# A Snapshot of the Global Drinking Water Virome: Diversity and Metabolic Potential Vary with Residual Disinfectant Use

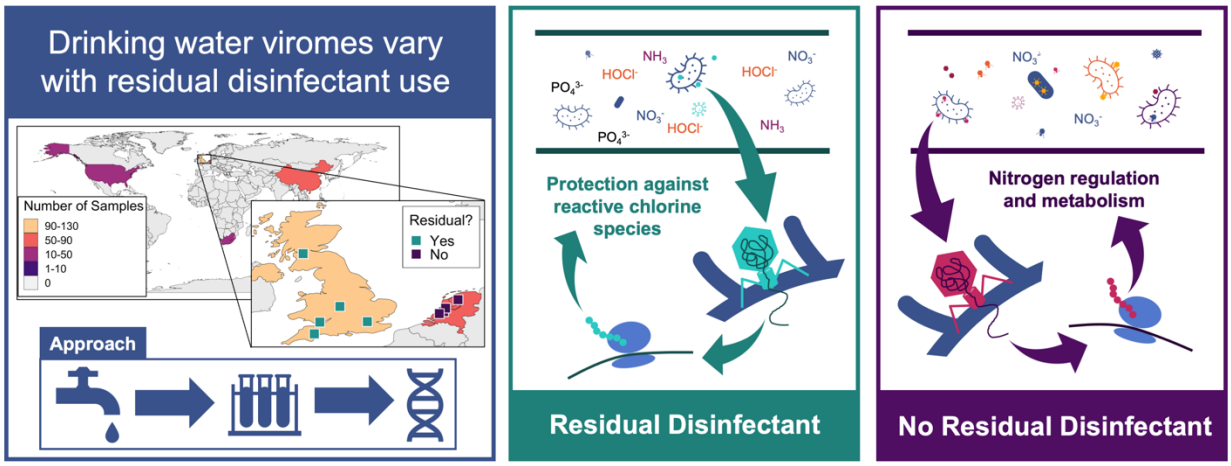
**Authors**
Bridget Hegarty[1]
Zihan Dai[2]
Lutgarde Raskin[1]
Ameet Pinto[3]
Krista Wigginton[1,+]
Melissa Duhaime[4,*]

**Corresponding Authors**
+ Mailing address: Department of Civil and Environmental Engineering, 1351 Beal Ave., 181 Environmental and Water Resources Engineering Building, Ann Arbor, MI 48109-2125, USA
Phone: +1 (734) 763-9661
Email: kwigg@umich.edu

* Mailing address: Department of Ecology and Evolutionary Biology, 1105 N University Ave., 4068 Biological Sciences Building, Ann Arbor, MI 48109-1085, USA.
Phone: +1 (734) 764-6219;
E-mail: duhaimem@umich.edu

[1] Civil and Environmental Engineering; University of Michigan
[2] Key Laboratory of Drinking Water Science and Technology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China
[3] University of Chinese Academy of Sciences, Beijing 100049, China
[3] School of Civil and Environmental Engineering; Georgia Institute of Technology.
[4] Ecology and Evolutionary Biology; University of Michigan

# Graphical Abstract




# Abstract

Viruses are important drivers of microbial community ecology and evolution, influencing microbial mortality, metabolism, and horizontal gene transfer. However, the effects of viruses remain largely unknown in many environments, including in drinking water systems. Drinking water metagenomic studies have offered a whole community perspective of bacterial impacts on water quality, but have not yet considered the influences of viruses. In this study, we address this gap by mining viral DNA sequences from publicly available drinking water metagenomes from distribution systems in six countries around the world. These datasets provide a snapshot of the taxonomic diversity and metabolic potential of the global drinking water virome, and provide an opportunity to investigate the effects of geography, climate, and drinking water treatment practices on viral diversity. Both environmental conditions and differences in sample processing were found to influence the viral composition. Using free chlorine as the residual disinfectant was associated with clear differences in viral taxonomic diversity and metabolic potential, with significantly fewer viral populations and less even viral community structures than observed in distribution systems without residual disinfectant. Additionally, significantly more viral-encoded genes involved in mitigating oxidative stress were observed in systems that use free chlorine, while significantly more viral genes involved in nitrogen metabolism were observed in systems that do not. Through this study, we have demonstrated that viral communities are diverse across drinking water systems and vary with the use of residual disinfectant. Our findings offer directions for future research developing a more robust understanding of how virus-bacteria interactions in drinking water distribution systems affect water quality.

**Keywords:** viruses, drinking water microbiome, disinfection, metagenomics

# 1. Introduction

In an effort to prevent the spread of waterborne diseases, previous research on viruses in drinking water has focused on tracking and eliminating human viral pathogens (Dong et al., 2010; Gall et al., 2015; Petrovich et al., 2020; Rao et al., 1981; Ye et al., 2012). Researchers have tracked specific human viruses, using qPCR to understand their distribution throughout drinking water treatment plants and distribution systems (Albinana-Gimenez et al., 2006; Samhan et al., 2015; Ye et al., 2012). Others have used bacterial viruses, commonly referred to as bacteriophages or phages, as a proxy for human viruses to understand the effect of treatment processes on viral persistence (Dong et al., 2010; Hill et al., 2005; Langenfeld et al., 2021). While research on the behavior of human viruses is essential for assessing drinking water quality, the impacts of phages have rarely been addressed. Phages influence microbial populations (Rohwer et al., 2009; Suttle, 2007) and biogeochemical cycling (Parmar et al., 2017), through mortality (Fuhrman and Noble, 1995; Guidi et al., 2016) and by facilitating horizontal gene transfer between microorganisms and across environments (Rohwer et al., 2009). Similarly, viruses likely impact both beneficial and detrimental bacteria and other microorganisms in drinking water systems that have both direct (e.g., causing illness) and indirect effects (e.g., nitrification and corrosion) on drinking water quality.

Metagenomic studies have become critical to evaluate the cumulative effects of the drinking water microbiome on water quality (Hull et al., 2019; Liu et al., 2013). Past studies have shown that many factors, ranging from treatment practices (Bautista-de los Santos et al., 2016; Dai et al., 2020; Pinto et al., 2012) to source waters (Assche et al., 2019; Stanish et al., 2016), influence the bacterial, archaeal, and eukaryotic inhabitants of drinking water. For instance, the use of residual disinfectants (e.g., free chlorine, chloramine) has been connected to the selection of opportunistic pathogens (Feazel et al., 2009; Kotlarz et al., 2019; Liu et al., 2019; Vaerewijck et al., 2005), the prevalence of antimicrobial resistance genes (Jia et al., 2019, 2015; Shi et al., 2012), and shaping microbial communities (Dai et al., 2020; Hull et al., 2019; Krishna et al., 2020; Wang et al., 2014). At most, these studies have quantified the fraction of viral sequences in a cellular metagenome (Gomez-Alvarez et al., 2012; Id et al., 2020; Stamps et al., 2018). This leaves fundamental information about drinking water viruses, such as their taxonomic or metabolic diversity and influence on the drinking water microbiome, unknown. To fill similar knowledge gaps in other environments, cellular metagenomes have been re-examined for DNA viruses, revealing novel viral diversity and offering insights into viral-host interactions (Paez-Espino et al., 2016; Roux et al., 2015b). Taking such an approach can provide an untapped opportunity to illuminate the effects of phages on drinking water quality and safety.

The overall goal of this work was to characterize viral communities in drinking water distribution systems. To do so, we identified viral sequences from publicly available metagenomic datasets from drinking water distribution systems from around the world to assess

viral taxonomic diversity and metabolic potential. We further investigated the effect of study characteristics (e.g., the type of filter used for sample processing and the DNA extraction method) and environmental conditions (e.g., water quality) on the structuring of the observed viral communities. This study lays the groundwork for understanding the impact of viruses on the microbial communities in drinking water distribution systems.

# 2. Methodology

## 2.1 Data Collection

Publicly available drinking water metagenomic reads were downloaded from NCBI using the SRA Toolkit v2.9.6 (Leinonen et al., 2011) (Table S1). Sequences deposited on MG-RAST were directly requested from researchers. Only water samples that were collected from drinking water distribution systems were included in our analysis, excluding samples taken from within treatment plants and distribution system biofilm samples. Based on these criteria, 166 samples were used. Sequencing for some samples was replicated, resulting in a total of 312 sequencing runs. Replicate sequencing sets from the same samples were pooled during subsequent analyses ("sample" Table S1). Table S1 provides information about sample processing, sequencing platform, and other metadata for each sample included in this study.

## 2.2 Sequence Read Processing

Sequence reads were filtered and trimmed using fastp v0.20.1 (Chen et al., 2018). All code is deposited on github and made freely available (https://github.com/DuhaimeLab/global_drinking_water_viromes_Hegarty_et_al_2021). The UniVec Core database (as of February 2021) was used to remove contamination from vectors (NCBI, 2016).

## 2.3 Assembly

The filtered and processed reads from all samples of a distribution system were co-assembled using MetaSPAdes v3.10.1 (Nurk et al., 2017). Assembly quality was assessed with QUAST v5.0.2 (Gurevich et al., 2013). Metagenome assembled genomes (MAGs) were reconstructed using MetaBAT2 v2.12.1 (Kang et al., 2019) when only a single sample was taken from a given distribution system. When multiple samples were from the same distribution system, CONCOCT (Alneberg et al., 2014) within Anvi'o v5.5 (Eren et al., 2021) was used for MAG reconstruction.

## 2.4 Viral Population Identification

To define viral populations, putative viral contigs were identified from contigs greater than 3 kb using a strategy based on VirSorter v1.0.6 (Roux et al., 2015a), VirSorter2 v2.1 (Guo et al., 2021), CheckV v0.7.0 (Nayfach et al., 2020), VIBRANT v1.2.1 (Kieft et al., 2020), and VirFinder v1.1 (Ren et al., 2018) predictions (Figure S1). Any contig classified (i) as category 1 or 2 by VirSorter, (ii) as high or medium by VIBRANT, (iii) as complete, high, or medium by CheckV, (iv) with a VirSorter2 score greater than 0.95, (v) with at least two hallmark viral genes from VirSorter2, or (vi) with a VirFinder score greater than 0.9 were considered viral. Additionally, any contig called viral by at least two of the following were considered viral: VirSorter categories 3-6, VIBRANT low, CheckV low, Virfinder score between 0.7 and 0.9, and VirSorter2 score between 0.5 and 0.95. If a contig had no viral genes and more than one host gene according to CheckV, it was removed from the viral list. Bacterial genes were excised from the edges of proviruses based on their identification by CheckV, VirSorter2, and VIBRANT. Based on the community-established standard for defining viral populations (Roux et al., 2019), all viral contigs were clustered (stampede-clustergenomes) (Roux and Bolduc, 2017) if they shared average nucleotide identity (ANI) of 95% across 85% of the contig length. The clusters are henceforth referred to as "viral populations" and the longest representative sequence in each cluster was used for subsequent analyses.

## 2.5 Read Mapping

Filtered and trimmed reads were mapped to the viral populations using Bowtie2 (Langmead and Salzberg, 2012) and quantified using SAM Tools v1.11 (Li et al., 2009). Due to large differences in sequencing depth, the viral reads for each sample were downsampled to 10,000 reads for alpha diversity analyses using seqtk v1-3 (Li, 2018).

