removal efficiency of specific human pathogens, including *Norovirus*, *Adenovirus*, *Enterovirus*, *Rotavirus*, *Hepatitis A* virus, *Hepatitis E* virus, and more recently, *SARS-CoV-2* (Plaza-Garrido et al., 2022). However, the selectivity of viral removal and its dependence on physicochemical factors remain unresolved (Plaza-Garrido et al., 2022). Our findings indicate that wastewater treatment significantly reduces viral species richness and evenness in effluents, while genetic diversity remains stable, suggesting a non-selective, stochastic viral removal process.

Furthermore, our results highlight the significant role of the catchment area in shaping viral species turnover, likely due to demographic factors such as dietary habits and health status (Rothman et al., 2021). This finding aligns with previous studies showing that viral communities are more similar within the same catchment area than across different treatment plants (Adriaenssens et al., 2021; Modin et al., 2022; Osunmakinde et al., 2021; Rothman et al., 2021).

In summary, the assembly of viral communities in wastewater is influenced by treatment processes and demographic factors within the catchment area. This demographic influence is exemplified by the significant differences between the RNA viromes in semi-rural populations, such as those served by San Pedro WWTP, and urban populations, such as those served by La Farfana WWTP. These findings highlight the complexity of wastewater viral dynamics and emphasize the need to consider spatial and temporal factors in future studies.

4.4. Viral indicators of wastewater influents and effluents

Wastewater is a major contributor to global pollution in surface water bodies (Farkas et al., 2018, 2020). Its microbiological characterization has shown high concentrations and a diverse population of viruses and microbes, some of which are pathogenic (Sidhu et al., 2018). Quantifying these viral and microbial contaminants is essential for assessing the impact of wastewater discharge and the effectiveness of wastewater treatment, particularly in the context of water scarcity and reuse (Fernandes et al., 2023; Rosario et al., 2009a; Sano et al., 2016). The persistence of viruses during and after wastewater treatment is influenced by several factors, including the health status of human communities contributing to the wastewater, which affects the viral load, diversity, shedding duration, and seasonal variability, as well as the effectiveness of treatment processes and the characteristics of the viruses. (Wang and Hong, 2020).

Significant differences in the viral composition between influents and effluents likely reflect varying resistance among viral groups to

treatment processes. However, our results indicate that genome type, virion morphology, and size alone do not fully explain viral susceptibility to treatment, suggesting the involvement of other influential factors. For instance, *Astroviridae*, which are associated with mammalian hosts, were more abundant in influents, aligning with similar studies conducted in different regions (El-Senousy et al., 2007). Conversely, resilient families like Reoviridae, including Rotavirus, appeared more frequently in effluents, possibly due to their stability and aggregation properties, which protect them during disinfection processes (Betancourt and Gerba, 2016). However, the stronger seasonality of rotavirus disease in Europe, North America, and Oceania may limit its use as a treatment effectiveness marker in these regions (Patel et al., 2013). Additionally, the detection of specific viral families, such as Picornavirales and Virgaviridae, in effluents highlights their resilience and potential as indicators of viral reduction efficiency in treatment plants (Dhakar and Geetanjali, 2022; Potapov et al., 2023). Nevertheless, Picornavirales, as markers for treatment effectiveness, have limitations concerning their taxonomic range, primarily due to the breadth of the class level. Therefore, it is recommended to use the specific sequences, even if they currently lack a more detailed taxonomic classification.

In conclusion, although the sensitivity of viruses to treatment is complex, the distinct abundance patterns of specific viral markers between influents and effluents emphasize their potential for monitoring and evaluating wastewater treatment performance.

4.5. Tracking viral genomes in space and time

Monitoring virus levels in wastewater from household sources like sinks, toilets, and showers provides insights into population-wide viral prevalence. This method, known as wastewater-based epidemiology (Corpuz et al., 2020), has advanced through viral metagenomics and computational biology, allowing broad, non-targeted tracking of viral genomes over time and space (Zhang et al., 2024). This method is valuable for identifying circulating viruses and predicting outbreaks, with past successes playing a pivotal role in programs such as the global eradication of poliovirus (Asghar et al., 2014), and the monitoring of human enteric viruses (Hellmér et al., 2014) and Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV2) (Li et al., 2022; Tisza et al., 2023; Tiwari et al., 2023; Yousif et al., 2023).

Our study utilized a genomic species-rank approach to monitor viral diversity and prevalence across treatment plants and periods, identifying key viral sequences. Sequences homologous to Pepper mild mottle virus (PMMoV) isolated from pepper seeds produced and harvested in Chile's Valparaíso Region in April 2019 (Berendsen and Schravesande, 2020) were consistently identified across all catchment areas and periods, underscoring the persistence and widespread distribution of local variants of PMMoV in wastewater. Such prevalent sequences could serve as viral markers for monitoring treatment efficiency (Dhakar and Geetanjali, 2022; Kitajima et al., 2018; Rosario et al., 2009b).

Additionally, our study revealed that genomes from the *Picornavirales* order present varying distribution patterns, possibly linked to their ecological origins among terrestrial or aquatic arthropod hosts that inhabit these environments (Solomon and Hewson, 2022). Conversely, genomes from the *Dicistroviridae* family, known to infect arthropod hosts as vectors (Zhao et al., 2022), were detected consistently across all locations. These findings suggest the presence of widespread arthropod vectors within or near wastewater catchment areas, indicating potential viral transmission routes through these vectors (Zhao et al., 2022). Meanwhile, the detection of genomes from the *Marnaviridae* family in specific locations aligns with the ecological niche of these viruses, which predominantly infect microeukaryotes from local aquatic environments (Lang et al., 2021).

Our findings also highlighted the prevalence of human-associated pathogens in wastewater, specifically segments or genomes from the *Reoviridae* and *Astroviridae* families. The widespread presence of Rotavirus A, a key member of the Reoviridae family, at all locations and times raises public health concerns due to its resilience and high effluent concentrations. *Rotavirus A* is the most common cause of hospitalization due to viral gastroenteritis worldwide and is also the most frequently identified *Rotavirus* type in Chile. It has been previously identified in wastewater influents using viral metagenomics (Guajardo-Leiva et al., 2020) and qPCR (Fumian et al., 2010), suggesting a high prevalence in the country.

Similarly, seasonal trends in the abundance of human *Astroviruses* (HAstVs) highlight the importance of ongoing monitoring to anticipate outbreak risks, considering their role in infantile gastroenteritis (Hata et al., 2015; Luchs et al., 2021). HAstVs are among the leading causes of infantile gastroenteritis globally, representing 2–10 % of all viral gastroenteritis cases (Hata et al., 2015; Luchs et al., 2021).

