

Effect of wastewater collection and concentration methods on assessment of viral diversity

Yabing Li^a, Brijen Miyani^a, Kevin L. Childs^b, Shin-Han Shiu^{b,c,d}, Irene Xagorarakis^{a,*}

^a Department of Civil and Environmental Engineering, Michigan State University, 1449 Engineering Research Ct, East Lansing, MI, United States

^b Department of Plant Biology, Michigan State University, East Lansing, MI, United States

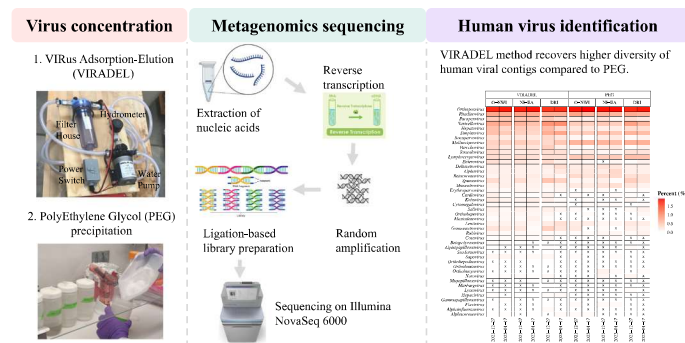
^c Department of Energy (DOE) Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI, United States

^d Department of Computational Mathematics, Science, and Engineering, Michigan State University, East Lansing, MI, United States

HIGHLIGHTS

- Municipal wastewater samples were collected and concentrated with the Virus Adsorption-Elution (VIRADEL) and the PolyEthylene Glycol (PEG) precipitation methods.
- Nucleic acid from concentrated samples was analyzed using metagenomics methods to assess human viral diversity.
- More viral reads and human viral-related contigs were obtained in samples concentrated with the VIRADEL method as compared to samples concentrated with the PEG precipitation method.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Warish Ahmed

Keywords:

Wastewater surveillance

Metagenomics

Human virus

Public health, COVID-19 outbreak

ABSTRACT

Monitoring of potentially pathogenic human viruses in wastewater is of crucial importance to understand disease trends in communities, predict potential outbreaks, and boost preparedness and response by public health departments. High throughput metagenomic sequencing opens an opportunity to expand the capabilities of wastewater surveillance. However, there are major bottlenecks in the metagenomic enabled wastewater surveillance, including the complexities in selecting appropriate sampling and concentration/virus enrichment methods as well as in bioinformatic analysis of complex samples with low human virus concentrations. To evaluate the abilities of two commonly used sampling and concentration methods in virus identification, virus communities concentrated with Virus Adsorption-Elution (VIRADEL) and PolyEthylene Glycol (PEG) precipitation were compared for three interceptor sites. Results indicated that more viral reads were obtained by the VIRADEL concentration method, with 2.84 ± 0.57 % viral reads in the sample. For samples concentrated with PEG, the average proportion of viral reads in the sample was 0.63 ± 0.19 %. In all wastewater samples, bacteriophage affiliated with the families *Siphoviridae*, *Myoviridae* and *Podoviridae* were found to be the abundant populations. Comparison against a custom Swiss-Prot human virus database indicated that the relatively abundant human viruses (average proportions in human virus community greater than 1.00 %) in samples concentrated with the VIRADEL method were *Orthopoxvirus*, *Rhadinovirus*, *Parapoxvirus*, *Varicellovirus*,

* Corresponding author.

E-mail address: xagorara@msu.edu (I. Xagorarakis).

<https://doi.org/10.1016/j.scitotenv.2023.168128>

Received 30 June 2023; Received in revised form 23 October 2023; Accepted 24 October 2023

Available online 31 October 2023

0048-9697/© 2023 Elsevier B.V. All rights reserved.

Hepatovirus, Simplexvirus, Molluscipoxvirus, Parechovirus, Lymphocryptovirus, and Spumavirus. In samples concentrated with the PEG method, fewer human viruses were found to be relatively abundant. These were Orthopoxvirus, Rhadinovirus, Varicellovirus, Simplexvirus, Molluscipoxvirus, Lymphocryptovirus, and Betacoronavirus. Contigs of Betacoronavirus, which contains severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), were identified in VIRADEL and PEG samples. Our study demonstrates the feasibility of using metagenomics in wastewater surveillance as a first screening tool and the need for selecting the appropriate virus concentration methods and optimizing bioinformatic approaches in analyzing metagenomic data of wastewater samples.

1. Introduction

Hundreds of viruses are found to infect humans, and new human virus species, or strains continue to be identified (Bibby and Peccia, 2013; Woolhouse et al., 2012). It has been confirmed that multiple enteric, respiratory, bloodborne and vector-borne viruses can be detected in wastewater samples (Levy et al., 2023; McCall et al., 2020; McCall et al., 2021; Miyani et al., 2021a; O'Brien and Xagorarakis, 2019; Silverman and Boehm, 2021). Most recently, the application of next-generation sequencing in wastewater-based epidemiology offers the opportunity of detection of known and new viral pathogens. However, identifying the diversity of human viruses in wastewater is still challenging, as compared to identifying viral diversity in clinical samples. Because the concentration of viruses in wastewater is much lower compared to clinical samples and the suspension contains a multitude of other microorganisms and chemicals, accurate and sensitive detection depends on effectively sampling and concentrating viruses from wastewater.

All sampling, concentration/virus enrichment and subsequent sequencing methods for viral metagenomics analysis involve physical and chemical stresses that affect the detection of viruses. Viruses differ drastically in morphology, genetic variability, chemical and mechanical resistivity (Pantaleo and Chiumenti, 2018). This diversity contributes to varied sensitivity to these treatments that greatly influence viral detection. Therefore, every process from sampling to data analysis of viral metagenomics may introduce bias (Parras-Moltó et al., 2018; Thomas et al., 2012). Additionally, the small relative abundance of human viruses compared to plant viruses, insect and other animal viruses, bacteriophage, bacteria, and other organisms in wastewater samples poses a challenge for sequencing and bioinformatic analyses. Specifically, the amount of viral genetic material extracted from environmental samples like wastewater is frequently insufficient to construct libraries compared to other microorganisms (Kim and Bae, 2011; Polson et al., 2011).

These challenges could be overcome by a combination of an improved virus-specific sampling and concentration/enrichment methods (McCall et al., 2020). Multiple different virus concentration methods have been applied in recovering viruses from water and wastewater samples (Bofill-Mas and Rusiñol, 2020; Xagorarakis et al., 2014). These methods are based on electronegative or electropositive filtration, flocculation/precipitation with organic/inorganic flocculants, and ultracentrifugation and centrifugal ultrafiltration. Among them, two commonly used methods for virus recovery from wastewater are Virus Adsorption-Elution (VIRADEL) and PolyEthylene Glycol (PEG) precipitation-based methods. The VIRADEL method is established to recover human enteric viruses and is used prevalently for its high recovery rate (Fout et al., 2001; Fout, 1996; Li et al., 2022b; Lu et al., 2020; Miyani et al., 2020; Miyani et al., 2021b; O'Brien et al., 2017; Zhao et al., 2022). Concentrating viruses using the electropositive NanoCeram column filters, the primary step of VIRADEL method, is suggested as the cost-effective approach to facilitate the binding of the negatively charged viruses (Haramoto et al., 2018). Following the primary concentration, flocculation, and centrifugation in the VIRADEL method, bacteria with relatively large sizes are excluded by further filtration with the 0.45 µm and 0.22 µm syringe filters. These processes will help decrease the bacterial contamination and improve virus

detectability. The PEG method on the other hand has been used for several decades due to its simplicity, rapidity, resource economy and generation of stable viral particles in low-temperature and high-salt environments (Alexander et al., 2020). The PEG method involved the collection of the composite samples followed by precipitation. Composite samples collected over multiple hours are believed to be more representative than grab samples considering the influences of different parameters (e.g., organic matter in the wastewater, wastewater temperature, wastewater pH, rainfalls) on virus occurrence throughout the day (Schmidt, 2020).

Virus concentration methods have been tested to assess their efficiency in recovering SARS-CoV-2 (Ahmed et al., 2020; McMinin et al., 2023). Evaluation of several virus concentration methods showed that the adsorption-extraction methods with minimal or without pretreatment are the most efficient methods to recover SARS-CoV-2 (Ahmed et al., 2020). Assessment of two volumetrically different concentration methods indicated the larger volume method (2 L) outperformed the small volume method (100 mL) for SARS-CoV-2 detection (McMinin et al., 2023). As for the effects of virus concentration methods on studying viral diversity with metagenomics, compositions of viral community recovered with an ultracentrifugation method and a Skimmed Milk organic Flocculation (SMF) method were compared with an untargeted metagenomics method. The results indicated that for most viral families, such as Adenoviridae, Caliciviridae, Parvoviridae, Circoviridae, Astroviridae and Picornaviridae, members of which include the typical human viruses, similar results were obtained for both the ultrafiltration and SMF methods. A low number of contigs of Anelloviridae, Alloherpesviridae, Geminiviridae, Hepeviridae, Totiviridae, Geminiviridae, and Polyomaviridae families were found by the ultrafiltration method. Viral families such as Luteoviridae, Nanoviridae, and Baculoviridae were detected with a few contigs by the SMF method (Fernandez-Cassi et al., 2018). However, little is known about efficiency of virus concentration methods like VIRADEL and PEG precipitation on the diversity of detected human viruses.