## 2.6 Taxonomy

The open reading frames (ORFs) from Distilled and Refined Annotation of Metabolism (DRAM) (Shaffer et al., 2020) were clustered with all full-length bacterial and archaeal virus genomes from NCBI's viral reference dataset ('Viral RefSeq', v.85, January 2018) using vConTACT2 (Bin Jang et al., 2019). Clusters with at least one taxonomic assignment from vConTACT2 were pulled for visualization with Cytoscape v3.8.2 (Shannon et al., 2003). The lowest common ancestor was identified for each vContact2 cluster with a taxonomic assignment.

## 2.7 Alpha and Beta Diversity

Only contigs with reads covering at least 3 kb of their length were included in diversity analyses. Alpha diversity measures (i.e., richness, Shannon Dissimilarity (henceforth, called "Shannon diversity")) were calculated using the vegan v2.5-7 package (Oksanen et al., 2020) in R v4.0.2

based on the downsampled reads. All analyses were performed in R v4.0.2 (R Core Team, 2021) using RStudio v1.3.595 (R Studio Team, 2020). When more than one sequencing run was performed for a given sample, the average alpha diversity value per sample for each metric was calculated.

Reads per kilobase million (RPKM) was determined for each viral population and used to calculate the Bray Curtis distance between samples in R using the vegan package (vegan v2.5-7) (Oksanen et al., 2020) and then Nonmetric Multidimensional Scaling (NMDS) ordination and hierarchical clustering were performed in R using hclust (method="Ward.D2"). Permutational Multivariate Analysis of Variance (PERMANOVA) using the adonis function (vegan) was used to test the effect of processing procedures, geography, and residual disinfectant use on the full community structure (Bray Curtis dissimilarities ~ continent * Köppen zone * filter pore size for biomass collection * DNA extraction method; with stratas for 'study'). The effect of 'study' and residual disinfectant were also tested in isolation (Bray Curtis dissimilarities ~ study; Bray Curtis dissimilarities ~ residual disinfectant).

To avoid the confounding effects of study design and climate on diversity analyses, only the Dai et al. (2020) samples collected in the temperate oceanic climate zone (Cfb Köppen climate classification scheme) were used to evaluate the effects of residual disinfectant use on taxonomic diversity and metabolic potential (Figure 1 inset). This study was chosen because it provided detailed water quality parameters and was the only study to include samples from distribution systems operated without a residual disinfectant. The six samples taken from the warm summer continental zone (Dfb Köppen climate classification) were excluded because they were all taken from distribution systems that use a residual disinfectant, complicating the isolation of climate's effect from that of residual disinfectant. Additionally, only residual disinfectant distribution systems that used free chlorine (and not chloramines) were included to facilitate comparison with the results presented by Dai et al. (2020). Differences in alpha diversity between distribution systems based on residual disinfectant use were tested using the Wilcoxon rank sum test. The envfit function (vegan) was used to determine the correlation strength between water quality parameters from Dai et al. (2020) (Table S2) and the two NMDS axes. PERMANOVA was used to test the effect on the structuring of the viral community of the Dai et al. (2020) samples of the parameters with a statistically significant correlation with the NMDS axes (Bray Curtis dissimilarities ~ chlorine + phosphate + temperature + pH + conductivity + total organic carbon (TOC) + dissolved oxygen (DO) + drinking water distribution system). Parameters were considered significantly correlated if p-value $< 0.05$.

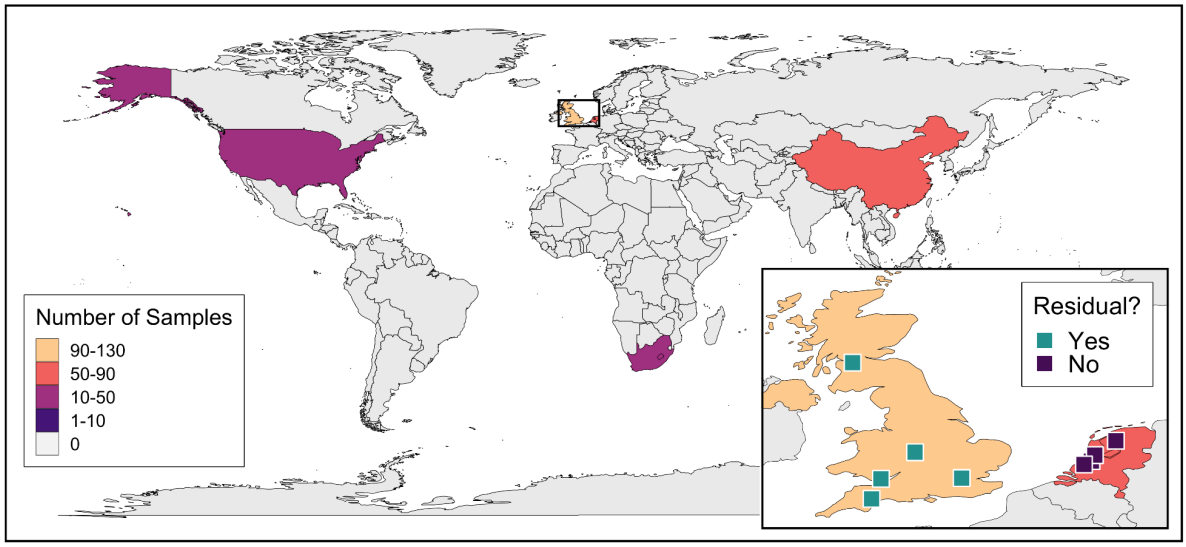



*Figure 1: Depiction of the number of drinking water distribution system samples per country included in this meta-analysis. The color of the country represents the sample number bin. Two samples taken from Singapore are not visible due to the map's scale. Countries with no drinking water metagenomic studies are in grey. Inset map illustrates the locations of the distribution system samples from the Dai et al. (2020) study that were used for subsequent analyses of the differences based on distribution system residual disinfectant use.*

## 2.8 Metabolic Potential Analyses

Viral contig gene functions were assigned using DRAM based on the KEGG and Pfam databases (Shaffer et al., 2020). Read coverages of the open reading frames predicted by DRAM were calculated using FeatureCounts (Liao et al., 2014) from the Subread package (Liao et al., 2013). Bray Curtis distances between samples were calculated using the vegan package in R and then NMDS ordination and hierarchical clustering using hclust (method="Ward.D2") were performed. The envfit function (vegan) was used to correlate the measured water quality parameters from the Dai et al. (2020) study with the two NMDS axes. The water quality parameters that were statistically significantly correlated with the two NMDS axes were used in a PERMANOVA model to evaluate the effect of water quality parameters on the abundance of KEGG orthologies (KOs), which are molecular function classes that represent the metabolic potential of the viromes. The following equation was used in the adonis function (vegan): Bray Curtis dissimilarities ~ chlorine + phosphate + temperature + pH + drinking water distributions system. DESeq2 (Love et al., 2014) was used to identify KOs identified in the Dai et al. (2020) samples that were differentially abundant between distribution systems based on whether or not they used a residual disinfectant. KOs were considered differentially abundant if their Benjamini-Hochberg adjusted p-value was less than 0.05. KEGG pathway modules were considered more abundant in a condition if the difference in normalized read counts was statistically significant between the two conditions (Welch's Two Sample t-test p-value $\leq 0.05$). Pathways were visualized using Pathview v1.30.1 in R (Luo and Brouwer, 2013).

# 3 Results/Discussion

To produce a snapshot of the virome in drinking water distribution systems, we mined publicly available drinking water metagenomes (Chao et al., 2013; Dai et al., 2020; Douterelo et al., 2018; Garner et al., 2018; Huang et al., 2014; Jia et al., 2015; Ma et al., 2019, 2017; Pinto et al., 2016; Potgieter et al., 2020; Sevillano et al., 2021; Zhang et al., 2019). The majority of the samples were collected in Europe (31%) and Asia (35%), with the remainder obtained from North America (23%) and Africa (11%) (Figure 1). In total, 166 samples were collected from seventy-nine distribution systems, with eighteen samples from systems with no residual disinfectant and 102 from systems with a residual disinfectant. Forty-six samples are presumed to be from distribution systems using a residual disinfectant based on the typical disinfection choice of their country of origin (i.e., China, Singapore, South Africa, United Kingdom, United States), but this could not be confirmed (Table S1).

All samples from the same drinking water distribution system were co-assembled (Tables S1 and S3), resulting in 729,724 contigs larger than 3 kb. We then mined the assembled contigs for viral sequences, identifying 106,461 viral sequences (or 15% of all the contigs longer than 3 kb). While this proportion is typically not reported in studies that have mined metagenomes for viral sequences, it is in line with the upper range (0.3-16%) of the percentage of viral sequences observed in a marine metagenome study (Delong et al., 2006). To facilitate comparisons of the viral populations across distribution systems, the viral contigs were clustered into 81,265 viral populations with an ANI of 95% over 85% of the contig (Roux et al., 2019). These viral populations allowed us to characterize the taxonomic diversity and metabolic potential of the drinking water virome, as well as the geographic and treatment-based factors shaping these communities.

## 3.1 Drinking water viral populations are predominantly novel and heterogeneous between samples

Of the 81,265 viral populations identified across all distribution systems, none were found in every sample. Eighty-nine high ubiquity viral populations were found in more than 90% of the samples (red samples in Figure 2B). Other studies have similarly found a high degree of heterogeneity within the viromes collected from similar environments (e.g., within ecological zones in marine environments) (Gregory et al., 2019; Paez-Espino et al., 2016).