Overall, wastewater viral surveillance provides valuable data for public health management. Identifying resilient viral taxa and monitoring their prevalence enables proactive health interventions, especially as water reuse becomes increasingly critical and the threat of viral epidemics rises. Despite progress in wastewater virome characterization through RNA metagenomics, our study has limitations. High-throughput sequencing (HTS) can bias results favoring more prevalent or easily assembled genomes, which may lead to underestimating low-abundance or highly divergent viruses. Moreover, the lack of quantitative infectivity data limits our ability to assess their impacts on humans, animals, and environmental health. While viral richness decreases after treatment, insufficient mechanistic analysis limits our comprehension of selective removal. Additionally, variability across time and space highlights the need for high-resolution longitudinal studies to enhance our understanding of wastewater viral dynamics and their implications for epidemiology.

5. Conclusions

This study offers a comprehensive analysis of RNA viral communities in wastewater, holding crucial implications for public health, environmental management, and the silvo-agricultural industry. Through High Throughput Sequencing (HTS), we characterized RNA virus diversity and dynamics across various wastewater treatment plants, revealing a dominance of plant viruses, particularly from the *Tobamovirus* genus, along with significant populations of animal viruses and bacteriophages.

Our findings highlight wastewater's potential as an essential resource for monitoring viral pathogens and assessing wastewater treatment efficacy, even in developing countries. Certain viral species, such as those from the *Virgaviridae*, *Reoviridae* and *Picornavirales* families, persist, suggesting their use as indicators of treatment effectiveness. Additionally, the stochastic nature of viral removal reflects influences from viral properties and demographic and environmental factors in catchment areas.

The diversity of RNA viruses in wastewater underscores the need for continuous monitoring, especially given global health challenges and water scarcity. Our genomic species-rank approach demonstrates the potential of wastewater-based epidemiology for early outbreak detection and the development of public health strategies.

In conclusion, our research contributes to a better understanding of the wastewater virome, offering valuable insights that may inform public health surveillance, treatment technologies, and the safe use of reclaimed water in agriculture.

CRediT authorship contribution statement

Sergio Guajardo-Leiva: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Beatriz Díez:** Writing – review & editing, Resources, Funding acquisition, Conceptualization. **Cecilia Rojas-Fuentes:** Writing – review & editing, Software, Methodology, Investigation, Formal analysis,

Data curation. **Jonás Chnaiderman:** Writing – review & editing, Investigation, Data curation. **Eduardo Castro-Nallar:** Writing – review & editing, Software, Resources. **Valentina Catril:** Writing – review & editing, Methodology. **Manuel Ampuero:** Writing – review & editing, Methodology. **Aldo Gaggero:** Writing – review & editing, Writing – original draft, Supervision, Resources, Funding acquisition, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the Writing process

During the preparation of this work the authors used ChatGPT in order to improve the cohesion and redundancy of this manuscript. After using this tool, the authors reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Funding

This work was financially supported by ANID-FONDECYT Regular 1181656, ANID-Anillo ATE220007, ANID-FONDECYT de Postdoctorado 3210547, ANID-FONDECYT de Iniciación 11241231, Center for Climate and Resilience Research (CR2) (ANID-FONDAP 1523A0002), the Millennium Institute Center for Genome Regulation (ANID-MILENIO ICN2021_044) and National Laboratory for High Performance Computing NLHPC (ECM-02).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Sergio Guajardo-Leiva reports financial support was provided by National Agency for Research and Development. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We extend our gratitude to the personnel of Aguas Andinas WWTP, San Pedro WWTP, and Analisis Ambientales (www.anam.cl) for their assistance in providing wastewater samples.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2025.121509.

Data availability

Raw reads are available at NCBI SRA, bioproject number PRJNA1196808, assembled data are available at https://doi.org/10.6084/m9.figshare.27885864.v2; https://doi.org/10.6084/m9.figshare.27885867.v1

References

Adriaenssens, E.M., Farkas, K., Harrison, C., Jones, D.L., Allison, H.E., McCarthy, A.J., 2018. Viromic analysis of wastewater input to a river catchment reveals a diverse assemblage of RNA viruses. mSystems 3, e00025. https://doi.org/10.1128/msystems.00025-18, 18.

Adriaenssens, E.M., Farkas, K., McDonald, J.E., Jones, D.L., Allison, H.E., McCarthy, A.J., 2021. Tracing the fate of wastewater viruses reveals catchment-scale virome diversity and connectivity. Water Res. 203, 117568. https://doi.org/10.1016/j.watres.2021.117568.

Al-Hazmi, H.E., Mohammadi, A., Hejna, A., Majtacz, J., Esmaeili, A., Habibzadeh, S., Saeb, M.R., Badawi, M., Lima, E.C., Mąkinia, J., 2023. Wastewater reuse in agriculture: prospects and challenges. Environ. Res. 236, 116711. https://doi.org/10.1016/j.envres.2023.116711.

Al-Hazmi, H.E., Shokrani, H., Shokrani, A., Jabbour, K., Abida, O., Khadem, S.S.M., Habibzadeh, S., Sonawane, S.H., Saeb, M.R., Bonilla-Petriciolet, A., Badawi, M., 2022. Recent advances in aqueous virus removal technologies. Chemosphere 305, 135441. https://doi.org/10.1016/j.chemosphere.2022.135441.

Andersen, K.S., Kirkegaard, R.H., Karst, S.M., Albertsen, M., 2018. ampvis2: an R package to analyse and visualise 16S rRNA amplicon data. Biorxiv, 299537. https://doi.org/10.1101/299537.

Asghar, H., Diop, O.M., Weldegebriel, G., Malik, F., Shetty, S., Bassioni, L.E., Akande, A.O., Maamoun, E.A., Zaidi, S., Adeniji, A.J., Burns, C.C., Deshpande, J., Oberste, M.S., Lowther, S.A., 2014. Environmental surveillance for polioviruses in the global polio eradication initiative. J. Infect. Dis. 210, S294–S303. https://doi.org/10.1093/infdis/jiu384.

Babaian, A., Edgar, R., 2022. Ribovirus classification by a polymerase barcode sequence. PeerJ 10, e14055. https://doi.org/10.7717/peerj.14055.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. https://doi.org/10.1089/cmb.2012.0021.

Berendsen, S.M.H., Schravesande, W.E.W., 2020. Complete genome sequence of a novel genotype of pepper mild mottle virus infecting pepper in Chile. Microbiol. Resour. Announc 9, e01183. https://doi.org/10.1128/mra.01183-20, 20.

Betancourt, W.Q., Gerba, C.P., 2016. Rethinking the significance of reovirus in water and wastewater. Food environ. Virology 8, 161–173. https://doi.org/10.1007/s12560-016-9250-8.

Boros, Á., Polgár, B., Pankovics, P., Fenyvesi, H., Engelmann, P., Phan, T.G., Delwart, E., Reuter, G., 2018. Multiple divergent picobirnaviruses with functional prokaryotic Shine-Dalgarno ribosome binding sites present in cloacal sample of a diarrheic chicken. Virology 525, 62–72. https://doi.org/10.1016/j.virol.2018.09.008.

Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60. https://doi.org/10.1038/nmeth.3176.

Calgua, B., Rodriguez-Manzano, J., Hundesa, A., Suñen, E., Calvo, M., Bofill-Mas, S., Girones, R., 2013. New methods for the concentration of viruses from urban sewage using quantitative PCR. J. Virol. Methods 187, 215–221. https://doi.org/10.1016/j.jviromet.2012.10.012.

Callanan, J., Stockdale, S.R., Adriaenssens, E.M., Kuhn, J.H., Rumnieks, J., Pallen, M.J., Shkoporov, A.N., Draper, L.A., Ross, R.P., Hill, C., 2021. Leviviricetes: expanding and restructuring the taxonomy of bacteria-infecting single-stranded RNA viruses. Microb. Genom. 7 (686). https://doi.org/10.1099/mgen.0.000686.

Callanan, J., Stockdale, S.R., Shkoporov, A., Draper, L.A., Ross, R.P., Hill, C., 2020. Expansion of known ssRNA phage genomes: from tens to over a thousand. Sci. Adv. 6, eaay5981. https://doi.org/10.1126/sciadv.aay5981.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinf. 10, 421. https://doi.org/10.1186/1471-2105-10-421, 421.

Cantalupo, P.G., Calgua, B., Zhao, G., Hundesa, A., Wier, A.D., Katz, J.P., Grabe, M., Hendrix, R.W., Girones, R., Wang, D., Pipas, J.M., 2011. Raw sewage harbors diverse viral populations. mBio 2, e00180. https://doi.org/10.1128/mbio.00180-11, 11.

Corpuz, M.V.A., Buonerba, A., Vigliotta, G., Zarra, T., Ballesteros, F., Campiglia, P., Belgiorno, V., Korshin, G., Naddeo, V., 2020. Viruses in wastewater: occurrence, abundance and detection methods. Sci. Total Environ. 745, 140910. https://doi.org/10.1016/j.scitotenv.2020.140910, 140910.

Delmas, B., Attoui, H., Ghosh, S., Malik, Y.S., Mundt, E., Vakharia, V.N., Consortium, I.R., 2018. ICTV virus taxonomy profile: Picobirnaviridae. J. Gen. Virol. 100, 133–134. https://doi.org/10.1099/jgv.0.001186.

Dhakar, V., Geetanjali, A.S., 2022. Role of pepper mild mottle virus as a tracking tool for fecal pollution in aquatic environments. Arch. Microbiol. 204 (513). https://doi.org/10.1007/s00203-022-03121-3.

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460–2461. https://doi.org/10.1093/bioinformatics/btq461.

Edgar, R.C., Taylor, J., Lin, V., Altman, T., Barbera, P., Meleshko, D., Lohr, D., Novakovsky, G., Buchfink, B., Al-Shayeb, B., Banfield, J.F., Peña, M. de la, Korobeynikov, A., Chikhi, R., Babaian, A., 2022. Petabase-scale sequence alignment catalyses viral discovery. Nature 602, 142–147. https://doi.org/10.1038/s41586-021-04332-2.

El-Senousy, W.M., Guix, S., Abid, I., Pintó, R.M., Bosch, A., 2007. Removal of astrovirus from water and sewage treatment plants, evaluated by a competitive reverse transcription-PCR. Appl. Environ. Microbiol. 73, 164–167. https://doi.org/10.1128/aem.01748-06.

Farkas, K., Cooper, D.M., McDonald, J.E., Malham, S.K., Rougemont, A., de Jones, D.L., 2018. Seasonal and spatial dynamics of enteric viruses in wastewater and in riverine and estuarine receiving waters. Sci. Total Environ. 634, 1174–1183. https://doi.org/10.1016/j.scitotenv.2018.04.038.

Farkas, K., Walker, D.I., Adriaenssens, E.M., McDonald, J.E., Hillary, L.S., Malham, S.K., Jones, D.L., 2020. Viral indicators for tracking domestic wastewater contamination in the aquatic environment. Water Res. 181, 115926. https://doi.org/10.1016/j.watres.2020.115926.

Fernandes, L.S., Galvão, A., Santos, R., Monteiro, S., 2023. Impact of water reuse on agricultural practices and human health. Environ. Res. 216, 114762. https://doi.org/10.1016/j.envres.2022.114762.

Fernandez-Cassi, X., Timoneda, N., Martínez-Puchol, S., Rusiñol, M., Rodriguez-Manzano, J., Figuerola, N., Bofill-Mas, S., Abril, J.F., Girones, R., 2018. Metagenomics for the study of viruses in urban sewage as a tool for public health surveillance. Sci. Total Environ. 618, 870–880. https://doi.org/10.1016/j.scitotenv.2017.08.249.

Fumian, T.M., Leite, J.P.G., Castello, A.A., Gaggero, A., Caillou, M.S.L. de, Miagostovich, M.P., 2010. Detection of rotavirus A in sewage samples using multiplex qPCR and an evaluation of the ultracentrifugation and adsorption-elution methods for virus concentration. J. Virol. Methods 170, 42–46. https://doi.org/10.1016/j.jviromet.2010.08.017.

Ghosh, S., Malik, Y.S., 2021. The true host/s of picobirnaviruses. Front. Vet. Sci. 7, 615293. https://doi.org/10.3389/fvets.2020.615293.

Gonzales-Gustavson, E., Cárdenas-Youngs, Y., Calvo, M., Silva, M.F.M. da, Hundesa, A., Amorós, I., Moreno, Y., Moreno-Mesonero, L., Rosell, R., Ganges, L., Araujo, R., Girones, R., 2017. Characterization of the efficiency and uncertainty of skimmed milk flocculation for the simultaneous concentration and quantification of water-borne viruses, bacteria and protozoa. J. Microbiol. Methods 134, 46–53. https://doi.org/10.1016/j.mimet.2017.01.006.

Guajardo-Leiva, S., Chnaiderman, J., Gaggero, A., Díez, B., 2020. Metagenomic insights into the sewage RNA virosphere of a large city. Viruses 12 (1050). https://doi.org/10.3390/v12091050.

Guajardo-Leiva, S., Santos, F., Salgado, O., Regeard, C., Quillet, L., Díez, B., 2021. Unveiling ecological and genetic novelty within lytic and lysogenic viral communities of hot spring phototrophic microbial mats. Microbiol. Spectr. 9, e00694. https://doi.org/10.1128/spectrum.00694-21, 21.

Hameed, A.S.S., Ninawe, A.S., Nakai, T., Chi, S.C., Johnson, K.L., Consortium, I.R, 2018. ICTV Virus Taxon. Profile: Nodaviridae. J Gen Virol 100, 3–4. https://doi.org/10.1099/jgv.0.001170.

Hata, A., Katayama, H., Kitajima, M., Furumai, H., 2015. Wastewater analysis indicates that genetically diverse Astroviruses, including strains belonging to novel clades MLB and VA, are circulating within Japanese populations. Appl. Environ. Microbiol. 81, 4932–4939. https://doi.org/10.1128/aem.00563-15.

