Important declarations

Please remove this info from manuscript text if it is also present there.

Associated Data

Data supplied by the author:
1. Github: 2. 385737018 3. https://git.io/JRmfe

Required Statements

Competing Interest statement:
The authors declare that they have no competing interests.

Funding statement:
Research start-up funds grant No. 322388 were assigned to Miguel Uyaguari-Diaz at the University of Manitoba. Collaborative grant No. 52622 (Drs. Uyaguari and Yuan) was awarded by the Faculty of Science, University of Manitoba.

Quantitation of human enteric viruses as alternative indicators of fecal pollution to evaluate wastewater treatment processes

Audrey Garcia [Equal first author, 1], **Tri Le** [Equal first author, 1], **Paul Jankowski** [1], **Kadir Yanaç** [2], **Qiuyan Yuan** [2], **Miguel I Uyaguari-Diaz** [Corresp. 1]

[1] Department of Microbiology, University of Manitoba, Winnipeg, Manitoba, Canada

[2] Department of Civil Engineering, University of Manitoba, Winnipeg, Manitoba, Canada

Corresponding Author: Miguel I Uyaguari-Diaz
Email address: Miguel.Uyaguari@umanitoba.ca

We investigated the potential use and quantitation of human enteric viruses in municipal wastewater samples of Winnipeg (Manitoba, Canada) as alternative indicators of contamination and evaluated the processing stages of the wastewater treatment plant. During the fall 2019 and winter 2020 seasons, samples of raw sewage, activated sludge, effluents, and biosolids (sludge cake) were collected from the North End Sewage Treatment Plant (NESTP), which is the largest wastewater treatment plant in the City of Winnipeg. DNA and RNA enteric viruses, as well as the *uidA* gene found in *Escherichia coli* were targeted in the samples collected from the NESTP. Total nucleic acids from each wastewater treatment sample were extracted using a commercial spin-column kit. Enteric viruses were quantitated in the extracted samples via quantitative PCR using TaqMan assays. The average gene copies assessed in the raw sewage were not significantly different (p-values ranged between 0.0547 and 0.7986) than the average gene copies assessed in the effluents for Adenovirus and crAssphage (DNA viruses), Pepper Mild Mottle Virus (RNA virus), and *uidA* in terms of both volume and biomass. A significant reduction of these enteric viruses was observed consistently in activated sludge samples compared with those for raw sewage. Corresponding reductions in gene copies per volume and gene copies per biomass were also seen for *uidA* but were not statistically significant (p-value = 0.8769 and p-value = 0.6353, respectively). The higher gene copy numbers of enteric viruses and *E. coli* observed in the effluents may be associated with the 12-hour hydraulic retention time in the facility. Enteric viruses found in gene copy numbers were at least one order of magnitude higher than the *E. coli* marker *uidA*. This indicate that enteric viruses may survive the wastewater treatment process and viral-like particles are being released into the aquatic environment. Our results suggest that Adenovirus, crAssphage, and Pepper mild mottle virus can be used as complementary viral indicators of human fecal

pollution.

Quantitation of human enteric viruses as alternative indicators of fecal pollution to evaluate wastewater treatment processes

Audrey Garcia[1*], Tri Le[1*], Paul Jankowski[1a], Kadir Yanaç[2], Qiuyan Yuan[2] and Miguel Uyaguari-Díaz[1^]

[1] Department of Microbiology, Faculty of Science, University of Manitoba, Winnipeg, MB, Canada R3T 2N2

[2] Department of Civil Engineering, Price Faculty of Engineering, University of Manitoba, Winnipeg, MB, Canada R3T 2N2

[*] These authors contributed equally to this work.

[a] Current address: Department of Medical Microbiology and Infectious Diseases, University of Manitoba, Winnipeg, Manitoba, Canada

[^] Corresponding Author:

Miguel Uyaguari-Díaz

45 Chancellors Circle, Winnipeg, Manitoba, R3T 2N2, Canada

Email address: Miguel.Uyaguari@umanitoba.ca

ABSTRACT

We investigated the potential use and quantitation of human enteric viruses in municipal wastewater samples of Winnipeg (Manitoba, Canada) as alternative indicators of contamination and evaluated the processing stages of the wastewater treatment plant. During the fall 2019 and winter 2020 seasons, samples of raw sewage, activated sludge, effluents, and biosolids (sludge cake) were collected from the North End Sewage Treatment Plant (NESTP), which is the largest wastewater treatment plant in the City of Winnipeg. DNA and RNA enteric viruses, as well as the *uidA* gene found in *Escherichia coli* were targeted in the samples collected from the NESTP. Total nucleic acids from each wastewater treatment sample were extracted using a commercial spin-column kit. Enteric viruses were quantified in the extracted samples via quantitative PCR using TaqMan assays.

The average gene copies assessed in the raw sewage were not significantly different (p-values ranged between 0.0547 and 0.7986) than the average gene copies assessed in the effluents for Adenovirus and crAssphage (DNA viruses), Pepper Mild Mottle Virus (RNA virus), and *uidA* in terms of both volume and biomass. A significant reduction of these enteric viruses was observed consistently in activated sludge samples compared with those for raw sewage. Corresponding reductions in gene copies per volume and gene copies per biomass were also seen for *uidA* but were not statistically significant (p-value = 0.8769 and p-value = 0.6353, respectively). The higher gene copy numbers of enteric viruses and *E. coli* observed in the effluents may be associated with the 12-hour hydraulic retention time in the facility. Enteric viruses found in gene copy numbers were at least one order of magnitude higher than the *E. coli* marker *uidA*, indicating that enteric viruses may survive the wastewater treatment process and viral-like particles are being released into the aquatic environment. Our results suggest that Adenovirus,

crAssphage, and Pepper mild mottle virus can be used as complementary viral indicators of human fecal pollution.