In the metropolitan area of Detroit, Michigan, wastewater surveillance has been applied to detect multiple human virus occurrences (McCall et al., 2020; McCall et al., 2021; Miyani et al., 2021a; Xagorarakis et al., 2014). In this study, the VIRADEL and PEG methods were used to collect and concentrate 12 untreated wastewater samples from three interceptors of the Great Lakes Water Authority (GLWA) wastewater treatment facility during December 2021 and January 2022 to assess the effect of wastewater collection and virus concentration methods on identification of human viruses with metagenomics.

2. Material and methods

2.1. Study area and sample collection

The Water Resource Recovery Facility (WRRF) is the wastewater system of the GLWA in Detroit, Michigan. The WRRF is the largest single-site wastewater treatment facility in North America and serves the three most populous counties in Michigan: Wayne, Oakland, and Macomb. It receives wastewater via three main interceptors including the Detroit River Interceptor (DRI), the North Interceptor-East Arm (NI-EA), and Oakwood-Northwest-Wayne County Interceptor (O-NWI). The three interceptors serve approximately 492,000 (DRI), 1,482,000 (NI-

EA) and 840,600 (O-NWI) people, based on 2020 population estimates provided by the Southeast Michigan Council of Governments. The WRRF is a semi-combined sewer-shed system, which collects and treats stormwater along with residential, industrial, and commercial waste, depending on service areas. To evaluate viral community diversity obtained by different sample collection and virus concentration methods, 12 untreated wastewater samples were collected from the three interceptor sites on December 27, 2021, and January 17, 2022.

2.2. Collection and concentration of viruses from wastewater samples

2.2.1. Sample collection and VIRADEL concentration

Viruses were isolated from wastewater using electropositive NanoCeram column filters (Argonide, Sanford, FL, USA) based on an EPA protocol (Miyani et al., 2021b; USEPA, 2001; Xagorarakis et al., 2014). Specifically, depending on the quantity of suspended solids in the wastewater, approximately 20 to 50 L of raw wastewater were passed through NanoCeram electropositive cartridge filters at a rate not more than 11 L/min. Flow meter readings were recorded at the beginning and termination of each sampling event to measure the total volume raw wastewater that passed through the filter. The filters containing viruses were placed in separated and sealed plastic bags on ice and were transported to the laboratory for downstream analysis within 48 h. The electropositive NanoCeram column filters were eluted with 1 L beef extract for 2 min. After which, pH of the beef solution was adjusted to 3.5 ± 0.1 and then flocculated for 30 min before centrifugation at 2500 g (4°C) for 15 min. Supernatant was discarded, and pellets were resuspended in 30 mL of sodium phosphate (0.15 M). The pH of the resuspended solution was adjusted to a range of 9.0–9.5. A second round of centrifugation was carried out at 7000 g (4°C) for 10 min. The supernatant was collected and adjusted to a pH of approximately 7.25. Filtration was performed on the samples with 0.45 μm and 0.22 μm syringe filters to eliminate large bacterial contamination. The final filtered solution was aliquoted into multiple 2 mL Corning tubes and stored in a -80°C freezer until nucleic acid extraction was performed.

2.2.2. Composite sample collection and PEG concentration

Autoclavable polythene plastic bottles were used to collect 1 L of 24-h composite samples from interceptor sites on December 27, 2021, and January 17, 2022. The bottles were labelled, placed in individual sealed plastic bags, stored on ice, and were transported to the laboratory for processing within 48 h. At the laboratory, each sample was mixed gently, and a 100 mL sample was transferred into a 500 mL autoclavable centrifuge bottle. To force viruses to aggregate into clusters, 1.17 g NaCl (0.2 M) and 8 g PEG (8 %) were added into the centrifuge bottle, and the

samples were mixed at 110 rpm at 4°C for two hours on a magnetic stirring plate. The sample was then centrifuged at 4°C at a speed of 4700 g for 45 min. The supernatant was carefully removed without disturbing the pellet and discarded. About 1–5 mL of solution was retained with the pellet. The pellet was resuspended by gentle mixing, and the final solution was aliquoted into 2–3 vials and stored in -80°C freezer for the following nucleic acid extraction.

Steps of the two sample collection and virus concentration methods (VIRADEL and PEG) were summarized as Fig. 1.

2.3. Metagenomic sequencing

2.3.1. Extraction of the nucleic acid and random amplification

Viral nucleic acid was extracted using QIAGEN QIAamp Viral RNA QIAGEN kits (QIAGEN, Hilden, Germany), following the manufacturer's protocol with the volume of final eluting reagent (buffer AVE) modified from 60 μL to 140 μL (Li et al., 2022b; Miyani et al., 2021a; Zhao et al., 2022). To ensure enough sample for the final metagenomic library, extracts of duplicate samples were pooled together in the subsequent steps. A random-primer protocol developed to identify viral pathogens was applied to perform the amplification (Wang et al., 2002; Wang et al., 2003). RNA was reverse-transcribed with primer-A (5'- GTTCCAGT-CACGATCNNNNNNNNNN), and second-strand DNA synthesis was carried out with Sequenase (Version 2.0 DNA Polymerase, Thermo Fisher Scientific). Primer-B (5'- GTTCCAGT-CACGATC) was used for the subsequent PCR amplification for 40 cycles (McCall et al., 2020; Wang et al., 2002; Wang et al., 2003).

2.3.2. Next generation sequencing

Viral cDNA for the wastewater samples ($n = 12$) were sent to the MSU Research Technology Support Facility's Genomics Core for library preparation and sequencing. Sequencing was performed on the Illumina NovaSeq 6000 platform. Details related to the library preparation for whole metagenome shotgun sequencing were shown in the supplementary information S1. Quality of the raw reads were assessed using FastQC (Andrews, 2010). Quality scores for more than 90 % of both R1 and R2 reads in every sample were higher than 30. A total of 1.16 billion reads were obtained for the 12 samples and the averaging number of reads for each sample was 97.0 million. The 12 samples had an average yield of 32.3 Gb. Adapters and low-quality reads were trimmed with Trimmomatic (v. 0.39, parameters: phred33 TruSeq3-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:35) (Bolger et al., 2014).

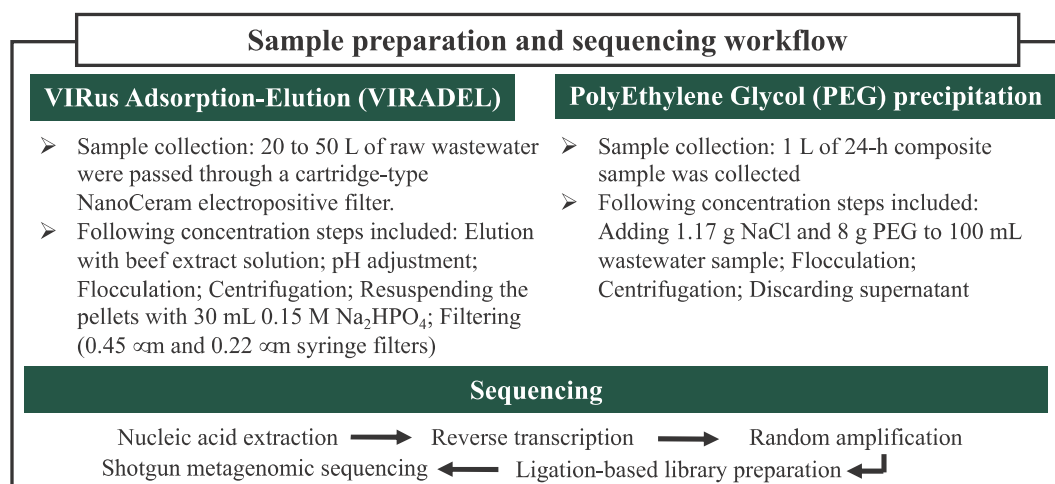


Fig. 1. Steps of the two sample collection and virus concentration methods (VIRADEL and PEG) and processes of the metagenomic sequencing.