Hellmér, M., Paxéus, N., Magnius, L., Enache, L., Arnholm, B., Johansson, A., Bergström, T., Norder, H., 2014. Detection of pathogenic viruses in sewage provided early warnings of hepatitis A virus and Norovirus outbreaks. Appl. Environ. Microbiol. 80, 6771–6781. https://doi.org/10.1128/aem.01981-14.

Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., Tappu, R., 2016. MEGAN community edition - interactive exploration and analysis of large-scale microbiome sequencing data. PLoS Comput. Biol. 12, e1004957. https://doi.org/10.1371/journal.pcbi.1004957.

Jang, H.B., Bolduc, B., Zablocki, O., Kuhn, J.H., Roux, S., Adriaenssens, E.M., Brister, J. R., Kropinski, A.M., Krupovic, M., Lavigne, R., Turner, D., Sullivan, M.B., 2019. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nat. Biotechnol. 37, 632–639. https://doi.org/10.1038/s41587-019-0100-8.

Kitajima, M., Iker, B.C., Pepper, I.L., Gerba, C.P., 2014. Relative abundance and treatment reduction of viruses during wastewater treatment processes — identification of potential viral indicators. Sci. Total Environ. 488, 290–296. https://doi.org/10.1016/j.scitotenv.2014.04.087.

Kitajima, M., Sassi, H.P., Torrey, J.R., 2018. Pepper mild mottle virus as a water quality indicator. Npj Clean Water 1 (19). https://doi.org/10.1038/s41545-018-0019-5.

Kopylova, E., Noé, L., Touzet, H., 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28, 3211–3217. https://doi.org/10.1093/bioinformatics/bts611.

Krishnamurthy, S.R., Janowski, A.B., Zhao, G., Barouch, D., Wang, D., 2016. Hyperexpansion of RNA bacteriophage diversity. PLoS Biol. 14, e1002409. https://doi.org/10.1371/journal.pbio.1002409.

Krishnamurthy, S.R., Wang, D., 2018. Extensive conservation of prokaryotic ribosomal binding sites in known and novel picobirnaviruses. Virology 516, 108–114. https://doi.org/10.1016/j.virol.2018.01.006.

Lang, A.S., Rise, M.L., Culley, A.I., Steward, G.F., 2009. RNA viruses in the sea. FEMS Microbiol. Rev. 33, 295–323. https://doi.org/10.1111/j.1574-6976.2008.00132.x.

Lang, A.S., Vlok, M., Culley, A.I., Suttle, C.A., Takao, Y., Tomaru, Y., Consortium, I.R., 2021. ICTV virus taxonomy profile: Marnaviridae 2021. J. Gen. Virol. 102, 001633. https://doi.org/10.1099/jgv.0.001633.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923.

Li, L., Mazurowski, L., Dewan, A., Carine, M., Haak, L., Guarin, T.C., Dastjerdi, N.G., Gerrity, D., Mentzer, C., Pagilla, K.R., 2022. Longitudinal monitoring of SARS-CoV-2 in wastewater using viral genetic markers and the estimation of unconfirmed COVID-19 cases. Sci. Total Environ. 817, 152958. https://doi.org/10.1016/j.scitotenv.2022.152958.

Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22, 1658–1659. https://doi.org/10.1093/bioinformatics/btl158.

Liu, C., Cui, Y., Li, X., Yao, M., 2021. microeco: an R package for data mining in microbial community ecology. FEMS Microbiol. Ecol. 97 (2), fiaa255. https://doi.org/10.1093/femsec/fiaa255.

Lofrano, G., Brown, J., 2010. Wastewater management through the ages: a history of mankind. Sci. Total Environ. 408, 5254–5264. https://doi.org/10.1016/j.scitotenv.2010.07.062.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15 (550). https://doi.org/10.1186/s13059-014-0550-8.

Luchs, A., Tardy, K., Tahmasebi, R., Morillo, S.G., Milagres, F.A. de P., Morais, V. dos S., Brustulin, R., Teles, M. da A.R., Azevedo, L.S. de, Souza, E.V. de, Medeiros, R.S., Souza, E.V. de, Araújo, E.L.L., Witkin, S.S., Deng, X., Delwart, E., Sabino, E.C., Leal, E., Costa, A.C. da, 2021. Human astrovirus types 1, 4 and 5 circulating among children with acute gastroenteritis in a rural Brazilian state, 2010-2016. Arch. Virol. 166, 3165–3172. https://doi.org/10.1007/s00705-021-05206-8.

Martínez-Puchol, S., Rusiñol, M., Fernández-Cassi, X., Timoneda, N., Itarte, M., Andrés, C., Antón, A., Abril, J.F., Girones, R., Bofill-Mas, S., 2020. Characterisation of the sewage virome: comparison of NGS tools and occurrence of significant pathogens. Sci. Total Environ. 713, 136604. https://doi.org/10.1016/j.scitotenv.2020.136604.

McMurdie, P.J., Holmes, S., 2013. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8, e61217. https://doi.org/10.1371/journal.pone.0061217.

Modin, O., Fuad, N., Abadikhah, M., I'Ons, D., Ossiansson, E., Gustavsson, D.J.I., Edefell, E., Suarez, C., Persson, F., Wilén, B.-M., 2022. A relationship between phages and organic carbon in wastewater treatment plant effluents. Water Res. X 16, 100146. https://doi.org/10.1016/j.wroa.2022.100146.

Mousazadeh, M., Kabdaşlı, I., Khademi, S., Sandoval, M.A., Moussavi, S.P., Malekdar, F., Gilhotra, V., Hashemi, M., Dehghani, M.H., 2022. A critical review on the existing wastewater treatment methods in the COVID-19 era: what is the potential of advanced oxidation processes in combatting viral especially SARS-CoV-2? J. Water Process Eng. 49, 103077. https://doi.org/10.1016/j.jwpe.2022.103077.

Muniesa, M., Imamovic, L., Jofre, J., 2011. Bacteriophages and genetic mobilization in sewage and faecally polluted environments. Microb. Biotechnol. 4, 725–734. https://doi.org/10.1111/j.1751-7915.2011.00264.x.

Nayfach, S., Camargo, A.P., Schulz, F., Eloe-Fadrosh, E., Roux, S., Kyrpides, N.C., 2021. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat. Biotechnol. 39, 578–585. https://doi.org/10.1038/s41587-020-00774-7.

Nepusz, T., Yu, H., Paccanaro, A., 2012. Detecting overlapping protein complexes in protein-protein interaction networks. Nat. Methods 9, 471–472. https://doi.org/10.1038/nmeth.1938.