INTRODUCTION

The human fecal waste present in raw sewage (RS) contains pathogens that can cause numerous diseases. This can have a huge negative impact to public, aquatic health, and the economy (Stachler, et al., 2017). Wastewater treatment plants (WWTPs) serve as protective barriers between communities and the environment by reducing the organic matter present in wastewater. Water quality is currently assessed using traditional markers such as coliforms and *Escherichia coli*, leaving other microbes such as viruses largely unexplored. The North End Sewage Treatment Plant (NESTP) in Winnipeg, Manitoba handles approximately 70% of the city's wastewater treatment, serving over 400,000 people (City of Winnipeg, Water and Waste Department, 2020). The treatment process at the NESTP first involves RS undergoing primary treatment to remove solids. During the next treatment cycle, activated sludge (AS), a heterotrophic cocktail of bacteria and protozoa, degrades organic matter present in solid waste. The activated sludge (also known as biological treatment or secondary treatment) is the most widely used process around the world to treat municipal wastewater (Racz et al., 2010; Scholz, 2016), and its use will likely continue due to its low cost and high efficiency. After the biological treatment, wastewater is UV-disinfected and discharged as effluents (EF) into the river (City of Winnipeg, Water and Waste Department, 2020). Approximately 200 million liters of EF are discharged per day (City of Winnipeg, Water and Waste Department, 2020).

The main indicator of biological contamination used in wastewater treatment screening is *E. coli*, a fecal coliform bacterium (Hood et al., 1983). It is present in the gut of humans and warm-

blooded animals, and widely used as the main indicator of fecal pollution during the wastewater treatment process. Nevertheless, the use of only fecal bacteria indicator in wastewater excludes other possible pathogen groups present, such as human enteric viruses. Targeting these viruses in EF could be an alternative method to monitor the wastewater treatment process. Within this context, Dutilh et al. (2014) targeted the DNA crAssphage genome in a human fecal sample. With further bioinformatics testing, it was predicted that the crAssphage genome is highly abundant, having been identified in 73% of human fecal metagenomes surveyed (Dutilh, et al., 2014). In a study conducted by Zhang et al. (2006), the most abundant fecal virus found in dry weight fecal matter was the plant RNA virus, Pepper mild mottle virus (PMMV).

In the present study, samples of RS, AS, EF, and biosolids/sludge cake (SC) from the NESTP were collected (during fall 2019 and winter 2020) to investigate the potential of quantitating human enteric viruses in wastewater samples as complementary indicators of contamination to evaluate the processing stages of wastewater treatment. DNA enteric viruses in this study include human Adenovirus (AdV) and cross-assembly phage (crAssphage), while RNA enteric viruses include PMMV, Noroviruses (NoV) of the genogroups GI and GII, Astrovirus (AstV), Sapovirus (SaV), and Rotavirus (RoV). We also studied the presence of a molecular marker for *E. coli*, the *uidA* gene, in the samples collected from the NESTP. An overview of the workflow is illustrated in *Fig. 1*.

MATERIALS AND METHODS

Sample Collection. A liter of RS, AS, EF, and 1 kg of SC were collected from the NESTP during each sampling event. Each sample was sealed in a 1-L sterile polyethylene container lined with a sterile plastic bag. Samples were collected on October 22nd, 2019 (Event 1) and November

28th, 2019 (Event 2) in the fall season. In the winter season, samples were collected on December 18th, 2019 (Event 3) and February 6th, 2020 (Event 4). SC samples were collected earlier in the day during Events 3 and 4. All samples were kept at 4°C and processed within 24 hours of collection.

Ultrafiltration of Wastewater Samples. Each wastewater treatment sample (RS, AS, and EF), including Millipore Milli-Q water as a negative control, was first filtered via a funnel and cheesecloth to remove any solid waste or debris. Next, 140 mL of each wastewater sample was concentrated using an ultrafiltration method with Centricon Plus-70 filter units of 30 KDa molecular-weight cutoff (Millipore Corporation, Billerica, MA, USA). The ultrafiltration process used a sterile glass pipette, where 70 mL of each wastewater sample was added into their correspondingly labeled sample filter cup pre-assembled with the filtrate collection cup. Each assembly was then sealed with a cap. The Centricon Plus-70 assemblies were placed into a swinging bucket rotor and centrifuged at 3000 x g for 30 minutes at 20°C. Subsequently, the filtrate was discarded, and the remaining 70 mL of the samples was added into their correspondingly labeled sample filter cup pre-assembled with the filtrate collection cup. Samples were spun at the same speed and temperature for 45 minutes. After centrifugation, the sample filter cup was separated from the filtrate collection cup. The concentration collection cup was then turned upside down and placed on top of the sample filter cup. The device was carefully inverted and placed into the centrifuge. Centricon Plus-70 filter units were centrifuged at 800 x g for 2 minutes at 20°C. After this step, the concentrated sample was collected from the concentration cup via a micropipette. The final volume was measured for each wastewater sample. If needed, 10 mM Tris-HCl, pH 8.5 buffer (Qiagen Sciences, Maryland, MD) was added

111 to the concentrate to make up a total volume of 250 µL. If the final volume of the concentrate
112 was over 250 µL, Tris buffer was not added. Aliquots containing 250 µL were made and stored
113 at 4°C and processed within 24h.
114
115 **Sludge Cake Preparation for Ultrafiltration.** To remove cells from the SC samples, a 1X
116 phosphate-buffered solution (PBS) with 0.15M NaCl, 0.05% Tween-20, and pH 7.5
117 was used. Approximately 30 g of SC sample per sampling event (Events 3 and 4) was collected
118 and divided into six Falcon tubes for each event (~5-6 g per tube). Approximately 30 mL of PBS
119 was added to each tube. The Falcon tubes filled with SC samples were homogenized at constant
120 agitation for 15 minutes at 2500 rpm in a vortex mixer. These tubes were then centrifuged at a
121 speed of 4500 x g for 50 minutes. The supernatant from each tube was subsequently recovered
122 and transferred to a new sterile Falcon tube. For each sample event, 140 mL of supernatant was
123 used for ultrafiltration as described previously.
124
125 **Nucleic Acid (DNA/RNA) Extraction and Fluorometric Assessment.** Once the final volume
126 of concentrate was collected from each wastewater sample, the sample was pretreated with
127 InhibitEX buffer (Qiagen Sciences, Maryland, MD) as indicated by the manufacturer. Then,
128 QIAamp MinElute virus spin kit (Qiagen Sciences, Maryland, MD) was used to extract total
129 nucleic acids from each wastewater sample as per the manufacturer's instructions, which
130 included the use of Qiagen Protease and carrier RNA (Qiagen Sciences, Maryland, MD).
131 Samples were eluted in 75 µL of Buffer AVE (Qiagen Sciences, Maryland, MD), quantified, and
132 stored at -80°C for downstream processes. The nucleic acid concentration and purity were
133 assessed using Qubit dsDNA high sensitivity and RNA assay kits in a Qubit 4 fluorometer

(Invitrogen, Carlsbad, CA, USA). Qubit results can be found in Supplementary Materials (*Table S1*).

qPCR Primers, Probes, and gBlocks Gene Fragments. *Table 1* summarizes the primers and probes used in this study. Forward and reverse primers listed in *Table 1* were used in the Primer-BLAST tool to extract gene target regions (Ye, et al., 2012). Extracted regions were then uploaded to the Geneious software to verify oligonucleotide sequences associated to the flanking regions and probe. The generated sequences were sent to Integrated DNA Technologies (IDT, Inc., Coralville, Iowa, USA) to generate the desired gBlocks constructs. IDT manufactured all the primers used for qPCR, as well as the probes Ast-P, Ring1a.2, and Ring 2.2 (*Table 1).* However, probes SaV124TP, SaV5TP, Tampere NSP3, AdV-P, PMMV-Probe, and 056P1 were manufactured by Life Technologies (Carlsbad, CA, USA).