2.4. Bioinformatic analyses and taxonomic alignments

2.4.1. Assessment of different metagenomic assemblers

Two metagenome assembly programs IDBA-UD (v. 1.1.3) and MEGAHIT (v. 1.2.9) were assessed to select the appropriate assembler in the metagenomics analysis of wastewater using samples collected from the NI-EA interceptor on December 27, 2021. Genomes of six viral species were selected to create simulated, artificial sequence reads as spike-ins due to their frequent occurrences in the wastewater systems: *Adenovirus 40/41*, *Human herpesvirus 3*, *Coxsackievirus*, *Hepatitis A virus*, *SARS-Coronavirus* and *Norovirus GII* (Dunkin et al., 2018; Li et al., 2022a; McCall et al., 2020; O'Brien and Xagorarakis, 2019; Wollants et al., 2022; Xagorarakis et al., 2014) (Table S1). Mason (v. 2.0.0) software tool was used to generate simulated Illumina paired end reads for each reference genome (Holtgrewe, 2010). Four fold-coverages ($5\times$, $10\times$, $20\times$ and $50\times$) of added simulated viral reads were used when mixing with the real metagenomic reads. MetaQUAST was used to assess the performance of each assembler, and the results were summarized in Table S3.

Running time of the compared processes were summarized in Table S2. MEGAHIT had a much shorter computational time and require fewer resources compared to IDBA-UD. Assembly integrity and accuracy measures (Table S3) show that MEGAHIT and IDBA-UD have comparable performance high throughput sequencing data of wastewater samples. Considering the time and computational resources, downstream bioinformatic analyses were based on the contigs assembled by MEGAHIT.

2.4.2. Human virus identification

Trimmed reads were aligned against the National Center for Biotechnology Information (NCBI) BLAST non-redundant database directly using Kaiju (v. 1.9.0) to determine the proportions of viral reads in the samples. A sensitive run mode "Greedy" was used, and the cutoff for *E*-value was set as 10^{-3} (Menzel et al., 2016). Viral reads were extracted and compared against a custom Swiss-Prot human virus database using BLASTX using a maximum *E*-value of 10^{-5} to get the count of human viral reads in each sample. Retrieval of the custom Swiss-Prot protein database of human viruses was shown in detail in the supplementary information S2. To be brief, Swiss-Prot protein database was screened for proteins of viruses with human as the host.

To achieve better accuracy in taxonomic mapping, long contiguous sequences (contigs) generated by the assembly process instead of the unassembled trimmed reads were used to identify viral and human viral composition (Ayling et al., 2019). To obtain the virus composition in the

wastewater samples, the assembled contigs were aligned against the NCBI RefSeq virus database (retrieved on December 1, 2022) with DIAMOND blastx using a maximum *E*-value of 10^{-3} (Bağcı et al., 2021; Bibby et al., 2011; McCall et al., 2020). In order to improve the discovery of human viruses, reduce ambiguity in the taxonomic alignment, and decrease the chance of false negative hits (McCall et al., 2020), the assembled contigs were also aligned against a custom Swiss-Prot human virus database using BLASTX using a maximum *E*-value of 10^{-5} .

2.4.3. Taxonomic annotation

Taxonomic assignment of the hit reads were performed with MEGAN the Community Edition (v. 6.22.2) (Huson et al., 2016). Counts per 100 million reads of each human virus in VIRADEL and PEG samples were calculated and compared. Meanwhile, compositions of viral contigs at the family level and human viral contigs at the genus level were further analyzed with MEGAN the Community Edition (v. 6.22.2) (Huson et al., 2016).

A bioinformatic workflow for identify human virus occurrence in wastewater samples with a metagenomics-enabled surveillance approach was shown in Fig. 2. Explanations and selections of the related parameters are provided as the supplementary table S4.

2.5. Statistical analyses and data visualization

Data were organized and relative abundances of each human virus were calculated using Microsoft Excel. Non-metric multidimensional scaling (NMDS) analysis was performed with RStudio. Illustrations including the violin plots, bubble plots, heatmaps and plots for NMDS were visualized in RStudio. Packages including "dplyr", "ggplot2", "vegan", "BiodiversityR", "tidyverse" were used in the statistical analysis and data visualization.

3. Results

3.1. More viral reads and contigs were recovered with the VIRADEL method

Sequencing yields, proportion of viral reads after trimming, proportions of viral contigs after assembly and the proportions of human viral contigs in the wastewater samples are summarized in Table 1. Yields of the sequencing data for VIRADEL and PEG samples were 30.6 ± 6.17 Gbp and 34.0 ± 2.63 Gbp, respectively (Fig. 3A). After trimming, the count of clean reads for VIRADEL samples ranged from 17.6 million

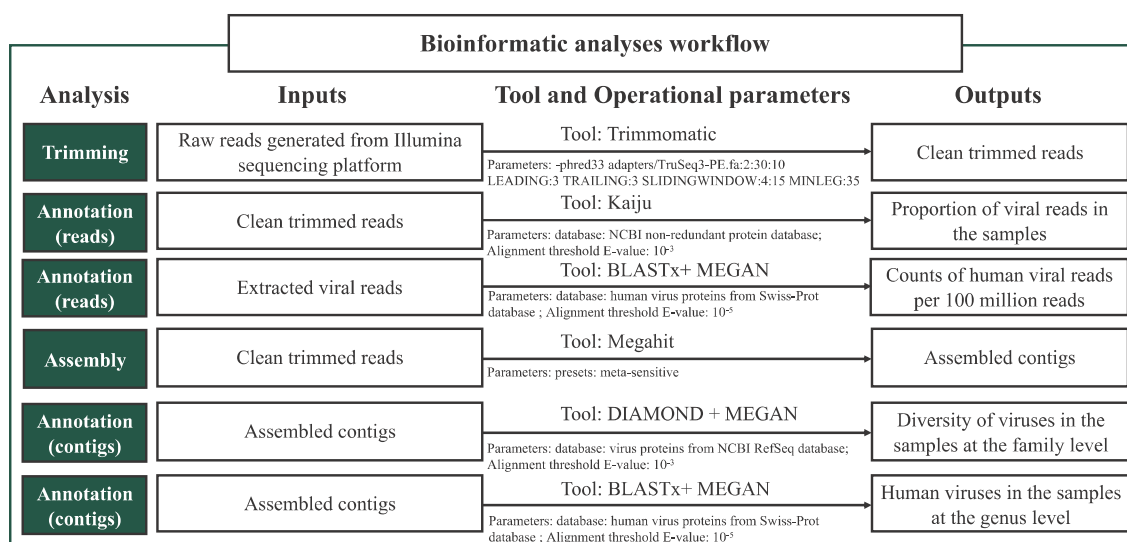


Fig. 2. The schematic workflow of the bioinformatic analyses.

Table 1

Sequencing yield, number of trimmed reads, number of trimmed viral reads, proportion of viral reads in the sample, number of contigs assembled, number of viral contigs, proportion of viral contigs, number of human virus related contigs, and proportion of human virus related contigs identified in the untreated wastewater samples collected and concentrated with VIRADEL and PEG precipitation methods.

Location	Date	Collection and concentration method	Yield (Gbp)	Number of trimmed reads	Number of trimmed viral reads	Proportion of viral reads (%) ^a	Number of contigs assembled	Number of viral contigs	Proportion of viral contigs (%) ^b	Number of human virus related contigs	Proportion of human viral contigs (%) ^c
ONWI	12/27/21	VIRADEL	39.73	123,186,605	4,155,228	3.37	1,966,935	407,881	20.7	7173	0.36
		PEG	29.72	93,390,723	700,211	0.75	479,616	43,584	9.09	1222	0.25
	1/17/22	VIRADEL	25.68	79,693,080	2,157,281	2.71	2,070,233	387,848	18.7	6078	0.29
		PEG	33.37	104,800,788	358,302	0.34	542,035	47,378	8.74	1323	0.24
NI-EA	12/27/21	VIRADEL	28.52	86,181,575	2,637,453	3.06	1,396,782	297,141	21.3	4858	0.35
		PEG	33.81	108,938,846	705,849	0.65	500,374	38,841	7.76	1155	0.23
	1/17/22	VIRADEL	31.39	99,283,275	3,139,034	3.16	2,418,889	489,628	20.2	6968	0.29
		PEG	37.78	119,600,043	612,979	0.51	561,138	46,381	8.27	1241	0.22
DRI	12/27/21	VIRADEL	35.36	110,483,023	3,268,418	2.96	1,338,903	291,942	21.8	6174	0.46
		PEG	35.26	111,063,640	1,008,526	0.91	600,368	54,228	9.03	1745	0.29
	1/17/22	VIRADEL	23.19	17,577,429	309,190	1.76	255,096	43,881	17.2	1019	0.40
		PEG	33.94	109,346,344	665,796	0.61	618,425	51,421	8.31	1576	0.25

^a - Proportion of viral reads in the sample. Number of trimmed viral reads divided by the number of trimmed reads.

^b - Proportion of viral contigs in the sample. Number of viral contigs divided by the number of contigs assembled.