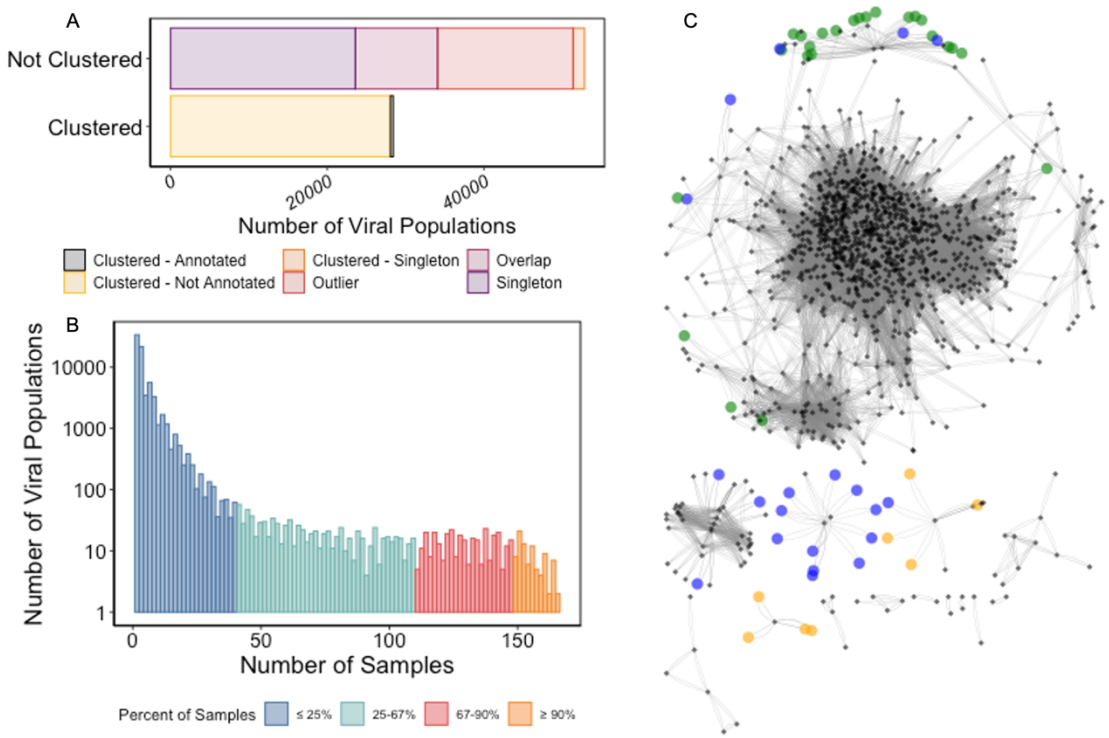



*Figure 2: (A) Status of drinking water viral populations after taxonomic clustering by vConTACT2. Viral populations that were given a taxonomic assignment by vConTACT2 are in black. (B) Histogram of the number of samples in which a specific viral population was detected. Bars are colored based on the proportion of samples in which the specific viral populations were detected. To visualize the low numbers of the high ubiquity viral populations, a log scale was used for the y-axis. (C) Network of clustered viral populations found in more than 25% of samples (teal, red, and orange in B), as well as connected viral genomes with taxonomic annotations from vConTACT2. Drinking water viral populations are black and viral genomes are colored based on their taxonomic assignment (family): Siphoviridae (green), Myoviridae (blue), and Podoviridae (light orange).*

Using vConTACT2 to examine the evolutionary relatedness of the viral populations, 9,874 clusters (at approximately the genus level) were formed from 28,426 (35%) of the viral populations (Figure S2); the remaining 52,839 viral populations were not clustered (Figure 2A). Out of the 28,426 viral populations clustered by vConTACT2, only 361 were given a definitive taxonomic assignment per the International Committee on Taxonomy of Viruses (ICTV) conventions (Figure S3). The remaining populations were given operational identifiers specific to this study. Similarly to other environments (Parmar et al., 2017), 347 viral populations were assigned to the order *Caudovirales*. *Siphoviridae*, *Myoviridae*, and *Podoviridae* were the three most prevalent families of viruses identified in these samples, mirroring the dominant families in many other freshwater environments (Moon et al., 2020; Saparbaevna et al., 2017; Skvortsov et al., 2016; Zaouri et al., 2020). Of the most ubiquitous viral populations (i.e., the 1,242 viral

populations found in at least a quarter of the samples), most did not have a taxonomic assignment (Figure 2C).

To date, most viral populations described in environmental viromics studies have been novel viruses, reflecting the incomplete nature of the viral databases, as well as the fragmented viral genomes included in these datasets. For instance, in recent studies of the Arctic Ocean (Gregory et al., 2019), honey bees (Deboutte et al., 2020), and New York City wastewater (Gulino et al., 2020), vConTACT2 only assigned a taxonomy to 9.8%, 9.5%, and 4% of their viral populations, respectively. Since many viruses are strongly tied to specific habitats (Paez-Espino et al., 2016; Roux et al., 2012), lack of representation of a habitat in existing viral databases increases the likelihood that novel viruses will be described when sequences from such an undersampled environment are analyzed. It is therefore unsurprising that only 2% of the viral populations in the drinking water distribution systems samples analyzed in our study had a taxonomic assignment. In addition to the lack of drinking water virome studies to date, groundwater and surface water, which are typically used as source waters for drinking water production, also have been less frequently studied than marine aquatic environments (Kallies et al., 2019; Kothari et al., 2021; Malki et al., 2020; Moon et al., 2020; Paez-Espino et al., 2016; Roux et al., 2012; Skvortsov et al., 2016).

## 3.2 Environmental conditions and experimental design influence drinking water viral communities

To evaluate the possible effects of experimental design and environmental conditions known for each sample, we tested the effect of the continent where samples were collected, the Köppen climate zone, the DNA extraction method, and the filter pore size for biomass collection. While our meta-analysis did not find filter pore size to impact the structure of viral communities (PERMANOVA; Table S4), other studies have observed differences between viruses sequenced from 0.1-20 µm and <0.1 µm fractions (Williamson et al., 2012). As viruses typically range from 0.02-0.3 µm, the relatively large filter pore size (0.22 µm and 0.45 µm) used across all studies in our analysis may explain why filter pore size did not affect viral community structure. The other three parameters tested, as well as the interaction of the continent and Köppen climate zone, did have a statistically significant effect on the viral community's structure (PERMANOVA; Table S4). Combined, the parameters and their interactions tested explain slightly over a quarter of the variability in the system (residual PERMANOVA $R^2$=0.71; Table S4). This could reflect the unbalanced nature of the metadata (e.g., many more samples used the Maxwell 16 DNA extraction system than the other three methods; Figure S4), as well as the importance of additional factors that may affect viral community structure (e.g., distribution system water quality).

We additionally included 'study' as a test parameter to evaluate the possible effects of parameters beyond simply those explicitly reported in the original metagenomics studies. The effect of 'study' on the viral community (PERMANOVA $R^2$=0.22, p-value=0.0001; Figure 3A) likely encompasses parameters tested above (e.g., DNA extraction kit, continent), as well as other parameters not reported (e.g., water source, cDNA library preparation). In a previous meta-analysis of drinking water bacterial community structure, 'study'-specific effects explained nearly a third of the variability among samples (Bautista-de los Santos et al., 2016). The collection and processing of samples are known to have a substantial impact on the viral populations identified (Brinkman et al., 2018; Dias et al., 2020; Hurwitz et al., 2013; Percival and Wyn-Jones, 2013). The low biomass of drinking water samples may make them particularly prone to influence by 'study'-specific differences resulting from contamination during sample processing (Bautista-de los Santos et al., 2016). Alternatively, as most studies only examined one distribution system or focused on a single region, the strong 'study'-specific effects may be due to other factors, such as differences in the communities of the source waters based on geography, climate, and source (e.g., groundwater, surface water type). Such parameters have been shown to influence the viral communities of other diverse environments (Paez-Espino et al., 2016).

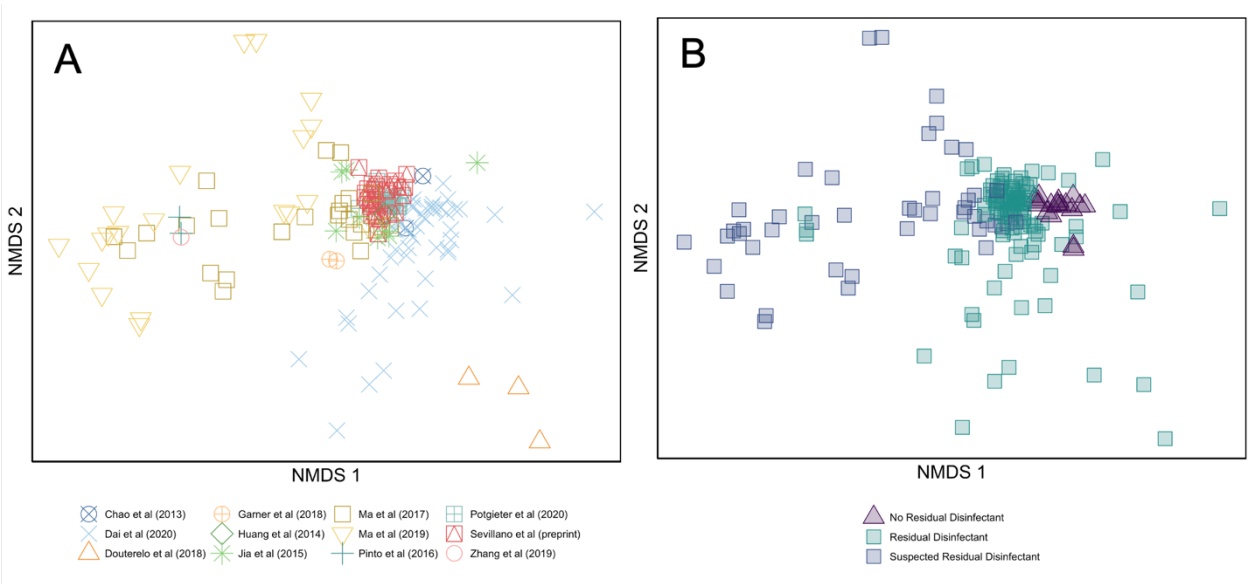



*Figure 3: Overview of the viral community structure. NMDS ordination (stress = 0.15) with coloring based on (A) 'study' and (B) residual disinfectant. The plots visualize the rank-based ordination of the samples' Bray Curtis dissimilarities of the viral abundances.*

Water quality is affected by parameters such as the type of source water (e.g., surface water versus groundwater) and residual disinfectant use, which impact the drinking water bacterial community in turn (Bautista-de los Santos et al., 2016; Dai et al., 2020). As bacterial community structure is a critical driver of bacteriophage community structure (Gregory et al., 2019; Jonge et al., 2018), we hypothesized that residual disinfectant use would affect viral populations. Some separation based on residual disinfectant can be seen in the NMDS plot of all the samples

(PERMANOVA $R^2$ = 0.04, p-value = 0.001; Figure 3B). To evaluate this further, we focused on the metagenomes provided in the study by Dai et al. (2020) to control for the strong effects of 'study' described above; using this approach, we observed a more defined separation based on residual disinfectant (PERMANOVA $R^2$=0.13, p-value=0.001; Figure 4A).