Neri, U., Wolf, Y.I., Roux, S., Camargo, A.P., Lee, B., Kazlauskas, D., Chen, I.M., Ivanova, N., Allen, L.Z., Paez-Espino, D., Bryant, D.A., Bhaya, D., Consortium, R.V. D., Narrowe, A.B., Probst, A.J., Sczyrba, A., Kohler, A., Séguin, A., Shade, A., Campbell, B.J., Lindahl, B.D., Reese, B.K., Roque, B.M., DeRito, C., Averill, C., Cullen, D., Beck, D.A.C., Walsh, D.A., Ward, D.M., Wu, D., Eloe-Fadrosh, E., Brodie, E.L., Young, E.B., Lilleskov, E.A., Castillo, F.J., Martin, F.M., LeCleir, G.R., Attwood, G.T., Cadillo-Quiroz, H., Simon, H.M., Hewson, I., Grigoriev, I.V., Tiedje, J. M., Jansson, J.K., Lee, J., VanderGheynst, J.S., Dangl, J., Bowman, J.S., Blanchard, J. L., Bowen, J.L., Xu, J., Banfield, J.F., Deming, J.W., Kostka, J.E., Gladden, J.M., Rapp, J.Z., Sharpe, J., McMahon, K.D., Treseder, K.K., Bidle, K.D., Wrighton, K.C., Thamatrakoln, K., Nusslein, K., Meredith, L.K., Ramirez, L., Buee, M., Huntemann, M., Kalyuzhnaya, M.G., Waldrop, M.P., Sullivan, M.B., Schrenk, M.O., Hess, M., Vega, M.A., O'Malley, M.A., Medina, M., Gilbert, N.E., Delherbe, N., Mason, O.U., Dijkstra, P., Chuckran, P.F., Baldrian, P., Constant, P., Stepanauskas, R., Daly, R.A., Lamendella, R., Gruninger, R.J., McKay, R.M., Hylander, S., Lebeis, S.L., Esser, S.P., Acinas, S.G., Wilhelm, S.S., Singer, S.W., Tringe, S.S., Woyke, T., Reddy, T.B.K., Bell, T.H., Mock, T., McAllister, T., Thiel, V., Denef, V.J., Liu, W.-T., Martens-Habbena, W., Liu, X.-J.A., Cooper, Z.S., Wang, Z., Krupovic, M., Dolja, V.V., Kyrpides, N.C., Koonin, E.V., Gophna, U., 2022. Expansion of the global RNA virome reveals diverse clades of bacteriophages. Cell 185, 4023–4037. https://doi.org/10.1016/j.cell.2022.08.023 e18.

Ng, T.F.F., Marine, R., Wang, C., Simmonds, P., Kapusinszky, B., Bodhidatta, L., Oderinde, B.S., Wommack, K.E., Delwart, E., 2012. High variety of known and new RNA and DNA viruses of diverse origins in untreated sewage. J. Virol. 86, 12161–12175. https://doi.org/10.1128/jvi.00869-12.

Nieuwenhuijse, D.F., Munnink, B.B.O., Phan, M.V.T., Hendriksen, R.S., Bego, A., Rees, C., Neilson, E.H., Coventry, K., Collignon, P., Allerberger, F., Rahube, T.O., Oliveira, G., Ivanov, I., Sopheak, T., Vuthy, Y., Yost, C.K., Tabo, D., Cuadros-Orellana, S., Ke, C., Zheng, H., Baisheng, L., Jiao, X., Donado-Godoy, P., Coulibaly, K.J., Hrenovic, J., Jergović, M., Karpíšková, R., Elsborg, B., Legesse, M., Eguale, T., Heikinheimo, A., Villacis, J.E., Sanneh, B., Malania, L., Nitsche, A., Brinkmann, A., Saba, C.K.S., Kocsis, B., Solymosi, N., Thorsteinsdottir, T.R., Hatha, A.M., Alebouyeh, M., Morris, D., O'Connor, L., Cormican, M., Moran-Gilad, J., Battisti, A., Alba, P., Shakenova, Z., Kiiyukia, C., Ng'eno, E., Raka, L., Bērziņš, A., Avsejenko, J., Bartkevics, V., Penny, C., Rajandas, H., Parimannan, S., Haber, M.V., Pal, P., Schmitt, H., Passel, M., van Schans, M.G.M., van de, Zuidema, T., Jeunen, G.-J., Gemmell, N., Fashae, K., Wester, A.L., Holmstad, R., Hasan, R., Shakoor, S., Rojas, M.L.Z., Wasyl, D., Bosevska, G., Kochubovski, M., Radu, C., Gassama, A., Radosavljevic, V., Tay, M.Y.F., Zuniga-Montanez, R., Wuertz, S., Gavačová, D., Trkov, M., Keddy, K., Esterhuyse, K., Cerdà-Cuéllar, M., Pathirage, S., Larsson, D.G.J., Norrgren, L., Örn, S., Heijden, T.V. der, Kumburu, H. H., RodaHusman, A.M. de, Njanpop-Lafourcade, B.-M., Bidjada, P., Nikiema-Pessinaba, S.C., Levent, B., Meschke, J.S., Beck, N.K., Dang, C.V., Tran, D.M.N., Phuc, N.D., Kwenda, G., Munk, P., Venkatakrishnan, S., Aarestrup, F.M., Cotten, M., Koopmans, M.P.G., 2020. Setting a baseline for global urban virome surveillance in sewage. Sci Rep-uk 10, 13748. https://doi.org/10.1038/s41598-020-69869-0.

O'Brien, E., Xagoraraki, I., 2020. Removal of viruses in membrane bioreactors. J. Environ. Eng. 146, 03120007. https://doi.org/10.1061/(asce)ee.1943-7870.0001743.

Ochar, K., Ko, H.-C., Woo, H.-J., Hahn, B.-S., Hur, O., 2023. Pepper mild mottle virus: an infectious pathogen in pepper production and a potential indicator of domestic water quality. Viruses 15 (282). https://doi.org/10.3390/v15020282.

Oksanen, J., 2020. Vegan: Community Ecology Package. R Package version 2.5-7.

Ondov, B.D., Treangen, T.J., Melsted, P., Mallonee, A.B., Bergman, N.H., Koren, S., Phillippy, A.M., 2016. Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol. 17 (132). https://doi.org/10.1186/s13059-016-0997-x.