Quantitative PCR Assays. Taqman Environmental Master Mix 2.0 (Life Technologies, Carlsbad, CA, USA) was used for assays involving DNA enteric viruses and *uidA*, while 4x Taqman Fast Virus 1-Step Master Mix (Life Technologies, Carlsbad, CA, USA) was used for RNA enteric viruses. Each 10 µl qPCR reaction contained 500 nM of each of the forward primer and the reverse primer and 250 nM of its designated probe when targeting both DNA and RNA viruses. Five µl of Environmental Master Mix was utilized in each qPCR reaction for targeting DNA viruses, while 2.5 µl of 4x Fast Virus Master Mix was used in each qPCR reaction for targeting RNA viruses. The *uidA* qPCR reaction consisted of 5 µl of Environmental Master Mix, 400 nM of each primer, and 100 nM of probe. All qPCR reactions used 2 µl of template. Each qPCR reaction was performed in triplicates on an ABI QuantStudio 5 PCR system (Applied Biosystems, Foster City, CA, USA). The DNA enteric viruses (AdV and crAssphage) and *uidA*

158 were subjected to the following conditions: 50.0°C for 2 minutes and 95.0°C for 10 minutes
159 followed by 40 cycles of 95.0°C for 15 seconds and 60.0°C for 1 minute. The RNA enteric
160 viruses (SaV, RoV, AstV, GI and GII NoV, PMMV) were subjected to the following conditions:
161 50.0°C for 5 minutes and 95.0°C for 20 seconds followed by 40 cycles of 95.0°C for 3 seconds
162 and 60.0°C for 30 seconds. Raw qPCR output files can be found on GitHub
163 (https://git.io/J8VJ6).
164
165 **Assessment of Ultrafiltration for Viral Recovery Efficiency.** Armored RNA (Asuragen, Inc.,
166 Austin, TX, USA), an artificial virus packed with a 1000-bp single-stranded fragment and
167 encapsulated in a protein coat, was used to assess recovery efficiency of the ultrafiltration
168 method employed herein. We spiked in 16000 copies of Armored RNA into 7.5 mL of
169 representative RS, AS, and EF samples from the NESTP. For the SC sample, 1.25 g of solid SC
170 was dissolved in 7.5 mL of PBS 1x then homogenized by vortexing at 2500 rpm for 15 minutes
171 and centrifuged at 4500 x g for 50 minutes. The supernatant was transferred to a new Falcon tube
172 to be undergoing the same treatment as the RS, AS, and EF samples. The 7.5-mL MilliQ
173 negative control also spiked with 16000 copies of Armored RNA. These five samples were first
174 filtered through cheesecloth. 0.5 mL was aliquoted from each filtrate for subsequent assessment
175 of recovery efficiency. The remaining volumes were subject to ultrafiltration using the Amicon®
176 Ultra-15 Centrifugal Filter Unit (Millipore Corporation, Billerica, MA, USA). Again, 0.5 mL of
177 each flowthrough was stored for efficiency evaluation. Nucleic acid extraction of the retentate
178 was performed in a manner similar to that described above. The final elution volume was 30 µL.
179 Primers (381F: 5'- AGCCTGTCAATACCTGCACC-3' and 475R: 5'-
180 CACGCTTAGATCTCCGTGCT-3'), and probe (420P: 5' Cy5-

AGAGTATGAGAGGTCGACGA-TAO 3') were designed using Primer design tool of Geneious Prime version 2021.1.1 (https://www.geneious.com) and targeted a 95-bp region within the Armored RNA genome. This targeted 95-bp fragment was sent to Integrated DNA Technologies (IDT, Inc., Coralville, Iowa) to synthetize a gBlock construct. Serial dilutions of this synthetic fragment were used to generate standards and quantify gene copy numbers (GCNs) of Armored RNA via quantitative reverse transcription PCR (RT-qPCR). DNA quantification was also performed via RT-qPCR. Standards, samples, and non-template controls were run in triplicates. Thermal cycling reactions were performed at 50°C for 5 minutes, followed by 45 cycles at 95°C for 10 seconds and 60°C for 30 seconds on a QuantStudio 5 Real-Time PCR System (Life Technologies, Carlsbad, CA, USA). For RNA assays, each 10-µl RT-qPCR mixture consisted of 2.5 µL 4X TaqMan Fast Virus 1-Step Master Mix (Life Technologies, Carlsbad, CA, USA), 400 nM each primer, 200 nM probe, and 2.5 µl of template, as well as ultrapure DNAse/RNAse free distilled water (Promega Corporation, Fitchburg, WI, USA). For DNA assays, 5.0 µL Master Mix was used.

Assessment of Gene Copy Numbers by Volume and Biomass. Gene copy numbers (GCNs) were expressed in terms of sample (per mL or g of sample) and biomass (per ng of DNA or RNA). GCNs per mL of sample were calculated as previously described by Ritalahti et al. (2006). When calculating GCNs per mL of sample, the final volume recovered after filtering 140 mL of wastewater sample was used in the formula. For the SC samples, the mass of SC collected was used in the formula to produce results in GCNs per g of sample.