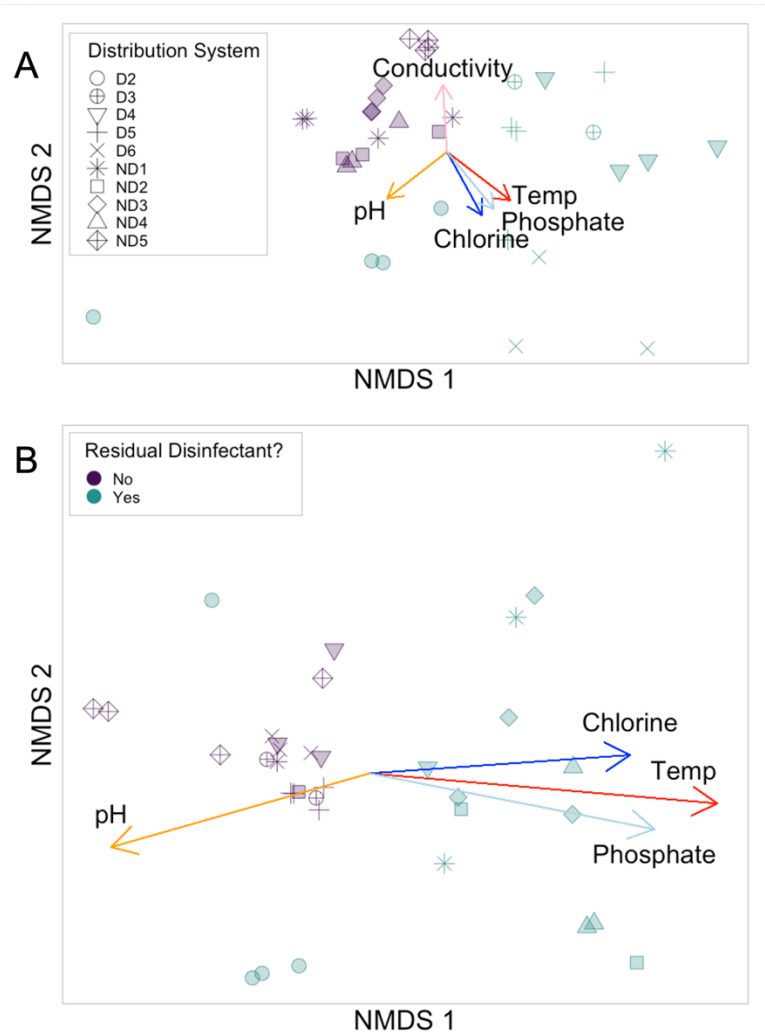



*Figure 4: Diversity of viral populations and metabolic potential cluster based on residual disinfectant use and water quality. NMDS ordination of the Bray Curtis dissimilarities of the (A) viral community structure (stress = 0.14) and (B) metabolic potential (stress = 0.16). The shape of the points represents the distribution system, and the color represents the residual disinfectant. The NMDS ordination is overlaid with the gradient of the fit between significantly correlated environmental parameters (p-value ≤ 0.001) and the NMDS axes. Environmental features that were not highly correlated are not shown. Note that only samples from the Cfb Köppen climate zone and samples with chlorine as the residual disinfectant from the Dai et al. (2020) study were included.*

Beta diversity measures (Bray Curtis dissimilarities) suggested that differences in water quality have a strong effect on the structuring of drinking water viromes (Figure 4A). These include both water quality parameters that were statistically different (i.e., water temperature, chlorine and

phosphate concentrations, DO, and conductivity) and those that were not statistically different (i.e., pH and TOC) between the distribution systems that used a residual disinfectant and those systems that did not (Figure 4A; Table 1; Table S5). Chlorine concentration, DO, and conductivity were each also correlated with the nucleotide signatures of the full microbial communities, a proxy for community-relatedness (Dai et al., 2020), illustrating the importance of these parameters on both the cellular and viral communities of drinking water. As drinking water distribution system has a significant correlation with the viral community structure (PERMANOVA $R^2$=0.30, p-value=0.02; Table S5) and nearly a third of the variability is not explained by any of the tested parameters (residual PERMANOVA $R^2$=0.35; Table S5), other differences in water quality and treatment not measured also likely affected the drinking water viral community structure.

*Table 1 - Welch's Two Sample t-test p-values testing the difference in water quality parameters with residual disinfectant use and correlation of the environmental data from the Dai et al. (2020) study with the Bray-Curtis dissimilarities of the viral abundance data (taxonomic diversity) and with the Bray-Curtis dissimilarities of the KEGG abundances (metabolic potential). ** : p-value < 0.005; * : p-value < 0.05.*

| | | **taxonomic diversity** | | | | **metabolic potential** | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **p-value** | **NMDS1** | **NMDS2** | **$R^2$** | **pr(>R)** | **NMDS1** | **NMDS2** | **$R^2$** | **pr(>R)** |
| **Temp** | 0.000003 ** | 0.794 | -0.608 | 0.479 | 0.001 ** | 0.996 | -0.086 | 0.528 | 0.001 ** |
| **pH** | 0.131 | -0.789 | -0.614 | 0.433 | 0.002 ** | -0.962 | -0.274 | 0.313 | 0.003 ** |
| **Conductivity** | 0.00254 ** | -0.063 | 0.998 | 0.341 | 0.002 ** | -0.383 | 0.924 | 0.074 | 0.304 |
| **DO** | 0.00896 * | 0.067 | -0.998 | 0.189 | 0.032 * | 0.517 | -0.856 | 0.173 | 0.061 |
| **Chlorine** | 0.00000003 ** | 0.482 | -0.876 | 0.393 | 0.001 ** | 0.998 | 0.070 | 0.296 | 0.005 ** |
| **Phosphate** | 0.0000001 ** | 0.635 | -0.772 | 0.413 | 0.001 ** | 0.981 | -0.192 | 0.365 | 0.001 ** |
| **TOC** | 0.0553 | -0.368 | 0.930 | 0.217 | 0.019 * | -0.985 | 0.172 | 0.060 | 0.371 |
| **Ammonia** | 0.0132 * | -0.217 | -0.976 | 0.042 | 0.496 | -0.053 | -0.999 | 0.009 | 0.847 |
| **Nitrate** | 0.276 | 0.677 | 0.736 | 0.135 | 0.108 | 0.989 | -0.149 | 0.159 | 0.064 |

The structuring of the viral communities between distribution systems that use a residual disinfectant and those that do not varies with both water quality differences that reflect treatment (e.g., chlorine, phosphate, DO, ammonia), as well as season (e.g., temperature) and water source (e.g., temperature and conductivity). The lower water temperatures (Table 1) of the systems without a residual disinfectant may reflect either water source or season. As the systems using no residual disinfectant generally had higher conductivities (suggesting a groundwater source), the season of sampling may not have influenced the microbial community of these systems. While temporal trends have been shown in some drinking water distribution systems that primarily used surface water (Pinto et al., 2014; Potgieter et al., 2018; Prest et al., 2016), systems that use a

groundwater source exclusively have not exhibited differences in microbial community (Assche et al., 2019; Roeselers et al., 2015). Further studies are necessary to determine the distinct effects of season, water source, and residual disinfectant use on the viral communities of drinking water distribution systems.

## 3.3 Systems with residual disinfectant have less diverse and even viral communities

Out of the 81,265 viral populations, 13,562 remained after downsampling to achieve even sampling across all datasets for calculation of richness (number of observed viral populations per sample) and Shannon diversity (a measure of the uniformity of abundance, or evenness, within a sample). The average richness was 1,248 ± 589 viral populations and the average Shannon diversity was 5.10 ± 1.10. The average Shannon diversity observed is consistent with ranges reported for other freshwater sources (Aguirre De Cárcer et al., 2015; Tseng et al., 2013).

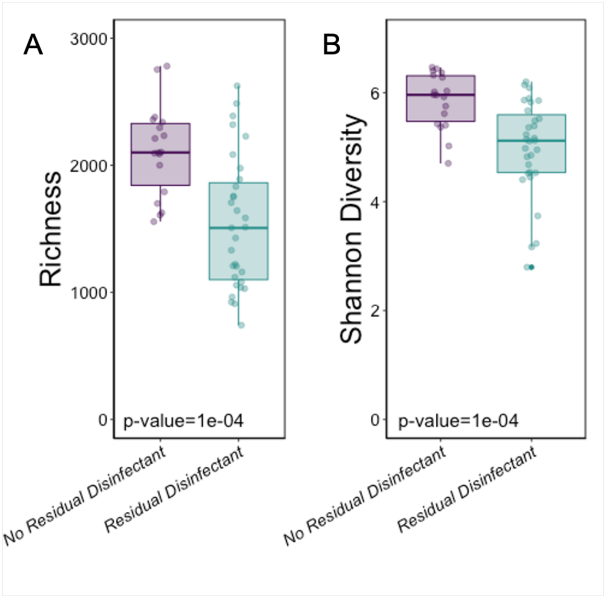



*Figure 5: Viral diversity varies between samples from distribution systems that use or do not use a residual disinfectant. Note that only samples from the Cfb Köppen climate zone and samples with chlorine as the residual disinfectant from the Dai et al. (2020) study were included. Panels are boxplots overlaid with points representing each sample's (A) observed viral richness and (B) Shannon diversity.*

Distribution systems with free chlorine as the residual disinfectant had less diverse and less even viral communities than distribution systems without a residual disinfectant (observed taxa Wilcoxon rank sum test p-value = 0.0003, Shannon diversity p-value = 0.0001; Figure 5). While there was no difference in richness between the *known* residual disinfectant samples and the *suspected* residual disinfectant systems (Wilcoxon rank sum test p-value = 0.4545), there was a strong difference in the number of identified taxa between the no residual disinfectant samples

and both the *suspected* and *known* residual disinfectant samples (Wilcoxon rank sum test p-value = $9.7*10^{-10}$) (Figure S5). This supported our prediction (based on country of origin) that samples from studies with presumed residual disinfectant usage were indeed likely obtained from systems operated with a residual disinfectant.

The lower richness and evenness of the viral communities in distribution systems using a residual disinfectant mirrors the pattern observed for bacterial communities in drinking water (Bautista-de los Santos et al., 2016; Dai et al., 2020). The disinfectant residuals may exert a stronger selective pressure than the lower phosphate and ammonia availability observed in distribution systems that do not use a residual disinfectant (Dai et al., 2020). Correlations between bacterial and viral diversity has been observed in in a range of environments (Anderson et al., 2017; Dalcin Martins et al., 2018; Yang et al., 2019). However, more research is necessary to determine whether this is because viruses depend on their hosts to replicate, or because high bacterial diversity requires phage predation, as has been observed in some systems (Rodriguez-Valera et al., 2009; Weinbauer et al., 2007).

## 3.4 The metabolic potential of drinking water viromes cluster based on residual disinfectant use

Viral genomes often include genes necessary for replication, as well as host-derived auxiliary metabolic genes (AMGs) thought to reprogram host cellular metabolism to support infection processes (Warwick-Dugdale et al., 2019). Characterizing the genes carried by drinking water viruses is, therefore, an essential step to understanding their influence on the cellular members of the drinking water microbiome. As is common in other environmental viromes (Deboutte et al., 2020; Gregory et al., 2019), most of the predicted viral genes identified in this study were not annotated with a metabolic function, i.e., they are novel and do not share homology with known genes in public databases. Specifically, only 13% have a KEGG orthology (KO) molecular function assignment. Of the reads that mapped to viral genes with KO annotations, 51.2% mapped to genes involved in carbohydrate metabolism (including central carbon metabolism). The next most common function based on read assignment was amino acid metabolism (13.2%), followed by metabolism of cofactors and vitamins (9.3%), unclassified metabolism (8.3%), and nucleotide metabolism (6.1%).

While the predominance of carbohydrate metabolism annotations in this study may in part represent KEGG database biases (carbohydrate metabolism related proteins represent 33.7% of the KEGG database), this observation mimics the metabolic pattern found in other viromes (Anderson et al., 2017; Gregory et al., 2019; Sullivan et al., 2005; Williamson et al., 2012) and the importance of these metabolic processes to bacterial hosts (Dinsdale et al., 2008). Carbohydrate metabolism genes carried by viruses are involved in reprogramming central carbon metabolism to favor energy production necessary for viral replication (Howard-Varona et al.,

2020). In the drinking water microbiome, these genes might contribute to the survival of the bacteriophages' hosts in the oligotrophic environment of drinking water. Genes related to both nucleotide and amino acid metabolism are commonly found in viral genomes due to their importance in viral replication (Hurwitz et al., 2015).