Osunmakinde, C., Selvarajan, R., Mamba, B., Msagati, T., 2021. Viral communities distribution and diversity in a wastewater treatment plants using high-throughput

sequencing analysis. Pol. J. Environ. Stud. 30, 3189–3201. https://doi.org/10.15244/pjoes/127888.
Patel, Manish M., Pitzer, Virginia, E. Alonso, Wladimir, J., Vera, David PharmD., Lopman, Ben, Tate, Jacqueline, Viboud, Cecile, Parashar, Umesh D., 2013. Global seasonality of rotavirus disease. Pediatr. Infect. Dis. J. 32 (4), e134–e147. https://doi.org/10.1097/INF.0b013e31827d3b68. April.
Paul, J.R., Trask, J.D., Gard, S., 1940. II. Poliomyelitic virus in urban sewage. J. Exp. Med. 71, 765–777. https://doi.org/10.1084/jem.71.6.765.
Plaza-Garrido, A., Limaico, M., Villamar-Ayala, C.A., 2022. Influence of wastewater treatment technologies on virus removal under a bibliometric-statistical analysis. J. Water Process Eng. 47, 102642. https://doi.org/10.1016/j.jwpe.2022.102642.
Polanco, J.A., Safarik, J., Dadakis, J.S., Johnson, C., Plumlee, M.H., 2023. Enteric virus removal by municipal wastewater treatment to achieve requirements for potable reuse. PLOS Water 2, e0000052. https://doi.org/10.1371/journal.pwat.0000052.
Potapov, S., Gorshkova, A., Krasnopeev, A., Podlesnaya, G., Tikhonova, I., Suslova, M., Kwon, D., Patrushev, M., Drucker, V., Belykh, O., 2023. RNA-seq virus fraction in lake baikal and treated wastewaters. Int. J. Mol. Sci. 24, 12049. https://doi.org/10.3390/ijms241512049.
Qiu, Y., Lee, B.E., Neumann, N., Ashbolt, N., Craik, S., Maal-Bared, R., Pang, X.L., 2015. Assessment of human virus removal during municipal wastewater treatment in Edmonton, Canada. J. Appl. Microbiol. 119, 1729–1739. https://doi.org/10.1111/jam.12971.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, D590–D596. https://doi.org/10.1093/nar/gks1219.
Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. https://doi.org/10.1093/bioinformatics/btp616.
Rosario, K., Nilsson, C., Lim, Y.W., Ruan, Y., Breitbart, M., 2009a. Metagenomic analysis of viruses in reclaimed water. Environ. Microbiol. 11, 2806–2820. https://doi.org/10.1111/j.1462-2920.2009.01964.x.
Rosario, K., Symonds, E.M., Sinigalliano, C., Stewart, J., Breitbart, M., 2009b. Pepper mild mottle virus as an indicator of fecal pollution. Appl. Environ. Microbiol. 75, 7261–7267. https://doi.org/10.1128/aem.00410-09.
Rothman, J.A., Loveless, T.B., Kapcia, J., Adams, E.D., Steele, J.A., Zimmer-Faust, A.G., Langlois, K., Wanless, D., Griffith, M., Mao, L., Chokry, J., Griffith, J.F., Whiteson, K. L., 2021. RNA viromics of southern California wastewater and detection of SARS-CoV-2 single-nucleotide variants. Appl. Environ. Microbiol. 87, e01448. https://doi.org/10.1128/aem.01448-21, 21.
Roux, S., Páez-Espino, D., Chen, I.-M.A., Palaniappan, K., Ratner, A., Chu, K., Reddy, T.B. K., Nayfach, S., Schulz, F., Call, L., Neches, R.Y., Woyke, T., Ivanova, N.N., Eloe-Fadrosh, E.A., Kyrpides, N.C., 2020. IMG/VR v3: an integrated ecological and evolutionary framework for interrogating genomes of uncultivated viruses. Nucleic Acids Res 49, gkaa946-. https://doi.org/10.1093/nar/gkaa946.
Runa, V., Wenk, J., Bengtsson, S., Jones, B.V., Lanham, A.B., 2021. Bacteriophages in biological wastewater treatment systems: occurrence, characterization, and function. Front. Microbiol. 12, 730071. https://doi.org/10.3389/fmicb.2021.730071.
Sanfaçon, H., Wellink, J., Gall, O.L., Karasev, A., Vlugt, R., van der, Wetzel, T., 2009. Secoviridae: a proposed family of plant viruses within the order Picornavirales that combines the families Sequiviridae and Comoviridae, the unassigned genera Cheravirus and Sadwavirus, and the proposed genus Torradovirus. Arch. Virol. 154, 899–907. https://doi.org/10.1007/s00705-009-0367-z.
Sanjuán, R., Domingo-Calap, P., 2016. Mechanisms of viral mutation. Cell. Mol. Life Sci. 73, 4433–4448. https://doi.org/10.1007/s00018-016-2299-6.
Sano, D., Amarasiri, M., Hata, A., Watanabe, T., Katayama, H., 2016. Risk management of viral infectious diseases in wastewater reclamation and reuse: review. Environ. Int. 91, 220–229. https://doi.org/10.1016/j.envint.2016.03.001.
Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504. https://doi.org/10.1101/gr.1239303.
Shi, M., Lin, X.-D., Tian, J.-H., Chen, L.-J., Chen, X., Li, C.-X., Qin, X.-C., Li, J., Cao, J.-P., Eden, J.-S., Buchmann, J., Wang, W., Xu, J., Holmes, E.C., Zhang, Y.-Z., 2016. Redefining the invertebrate RNA virosphere. Nature 540, 539–543. https://doi.org/10.1038/nature20167.
Sidhu, J.P.S., Sena, K., Hodgers, L., Palmer, A., Toze, S., 2018. Comparative enteric viruses and coliphage removal during wastewater treatment processes in a sub-tropical environment. Sci. Total Environ. 616, 669–677. https://doi.org/10.1016/j.scitotenv.2017.10.265.
Smertina, E., Hall, R.N., Urakova, N., Strive, T., Frese, M., 2021. Calicivirus non-structural proteins: potential functions in replication and host Cell manipulation. Front. Microbiol. 12, 712710. https://doi.org/10.3389/fmicb.2021.712710.
Smith, M.F., Maqsood, R., Sullins, R.A., Driver, E.M., Halden, R.U., Lim, E.S., 2024. Seasonality of respiratory, enteric, and urinary viruses revealed by wastewater genomic surveillance. mSphere 9, e00105–e00124. https://doi.org/10.1128/msphere.00105-24.
Solomon, C., Hewson, I., 2022. Putative invertebrate, plant, and wastewater derived ssRNA viruses in plankton of the anthropogenically impacted anacostia river, District of Columbia, USA. Microb. Environ. 37, ME21070. https://doi.org/10.1264/jsme2.me21070.
Starr, E.P., Nuccio, E.E., Pett-Ridge, J., Banfield, J.F., Firestone, M.K., 2019. Metatranscriptomic reconstruction reveals RNA viruses with the potential to shape carbon cycling in soil. P Natl Acad Sci Usa 116, 25900–25908. https://doi.org/10.1073/pnas.1908291116.
Symonds, E.M., Griffin, D.W., Breitbart, M., 2009. Eukaryotic viruses in wastewater samples from the United States. Appl. Environ. Microbiol. 75, 1402–1409. https://doi.org/10.1128/aem.01899-08.