^c - Proportion of human viral contigs in the sample. Number of human viral contigs divided by the number of contigs assembled.

to 123.2 million. For PEG samples, it ranged from 93.4 million to 119.6 million (Fig. 3B). Results of comparing the clean reads against the NCBI non-redundant database showed the proportion of viral reads in the VIRADEL samples ranged from 1.76 % to 3.37 %, with an averaging value as 2.84 ± 0.57 % (Fig. 3C). For PEG samples, the proportion of viral reads in the samples ranged from 0.34 % to 0.91 %, with an averaging value of 0.63 ± 0.19 % (Fig. 3C). Larger proportions of viral reads were obtained in the VIRADEL samples compared with those in the PEG samples (Fig. 3C). Similar results were obtained when comparing the trimmed reads against the Swiss-Prot human virus database. The proportion of human viral reads in the VIRADEL samples ranged from 0.00059 % to 0.00162 %, with an averaging value as 0.01210 ± 0.00034 % (Fig. 3D). For PEG samples, the proportion of viral reads in the samples was 0.00007 % to 0.00020 %, with an averaging value as 0.00013 ± 0.00005 % (Fig. 3D).

Virus compositions of assembled contigs to achieve better accuracy in taxonomic mapping were analyzed. Viral contigs were obtained by comparing the assembled contigs against the NCBI RefSeq virus database with DIAMOND blastx. Proportion of viral contigs obtained for VIRADEL samples ranged from 17.2 % to 21.8 %. For PEG samples, it was 7.76 % to 9.09 % (Fig. 3E). Comparison of the assembled contigs against the Swiss-Prot human virus database showed that the proportion of human viral contigs in VIRADEL samples was 0.29 % to 0.46 %, and for PEG samples, it was 0.22 % to 0.29 % (Fig. 3F).

Combined results showed that proportions of reads and contigs for both viruses and human viruses recovered with VIRADEL concentration method were higher than those recovered with PEG precipitation method (Fig. 3).

3.2. Impact of sampling methods on viral and human viral reads

Among the clean reads in VIRADEL samples, 61.2 % ~ 80.7 % of them were unclassified reads, and the proportion of reads that were classified as bacteria ranged from 16.2 % to 36.9 %. As for the clean reads in PEG samples, 38.3 % ~ 51.2 % of them were unclassified reads, and proportion of reads that were classified as bacteria ranged from 48.4 % to 60.8 %. The average proportion of viral reads in VIRADEL samples were 2.84 ± 0.57 %. For PEG samples it was 0.34 % to 0.91 %. A

smaller proportion of bacterial reads and a greater proportion of viral reads were identified in the VIRADEL samples compared with the PEG samples, indicating a greater efficiency of VIRADEL method in removing bacteria and concentrating viruses (Fig. 4A).

Trimmed reads that were classified as viruses were extracted and compared against the custom Swiss-Prot human virus database to dissect the composition of human viral reads recovered by VIRADEL and PEG precipitation methods. Proportions of human viral reads in both VIRADEL and PEG samples were very low (VIRADEL: 0.00059 % ~ 0.00162 %; PEG: 0.00007 % ~ 0.00020 %). Even with low proportions of discovery, more human viral reads were detected with VIRADEL method (Fig. 3D). Further taxonomic classification of the identified human viral reads showed more read counts per 100 million reads were found in VIRADEL samples compared with those in PEG sample for almost all the human viruses identified such as *Mamastrovirus*, *Cyclovirus*, *Bocaparvovirus*, *Varicellovirus*, *Sapovirus*, and *Orthopoxvirus* (Fig. 4B).

3.3. VIRADEL method recovers similar abundant viruses but higher diversity of human viral contigs compared to PEG

To improve the discovery of human viruses, reduce ambiguity in the taxonomic alignment, and decrease the chance of false negative hits (McCall et al., 2020), assembled contigs were compared against the RefSeq virus database and the custom Swiss-Prot human virus database to identify human virus diversity in wastewater samples.

Virus compositions identified for both the VIRADEL and PEG samples based on RefSeq virus database matches were analyzed at the family level. The three most abundant families were *Myoviridae*, *Siphoviridae* and *Podoviridae* (Fig. 5A). While the abundant viral families identified in VIRADEL and PEG samples are similar, NMDS analysis revealed the divergent virus populations concentrated by these two methods (Fig. 5B), demonstrating the impact of sampling methods and representation of viral taxa. Compared to the timing of sample collection, which is expected to have a strong batch effect, the sample collection methods have an even stronger effect since they are readily separable on the NMDS1 axis.

The averaging proportion of human viral contigs in VIRADEL samples was 0.36 ± 0.07 %, and for PEG samples, it was 0.25 ± 0.02 %.

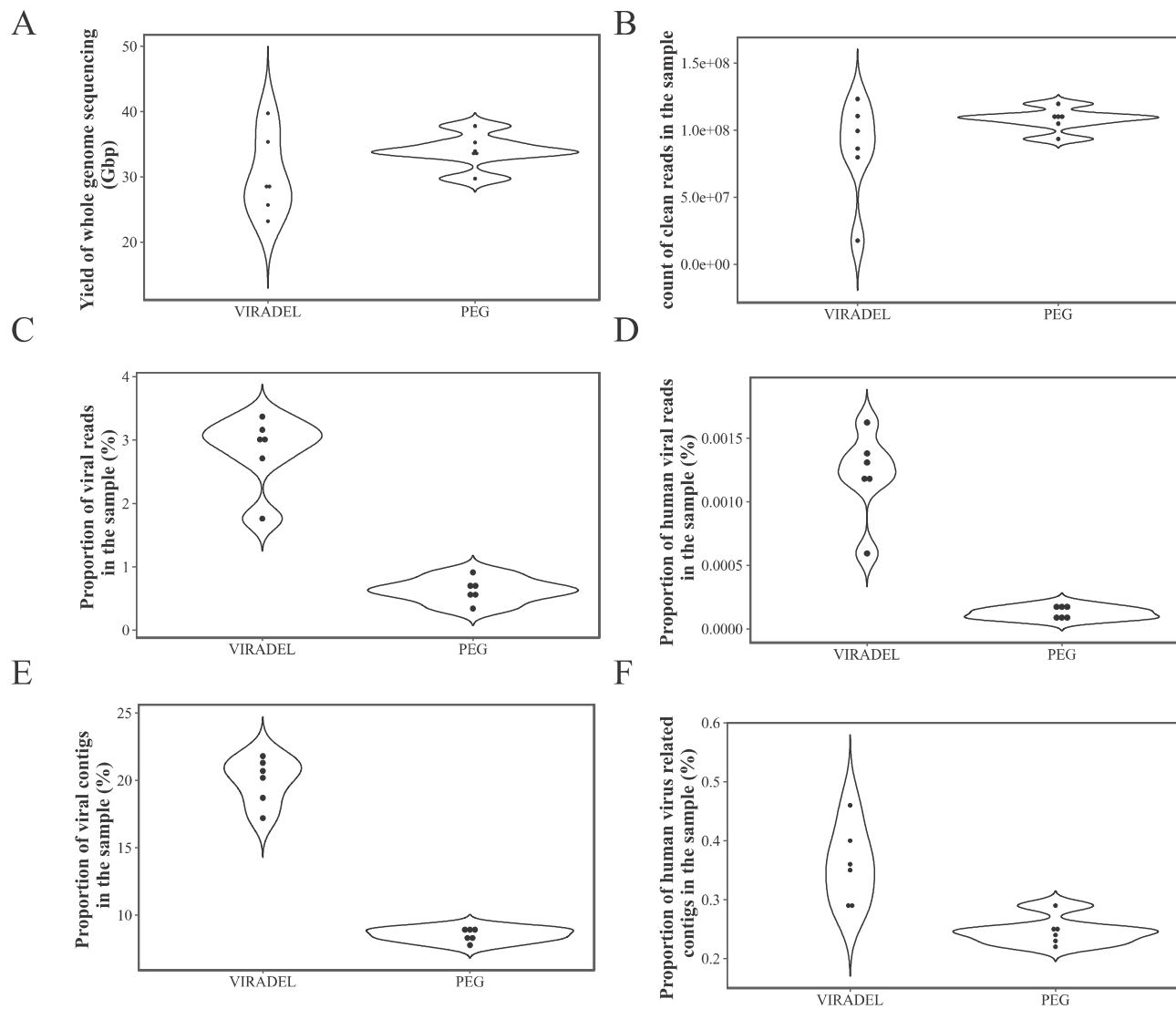


Fig. 3. A: Yields of whole genome sequencing data, B: Count of the clean reads, C: Proportion of viral reads, D: Proportion of human viral reads, E: Proportion of viral contigs, F: Proportion of human health related contigs in the wastewater samples concentrated with VIRADEL and PEG precipitation methods.