As with viral assemblage beta diversity, samples clustered primarily by disinfectant status when differences in metabolic potential (according to KO abundances) were considered (Figure 4B). In line with observed trends in viral assemblage beta diversity, chlorine and phosphate concentration, temperature, and pH were correlated with the metabolic potential profiles of samples from systems using residual disinfectant (Figure 4B and Table 1; PERMANOVA Table S6). Similarly, much of the variability was likely due to other water quality and treatment specific differences, due to the strong correlation with drinking water distribution system (PERMANOVA $R^2$=0.39, p-value=0.001) and large degree of unexplained variability (residual PERMANOVA $R^2$=0.28). These differences in metabolic potential were not driven by differences in the number (Figure S6) or presence of specific (Figure S7) KOs. This contrasts with the viral populations, which were more diverse in distribution systems with a residual disinfectant.

## 3.5 Differences in viral metabolic potential parallel contrasting bacterial survival mechanisms based on residual disinfectant use

The distribution of viral encoded metabolic genes across drinking water distribution systems with and without disinfectant residuals offers insights into how viruses in drinking water may impact host metabolisms in a habitat-specific manner, either through AMG-driven reprogramming of host metabolism during infection or by expanding the genomic repertoire of their hosts via horizontal gene transfer. Overall, there were 1,695 differentially abundant molecular functions (identified as KOs; Table S7), with twice as many identified from distribution systems that use free chlorine as a residual disinfectant than from those that do not (1,158 versus 537 KOs, respectively). Of the differentially abundant viral gene orthologs with KO assignments, 406 were assigned to KEGG metabolism pathway modules. Differences in pathway module abundance based on residual disinfectant (Table S8) point to differing evolutionary pressures in drinking water distribution systems based on residual disinfectant use, in particular oxidative stress (residual disinfectant systems) and very low nutrient availability (no residual disinfectant).

To survive in systems with a residual disinfectant, microorganisms must mitigate the effects of reactive chlorine species (Gray et al., 2013). Genes essential for this include chaperones, particularly the redox-regulated protein Hsp33 (K04083) (Wholey and Jakob, 2012), which prevent the aggregation of unfolded proteins. Other relevant genes include those responsible for the biosynthesis of low molecular weight thiols, including glutathione and coenzyme A, which

maintain the reducing potential of the cytosol (Chesney and Eaton, 1996). Glutathione-related genes occur often in bacterial genomes from drinking water distribution systems treated with chlorine (Dai et al., 2020; Douterelo et al., 2018; Gomez-Alvarez et al., 2012). Similarly, glutathione synthase (GSS, K01920, K21456; Figure S8) and genes responsible for four of the five steps of coenzyme A biosynthesis (M00120; Figure S9) were all overabundant in the viruses of distribution systems that use a residual disinfectant compared to distribution systems that do not use a residual disinfectant (Figure 6A; Table S9). In other environments, genes linked to surviving oxidative stress more broadly have been found in phage genomes (Huang et al., 2021; Kavagutti et al., 2019) and have been shown to be beneficial (Holmgren, 1989; Russel and Model, 1986). While viral genes are implicated in combating oxidative stress in distributions systems using a residual disinfectant, additional research is needed to assess the role these genes take in mitigating the harmful effects of reactive chlorine species, specifically.

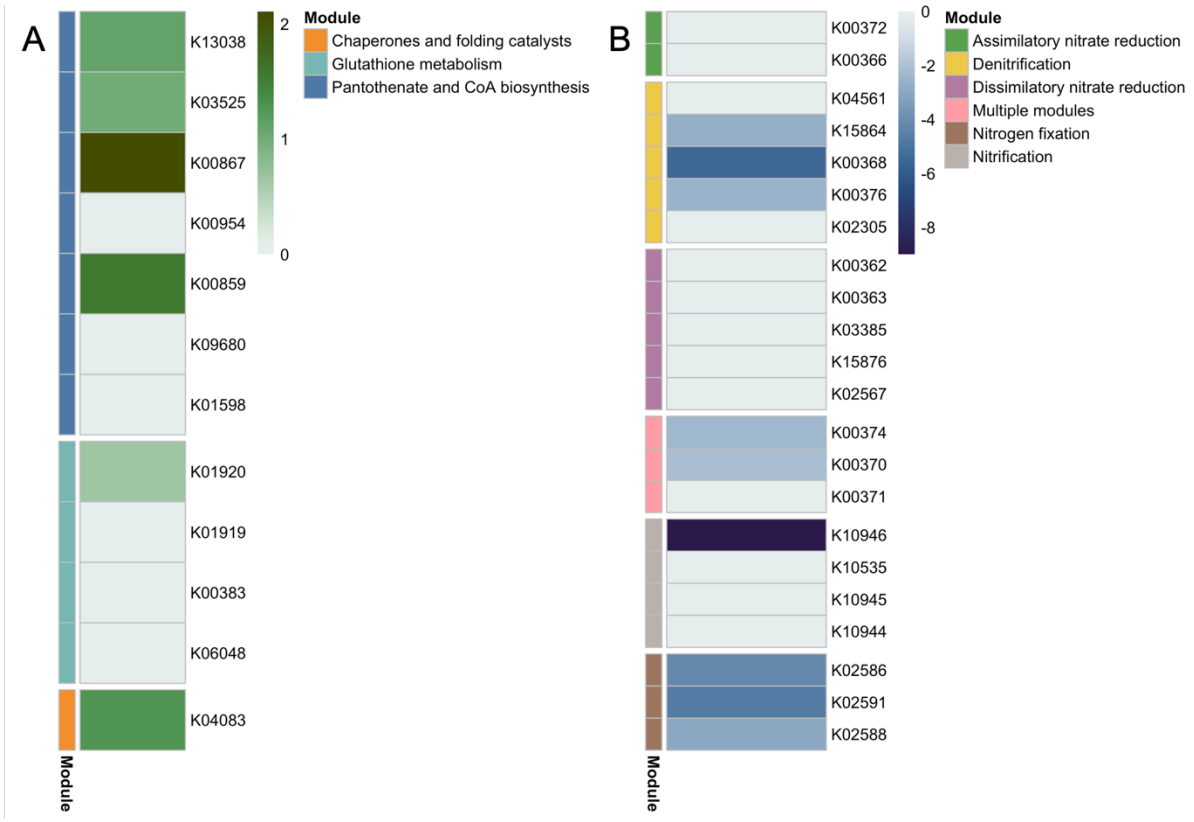



*Figure 6 – Heatmaps of the abundance ($log_2$Fold change) between the samples from distribution systems that use or do not a residual disinfectant. (A) In drinking water distribution systems with a residual disinfectant, KOs related to surviving oxidative stress, such as caused by chlorination, are overabundant. (B) Alternatively, genes related to nitrogen metabolism are overabundant in drinking water distribution systems with no residual disinfectant. Colored bars alongside each heatmap indicate KEGG module membership.*

In distribution systems that do not use a residual disinfectant, limiting nutrient levels is an essential strategy to reduce bacterial growth (Kooij and Wielen, 2013). In the Dai et al. (2020)

samples analyzed here, though levels of nitrate were not statistically different based on residual disinfectant use, levels of ammonia were (Table 1). As such, we hypothesized that KOs related to nitrogen limitation (Figure S10) would be correlated with residual disinfectant use. Considering the six KEGG modules of nitrogen metabolism, all of the differentially abundant KOs ($p_{adj} \leq 0.05$) were more abundant in the distribution systems without a residual disinfectant (Figure 6B; Table S10). Specifically considering nitrogen fixation, which was overabundant in the full microbial communities of distribution systems with no residual disinfectant (Dai et al., 2020), all three KOs found in the virome were also overabundant in the no systems without residual disinfectant (Figure 6B; Table S10).). In other nitrogen-limited environments, phages have been observed to carry more genes related to nitrogen cycling (Warwick-Dugdale et al., 2019). These trends in viral encoded nitrogen-related genes suggest that in drinking water distribution systems that limit nutrients to control bacterial growth rather than use a residual disinfectant, viruses may impact nutrient metabolism, either through rewiring host metabolism to support infection or by mediating horizontal gene transfer of nutrient-stress related genes (Warwick-Dugdale et al., 2019; Zimmerman et al., 2020).

## 3.6 Opportunities for Future Work

Metagenomic analyses allow researchers to make broad assessments of microbial communities and suggest areas for more targeted studies. The analyses conducted here suggest that residual disinfectant use impacts the diversity and metabolic potential of the drinking water virome. However, future bench-scale experiments will be necessary to separate the effects of residual disinfectant from other water quality and sampling differences between distribution systems observed in this study. Confounding methodological differences between studies, such as DNA extraction method, can mask or be correlated with the factors, including climate, geography, and treatment regime, that may shape the microbial communities of drinking water. While standardizing methodologies may be impossible and premature, coordination between researchers could still facilitate cross-study comparisons (Bautista-de los Santos et al., 2016; Hull et al., 2019). An excellent first step would be to make raw sequencing data public and report temperature, water quality characteristics (e.g., chlorine, phosphate, ammonia, nitrate, TOC concentrations), and biomass proxies (e.g., ATP levels, cell counts), as well as other metadata (e.g., source water characteristics, details of the treatment, season, location) (Bautista-de los Santos et al., 2016). Application of standardized spike-ins, such as “sequins” (Hardwick et al., 2018), to drinking water microbiomes would further facilitate between study comparisons, as well as reveal whether the changes in viral gene relative abundances described in this work mirror the trends in absolute abundance.

For the most robust assessments of the drinking water virome, future studies must specifically develop sample processing methods that enrich for viruses. As the drinking water samples of this meta-analysis were not collected with the intention of enriching for submicron viral particles, the

sequenced viruses may be biased towards particle-associated viruses, viruses actively infecting a cellular host, and free viruses trapped by the chosen filter. Future studies designed specifically to query the viral fraction of drinking water (including both DNA and RNA viruses) are necessary to confirm whether the patterns described here are reflective of the complete DNA virus community and provide additional insights into the free viruses of drinking water distribution systems. These studies will also determine the extent to which phage facilitated metabolic reprogramming or gene transfer is contributing to bacterial survival in drinking water systems. By connecting the drinking water virome's behavior with that of the bacterial community, improved treatment processes may become possible.

# 4 Conclusions

- Mining drinking water metagenomes for viral sequences reveals differences in the viral community based on geography and water quality parameters.
- More diverse and even viral communities are found in distribution systems that do not use a residual disinfectant.
- Viruses from distribution systems that use a residual disinfectant contain more genes for secondary metabolic processes, like glutathione synthase, that may help counter the effects of chlorine use.
- Viruses from distribution systems that do not use a residual disinfectant are enriched for nitrogen metabolism genes that may confer an evolutionary advantage during replication in the nitrogen depleted environment of distribution systems without a residual disinfectant.

# Acknowledgements

We would like to thank the Remaking Water Infrastructure Blue Sky team, particularly Sarah Potgieter, Christopher Anderson, and Katherine Dowdell, as well as the Duhaime and Wigginton labs for contributing their technical and intellectual expertise for bioinformatic experimental design and analyses and manuscript preparation. This work was supported by the Blue Sky Initiative of the University of Michigan College of Engineering and NSF CAREER funding to Ameet Pinto [NSF CBET 1749530].

# Citations

Aguirre De Cárcer, D., López-Bueno, A., Pearce, D.A., Alcamí, A., 2015. Biodiversity and distribution of polar freshwater DNA viruses. Sci. Adv.