Collection of Metadata for Sampling Events. To perform Principal Component Analysis (PCA) and Spearman's rank correlation analysis for EF samples, metadata pertinent to the sampling events was retrieved. Water quality parameters obtained from the NESTP were combined with their October 2019 monitoring data (City of Winnipeg, Water and Waste Department, 2019) to complete some of the missing fields. For each value not found in either document, data interpolation was performed by taking an average of the corresponding values for the days before and after the sampling event. In addition, the Government of Canada's historical weather database was utilized to obtain the mean temperature on the sampling dates and the total precipitation over three days before each sampling event (hereafter referred to as "precipitation") (Environment and Climate Change Canada, 2021). The values for all parameters were transformed using $\log_{10}$, except for precipitation due to the presence of zero values. These variables were used with $\log_{10}$-transformed GCNs per mL sample for AdV, crAssphage, PMMV, and *uidA* (targets with quantifiable qPCR readings for all replicates across all events) as input for downstream analyses (PCA and Spearman's rank correlation analysis).

Data Handling, Statistical Analysis, and Data Visualization. Various applications were employed to process data at different steps of the pipeline. Input data, such as output from the qPCR instrument, was subjected to rudimentary formatting and cleaning in Microsoft Excel, which was also used to calculate GCNs per mL or g sample and per ng nucleic acid.
R (R Core Team, 2021) and its integrated development environment RStudio (RStudio Team, 2021) were utilized to further process the data and perform statistical analyses and output visualizations. These operations included general linear models (and estimated pairwise differences) using the package *sasLM* version 0.6.0 (Bae, 2021), PCA (corresponding biplots

were created using the package *ggbiplot* version 0.55 (Vu, 2011)), and Spearman's correlation matrix using the package *Hmisc* version 4.5-0 (Harrell Jr., 2021). The package *reshape2* version 1.4.4 (Wickham, 2020) was used to reformat these correlation matrices to enhance compatibility with other data-handling tools. Information about other packages is provided in Supplementary Materials (*Table S2*). The R script used for analysis can be found on GitHub (https://git.io/J8VUl).

Another software involved in data visualization was Tableau. Specifically, it was used to generate boxplots for GCNs per mL or g sample and per ng nucleic acid, as well as the heatmap representing the above-mentioned Spearman correlation matrix.

For all tests, a p-value of 0.05 was assumed to be the minimum level of significance.

RESULTS

From our assessment of the sample processing method used in this study, the recovery efficiencies of Armored RNA as measured by RT-qPCR were between 14.03% and 15.94% for RS, 2.63-4.36% for AS, 12.36-18.74% for EF, and 2.40-5.45% for SC. Meanwhile, DNA recovery efficiencies were 32.48-40.87%, 20.96-45.22%, 14.14-20.15%, and 23.41-68.42%, respectively.

The GCN values for the DNA and RNA viruses and *uidA* were transformed into $\log_{10}$ form. These values were run through a general linear model Tukey-Kramer analysis, and the means of each wastewater processing stage for each target were analyzed. The GCNs were expressed in terms of volume (mL) or weight (g) of sample and biomass (ng of nucleic acids). The result for each GCN value for each replicate from the corresponding sampling event was visualized as one dot in the box plots. We followed cut-off Ct values established by the Molecular Microbiology &

Genomics Team at the British Columbia Centre for Disease Control (2017a, 2017b). With these values, the presence of DNA and RNA viral gene copies and *uidA* in the Milli-Q water (negative control) samples across all Events 1-4 were determined to be negative. The orange-dotted lines in *Figs. 2-6* indicate the mean of the number of gene copies of each wastewater treatment sample across all events.

The average GCNs assessed in the RS were not significantly different (p-values ranged between 0.0547 and 0.7986) compared to the average GCNs assessed in the EF for the DNA enteric viruses (AdV and crAssphage) (*Fig. 2*), PMMV (*Fig. 3*), and *uidA* (*E. coli*) (*Fig. 4*) in terms of both volume and biomass. However, the average GCNs of the DNA enteric viruses assessed in AS were significantly and consistently lower compared to RS. Corresponding reductions in gene copies per volume and gene copies per biomass were also seen for *uidA*, although these reductions were not statistically significant, with p-values being 0.8769 and 0.6353, respectively. For all the aforementioned targets, there was a relatively higher number of gene copies observed in the EF across all events compared to AS samples.

NoV GI and GII were also targets for our study. Boxplots of their GCNs across the different wastewater stages throughout Events 1-4 can be found in Supplementary Materials (*Fig. S1*). In Events 1 and 2 (Fall season), NoV GI was below qPCR detection limits for all samples (RS, AS, and EF). In addition, NoV GII GCNs for all samples collected in Event 2 and AS samples in Events 3 and 4 (Winter season) were also below the detection limits. Among the quantifiable samples, statistically significant GCN differences in terms of volume/mass and biomass were calculated for the pairs of AS-EF (p-values were 4.483×10^{-6} and 3.226×10^{-7}, respectively), AS-RS (1.658×10^{-6}, 1.091×10^{-5}), and AS-SC (1.481×10^{-9}, 4.083×10^{-7}). No significant differences were detected among treatments for GCNs of NoV GI. There was no significant

difference between the mean gene copies of NoV GII in the RS and EF samples in terms of volume (p-value = 0.7377), but the difference was significant in terms of biomass (p-value = 0.04905). The corresponding quantities of all the other sample pairs were statistically significant when looking at both the volume/mass and biomass perspectives, with p-values ranging from 1.304×10^{-8} to 0.0046, except for AS-RS GCN difference in terms of biomass (p-value = 0.0637).

RoV gene copies across the various wastewater treatment stages from Event 1 to 4 were also examined. The boxplots illustrating these results in terms of both sample and biomass can be found in the Supplementary Materials (*Fig. S2*). RoV GCNs were below detection limit for all samples collected in Events 1 and 2. Looking at the EF-SC pair, the mean GCNs differed significantly in terms of volume/mass (p-value = 2.649×10^{-7}) but not biomass (p-value = 0.4298). No significant GCN differences could be detected between RS and AS samples in terms of both volume (p-value = 0.4155) and biomass (p-value = 0.6662). The equivalent magnitudes for the remaining pairs per volume/mass and per biomass were statistically significant, with p-values being between 7.907×10^{-10} and 0.02433, respectively.

In the present study, there was no detection of gene copies for AstV and SaV (Sav1, Sav124, and Sav5) in any of the wastewater samples across all events. In addition, to eliminate the possibility of inhibitors or contaminants such as humic acids, additional qPCR tests using bovine serum albumin (data not shown) were conducted with environmental samples (including AS). No significant differences were observed between samples with and without the enzyme.