Compositions of the human viruses were normalized to the human virus community. Among the human viral contigs in both VIRADEL and PEG samples, the dominant one was *Orthopoxvirus* (Fig. 6A). Further composition analysis of *Orthopoxvirus* indicated more than 80 % of *Orthopoxvirus* contigs were unclassified, as for the classified contigs, the predominant one is the vaccinia virus. In addition to *Orthopoxvirus*, other relatively abundant human viruses (average proportions in human virus community greater than 1.00 %) in samples concentrated with the VIRADEL method were *Rhadinovirus*, *Parapoxvirus*, *Varicellovirus*, *Hepatovirus*, *Simplexvirus*, *Molluscipoxvirus*, *Parechovirus*, *Lymphocryptovirus*, and *Spumavirus*. While in samples concentrated with the PEG method, fewer human viruses were found to be relatively abundant. These were *Rhadinovirus*, *Varicellovirus*, *Simplexvirus*, *Molluscipoxvirus*, *Lymphocryptovirus*, and *Betacoronavirus* (Fig. 6A). Contigs affiliated with *Betacoronavirus* genus, which includes the SARS-COV-2 virus, were identified in all the samples, further targeted sequencing is needed to identify the *Betacoronavirus* members at the species or strain level. NMDS analysis with Bray-Curtis distance method using human viruses recovered by VIRADEL and PEG methods from wastewater samples revealed the divergent human virus populations recovered by these two methods (Fig. 6B).

Occurrence frequency of each human virus in the VIRADEL and PEG

samples was showed contigs related to several human viruses were identified more frequently in VIRADEL samples (Fig. 7). To be more specific, contigs related to genus *Orthopoxvirus*, *Rhadinovirus*, *Parapoxvirus*, *Varicellovirus*, *Hepatovirus*, *Simplexvirus*, *Bocaparvovirus*, *Molluscipoxvirus*, *Parechovirus*, *Roseolovirus*, *Lymphocryptovirus*, *Deltaretrovirus*, *Alphavirus*, *Betacoronavirus*, *Mamastrovirus*, *Spumavirus*, *Lentivirus* and *Rubivirus* were found in each sample concentrated with VIRADEL and PEG methods. Contigs related to genus *Enterovirus*, *Erythroparvovirus*, *Cytomegalovirus*, *Kobuvirus*, *Salivirus* and *Gammaretrovirus* were identified in every VIRADEL sample, while in PEG samples, their occurrence frequencies were 83.3 %, 66.7 %, 66.7 %, 33.3 %, 50.0 % and 66.7 %. In addition, there were some human viruses that were identified in VIRADEL samples (occurrence frequency was 50 % ~ 90 %) but that turned out to be undetected in PEG samples. The viruses missing from PEG samples were *Mastadenovirus*, *Hepacivirus*, *Orthohepadnavirus*, and *Betapolyomavirus* (Fig. 7).

4. Discussion

Although metagenomic-enabled surveillance allows comprehensive and deep investigation of viral pathogens in wastewater, some important uncertainties remain. Ongoing method development and

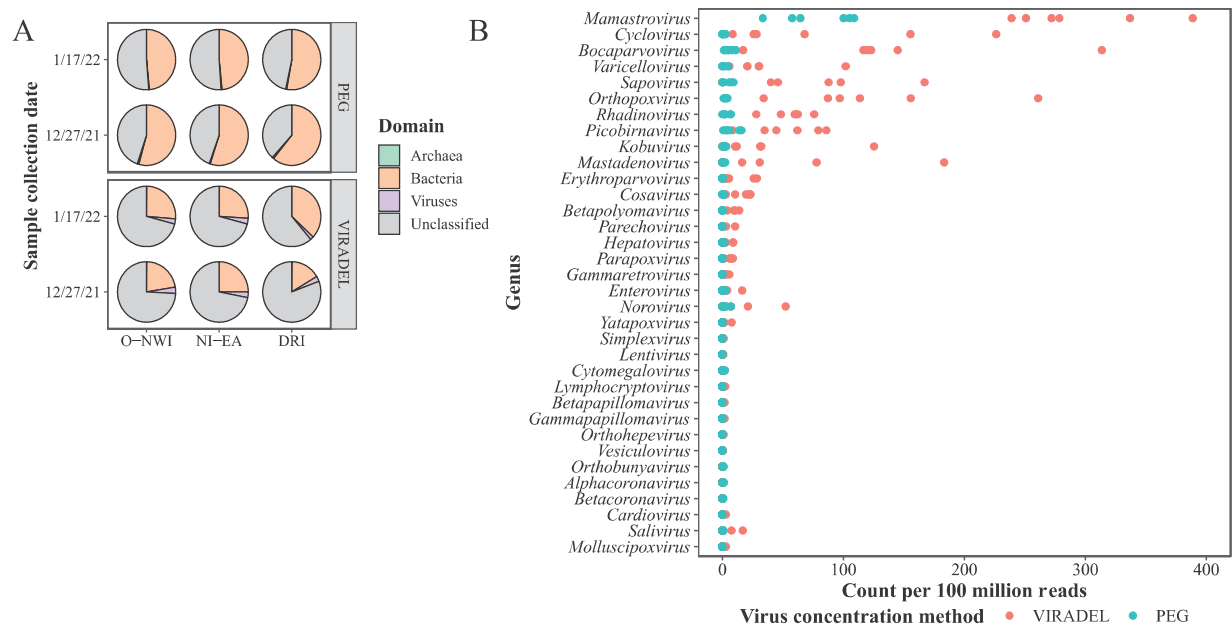


Fig. 4. A: Proportion of bacterial, archaeal, and viral reads in the wastewater samples; B: The read count of each human virus concentrated from the wastewater samples with VIRADEL and PEG samples, values are described as count per 100 million reads.

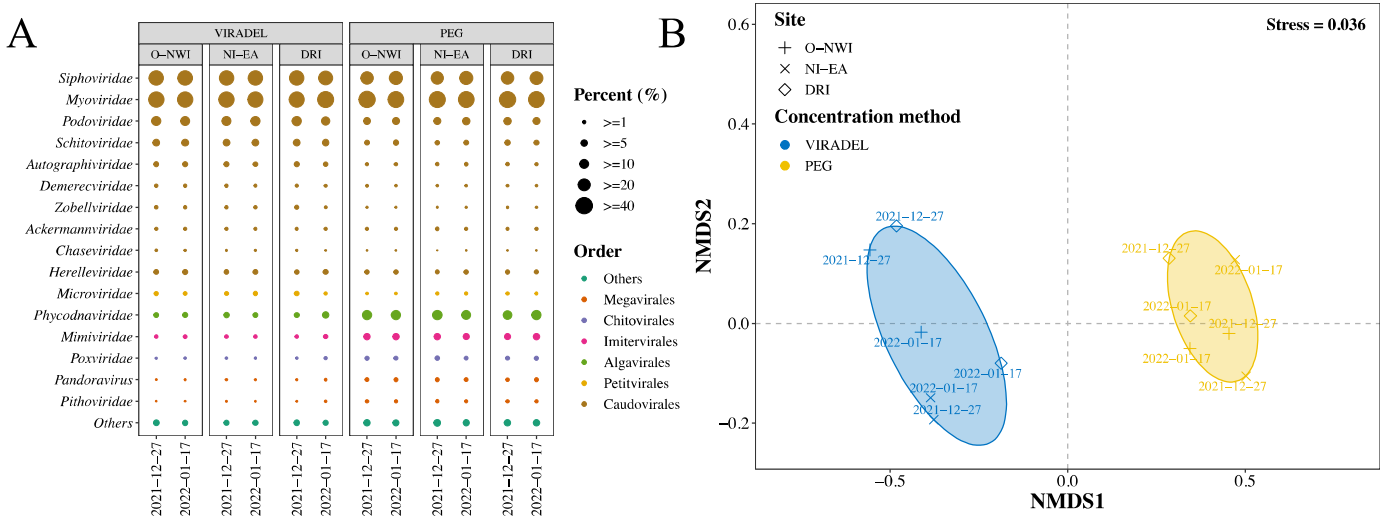


Fig. 5. A: Proportions of viral families recovered in the VIRADEL ($n = 6$) and PEG ($n = 6$) samples. Family with percent greater than 1 % in at least one sample were included, otherwise, was classified as Others. B: Non-metric multidimensional scaling (NMDS) analysis for viral community concentrated with VIRADEL and PEG precipitation methods. Distance method Bray-Curtis was used. The analyses were based on the virus composition at the family level.

optimization are essential to further broaden the capabilities of wastewater surveillance using metagenomics. Factors influencing the efficiency and accuracy of viral surveillance in wastewater can be traced back to the sampling stage. No single method is entirely effective for concentrating all human viruses from wastewater (Haramoto et al., 2018). With the electropositive NanoCeram column filters, the primary step of VIRADEL method, is suggested as the cost-effective approach to facilitate the binding of negatively charged viruses (Haramoto et al., 2018). Following primary concentration, processes such as flocculation, centrifugation, and filtration with 0.45 μm and 0.22 μm syringe filters, the VIRADEL method helps reduce the influence from the presence of bacteria and improve virus detectability. Our findings suggest that the VIRADEL concentration method, which involves collection of large volumes of grab samples followed by electropositive filtration, may be beneficial to recover more human viruses compared to PEG. Larger

proportions of reads and contigs of viruses and human viruses were found in wastewater samples collected and concentrated with the VIRADEL method.