Albinana-Gimenez, N., Clemente-Casares, P., Bofill-Mas, S., Hundesa, A., Ribas, F., Girones, R., 2006. Distribution of Human Polyomaviruses, Adenoviruses, and Hepatitis E Virus in the Environment and in a Drinking-Water Treatment Plant. Environ. Sci. Technol. 40, 7416–7422. https://doi.org/10.1021/es060343i

Alneberg, J., Bjarnason, B.S., Bruijn, I. De, Schirmer, M., Quick, J., Ijaz, U.Z., Lahti, L., Loman,

N.J., Andersson, A.F., Quince, C., 2014. Binning metagenomic contigs by coverage and composition 11. https://doi.org/10.1038/nmeth.3103

Anderson, C.L., Sullivan, M.B., Fernando, S.C., 2017. Dietary energy drives the dynamic response of bovine rumen viral communities. Microbiome 5, 155. https://doi.org/10.1186/s40168-017-0374-3

Assche, A. Van, Lievens, B., Crauwels, S., Brabanter, J. De, Willems, K.A., 2019. Characterization of the bacterial community composition in water of drinking water production and distribution systems in 1–11. https://doi.org/10.1002/mbo3.726

Bautista-de los Santos, Q.M., Schroeder, J.L., Sevillano-Rivera, M.C., Sungthong, R., Ijaz, U.Z., Sloan, W., T, Pinto, A.J., 2016. Emerging investigators series: microbial communities in full-scale drinking water distribution systems – a meta-analysis. Environ. Sci. Water Res 2, 631–644. https://doi.org/10.1039/c6ew00030d

Bin Jang, H., 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

Brinkman, N.E., Villegas, E.N., Garland, J.L., Keely, S.P., 2018. Reducing inherent biases introduced during DNA viral metagenome analyses of municipal wastewater. PLoS One 13, 1–23. https://doi.org/10.1371/journal.pone.0195350

Chao, Y., Ma, L., Yang, Y., Ju, F., Zhang, X.X., Wu, W.M., Zhang, T., 2013. Metagenomic analysis reveals significant changes of microbial compositions and protective functions during drinking water treatment. Sci. Rep. 3, 1–9. https://doi.org/10.1038/srep03550

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. fastp : an ultra-fast all-in-one FASTQ preprocessor 884–890. https://doi.org/10.1093/bioinformatics/bty560

Chesney, J.A., Eaton, J.W., 1996. Bacterial Glutathione : a Sacrificial Defense against Chlorine Compounds. J. Bacteriol. 178, 2131–2135.

Dai, Z., Sevillano-rivera, M.C., Calus, S.T., Santos, Q.M.B.L., Eren, A.M., Wielen, P.W.J.J. Van Der, Ijaz, U.Z., Pinto, A.J., 2020. Disinfection exhibits systematic impacts on the drinking water microbiome. Microbiome 1–19.

Dalcin Martins, P., Danczak, R.E., Roux, S., Frank, J., Borton, M.A., Wolfe, R.A., Burris, M.N., Wilkins, M.J., 2018. Viral and metabolic controls on high rates of microbial sulfur and carbon cycling in wetland ecosystems. Microbiome 6, 1–17. https://doi.org/10.1186/s40168-018-0522-4

Deboutte, W., Beller, L., Kwe, C., Maes, P., Graaf, D.C. De, 2020. Honey-bee – associated prokaryotic viral communities reveal wide viral diversity and a profound metabolic coding potential 117, 10511–10519. https://doi.org/10.1073/pnas.1921859117

Delong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J., Frigaard, N., Martinez, A., Sullivan, M.B., Edwards, R., Brito, B.R., Chisholm, S.W., Karl, D.M., 2006. Community Genomics Among Stratified Microbial Assemblages in the Ocean's Interior 311, 496–504.

Dias, R.S., Abe, A.E., Lima, H.S., Silva, L.C.F., de Paula, S.O., da Silva, C.C., 2020. Viral concentration methods for diversity studies in soil samples. Appl. Soil Ecol. 155, 103666. https://doi.org/10.1016/j.apsoil.2020.103666

Dinsdale, E.A., Edwards, R.A., Hall, D., Angly, F., Breitbart, M., Brulc, J.M., Furlan, M., Desnues, C., Haynes, M., Li, L., Mcdaniel, L., Moran, M.A., Nelson, K.E., Nilsson, C., Olson, R., Paul, J., Brito, B.R., Ruan, Y., Swan, B.K., Stevens, R., Valentine, D.L., Thurber, R.V., Wegley, L., White, B.A., Rohwer, F., 2008. Functional metagenomic

profiling of nine biomes 452. https://doi.org/10.1038/nature06810
Dong, Y., Kim, J., Lewis, G.D., 2010. Evaluation of methodology for detection of human adenoviruses in wastewater , drinking water , stream water and recreational waters. J. Appl. Microbiol. 108, 800–809. https://doi.org/10.1111/j.1365-2672.2009.04477.x
Douterelo, I., Calero-Preciado, C., Soria-Carrasco, V., Boxall, J.B., 2018. Whole metagenome sequencing of chlorinated drinking water distribution systems. Environ. Sci. Water Res. Technol. 4, 2080–2091. https://doi.org/10.1039/c8ew00395e
Eren, A.M., Kiefl, E., Shaiber, A., Veseli, I., Miller, S.E., Schechter, M.S., Fink, I., Pan, J.N., Yousef, M., Fogarty, E.C., Trigodet, F., Watson, A.R., Esen, Ö.C., Moore, R.M., Clayssen, Q., Lee, M.D., Kivenson, V., Graham, E.D., Merrill, B.D., Karkman, A., Blankenberg, D., Eppley, J.M., Sjödin, A., Scott, J.J., Vázquez-campos, X., Mckay, L.J., Mcdaniel, E.A., Stevens, S.L.R., Anderson, R.E., Fuessel, J., Fernandez-guerra, A., Maignien, L., Delmont, T.O., Willis, A.D., 2021. Community-led, integrated, reproducible multi-omics with anvi'o. Nat. Microbiol. 6, 3–6. https://doi.org/10.1038/s41564-020-00834-3
Feazel, L.M., Baumgartner, L.K., Peterson, K.L., Frank, D.N., Harris, J.K., Pace, N.R., 2009. Opportunistic pathogens enriched in showerhead biofilms. Proc. Natl. Acad. Sci. 106, 16393–16399. https://doi.org/10.1073/pnas.0908446106
Fuhrman, J.A., Noble, R.T., 1995. Viruses and protists cause similar bacterial mortality in coastal seawater 40, 1236–1242.
Gall, A.M., Mari, B.J., Lu, Y., Shisler, J.L., 2015. Waterborne Viruses : A Barrier to Safe Drinking Water. PLOS Pathog. 1–7. https://doi.org/10.1371/journal.ppat.1004867
Garner, E., Chen, C., Xia, K., Bowers, J., Engelthaler, D.M., Mclain, J., Edwards, M.A., Pruden, A., 2018. Metagenomic Characterization of Antibiotic Resistance Genes in Full- Scale Reclaimed Water Distribution Systems and Corresponding Potable Systems. Environ. Sci. Technol. 52, 6113–6125. https://doi.org/10.1021/acs.est.7b05419
Gomez-Alvarez, V., Revetta, R.P., Domingo, J.W.S., 2012. Metagenomic Analyses of Drinking Water Receiving Different Disinfection Treatments. AEM 78, 6095–6102. https://doi.org/10.1128/AEM.01018-12
Gray, M.J., Wholey, W.-Y., Jakob, U., 2013. Bacterial Responses to Reactive Chlorine Species. Annu. Rev. Microbiol. 141–160. https://doi.org/10.1146/annurev-micro-102912-142520.Bacterial
Gregory, A.C., Zayed, A.A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., Ardyna, M., Arkhipova, K., Carmichael, M., Cruaud, C., Dimier, C., Domínguez-Huerta, G., Ferland, J., Kandels, S., Liu, Y., Marec, C., Pesant, S., Picheral, M., Pisarev, S., Poulain, J., Tremblay, J.É., Vik, D., Acinas, S.G., Babin, M., Bork, P., Boss, E., Bowler, C., Cochrane, G., de Vargas, C., Follows, M., Gorsky, G., Grimsley, N., Guidi, L., Hingamp, P., Iudicone, D., Jaillon, O., Kandels-Lewis, S., Karp-Boss, L., Karsenti, E., Not, F., Ogata, H., Poulton, N., Raes, J., Sardet, C., Speich, S., Stemmann, L., Sullivan, M.B., Sunagawa, S., Wincker, P., Culley, A.I., Dutilh, B.E., Roux, S., 2019. Marine DNA Viral Macro- and Microdiversity from Pole to Pole. Cell 177, 1109-1123.e14. https://doi.org/10.1016/j.cell.2019.03.040
Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., Darzi, Y., Audic, S., Berline, L., Brum, J.R., Coelho, L.P., Cesar, J., Espinoza, I., Malviya, S., Sunagawa, S., Dimier, C., Kandels-lewis, S., Picheral, M., Poulain, J., 2016. Plankton networks driving carbon export in the oligotrophic ocean. Nature. https://doi.org/10.1038/nature16942
Gulino, K., Rahman, J., Badri, M., Morton, J., Bonneau, R., Ghedin, E., 2020. Initial Mapping of

the New York City Wastewater Virome. mSystems 5, 1–18. https://doi.org/10.1128/msystems.00876-19

Guo, J., Bolduc, B., Zayed, A.A., Varsani, A., Dominguez-huerta, G., Delmont, T.O., Pratama, A.A., Gazitúa, M.C., Vik, D., Sullivan, M.B., Roux, S., 2021. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses 1–13.

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST : quality assessment tool for genome assemblies. Bioinformatics 29, 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Hardwick, S.A., Chen, W.Y., Wong, T., Kanakamedala, B.S., Deveson, I.W., Ongley, S.E., Santini, N.S., Marcellin, E., Smith, M.A., Nielsen, L.K., Lovelock, C.E., Neilan, B.A., Mercer, T.R., 2018. Synthetic microbe communities provide internal reference standards for metagenome sequencing and analysis. Nat. Commun. 9, 1–10. https://doi.org/10.1038/s41467-018-05555-0

Hill, V., Polaczyk, A., Hahn, D., Narayanan, J., Cromeans, T., Roberts, J., Amburgey, J., 2005. Development of a Rapid Method for Simultaneous Recovery of Diverse Microbes in Drinking Water by Ultrafiltration with Sodium Polyphosphate and Surfactants. Appl. Environ. Microbiol. 21, 6878–6884. https://doi.org/10.1128/AEM.71.11.6878

Holmgren, A., 1989. Electron Transport to Reductive Enzymes 264, 13963–13966.