Thompson, L.R., Sanders, J.G., McDonald, D., Amir, A., Ladau, J., Locey, K.J., Prill, R.J., Tripathi, A., Gibbons, S.M., Ackermann, G., Navas-Molina, J.A., Janssen, S., Kopylova, E., Vázquez-Baeza, Y., González, A., Morton, J.T., Mirarab, S., Xu, Z.Z., Jiang, L., Haroon, M.F., Kanbar, J., Zhu, Q., Song, S.J., Kosciolek, T., Bokulich, N.A., Lefler, J., Brislawn, C.J., Humphrey, G., Owens, S.M., Hampton-Marcell, J., Berg-Lyons, D., McKenzie, V., Fierer, N., Fuhrman, J.A., Clauset, A., Stevens, R.L., Shade, A., Pollard, K.S., Goodwin, K.D., Jansson, J.K., Gilbert, J.A., Knight, R., Rivera, J.L.A., Al-Moosawi, L., Alverdy, J., Amato, K.R., Andras, J., Angenent, L.T., Antonopoulos, D.A., Apprill, A., Armitage, D., Ballantine, K., Bárta, J., Baum, J.K., Berry, A., Bhatnagar, A., Bhatnagar, M., Biddle, J.F., Bittner, L., Boldgiv, B., Bottos, E., Boyer, D.M., Braun, J., Brazelton, W., Brearley, F.Q., Campbell, A.H., Caporaso, J.G., Cardona, C., Carroll, J., Cary, S.C., Casper, B.B., Charles, T.C., Chu, H., Claar, D.C., Clark, R.G., Clayton, J.B., Clemente, J.C., Cochran, A., Coleman, M.L., Collins, G., Colwell, R.R., Contreras, M., Crary, B.B., Creer, S., Cristol, D.A., Crump, B.C., Cui, D., Daly, S.E., Davalos, L., Dawson, R.D., Defazio, J., Delsuc, F., Dionisi, H.M., Dominguez-Bello, M.G., Dowell, R., Dubinsky, E.A., Dunn, P.O., Ercolini, D., Espinoza, R.E., Ezenwa, V., Fenner, N., Findlay, H.S., Fleming, I.D., Fogliano, V., Forsman, A., Freeman, C., Friedman, E.S., Galindo, G., Garcia, L., Garcia-Amado, M.A., Garshelis, D., Gasser, R.B., Gerdts, G., Gibson, M.K., Gifford, I., Gill, R.T., Giray, T., Gittel, A., Golyshin, P., Gong, D., Grossart, H.-P., Guyton, K., Haig, S.-J., Hale, V., Hall, R.S., Hallam, S.J., Handley, K.M., Hasan, N.A., Haydon, S.R., Hickman, J.E., Hidalgo, G., Hofmockel, K.S., Hooker, J., Hulth, S., Hultman, J., Hyde, E., Ibáñez-Álamo, J.D., Jastrow, J.D., Jex, A.R., Johnson, L.S., Johnston, E.R., Joseph, S., Jurburg, S.D., Jurelevicius, D., Karlsson, A., Karlsson, R., Kauppinen, S., Kellogg, C.T.E., Kennedy, S.J., Kerkhof, L.J., King, G.M., Kling, G.W., Koehler, A.V., Krezalek, M., Kueneman, J., Lamendella, R., Landon, E.M., Lane-deGraaf, K., LaRoche, J., Larsen, P., Laverock, B., Lax, S., Lentino, M., Levin, I.I., Liancourt, P., Liang, W., Linz, A.M., Lipson, D.A., Liu, Y., Lladser, M.E., Lozada, M., Spirito, C.M., MacCormack, W.P., MacRae-Crerar, A., Magris, M., Martín-Platero, A. M., Martín-Vivaldi, M., Martínez, L.M., Martínez-Bueno, M., Marzinelli, E.M., Mason, O.U., Mayer, G.D., McDevitt-Irwin, J.M., McDonald, J.E., McGuire, K.L., McMahon, K.D., McMinds, R., Medina, M., Mendelson, J.R., Metcalf, J.L., Meyer, F., Michelangeli, F., Miller, K., Mills, D.A., Minich, J., Mocali, S., Moitinho-Silva, L., Moore, A., Morgan-Kiss, R.M., Munroe, P., Myrold, D., Neufeld, J.D., Ni, Y., Nicol, G. W., Nielsen, S., Nissimov, J.I., Niu, K., Nolan, M.J., Noyce, K., O'Brien, S.L., Okamoto, N., Orlando, L., Castellano, Y.O., Osuolale, O., Oswald, W., Parnell, J., Peralta-Sánchez, J.M., Petraitis, P., Pfister, C., Pilon-Smits, E., Piombino, P., Pointing, S.B., Pollock, F.J., Potter, C., Prithiviraj, B., Quince, C., Rani, A., Ranjan, R., Rao, S., Rees, A.P., Richardson, M., Riebesell, U., Robinson, C., Rockne, K.J., Rodriguezl, S.M., Rohwer, F., Roundstone, W., Safran, R.J., Sangwan, N., Sanz, V., Schrenk, M., Schrenzel, M.D., Scott, N.M., Seger, R.L., Seguin-Orlando, A., Seldin, L., Seyler, L.M., Shakhsheer, B., Sheets, G.M., Shen, C., Shi, Y., Shin, H., Shogan, B.D., Shutler, D., Siegel, J., Simmons, S., Sjöling, S., Smith, D.P., Soler, J.J., Sperling, M., Steinberg, P.D., Stephens, B., Stevens, M.A., Taghavi, S., Tai, V., Tait, K., Tan, C.L., Taş, N., Taylor, D.L., Thomas, T., Timling, I., Turner, B.L., Urich, T., Ursell, L.K., Lelie, D. van der, Treuren, W.V., Zwieten, L. van, Vargas-Robles, D., Thurber, R.V., Vitaglione, P., Walker, D.A., Walters, W.A., Wang, S., Wang, T., Weaver, T., Webster, N.S., Wehrle, B., Weisenhorn, P., Weiss, S., Werner, J.J., West, K., Whitehead, A., Whitehead, S.R., Whittingham, L.A., Willerslev, E., Williams, A.E., Wood, S.A., Woodhams, D.C., Yang, Y., Zaneveld, J., Zarraonaindia, I., Zhang, Q., Zhao, H., 2017. A communal catalogue reveals Earth's multiscale microbial diversity. Nature 551, 457–463. https://doi.org/10.1038/nature24621.
Tisza, M., Cregeen, S.J., Avadhanula, V., Zhang, P., Ayvaz, T., Feliz, K., Hoffman, K.L., Clark, J.R., Terwilliger, A., Ross, M.C., Cormier, J., Moreno, H., Wang, L., Payne, K., Henke, D., Troisi, C., Wu, F., Rios, J., Deegan, J., Hansen, B., Balliew, J., Gitter, A., Zhang, K., Li, R., Bauer, C.X., Mena, K.D., Piedra, P.A., Petrosino, J.F., Boerwinkle, E., Maresso, A.W., 2023. Wastewater sequencing reveals community and variant dynamics of the collective human virome. Nat. Commun. 14 (6878). https://doi.org/10.1038/s41467-023-42064-1.
Tiwari, A., Adhikari, S., Zhang, S., Solomon, T.B., Lipponen, A., Islam, MdA., Thakali, O., Sangkham, S., Shaheen, M.N.F., Jiang, G., Haramoto, E., Mazumder, P., Malla, B., Kumar, M., Pitkänen, T., Sherchan, S.P., 2023. Tracing COVID-19 trails in wastewater: a systematic review of SARS-CoV-2 surveillance with viral variants. Water 15 (1018). https://doi.org/10.3390/w15061018.
Valles, S.M., Chen, Y., Firth, A.E., Guérin, D.M.A., Hashimoto, Y., Herrero, S., Miranda, J. R. de, Ryabov, E., Consortium, I.R., 2017a. ICTV virus taxonomy profile: Dicistroviridae. J. Gen. Virol. 98, 355–356. https://doi.org/10.1099/jgv.0.000756.
Valles, S.M., Chen, Y., Firth, A.E., Guérin, D.M.A., Hashimoto, Y., Herrero, S., Miranda, J. R. de, Ryabov, E., Consortium, I.R., 2017b. ICTV virus taxonomy profile: iflaviridae. J. Gen. Virol. 98, 527–528. https://doi.org/10.1099/jgv.0.000757.
Velthuis, A.J.W. te, 2014. Common and unique features of viral RNA-dependent polymerases. Cell. Mol. Life Sci. 71, 4403–4420. https://doi.org/10.1007/s00018-014-1695-z.
Vlok, M., Lang, A.S., Suttle, C.A., 2019. Application of a sequence-based taxonomic classification method to uncultured and unclassified marine single-stranded RNA viruses in the order Picornavirales. Virus Evol 5, vez056. https://doi.org/10.1093/ve/vez056.
Wang, C., Hong, P.-Y., 2020. Genome-resolved metagenomics and antibiotic resistance genes analysis in reclaimed water distribution systems. Water 12 (3477). https://doi.org/10.3390/w12123477.
Wang, H., Sikora, P., Rutgersson, C., Lindh, M., Brodin, T., Björlenius, B., Larsson, D.G.J., Norder, H., 2018. Differential removal of human pathogenic viruses from sewage by