The PEG method applied in this work involved 24-h composite sampling followed by precipitation. Although composite samples collected over multiple hours are believed to be more representative than grab samples considering the influences of different parameters (e.g., organic matter in the wastewater, wastewater temperature, wastewater pH, rainfall) on virus occurrence throughout the day (Schmidt, 2020), small volumes processed for composite samples in most studies (within 30–250 mL, in this study it was 100 mL) due to the practicality of laboratory operation (Zheng et al., 2022) in the subsequent virus concentration step, may be one of the factors that contribute to the low viral amounts being captured.

Contigs related to *Orthopoxvirus*, *Rhadinovirus*, *Parapoxvirus*,

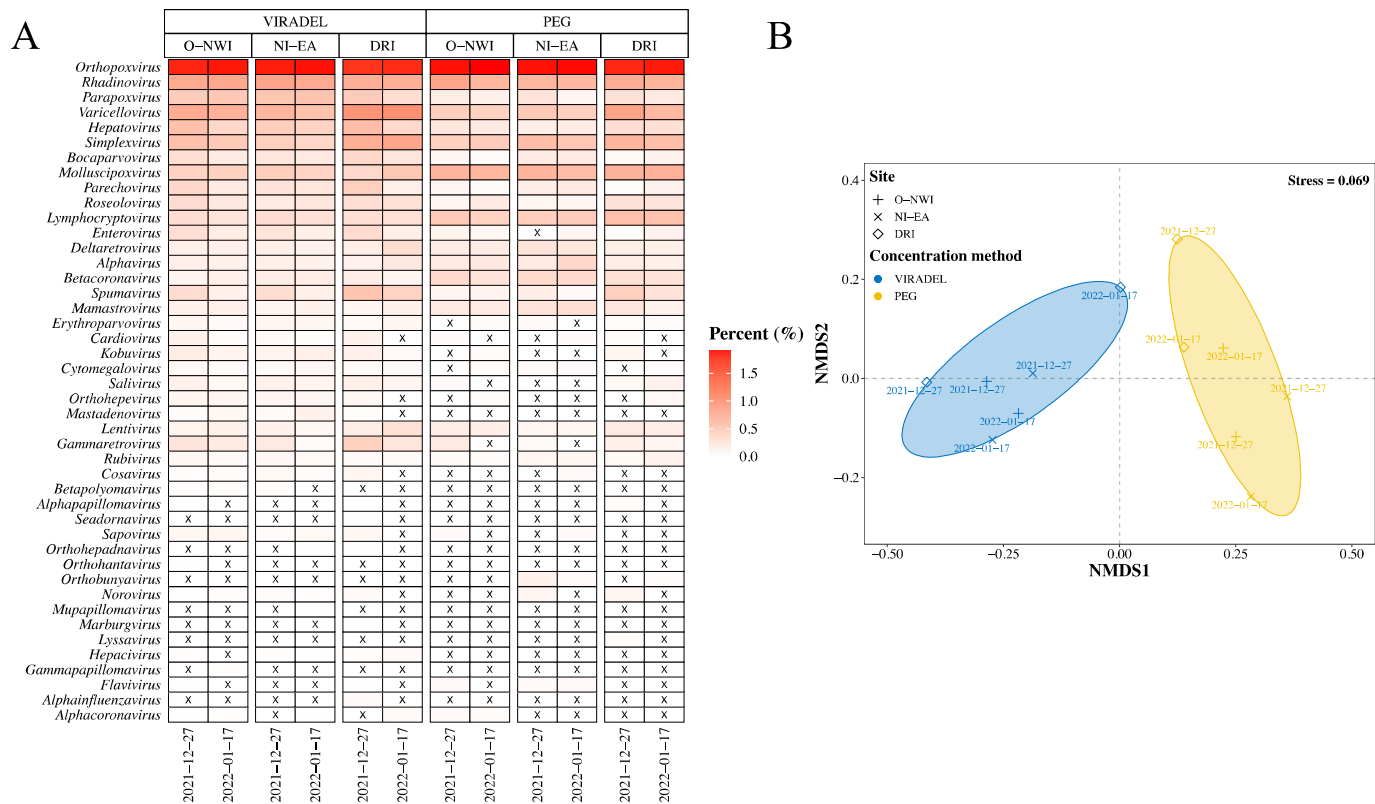


Fig. 6. A: Proportions (log-transformed) of human virus genera normalized to the human viruses identified in the wastewater samples concentrated with VIRADEL (n = 6) and PEG (n = 6) methods. Symbol “X” indicates the virus is not identified in the sample. B: Non-metric MultiDimensional Scaling (NMDS) analysis of the human virus composition at the genus level in the samples concentrated with VIRADEL and PEG precipitation methods, with distance method as Bray-Curtis. Normalized compositions with the human virus community were used.

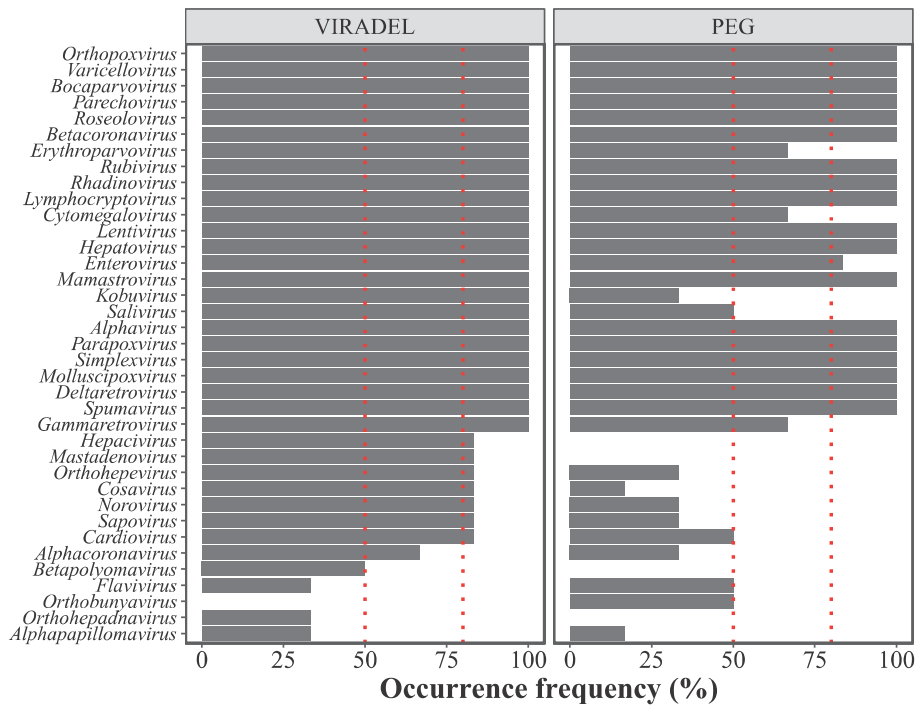


Fig. 7. Occurrence frequency (%) of the related human viruses in 6 VIRADEL samples and 6 PEG samples during the COVID-19 pandemic.

Varicellovirus, *Hepatovirus*, *Simplexvirus*, *Bocaparvovirus*, *Molluscipoxvirus*, *Parechovirus*, *Roseolovirus*, *Lymphocryptovirus*, *Deltaretrovirus*, *Alphavirus*, *Betacoronavirus*, *Mamastrovirus*, *Spumavirus*, *Lentivirus* and *Rubivirus* genera were found in each sample concentrated with VIRADEL and PEG methods. Both concentration methods facilitated the detection of human viruses such as *Betacoronavirus*, which includes SARS-CoV-2, the causative agent of COVID-19 (Li et al., 2022b; Zhao et al., 2022). The VIRADEL method allowed the detection of *Mastadenovirus*, *Hepacivirus*, *Orthohepadnavirus* and *Betapolyomavirus* genera even with a few contigs, which may be related to parameters such as the virus properties and wastewater sample types. For example, the virus hydrophobicity and morphology are important factors that can affect the recovery efficiency of them by different concentration methods (Shi et al., 2017). Meanwhile, the VIRADEL method targeted supernatant viruses in wastewater while the PEG precipitation method targeted viruses adsorbed on particles (Flood et al., 2021; Zhao et al., 2022). Wastewater parameters such as the particle sizes that are possibly different in these two types of samples may contribute to the varied recovery of human viruses. The efficiency of concentration methods for different types of human viruses and their recovery in wastewater remain to be assessed.

In addition to the sample collection and concentration processes, the following extracting and amplifying steps are also of great significance for virus identification in wastewater with metagenomics. For example, it is showed that the lysis-buffer-based extraction method, which is identical to the method used in this study, led to a higher SARS-CoV-2 recovery efficiencies compared to that with the acid-guanidinium-phenol-based method (Zheng et al., 2022). Evaluation of virus concentration and extraction methods from sewage with metagenomic sequencing showed that the largest viral richness was found in samples extracted with QIAamp Viral RNA Mini Kit, which is the extraction kit applied in this study (Hjelmsø et al., 2017). Furthermore, obtaining longer sequences or the nearly complete metagenome-assembled genomes is key to human virus identification (Ko et al., 2022), especially at the species or strain level as nearly complete viral genome assemblies enable more accurate taxonomic mapping and allow further comparative genomic analyses to be performed. For example, in this work, more viruses and human viruses were found when comparing the assembled reads against the viral and human viral databases.