Howard-Varona, C., Lindback, M.M., Bastien, G.E., Solonenko, N., Zayed, A.A., Jang, H., Andreopoulos, B., Brewer, H.M., Rio, T.G. del, Adkins, J.N., Paul, S., Sullivan, M.B., Duhaime, M.B., 2020. Phage-specific metabolic reprogramming of virocells. ISME J. 14, 881–895. https://doi.org/10.1038/s41396-019-0580-z

Huang, K., Zhang, X.X., Shi, P., Wu, B., Ren, H., 2014. A comprehensive insight into bacterial virulence in drinking water using 454 pyrosequencing and Illumina high-throughput sequencing. Ecotoxicol. Environ. Saf. 109, 15–21. https://doi.org/10.1016/j.ecoenv.2014.07.029

Huang, X., Jiao, N., Zhang, R., 2021. The genomic content and context of auxiliary metabolic genes in roseophages. Environ. Microbiol. 23, 3743–3757. https://doi.org/https://doi.org/10.1111/1462-2920.15412

Hull, N.M., Ling, F., Pinto, A.J., Albertsen, M., Jang, H.G., Hong, P.Y., Konstantinidis, K.T., LeChevallier, M., Colwell, R.R., Liu, W.T., 2019. Drinking Water Microbiome Project: Is it Time? Trends Microbiol. 27, 670–677. https://doi.org/10.1016/j.tim.2019.03.011

Hurwitz, B.L., Brum, J.R., Sullivan, M.B., 2015. Depth-stratified functional and taxonomic niche specialization in the ‘ core ’ and ‘ flexible ’ Pacific Ocean Virome 472–484. https://doi.org/10.1038/ismej.2014.143

Hurwitz, B.L., Deng, L., Poulos, B.T., Sullivan, M.B., 2013. Evaluation of methods to concentrate and purify ocean virus communities through comparative, replicated metagenomics. Environ. Microbiol. 15, 1428–1440. https://doi.org/10.1111/j.1462-2920.2012.02836.x

Id, K.D.B., Hasan, N.A., Id, M.B.L., Cotruvo, J.A., Rashed, M., Id, R.R.C., Huq, A., 2020. A comparative analysis of drinking water employing metagenomics 1–27. https://doi.org/10.1371/journal.pone.0231210

Jia, S., Shi, P., Hu, Q., Li, B., Zhang, T., Zhang, X., 2015. Bacterial Community Shift Drives Antibiotic Resistance Promotion during Drinking Water Chlorination. https://doi.org/10.1021/acs.est.5b03521

Jia, S., Wu, J., Ye, L., Zhao, F., Li, T., Zhang, X., 2019. Metagenomic assembly provides a deep

insight into the antibiotic resistome alteration induced by drinking water chlorination and its correlations with bacterial host changes. J. Hazard. Mater. 379, 120841. https://doi.org/10.1016/j.jhazmat.2019.120841

Jonge, P.A. De, Nobrega, F.L., Brouns, S.J.J., Dutilh, B.E., 2018. Molecular and Evolutionary Determinants of Bacteriophage Host Range. Trends Microbiol. 27, 51–63. https://doi.org/10.1016/j.tim.2018.08.006

Kallies, R., Hölzer, M., Toscan, R.B., Nunes, U., 2019. Evaluation of Sequencing Library Preparation Protocols for Viral Metagenomic Analysis from Pristine Aquifer Groundwaters. Viruses 11. https://doi.org/doi:10.3390/v11060484

Kang, D.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., Wang, Z., 2019. MetaBAT 2 : an adaptive binning algorithm for robust and ef fi cient genome reconstruction from metagenome assemblies 1–13. https://doi.org/10.7717/peerj.7359

Kavagutti, V.S., Andrei, A.-Ş., Mehrshad, M., Salcher, M.M., Ghai, R., 2019. Phage-centric ecological interactions in aquatic ecosystems revealed through ultra- deep metagenomics 1–15.

Kieft, K., Zhou, Z., Anantharaman, K., 2020. VIBRANT: Automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. Microbiome 8, 1–23. https://doi.org/10.1186/s40168-020-00867-0

Kooij, D. van der, Wielen, P.W.J.J. van der, 2013. Microbial Growth in Drinking Water Supplies. IWA Publishing.

Kothari, A., Roux, S., Zhang, H., Prieto, A., Soneja, D., Chandonia, J., Spencer, S., 2021. Ecogenomics of Groundwater Phages Suggests Niche Differentiation Linked to Specific Environmental Tolerance. mSystems 6, 1–12.

Kotlarz, N., Raskin, L., Zimbric, M., Errickson, J., Lipuma, J.J., Caverly, L.J., 2019. Retrospective Analysis of Nontuberculous Mycobacterial Infection and Monochloramine Disinfection of Municipal Drinking Water in Michigan 4, 1–8.

Krishna, K.C.B., Sathasivan, A., Listowski, A., 2020. Influence of treatment processes and disinfectants on bacterial community compositions and opportunistic pathogens in a full-scale recycled water distribution system. J. Clean. Prod. 274, 123034. https://doi.org/10.1016/j.jclepro.2020.123034

Langenfeld, K., Chin, K., Roy, A., Wigginton, K., Duhaime, M.B., 2021. Comparison of ultrafiltration and iron chloride flocculation in the preparation of aquatic viromes from contrasting sample types. PeerJ 9:e11111. https://doi.org/10.7717/peerj.11111

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

Leinonen, R., Sugawara, H., Shumway, M., 2011. The Sequence Read Archive 39, 2010–2012. https://doi.org/10.1093/nar/gkq1019

Li, H., 2018. Seqtk. https://doi.org/https://github.com/lh3/seqtk

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Subgroup, 1000 Genome Project Data Processing, 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liao, Y., Smyth, G.K., Shi, W., 2014. Sequence analysis featureCounts : an efficient general purpose program for assigning sequence reads to genomic features 30, 923–930. https://doi.org/10.1093/bioinformatics/btt656

Liao, Y., Smyth, G.K., Shi, W., 2013. The Subread aligner : fast , accurate and scalable read

mapping by seed-and-vote 41. https://doi.org/10.1093/nar/gkt214
Liu, G., Verberk, J.Q.J.C., Van Dijk, J.C., 2013. Bacteriology of drinking water distribution systems: An integral and multidimensional review. Appl. Microbiol. Biotechnol. 97, 9265–9276. https://doi.org/10.1007/s00253-013-5217-y
Liu, L., Xing, X., Hu, C., Wang, H., Lyu, L., 2019. Chemosphere Effect of sequential UV / free chlorine disinfection on opportunistic pathogens and microbial community structure in simulated drinking water distribution systems. Chemosphere 219, 971–980. https://doi.org/10.1016/j.chemosphere.2018.12.067
Love, M., Anders, S., Huber, W., 2014. Differential analysis of count data – the DESeq2 package.
Luo, W., Brouwer, C., 2013. Pathview : an R / Bioconductor package for pathway-based data integration and visualization 29, 1830–1831. https://doi.org/10.1093/bioinformatics/btt285
Ma, L., Li, B., Jiang, X.T., Wang, Y.L., Xia, Y., Li, A.D., Zhang, T., 2017. Catalogue of antibiotic resistome and host-tracking in drinking water deciphered by a large scale survey. Microbiome 5, 154. https://doi.org/10.1186/s40168-017-0369-0
Ma, L., Li, B., Zhang, T., 2019. New insights into antibiotic resistome in drinking water and management perspectives : A metagenomic based study of small-sized microbes. Water Res. 152, 191–201. https://doi.org/10.1016/j.watres.2018.12.069
Malki, K., Rosario, K., Sawaya, N.A., Székely, A.J., Tisza, M.J., Breitbart, M., 2020. Prokaryotic and Viral Community Composition of Freshwater Springs in Florida, USA. mBiom 11, e00436-20. https://doi.org/https://doi.org/10.1128/mBio.00436-20
Moon, K., Kim, S., Kang, I., Cho, J., 2020. Viral metagenomes of Lake Soyang , the largest freshwater lake in South Korea. Sci. Data 1–6. https://doi.org/10.1038/s41597-020-00695-9
Nayfach, S., Camargo, A.P., Schulz, F., Eloe-fadrosh, E., Roux, S., Kyrpides, N.C., 2020. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat. Biotechnol. https://doi.org/10.1038/s41587-020-00774-7
NCBI, N.C. for B.I., 2016. UniVec.
Nurk, S., Meleshko, D., Korobeynikov, A., Pevzner, P.A., 2017. metaSPAdes : a new versatile metagenomic assembler 824–834. https://doi.org/10.1101/gr.213959.116.4
Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Wagner, P.R.M., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2020. vegan: Community Ecology Package.
Paez-Espino, D., Eloe-Fadrosh, E.A., Pavlopoulos, G.A., Thomas, A.D., Huntemann, M., Mikhailova, N., Rubin, E., Ivanova, N.N., Kyrpides, N.C., 2016. Uncovering Earth's virome. Nature 536, 425–430. https://doi.org/10.1038/nature19094
Parmar, K.M., Gaikwad, S.L., Dhakephalkar, P.K., Kothari, R., Singh, R.P., 2017. Intriguing interaction of bacteriophage-host association: An understanding in the era of omics. Front. Microbiol. 8. https://doi.org/10.3389/fmicb.2017.00559
Percival, S.L., Wyn-Jones, P., 2013. Methods for the Detection of Waterborne Viruses, Second Edi. ed, Microbiology of Waterborne Diseases: Microbiological Aspects and Risks: Second Edition. Elsevier. https://doi.org/10.1016/B978-0-12-415846-7.00022-6
Petrovich, M.L., Zilberman, A., Kaplan, A., Eliraz, G.R., Wang, Y., Langenfeld, K., Duhaime, M., Wigginton, K., Poretsky, R., Avisar, D., Wells, G.F., 2020. Microbial and Viral Communities and Their Antibiotic Resistance Genes Throughout a Hospital Wastewater Treatment System. Front. Microbiol. 11, 1–13. https://doi.org/10.3389/fmicb.2020.00153
Pinto, A., Schroeder, J., Lunn, M., Sloan, W., Raskin, L., 2014. Spatial-Temporal Survey and

Occupancy-Abundance Modeling To Predict Bacterial Community Dynamics in the Drinking Water Microbiome. MBio 5, e01135-14. https://doi.org/10.1128/mBio.01135-14

Pinto, A.J., Marcus, D.N., Ijaz, U.Z., Bautista-de lose Santos, Q.M., Dick, G.J., Raskin, L., 2016. Metagenomic Evidence for the Presence of Comammox Nitrospira -Like Bacteria in a Drinking Water System . mSphere 1, 1–8. https://doi.org/10.1128/msphere.00054-15

Pinto, A.J., Xi, C., Raskin, L., 2012. Bacterial community structure in the drinking water microbiome is governed by filtration processes. Environ. Sci. Technol. 46, 8851–8859. https://doi.org/10.1021/es302042t

Potgieter, S., Pinto, A., Sigudu, M., du Preez, H., Ncube, E., Venter, S., 2018. Long-term spatial and temporal microbial community dynamics in a large-scale drinking water distribution system with multiple disinfectant regimes. Water Res. 139, 406–419. https://doi.org/10.1016/j.watres.2018.03.077

Potgieter, S.C., Dai, Z., Venter, S.N., Sigudu, M., Pinto, A.J., 2020. Microbial Nitrogen Metabolism in Chloraminated Drinking. mSphere2 5. https://doi.org/https://doi .org/10.1128/mSphere.00274-20. Editor

Prest, E.I., Weissbrodt, D.G., Hammes, F., Loosdrecht, M.C.M. Van, Vrouwenvelder, J.S., 2016. Long-Term Bacterial Dynamics in a Full-Scale Drinking Water Distribution System 1–20. https://doi.org/10.1371/journal.pone.0164445

Rao, C., Waghmare, S., Lakhe, S., 1981. Detection of viruses in drinking water by concentration on magnetic iron oxide. Appl. Environ. Microbiol. 42, 421–426.