conventional and ozone treatments. Int. J. Hyg Environ. Health 221, 479–488. https://doi.org/10.1016/j.ijheh.2018.01.012.

Wolf, Y.I., Silas, S., Wang, Y., Wu, S., Bocek, M., Kazlauskas, D., Krupovic, M., Fire, A., Dolja, V.V., Koonin, E.V., 2020. Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome. Nat Microbiol 5, 1262–1270. https://doi.org/10.1038/s41564-020-0755-4.

Woo, P.C.Y., Teng, J.L.L., Bai, R., Tang, Y., Wong, A.Y.P., Li, K.S.M., Lam, C.S.F., Fan, R. Y.Y., Lau, S.K.P., Yuen, K.-Y., 2019. Novel picobirnaviruses in respiratory and alimentary tracts of cattle and monkeys with large intra- and inter-host diversity. Viruses 11 (574). https://doi.org/10.3390/v11060574.

Yinda, C.K., Zell, R., Deboutte, W., Zeller, M., Conceição-Neto, N., Heylen, E., Maes, P., Knowles, N.J., Ghogomu, S.M., Ranst, M.V., Matthijnssens, J., 2017. Highly diverse population of Picornaviridae and other members of the Picornavirales, in Cameroonian fruit bats. BMC Genom. 18 (249). https://doi.org/10.1186/s12864-017-3632-7.

Yousif, M., Rachida, S., Taukobong, S., Ndlovu, N., Iwu-Jaja, C., Howard, W., Moonsamy, S., Mhlambi, N., Gwala, S., Levy, J.I., Andersen, K.G., Scheepers, C., Gottberg, A. von, Wolter, N., Bhiman, J.N., Amoako, D.G., Ismail, A., Suchard, M., McCarthy, K., 2023. SARS-CoV-2 genomic surveillance in wastewater as a model for monitoring evolution of endemic viruses. Nat. Commun. 14 (6325). https://doi.org/10.1038/s41467-023-41369-5.

Zayed, A.A., Wainaina, J.M., Dominguez-Huerta, G., Pelletier, E., Guo, J., Mohssen, M., Tian, F., Pratama, A.A., Bolduc, B., Zablocki, O., Cronin, D., Solden, L., Delage, E., Alberti, A., Aury, J.-M., Carradec, Q., Silva, C., da, Labadie, K., Poulain, J., Ruscheweyh, H.-J., Salazar, G., Shatoff, E., Coordinators, T.O., Bundschuh, R., Fredrick, K., Kubatko, L.S., Chaffron, S., Culley, A.I., Sunagawa, S., Kuhn, J.H., Wincker, P., Sullivan, M.B., Acinas, S.G., Babin, M., Bork, P., Boss, E., Bowler, C., Cochrane, G., Vargas, C. de, Gorsky, G., Guidi, L., Grimsley, N., Hingamp, P., Iudicone, D., Jaillon, O., Kandels, S., Karp-Boss, L., Karsenti, E., Not, F., Ogata, H., Poulton, N., Pesant, S., Sardet, C., Speich, S., Stemmann, L., Sullivan, M.B., Sungawa, S., Wincker, P., 2022. Cryptic and abundant marine viruses at the evolutionary origins of Earth's RNA virome. Science 376, 156–162. https://doi.org/10.1126/science.abm5847.

Zell, R., 2018. Picornaviridae-the ever-growing virus family. Arch. Virol. 163, 299–317. https://doi.org/10.1007/s00705-017-3614-8.

Zell, R., Groth, M., Selinka, L., Selinka, H.-C., 2022. Picorna-like viruses of the Havel River, Germany. Front. Microbiol. 13, 865287. https://doi.org/10.3389/fmicb.2022.865287.

Zhang, Z., He, F., Yi, L., Deng, Z., Wang, R., Shen, L., Fu, S., 2024. Wastewater surveillance together with metaviromic data revealed the unusual resurgence of infectious diseases after the first wave of the COVID-19 outbreak. J. Hazard Mater. 473, 134635. https://doi.org/10.1016/j.jhazmat.2024.134635.

Zhao, M., Yue, C., Yang, Z., Li, Y., Zhang, D., Zhang, J., Yang, S., Shen, Q., Su, X., Qi, D., Ma, R., Xiao, Y., Hou, R., Yan, X., Li, L., Zhou, Y., Liu, J., Wang, X., Wu, W., Zhang, W., Shan, T., Liu, S., 2022. Viral metagenomics unveiled extensive communications of viruses within giant pandas and their associated organisms in the same ecosystem. Sci. Total Environ. 820, 153317. https://doi.org/10.1016/j.scitotenv.2022.153317.