Despite the need for further optimization, current methods reveal a wide diversity of human-related viral contigs in wastewater samples collected and concentrated from municipal wastewater in the Detroit area during the COVID-19 pandemic. This methodology may serve as a screening method for identifying viruses of potential concern for further investigation. Since the COVID-19 pandemic the metropolitan Detroit wastewater surveillance program focused on SARS-CoV-2 and has been proven to be a useful tool in providing early warnings of disease prevalence (Li et al., 2022b; Miyani et al., 2021b; Zhao et al., 2023; Zhao et al., 2022). To demonstrate the general utility of viral monitoring through wastewater surveillance it is necessary to study potentially infectious viruses other than SARS-CoV-2 with innovative technologies like metagenomics. Indeed, this work shows that in addition to coronaviruses, multiple other human related viral contigs were found in the samples. Among the identified virus genera, *Orthopoxvirus*, was found to be the most abundant and one of the most frequently identified human-related viruses in all the samples. Most of the assigned contigs within genus *Orthopoxvirus* are classified as vaccinia virus (VACV) in this study, which is consistent with our previous work (McCall et al., 2020). *Betacoronavirus* was identified in the wastewater samples, which is expected since the samples were collected during the COVID-19 pandemic period. The same samples were analyzed with ddPCR, and SARS-CoV-2 occurrence was confirmed (Li et al., 2022b; Zhao et al., 2023; Zhao et al., 2022). *Bocaparvovirus* and *Erythroparvovirus*, two respiratory human viruses belonging to family *Parvoviridae*, were found to be prevalent in the wastewater samples. Beyond contigs related to respiratory viruses, contigs related to human viruses potentially transmitted with fecal-oral route were also observed with high occurrences in our samples. These

include *Mamastrovirus*, *Norovirus*, *Orthohepevirus*, *Hepatovirus*, *Enterovirus*, *Kobuvirus*, *Salivirus*, and *Cosavirus* related contigs. Other than respiratory and fecal-oral transmission, contigs potentially related to bloodborne viruses were detected in the study. *Lentivirus* was found to be prevalent. Also, *Orthohepadnavirus*, *Hepacivirus* and *Lymphocryptovirus* were detected. Furthermore, *Flavivirus* and *Alphavirus* were detected in multiple samples. Within genus *Flavivirus* and *Alphavirus* there are some mosquito-transmitted viruses.

Random amplification and untargeted metagenomics enable broad surveillance of pathogens at the genus level and can be used to identify possible novel pathogens. However, the complex wastewater matrix and the high diversity of viral pathogen types make it challenging to recover virus communities with high abundance and resolution. Identification of contigs related to human viruses of concern does not necessarily indicate presence of infectious species in the community, however it provides critical information for further human virus surveillance studies with an integrated workflow that may include targeted PCR, Sanger sequencing, and phylogenetic analysis.

5. Conclusions

- Comparison of human viruses concentrated with two methods revealed that more viral reads, contigs, and human viral-related contigs were obtained in almost all wastewater samples collected and concentrated with the VIRADEL method as compared to samples collected and concentrated with the PEG precipitation method.
- Prevalent identification of human viruses in metro Detroit revealed the application of metagenomics in wastewater surveillance as a first screening tool.
- Optimization of the current methods are essential to further broaden the capabilities of wastewater surveillance using metagenomics.

CRedit authorship contribution statement

Yabing Li: Methodology, Investigation, Formal analysis, Writing – original draft, Writing – review & editing. **Brijen Miyani:** Investigation, Writing – original draft, Writing – review & editing. **Kevin L. Childs:** Methodology, Writing – review & editing. **Shin-Han Shiu:** Methodology, Writing – review & editing. **Irene Xagoraki:** Conceptualization, Methodology, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

Funding for this work was provided by the Great Lakes Water Authority (GLWA). We thank CDM Smith, the City of Detroit and the local Health Departments for their support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2023.168128>.

References

- Ahmed, W., Bertsch, P.M., Bivins, A., Bibby, K., Farkas, K., Gathercole, A., Haramoto, E., Gyawali, P., Korajkic, A., McMinn, B.R., Mueller, J.F., Simpson, S.L., Smith, W.J.M.,