Ren, J., Song, K., Deng, C., Ahlgren, N.A., Fuhrman, J.A., Li, Y., Xie, X., Sun, F., 2018. Identifying viruses from metagenomic data by deep learning.

Rodriguez-Valera, F., Rodriguez-, B., Thingstad, T.F., Rohwer, F., Mira, A., 2009. Explaining microbial population genomics through phage predation. Nat. Rev. Microbiol. 7. https://doi.org/10.1038/nrmicro2235

Roeselers, G., Coolen, J., Wielen, P.W.J.J. Van Der, Jaspers, M.C., Atsma, A., Graaf, B. De, Schuren, F., 2015. Microbial biogeography of drinking water : patterns in phylogenetic diversity across space and time 17, 2505–2514. https://doi.org/10.1111/1462-2920.12739

Rohwer, F., Prangishvili, D., Lindell, D., 2009. Roles of viruses in the environment. Environ. Microbiol. 11, 2771–2774. https://doi.org/10.1111/j.1462-2920.2009.02101.x

Roux, S., Adriaenssens, E.M., Dutilh, B.E., Koonin, E. V, Kropinski, A.M., Krupovic, M., Kuhn, J.H., Lavigne, R., Brister, J.R., Varsani, A., Amid, C., Aziz, R.K., Bordenstein, S.R., Bork, P., Breitbart, M., Cochrane, G.R., Daly, R.A., Desnues, C., Duhaime, M.B., Emerson, J.B., Enault, F., Fuhrman, J.A., Hingamp, P., Hugenholtz, P., Hurwitz, B.L., Ivanova, N.N., Labonté, J.M., Lee, K., Malmstrom, R.R., Martinez-garcia, M., Mizrachi, I.K., Ogata, H., Rodriguez-valera, F., Rosario, K., Schriml, L., Schulz, F., Steward, G.F., Sullivan, M.B., Sunagawa, S., Suttle, C.A., Temperton, B., Tringe, S.G., Thurber, R.V., Webster, N.S., Whiteson, K.L., Wilhelm, S.W., Wommack, K.E., Woyke, T., Wrighton, K.C., Yilmaz, P., Yoshida, T., 2019. Minimum Information about an Uncultivated Virus Genome ( MIUViG ). Nat. Biotechnol. 37, 29–37. https://doi.org/10.1038/nbt.4306

Roux, S., Bolduc, B., 2017. stampede-clustergenomes. https://doi.org/https://bitbucket.org/MAVERICLab/stampede-clustergenomes/src/master/

Roux, S., Enault, F., Hurwitz, B.L., Sullivan, M.B., 2015a. VirSorter: Mining viral signal from microbial genomic data. PeerJ 2015, 1–20. https://doi.org/10.7717/peerj.985

Roux, S., Enault, F., Robin, A., Ravet, V., Personnic, S., Theil, S., Colombet, J., Sime-Ngando, T., Debroas, D., 2012. Assessing the Diversity and Specificity of Two Freshwater Viral

Communities through Metagenomics 7. https://doi.org/10.1371/journal.pone.0033641
Roux, S., Hallam, S.J., Woyke, T., Sullivan, M.B., 2015b. Viral dark matter and virus – host interactions resolved from publicly available microbial genomes 1–20. https://doi.org/10.7554/eLife.08490
Russel, M., Model, P., 1986. The Role of Thioredoxin in Filamentous Phage Assembly. J. Biol. Chem. 261, 14997–15005. https://doi.org/10.1016/S0021-9258(18)66819-X
Samhan, F.A., Kronlein, M.R., Fakher, U., Kronlein, C., Robert, D., Hashsham, S.A., 2015. Detection and Occurrence of Indicator Organisms and Pathogens 87, 883–900. https://doi.org/10.2175/106143015X14338845155147
Saparbaevna, M., Aizhan, A., Turmagambetova, S., Gennadievich, P., Andrey, A., Bogoyavlenskiy, P., Berezin, V.E., 2017. Comparative study of viromes from freshwater samples of the Ile-Balkhash region of Kazakhstan captured through metagenomic analysis. VirusDisease 28, 18–25. https://doi.org/10.1007/s13337-016-0353-5
Sevillano, M., Vosloo, S., Cotto, I., Dai, Z., Jiang, T., Santiago, J.M., Padilla, I.Y., Rosario-pabon, Z., Vega, C.V., Alshawabkeh, A., Gu, A., Pinto, A.J., Juan, S., 2021. Spatial-temporal targeted and non-targeted surveys to assess microbiological composition of drinking water in Puerto Rico following Hurricane Maria. preprint.
Shaffer, M., Borton, M.A., McGivern, B.B., Zayed, A.A., La Rosa, S.L. 0003 3527 8101, Solden, L.M., Liu, P., Narrowe, A.B., Rodríguez-Ramos, J., Bolduc, B., Gazitúa, M.C., Daly, R.A., Smith, G.J., Vik, D.R., Pope, P.B., Sullivan, M.B., Roux, S., Wrighton, K.C., 2020. DRAM for distilling microbial metabolism to automate the curation of microbiome function. Nucleic Acids Res. 48, 8883–8900. https://doi.org/10.1093/nar/gkaa621
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. 2498–2504. https://doi.org/10.1101/gr.1239303.metabolite
Shi, P., Jia, S., Zhang, X., Zhang, T., Cheng, S., Li, A., 2012. Metagenomic insights into chlorination effects on microbial antibiotic resistance in drinking water. Water Res. 47, 111–120. https://doi.org/10.1016/j.watres.2012.09.046
Skvortsov, T., Leeuwe, C. De, Quinn, J.P., Mcgrath, J.W., Christopher, C., Allen, R., Mcelarney, Y., Watson, C., Arkhipova, K., Lavigne, R., 2016. Metagenomic Characterisation of the Viral Community of Lough Neagh , the Largest Freshwater Lake in Ireland. PLoS One 11, e0140361. https://doi.org/10.1371/journal.pone.0150361
Stamps, B.W., Leddy, M.B., Plumlee, M.H., Hasan, N.A., Colwell, R.R., Spear, J.R., 2018. Characterization of the microbiome at the world's largest potable water reuse facility. Front. Microbiol. 9, 1–16. https://doi.org/10.3389/fmicb.2018.02435
Stanish, L.F., Hull, N.M., Robertson, C.E., Harris, J.K., Stevens, J., Spear, J.R., Pace, N.R., 2016. Factors Influencing Bacterial Diversity and Community Composition in Municipal Drinking Waters in the Ohio River Basin , USA 1–21. https://doi.org/10.1371/journal.pone.0157966
Sullivan, M.B., Coleman, M.L., Weigele, P., Rohwer, F., Chisholm, S.W., 2005. Three Prochlorococcus Cyanophage Genomes : Signature Features and Ecological Interpretations. PLoS Biol. 3. https://doi.org/10.1371/journal.pbio.0030144
Suttle, C.A., 2007. V lo Marine viruses — major players in the global ecosystem 5. https://doi.org/10.1038/nrmicro1750
Team, R.C., 2021. R: A language and environment for statistical computing.

https://doi.org/https://www.r-project.org/
Team, Rs., 2020. RStudio: Integrated Development for R. https://doi.org/https://www.rstudio.com/
Tseng, C.H., Chiang, P.W., Shiah, F.K., Chen, Y.L., Liou, J.R., Hsu, T.C., Maheswararajah, S., Saeed, I., Halgamuge, S., Tang, S.L., 2013. Microbial and viral metagenomes of a subtropical freshwater reservoir subject to climatic disturbances. ISME J. 7, 2374–2386. https://doi.org/10.1038/ismej.2013.118
Vaerewijck, M.J.M., Huys, G., Carlos, J., Swings, J., 2005. Mycobacteria in drinking water distribution systems : ecology and significance for human health 29, 911–934. https://doi.org/10.1016/j.femsre.2005.02.001
Wang, H., Proctor, C.R., Edwards, M.A., Pryor, M., Domingo, J.W.S., Ryu, H., Camper, A.K., Olson, A., Pruden, A., 2014. Microbial Community Response to Chlorine Conversion in a Chloraminated Drinking Water Distribution System.
Warwick-Dugdale, J., Buchholz, H.H., Allen, M.J., Temperton, B., 2019. Host-hijacking and planktonic piracy : how phages command the microbial high seas. Virol. J. 1, 1–13. https://doi.org/10.1186/s12985-019-1120-1
Weinbauer, M.G., Hornák, K., Jezbera, J., Dolan, J.R., Šimek, K., 2007. Synergistic and antagonistic effects of viral lysis and protistan grazing on bacterial biomass , production and diversity. Environ. Microbiol. 9, 777–788. https://doi.org/10.1111/j.1462-2920.2006.01200.x
Wholey, W., Jakob, U., 2012. Hsp33 confers bleach resistance by protecting elongation factor Tu against oxidative degradation in Vibrio cholerae 83, 981–991. https://doi.org/10.1111/j.1365-2958.2012.07982.x
Williamson, S.J., Allen, L.Z., Lorenzi, H.A., Fadrosh, D.W., Brami, D., Thiagarajan, M., McCrow, J.P., Tovchigrechko, A., Yooseph, S., Venter, J.C., 2012. Metagenomic Exploration of Viruses throughout the Indian Ocean. PLoS One 7. https://doi.org/10.1371/journal.pone.0042047
Yang, Y., Gu, X., Te, S.H., Goh, S.G., Mani, K., He, Y., Gin, K.Y.H., 2019. Occurrence and distribution of viruses and picoplankton in tropical freshwater bodies determined by flow cytometry. Water Res. 149, 342–350. https://doi.org/10.1016/j.watres.2018.11.022
Ye, X.Y., Xing, M., Zhang, Y.L., Xiao, W.Q., Huang, X.N., Cao, Y.G., Gu, K.D., 2012. Real-Time PCR Detection of Enteric Viruses in Source Water and Treated Drinking Water in Wuhan , China. Curr. Microbiol. 65, 244–253. https://doi.org/10.1007/s00284-012-0152-1
Zaouri, N., Jumat, M.R., Cheema, T., Hong, P.Y., 2020. Metagenomics-based evaluation of groundwater microbial profiles in response to treated wastewater discharge. Environ. Res. 180, 108835. https://doi.org/10.1016/j.envres.2019.108835
Zhang, H., Chang, F., Shi, P., Ye, L., Zhou, Q., Pan, Y., Li, A., 2019. Antibiotic Resistome Alteration by Different Disinfection Strategies in a Full-Scale Drinking Water Treatment Plant Deciphered by Metagenomic Assembly. Environ. Sci. Technol. 53, 2141–2150. https://doi.org/10.1021/acs.est.8b05907
Zimmerman, A.E., Varona, C.H.-, Needham, D.M., John, S.G., Worden, A.Z., Sullivan, M.B., Waldbauer, J.R., Coleman, M.L., 2020. Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. Nat. Rev. Microbiol. 18. https://doi.org/10.1038/s41579-019-0270-x