To investigate any potential relationship between collected data for EF samples, PCA was performed with $\log_{10}$-transformed variables. We found that three components (PC1, PC2, and PC3) explained 99.14% of the variance between variables. A summary of the weight of

components is included in the Supplementary Materials (*Table S3*). PC1 and PC2 were used to create the biplot in *Fig. 5*. Biplots for PC1 versus PC3 (*Fig. S3*) and PC2 versus PC3 (*Fig. S4*) are included in the Supplementary Materials.

Overall, based on the biplot of PC1 and PC2, samples from the four events were distinct from one another, as point clusters of the four events can be seen occupying different quadrants. PC1, explaining 54.9% of the observed variance, received a notable and positive contribution from COD, cBOD, BOD, and TOC. Strongly negative contributors to PC1 were mean temperature, grab filtered UVT, NO_x-N, and TS. These observations were supported by subsequent Spearman's rank correlation analysis (*Fig. 6*), as COD, cBOD, BOD, and TOC demonstrated strongly positive correlations with one another (rho ranging between 0.8000 and 0.9487) (p-value < 0.005) and strongly negative correlations with mean temperature, grab filtered UVT, NO_x-N, and TS (rho ranging between -1.000 and -0.8000) (p-value < 0.005). PC2 explained 31.9% of the variance between sampling events and showed a strong contribution from crAssphage, *uidA*, and grab temperature. This observation was also supported by the Spearman's rank correlation analysis showing these variables having strongly positive correlation with one another (rho ranging between 0.7169 and 0.9218) (p-value < 0.0100). Additionally, in the biplot, the axes representing *E. coli* and fecal coliform specifically pointed towards the same quadrant, which was reflected in their moderately positive Spearman's coefficient (0.6325) (p-value = 0.0273). However, it is worth noting that *uidA* and *E. coli* exhibited a moderately weak negative correlation (rho = -0.3073), although it was not statistically significant (p-value = 0.3313). The two parameters with the strongest contribution against PC2 were grab pH and turbidity, which was illustrated by the strongly positive Spearman's coefficient heatmap (rho = 0.8000) (p-value = 0.0018).

318

319 **DISCUSSION**

320 The ultrafiltration method used in this study was assessed and the recovery efficiencies among
321 all samples for Armored RNA were estimated to be between 2.40-18.74% for RNA. This range
322 was comparable to other methods to concentrate viral particles such as JumboSep (13.38% ±
323 9.11%) or skimmed milk flocculation (15.27% ± 3.32%), spiked-in wastewater samples, and
324 using Armored RNA as internal control (Yanaç and Uyaguari, unpublished results). Viral
325 particles may have been sorbed to biosolids present in wastewater samples, which were filtered
326 out during the processing stage. In this context, matrix has a significant effect for recovery of
327 viral particles. When compared to other environmental matrices such as surface water samples,
328 recovery efficiency is higher using ultrafiltration (tangential flow filtration) (32.6% ± 11.81%)
329 and skimmed milk flocculation (42.64% ± 15.12%) (Francis and Uyaguari, unpublished results).
330 Water with high turbidity and amounts of suspended solids tend to saturate filters and impact the
331 recovery of viral particles (Aslan et al., 2011; Karim et al., 2009; Uyaguari-Diaz, et al., 2016).
332 Additionally, the flow-through from ultrafiltration is another potential source of lost nucleic acid.
333 The GCNs were expressed in terms of biomass and volume (except for SC, which was expressed
334 in g s of sample). The higher abundance and more stable signal over time of GCNs of AdV and
335 crAssphage (*Fig. 2*) as well as PMMV (*Fig. 3*) relative to the results of other assays make these
336 target more representative for conducting comparisons with *E. coli*. This persistent presence is
337 consistent with various longitudinal studies previously performed (Ballesté, et al., 2019; Farkas,
338 et al., 2018; Farkas, et al., 2019; Hamza et al., 2019; Nour, et al., 2021; Schmitz et al., 2016;
339 Tandukar et al., 2020; Worley-Morse et al., 2019).

340 A reduction of AdV, crAssphage, PMMV, and *uidA* GCNs was observed consistently in AS

samples (*Figs. 2-4*). This could be a result of viral particles being sorbed to larger fractions of organic matter that had been filtered by cheesecloth early in the sample-handling process or retained in the filtration devices as previously described. It is important to mention that samples were collected within a 2-hour period from RS→AS→EF consecutively within each sampling event. The higher GCNs of viruses and *E. coli* observed in the EF may be associated with the hydraulic retention time (12 hours) in the facility and may not reflect wastewater treatment profiles at the time of collection. In other words, the EF samples may not have been the corresponding RS samples collected earlier. The ideal situation would have seen the former being collected 12 hours after the latter. It is best that similar logistical issues be accounted for in future studies. Other variables to consider are the overflow of sewage from rainy events and fluctuations in mixed liquor-suspended solids (Pérez et al., 2019). In our study, there were 4.6 mm of precipitation for Event 1, which may have affected the results. In the PCA analysis (*Fig. 5*), the vector for precipitation sharply denotes data points representing Event 1, indicating a possible relationship. Precipitation was also found to have positive correlations with grab flow (rho = 0.7746) and raw flow (rho = 0.7746) (*Fig. 6*). Nonetheless, further studies and/or more replications are needed to corroborate the potential link between precipitation and microbial counts.

Moreover, the duration of anaerobic sludge digestion is 15 days (City of Winnipeg, Water and Waste Department, 2020). In this context, GCNs of *uidA* in the SC were significantly reduced by anaerobic digestion (*Fig. 4*). This may explain why the gene copies of *uidA* in terms of biomass were lower in SC compared to all treatments (p-value < 0.02). The average gene copies across all wastewater stages (RS, AS, and EF) for *uidA* were not significantly different in terms of both volume and biomass. When compared to *uidA*, enteric viruses were found to be at least one order

of magnitude more abundant than the *E. coli* marker. Similar studies have reported *uidA* in RS at copy numbers nearly 10^4 times that in this report (Jikumaru, et al., 2020; Mbanga, Abia, Amoako, & Essack, 2020). This indicates the possibility of much having been lost during the sample treatment process.

GCNs of crAssphage in terms of biomass in SC were significantly higher than RS (p-value = 0.0040) and AS (p-value = 5.877×10^{-5}) (*Fig. 2*). For PMMV, SC samples had significantly more GCNs in terms of biomass than samples from other parts of the wastewater treatment process (p-values ranged from 1.487×10^{-5} to 0.03788) (*Fig. 3*). Since SC is the by-product of RS and AS using anaerobic digestion, this may indicate that the presence of crAssphage and PMMV was lower in the wastewater being treated in the AS, but higher in the solids. On the other hand, GCNs of AdV in terms of biomass were not significantly different between the AS and SC samples (*Fig. 2*). Meanwhile, plant viruses such as PMMV remain more stable (in terms of biomass) during these digestion processes (Jumat, et al., 2017).