- Symonds, E.M., Thomas, K.V., Verhagen, R., Kitajima, M., 2020. Comparison of virus concentration methods for the RT-qPCR-based recovery of murine hepatitis virus, a surrogate for SARS-CoV-2 from untreated wastewater. *Sci. Total Environ.* 739, 139960 <https://doi.org/10.1016/j.scitotenv.2020.139960>.
- Alexander, M.R., Rootes, C.L., van Vuren, P.J., Stewart, C.R., 2020. Concentration of infectious SARS-CoV-2 by polyethylene glycol precipitation. *J. Virol. Methods* 286, 113977. <https://doi.org/10.1016/j.jviromet.2020.113977>.
- Andrews, S., 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.
- Ayling, M., Clark, M.D., Leggett, R.M., 2019. New approaches for metagenome assembly with short reads. *Brief. Bioinform.* 21 (2), 584–594. <https://doi.org/10.1093/bib/bbz020>.
- Bağcı, C., Patz, S., Huson, D.H., 2021. DIAMOND+ MEGAN: fast and easy taxonomic and functional analysis of short and long microbiome sequences. *Current protocols* 1 (3), e59. <https://doi.org/10.1002/cpz1.59>.
- Bibby, K., Peccia, J., 2013. Identification of viral pathogen diversity in sewage sludge by metagenome analysis. *Environ. Sci. Technol.* 47 (4), 1945–1951. <https://doi.org/10.1021/es305181x>.
- Bibby, K., Viau, E., Peccia, J., 2011. Viral metagenome analysis to guide human pathogen monitoring in environmental samples. *Lett. Appl. Microbiol.* 52 (4), 386–392.
- Bofill-Mas, S., Rusinöl, M., 2020. Recent trends on methods for the concentration of viruses from water samples. *Current Opinion in Environmental Science & Health* 16, 7–13. <https://doi.org/10.1016/j.coesh.2020.01.006>.
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30 (15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Dunkin, N., Weng, S., Coulter, C.G., Jacangelo, J.G., Schwab, K.J., 2018. Impacts of virus processing on human norovirus GI and GII persistence during disinfection of municipal secondary wastewater effluent. *Water Res.* 134, 1–12. <https://doi.org/10.1016/j.watres.2018.01.053>.
- Fernandez-Cassi, X., Timoneda, N., Martinez-Puchol, S., Rusinöl, 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>.
- Flood, M.T., D'Souza, N., Rose, J.B., Aw, T.G., 2021. Methods evaluation for rapid concentration and quantification of SARS-CoV-2 in raw wastewater using droplet digital and quantitative RT-PCR. *Food and Environmental Virology* 13 (3), 303–315. <https://doi.org/10.1007/s12560-021-09488-8>.
- Fout, G., Dahling, D., Safferman, R., 2001. Concentration and Processing of Waterborne Viruses by Positive Charge 1MDS Cartridge Filters and Organic Flocculation. Environmental Protection Agency (EPA) Document No. EPA/600/4-84/013. EPA, Cincinnati.
- Fout, G.S., 1996. ICR Microbial Laboratory Manual.
- Haramoto, E., Kitajima, M., Hata, A., Torrey, J.R., Masago, Y., Sano, D., Katayama, H., 2018. A review on recent progress in the detection methods and prevalence of human enteric viruses in water. *Water Res.* 135, 168–186. <https://doi.org/10.1016/j.watres.2018.02.004>.
- Hjelmso, M.H., Hellmér, M., Fernandez-Cassi, X., Timoneda, N., Lukjancenko, O., Seidel, M., Elsässer, D., Aarestrup, F.M., Löfström, C., Bofill-Mas, S., 2017. Evaluation of methods for the concentration and extraction of viruses from sewage in the context of metagenomic sequencing. *PLoS One* 12 (1), e0170199.
- Holtgrewe, M., 2010. Mason – a read simulator for second generation sequencing data.
- Huson, D.H., Beier, S., Flade, I., Górka, 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 (6), e1004957. <https://doi.org/10.1371/journal.pcbi.1004957>.
- Kim, K.-H., Bae, J.-W., 2011. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. *Appl. Environ. Microbiol.* 77 (21), 7663–7668. <https://doi.org/10.1128/AEM.00289-11>.
- Ko, K.K.K., Chng, K.R., Nagarajan, N., 2022. Metagenomics-enabled microbial surveillance. *Nat Microbiol* 7 (4), 486–496. <https://doi.org/10.1038/s41564-022-01089-w>.
- Levy, J.I., Andersen, K.G., Knight, R., Karthikeyan, S., 2023. Wastewater surveillance for public health. *Science* 379 (6627), 26–27.
- Li, J., Verhagen, R., Ahmed, W., Metcalfe, S., Thai, P.K., Kaserzon, S.L., Smith, W.J., Schang, C., Simpson, S.L., Thomas, K.V., 2022a. In situ calibration of passive samplers for viruses in wastewater. *ACS ES&T Water*. <https://doi.org/10.1021/acsestwater.1c00406>.
- Li, Y., Miyani, B., Zhao, L., Spooner, M., Gentry, Z., Zou, Y., Rhodes, G., Li, H., Kaye, A., Norton, J., Xagorarakis, I., 2022b. Surveillance of SARS-CoV-2 in nine neighborhood sewersheds in Detroit Tri-County area, United States: assessing per capita SARS-CoV-2 estimations and COVID-19 incidence. *Sci. Total Environ.* 851, 158350 <https://doi.org/10.1016/j.scitotenv.2022.158350>.
- Lu, D., Huang, Z., Luo, J., Zhang, X., Sha, S., 2020. Primary concentration – the critical step in implementing the wastewater based epidemiology for the COVID-19 pandemic: a mini-review. *Sci. Total Environ.* 747, 141245 <https://doi.org/10.1016/j.scitotenv.2020.141245>.
- McCall, C., Wu, H., Miyani, B., Xagorarakis, I., 2020. Identification of multiple potential viral diseases in a large urban center using wastewater surveillance. *Water Res.* 184, 116160 <https://doi.org/10.1016/j.watres.2020.116160>.
- McCall, C., Wu, H., O'Brien, E., Xagorarakis, I., 2021. Assessment of enteric viruses during a hepatitis outbreak in Detroit MI using wastewater surveillance and metagenomic analysis. *J. Appl. Microbiol.* 131 (3), 1539–1554. <https://doi.org/10.1111/jam.15027>.
- McMinn, B.R., Korajkic, A., Pemberton, A.C., Kelleher, J., Ahmed, W., Villegas, E.N., Oshima, K., 2023. Assessment of two volumetrically different concentration approaches to improve sensitivities for SARS-CoV-2 detection during wastewater monitoring. *J. Virol. Methods* 311, 114645. <https://doi.org/10.1016/j.jviromet.2022.114645>.
- Menzel, P., Ng, K.L., Krogh, A., 2016. Fast and sensitive taxonomic classification for metagenomics with kaiju. *Nat. Commun.* 7 (1), 11257. <https://doi.org/10.1038/ncomms11257>.
- Miyani, B., Fonoll, X., Norton, J., Mehrotra, A., Xagorarakis, I., 2020. SARS-CoV-2 in Detroit wastewater. *Journal of Environmental Engineering* 146 (11), 06020004. [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0001830](https://doi.org/10.1061/(ASCE)EE.1943-7870.0001830).
- Miyani, B., McCall, C., Xagorarakis, I., 2021a. High abundance of human herpesvirus 8 in wastewater from a large urban area. *J. Appl. Microbiol.* 130 (5), 1402–1411. <https://doi.org/10.1111/jam.14895>.
- Miyani, B., Zhao, L., Spooner, M., Buch, S., Gentry, Z., Mehrotra, A., Norton, J., Xagorarakis, I., 2021b. Early warnings of COVID-19 second wave in Detroit. *J. Environ. Eng.* 147 (8) [https://doi.org/10.1061/\(asce\)ee.1943-7870.0001907](https://doi.org/10.1061/(asce)ee.1943-7870.0001907).
- O'Brien, E., Xagorarakis, I., 2019. A water-focused one-health approach for early detection and prevention of viral outbreaks. *One Health* 7, 100094. <https://doi.org/10.1016/j.onehlt.2019.100094>.
- O'Brien, E., Munir, M., Marsh, T., Heran, M., Lesage, G., Tarabara, V.V., Xagorarakis, I., 2017. Diversity of DNA viruses in effluents of membrane bioreactors in Traverse City, MI (USA) and La Grande motte (France). *Water Res.* 111, 338–345. <https://doi.org/10.1016/j.watres.2017.01.014>.
- Pantaleo, V., Chiumenti, M., 2018. *Viral Metagenomics*. Springer.
- Parras-Moltó, M., Rodríguez-Galet, A., Suárez-Rodríguez, P., López-Bueno, A., 2018. Evaluation of bias induced by viral enrichment and random amplification protocols in metagenomic surveys of saliva DNA viruses. *Microbiome* 6 (1), 1–18. <https://doi.org/10.1186/s40168-018-0507-3>.
- Polson, S.W., Wilhelm, S.W., Wommack, K.E., 2011. Unraveling the viral tapestry (from inside the capsid out). *ISME J.* 5 (2), 165–168. <https://doi.org/10.1038/ismej.2010.81>.
- Schmidt, C., 2020. Watcher in the wastewater. *Nat. Biotechnol.* 38 (8), 917–921.
- Shi, H., Pasco, E.V., Tarabara, V.V., 2017. Membrane-based methods of virus concentration from water: a review of process parameters and their effects on virus recovery. *Environ. Sci.: Water Res. Technol.* 3 (5), 778–792.
- Silverman, A.I., Boehm, A.B., 2021. Systematic review and meta-analysis of the persistence of enveloped viruses in environmental waters and wastewater in the absence of disinfectants. *Environ. Sci. Technol.* 55 (21), 14480–14493. <https://doi.org/10.1021/acs.est.1c03977>.
- Thomas, T., Gilbert, J., Meyer, F., 2012. Metagenomics - a guide from sampling to data analysis. *Microbial Informatics and Experimentation* 2 (1), 3. <https://doi.org/10.1186/2042-5783-2-3>.
- USEPA, 2001. Concentration and Processing of Waterborne Viruses Bypositive Charge 1MDS Cartridge Filters and Organic Flocculation. Chap.14 in *USEPA Manual of Methods of Virology*. USEPA, Washington, DC.
- Wang, D., Coscoy, L., Zylberberg, M., Avila, P.C., Boushey, H.A., Ganem, D., DeRisi, J.L., 2002. Microarray-based detection and genotyping of viral pathogens. *Proc. Natl. Acad. Sci.* 99 (24), 15687–15692. <https://doi.org/10.1073/pnas.242579699>.
- Wang, D., Urisman, A., Liu, Y.-T., Springer, M., Ksiazek, T.G., Erdman, D.D., Mardis, E.R., Hickenbotham, M., Magrini, V., Eldred, J., 2003. Viral discovery and sequence recovery using DNA microarrays. *PLoS Biol.* 1 (2), e2 <https://doi.org/10.1371/journal.pbio.0000002>.
- Wollants, E., Keyaerts, E., Cuypers, L., Bloemen, M., Thijssen, M., Ombelet, S., Raymenants, J., Beuselinck, K., Laenen, L. and Budts, L. (2022). Environmental circulation of adenovirus 40/41 and SARS-CoV-2 in the context of the emergence of acute hepatitis of unknown origin. *medRxiv*. doi:<https://doi.org/10.1101/2022.06.08.22276091>.
- Woolhouse, M., Scott, F., Hudson, Z., Howey, R., Chase-Topping, M., 2012. Human viruses: discovery and emergence. *Philos. Trans. R. Soc., B* 367 (1604), 2864–2871. <https://doi.org/10.1098/rstb.2011.0354>.
- Xagorarakis, I., Yin, Z., Svambayev, Z., 2014. Fate of viruses in water systems. *Journal of Environmental Engineering* 140 (7). [https://doi.org/10.1061/\(asce\)ee.1943-7870.0000827](https://doi.org/10.1061/(asce)ee.1943-7870.0000827).
- Zhao, L., Zou, Y., Li, Y., Miyani, B., Spooner, M., Gentry, Z., Jacobi, S., David, R.E., Withington, S., McFarlane, S., Faust, R., Sheets, J., Kaye, A., Broz, J., Gosine, A., Mobley, P., Busch, A.W.U., Norton, J., Xagorarakis, I., 2022. Five-week warning of COVID-19 peaks prior to the omicron surge in Detroit, Michigan using wastewater surveillance. *Sci. Total Environ.* 844, 157040 <https://doi.org/10.1016/j.scitotenv.2022.157040>.
- Zhao, L., Zou, Y., David, R.E., Withington, S., McFarlane, S., Faust, R.A., Norton, J., Xagorarakis, I., 2023. Simple methods for early warnings of COVID-19 surges: lessons learned from 21 months of wastewater and clinical data collection in Detroit, Michigan, United States. *Sci. Total Environ.* 864, 161152 <https://doi.org/10.1016/j.scitotenv.2022.161152>.
- Zheng, X., Deng, Y., Xu, X., Li, S., Zhang, Y., Ding, J., On, H.Y., Lai, J.C., Yau, C.I., Chin, A.W., 2022. Comparison of virus concentration methods and RNA extraction methods for SARS-CoV-2 wastewater surveillance. *Sci. Total Environ.* 824, 153687 <https://doi.org/10.1016/j.scitotenv.2022.153687>.