The higher presence of RoV gene copies in the EF (p-value = 0.0006592 in terms of sample and p-value = 0.001550 in terms of biomass) during the winter season (*Fig. S2*) may indicate a higher risk of transmission during cold seasons (Atabakhsh et al. 2020), since a greater presence of RoV in EF has been previously found during the winter season (Li, et al., 2011).

The negative results of SaV (Sav1, Sav124, and Sav5) across all wastewater treatment stages during the fall and winter season are consistent with Varela et al. (2018) where samples were retrieved from a wastewater treatment plant in Tunisia. Their results did not support the general belief that the peak of detection of SaV occurs during the cold and rainy months of the year. However, quantitative detection of SaV in wastewater and river water in Japan showed an

386 increased concentration of SaV in influents between winter and spring (December to May), but a
387 decrease in SaV concentration during the summer and autumn months (July to October)
388 (Haramoto et al., 2008). Yet another pattern of SaV presence was reported in France, as Sima et
389 al. (2011) found the virus to be readily detected in influents but had no clear variations in
390 numbers over the 9-month (October to June) duration of the study. Similarly, seasonal
391 differences in SaV concentrations were not statistically significant in a 3-year study conducted
392 by Song et al. (2021) in China between 2017 and 2019. As a result, there are other factors that
393 can influence wastewater SaV concentrations. For example, it has been hypothesized that
394 isoelectric point could affect how viruses and their different strains behave in bioreactors (Miura
395 et al., 2015). Monitoring over a time period longer than our current study would likely shed more
396 insight into the seasonal variation in the presence of SaV in wastewater.

397 The gene copies of NoV GI and GII were below the detection limit in many of the AS samples
398 (in terms of both volume and biomass), but still relatively abundant in the EF (*Fig. S1*). A
399 possible explanation for the greatly reduced viral GCNs in AS samples is the high efficiency
400 with which NoV GI and GII are removed, a notion supported by literature (Ibrahim et al., 2020;
401 Kitajima et al., 2014; Schmitz et al., 2016). Furthermore, considering the observation that these
402 viruses were found in abundance in SC samples, another contributing factor could be limitations
403 in the sample collection process, which might not have adequately retrieved the slurry part of the
404 sludge where the viruses are found in greater numbers as they might have sorbed to the larger
405 fractions of the sludge solids. The relative abundance of NoV GI and GII gene copies in RS and
406 EF during the winter months (December and February) and the absence of NoV GII in RS in
407 October may be due to seasonal variability (Pérez, Guerrero, Orellana, Figuerola, & Erijman,
408 2019). However, the presence of NoV GI and GII gene copies in RS during Event 4 (February) is

409 consistent with a study conducted by Flannery et al. (2012), in which the concentration of NoV
410 GI and GII gene copies in the influents of a wastewater treatment plant were significantly higher
411 during the winter months (January to March). This seasonal trend is also reflected colloquially
412 through the virus's sobriquet, the winter vomiting bug (Farkas, et al., 2021).

413 In a study conducted by El-Senousy et al. (2007), high numbers of AstV gene copies (per liter) in
414 sewage water samples (from the Greater Cairo area in Egypt) were observed at the end of
415 autumn and during the winter months, but the AstV concentrations tended to decrease as
416 temperatures increased. These results are different from our findings where there was no
417 detection of AstV in any of the wastewater treatment stages across all events. These results may
418 be due to seasonal variability (Pérez et al., 2019) and/or reflect the pattern of infection (Corpuz et
419 al., 2020) within the community under study.

420 Grab filtered UVT being inversely correlated with COD, cBOD, BOD, and TOC is consistent
421 with the widespread use of UV radiation to regulate microbial growth in a variety of medium,
422 including water (Raeiszadeh & Adeli, 2020). Furthermore, it had been suggested that UV is an
423 important influence to the survival of pathogens in wastewater environments, especially in cold
424 weather conditions, such as those found in Manitoba during the surveying period (Murphy,
425 2017). The NESTP uses UV disinfection. Further studies are needed to evaluate the survival of
426 enteric viruses in these reservoirs by using modification to the biological treatment and/or the
427 disinfection process. Some of these modifications include fixed bed reactors (Sizirici & Yildiz,
428 2020), biofilm systems such as membrane bioreactors, biofilters, biofiltration, and carriers (Zhao
429 et al., 2019). Other disinfection processes include the use of chlorine (liquid sodium hypochlorite
430 solution, solid calcium hypochlorite) or newer methods such as ozone (Mezzanotte et al., 2007;
431 Abou-Elela et al., 2012; Collivignarelli et al., 2018).

There is a possibility that viral GCNs quantified in the EF may represent an overestimation of the actual number of infectious viral particles since qPCR detects both infective and non-infective agents and UV treatment influences viral viability (Lizasoain et al., 2017). Thus, the interpretation of these results must be performed with caution. Future studies could incorporate culturable assays for a more complete and accurate evaluation. On the other hand, it is also possible that the non-enveloped enteric viruses (Corpuz, et al., 2020) studied here survived the wastewater treatment process. Non-enveloped viruses are more resilient than their enveloped counterparts in numerous environmental conditions and water treatment processes (La Rosa et al., 2020). This is because of the latter group's envelope, which contains receptors needed for infection; if the envelope is lysed, infection is not possible (La Rosa et al., 2020). Various publications have noted the resilience of non-enveloped viruses after wastewater treatment (Adefisoye et al., 2016; Campos & Lees, 2014; Farkas, et al., 2019; Fitzgerald, 2015; Fong et al., 2010; Li, et al., 2021; Prevost, et al., 2015; Ruggeri, et al., 2015; Varela, et al., 2018). In this context, we have detected GCNs of AdV, crAssphage, and PMMV in environmental surface waters receiving discharges from the NESTP, two other WWTPs, and other areas radiating away from the WWTPs within the city of Winnipeg (Francis and Uyaguari, unpublished data). Therefore, despite potential factors affecting interpretation, our results still reflect the presence of several non-enveloped enteric viruses in EF samples with reasonable quantitative accuracy.

CONCLUSION

Our study's primary goal was to identify human enteric viruses with the potential to become alternative indicators of fecal pollution. Towards that end, we propose AdV, crAssphage, and PMMV as more stable viral indicators of water quality due to their quantifiability illustrated in

this investigation and the literature. Regular monitoring of these organisms can be useful complements to current methods for assessing wastewater treatment processes. Such vigilance could be a helpful tool to assist public health efforts in the event of a viral outbreak. Additionally, our study indicated that enteric viruses may have survived the wastewater treatment process and viral-like particles are possibly being released into the aquatic environment. Therefore, in addition to such methods as UV radiation (which is currently used in the NESTP and was shown in our study to be inversely correlated with biological parameters), we also suggest that WWTPs consider implementing modifications and/or additions (disinfection processes) to their workflow to reduce the number of viral particles across different stages of the wastewater treatment process.

ABBREVIATIONS

AdV: Adenovirus

AS: activated sludge

AstV: Astrovirus

BOD: biochemical oxygen demand

cBOD: carbonaceous biochemical oxygen demand

COD: chemical oxygen demand

EF: effluents

GCN: gene copy number

NESTP: North End Sewage Treatment Plant

NH_4-N: ammonium-nitrogen

NoV: Norovirus

NO_x-N: nitrogen oxides - nitrogen

PCA: Principal Component Analysis

PMMV: Pepper mild mottle virus

PO_4-P: orthophosphate as phosphorus

RoV: Rotavirus

RS: raw sewage

RT-qPCR: quantitative reverse transcription PCR

SaV: Sapovirus

SC: sludge cake

sCOD: soluble chemical oxygen demand

TN: total nitrogen

TOC: total organic carbon

TP: total phosphorus

TS: total solids

TSS: total suspended solids

***uidA*:** β-d-glucuronidase gene

WWTP: wastewater treatment plant

AUTHOR CONTRIBUTIONS

AG performed the experiments, analyzed the data, prepared the figures and tables, and wrote and reviewed the drafts of the manuscript.

TL performed the validation experiments, analyzed the data, prepared the figures and tables, and wrote and reviewed the drafts of the manuscript.

PJ performed the experiments and reviewed the drafts of the manuscript.

KY performed the validation experiments and reviewed the drafts of the manuscript.

QY contributed the analysis tools and reviewed the drafts of the manuscript.

MUD designed the experiments, provided lead guidance during the experiments and analyses, contributed the analysis tools, and reviewed the drafts of the manuscript.

All authors read and approved of the final manuscript.

ACKNOWLEDGEMENTS

Special thanks to the City of Winnipeg and Palwinder Singh, graduate student, Department of Civil Engineering at the University of Manitoba (UoM) for sample collection. Research start-up funds grant No. 322388 were assigned to Miguel Uyaguari-Diaz at the UoM. We acknowledge The Faculty of Science, UoM, collaborative grant No. 52622 (Drs. Uyaguari and Yuan). This research was conducted at the University of Manitoba. "*The University of Manitoba campuses are located on original lands of Anishinaabeg, Cree, Oji-Cree, Dakota, and Dene peoples, and on the homeland of the Métis Nation*".

REFERENCES

Abou-Elela, S. I., El-Sayed, M. M. H., El-Gendy, A. S., & Abou-Taleb, E. M. (October 2012). Comparative study of disinfection of secondary treated wastewater using chlorine, UV and ozone. Journal of Applied Sciences Research, pp.5190-5197 ref.12

Adefisoye, M. A., Nwodo, U. U., Green, E., & Okoh, A. I. (2016). Quantitative PCR Detection and Characterisation of Human Adenovirus, Rotavirus and Hepatitis A Virus in Discharged Effluents of Two Wastewater Treatment Facilities in the Eastern Cape, South Africa. *Food and Environmental Virology, 8*, 262-274. doi:10.1007/s12560-016-9246-4

Aslan, A., Xagoraraki, I., Simmons, F., Rose, J., & Dorevitch, S. (2011, August 19). Occurrence of adenovirus and other enteric viruses in limited-contact freshwater recreational areas and bathing waters. *Journal of Applied Microbiology, 111*(5), 1250-1261. doi:10.1111/j.1365-2672.2011.05130.x

Bae, K.-S. (2021). sasLM: 'SAS' Linear Model. Retrieved from https://CRAN.R-project.org/package=sasLM

Ballesté, E., Pascual-Benito, M., Martín-Díaz, J., Blanch, A. R., Lucena, F., Muniesa, M., . . . García-Aljaro, C. (2019, May 15). Dynamics of crAssphage as a human source tracking marker in potentially faecally polluted environments. *Water Research, 155*, 233-244. doi:10.1016/j.watres.2019.02.042

Campos, C. J., & Lees, D. N. (2014, June). Environmental Transmission of Human Noroviruses in Shellfish Waters. *Applied and Environmental Microbiology, 80*(12), 3552-3561. doi:10.1128/AEM.04188-13

City of Winnipeg, Water and Waste Department. (2019, October). North End Water Pollution Control Centre Monitoring Data. Winnipeg, Manitoba, Canada. Retrieved July 21, 2021,

538 from
539 https://www.winnipeg.ca/waterandwaste/pdfs/sewage/ComplianceReporting/2019/oct/ne
540 wpcc.pdf
541 City of Winnipeg, Water and Waste Department. (2020, October 8). *Sewage Treatment Plants*.
542 Retrieved July 21, 2021, from City of Winnipeg:
543 https://www.winnipeg.ca/waterandwaste/sewage/treatmentPlant/default.stm#tab-north-
544 end-sewage-treatment-plant.
545 Collivignarelli, M.C., Abbà, A., Benigna, I., Sorlini, S., & Torretta, V. (2018). Overview of the
546 Main Disinfection Processes for Wastewater and Drinking Water Treatment
547 Plants. *Sustainability*, *10*, 86. https://doi.org/10.3390/su10010086.
548 Comber, S. D., Gardner, M. J., & Ellor, B. (2020, September). Seasonal variation of contaminant
549 concentrations in wastewater treatment works effluents and river waters. *Environmental*
550 *Technology, 41*(21), 2716-2730. doi:10.1080/09593330.2019.1579872
551 Corpuz, M. V., Buonerba, A., Vigliotta, G., Zarra, T., Ballesteros Jr, F., Campiglia, P., . . .
552 Naddeo, V. (2020, November 25). Viruses in wastewater: occurrence, abundance and
553 detection methods. *Science of the Total Environment, 745*.
554 doi:10.1016/j.scitotenv.2020.140910
555 Dutilh, B. E., Cassman, N., McNair, K., Sanchez, S. E., Silva, G. G., Boling, L., . . . Edwards, R.
556 A. (2014). A highly abundant bacteriophage discovered in the unknown sequences of
557 human faecal metagenomes. *Nature Communications, 5*(4498), 1-11.
558 doi:10.1038/ncomms5498
559 El-Senousy, W. M., Guix, S., Abid, I., Pintó, R. M., & Bosch, A. (2007, January). Removal of
560 astrovirus from water and sewage treatment plants, evaluated by a competitive reverse

561 transcription-PCR. *Applied and Environmental Microbiology, 73*(1), 164-7.
562 doi:10.1128/AEM.01748-06
563 Environment and Climate Change Canada. (2021). *Historical Data*. Retrieved July 21, 2021,
564 from Government of Canada:
565 https://climate.weather.gc.ca/historical_data/search_historic_data_e.html
566 Farkas, K., Adriaenssens, E. M., Walker, D. I., McDonald, J. E., Malham, S. K., & Jones, D. L.
567 (2019, June). Critical Evaluation of CrAssphage as a Molecular Marker for Human-
568 Derived Wastewater Contamination in the Aquatic Environment. *Food and*
569 *Environmental Virology, 11*(2), 113-119. doi:10.1007/s12560-019-09369-1
570 Farkas, K., Green, E., Rigby, D., Cross, P., Tyrrel, S., Malham, S. K., & Jones, D. L. (2021, May
571 27). Investigating awareness, fear and control associated with norovirus and other
572 pathogens and pollutants using best–worst scaling. *Scientific Reports, 11*.
573 Farkas, K., Marshall, M., Cooper, D., McDonald, J. E., Malham, S. K., Peters, D. E., . . . Jones,
574 D. L. (2018, November). Seasonal and diurnal surveillance of treated and untreated
575 wastewater for human enteric viruses. *Environmental Science and Pollution Research,*
576 *25*(33), 33391-33401. doi:10.1007/s11356-018-3261-y
577 Fitzgerald, A. (2015). *Review of Approaches for Establishing Exclusion Zones for Shellfish*
578 *Harvesting around Sewage Discharge Points - Desk Study to Inform Consideration of the*
579 *Possible Introduction of Exclusion Zones as a Control for Norovirus in Oysters.*
580 Technical Report, Aquatic Water Services Ltd. Retrieved July 26, 2021, from
581 https://webarchive.nationalarchives.gov.uk/20150418173120/http://www.food.gov.uk/sit
582 es/default/files/Exclusion%20Zones%20Project%20FS513404%20-
583 %20Technical%20Report%20FINAL.pdf

584 Fong, T.-T., Phanikumar, M. S., Xagoraraki, I., & Rose, J. B. (2010, February). Quantitative
585 detection of human adenoviruses in wastewater and combined sewer overflows
586 influencing a Michigan river. *Applied and Environmental Microbiology, 76*(3), 715-23.
587 doi:10.1128/AEM.01316-09
588 Frahm, E., & Obst, U. (2003). Application of the fluorogenic probe technique (TaqMan PCR) to
589 the detection of Enterococcus spp. and Escherichia coli in water samples. *Journal of*
590 *Microbiological Methods, 52*(1), 123-31. doi:10.1016/s0167-7012(02)00150-1
591 Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., & Hothorn, T. (2021). mvtnorm:
592 Multivariate Normal and t Distributions. Retrieved from http://CRAN.R-
593 project.org/package=mvtnorm
594 Hamza, H., Rizk, N. M., Gad, M. A., & Hamza, I. A. (2019, November). Pepper mild mottle
595 virus in wastewater in Egypt: a potential indicator of wastewater pollution and the
596 efficiency of the treatment process. *Archives of Virology, 164*(11), 2707-2713.
597 doi:10.1007/s00705-019-04383-x
598 Haramoto, E., Katayama, H., Phanuwan, C., & Ohgaki, S. (2008, March). Quantitative detection
599 of sapoviruses in wastewater and river water in Japan. *Letters in Applied Microbiology,*
600 *46*(3), 408-13. doi:10.1111/j.1472-765X.2008.02330.x
601 Harrell Jr., F. E. (2021). Hmisc: Harrell Miscellaneous. Retrieved from https://CRAN.R-
602 project.org/package=Hmisc
603 Hood, M. A., Ness, G. E., & Blake, N. J. (1983). Relationship among fecal coliforms,
604 Escherichia coli, and Salmonella spp. in shellfish. *Applied and environmental*
605 *microbiology, 45*(1), 122-126. doi:10.1128/aem.45.1.122-126.1983

606 Ibrahim, C., Hammami, S., Khelifi, N., Pothier, P., & Hassen, A. (2020). The Effectiveness of
607 Activated Sludge Procedure and UV-C 254 in Norovirus Inactivation in a Tunisian
608 Industrial Wastewater Treatment Plant. *Food and Environmental Virology, 12*, 250-259.
609 doi:10.1007/s12560-020-09434-0
610 Jikumaru, A., Ishii, S., Fukudome, T., Kawahara, Y., Iguchi, A., Masago, Y., . . . Suzuki, Y.
611 (2020, July). Fast, sensitive, and reliable detection of waterborne pathogens by digital
612 PCR after coagulation and foam concentration. *Journal of Bioscience and*
613 *Bioengineering, 130*(1), 76-81. doi:10.1016/j.jbiosc.2020.02.004
614 Jumat, M. R., Hasan, N. A., Subramanian, P., Heberling, C., Colwell, R. R., & Hong, P.-Y.
615 (2017). Membrane Bioreactor-Based Wastewater Treatment Plant in Saudi Arabia:
616 Reduction of Viral Diversity, Load, and Infectious Capacity. *Water, 9*(7).
617 doi:10.3390/w9070534
618 Kageyama, T., Kojima, S., Shinohara, M., Uchida, K., Fukushi, S., Hoshino, F. B., . . .
619 Katayama, K. (2003, April). Broadly Reactive and Highly Sensitive Assay for Norwalk-
620 Like Viruses Based on Real-Time Quantitative Reverse Transcription-PCR. *Journal of*
621 *Clinical Microbiology, 41*(4), 1548-1557. doi:10.1128/JCM.41.4.1548-1557.2003
622 Karim, M. R., Rhodes, E. R., Brinkman, N., Wymer, L., & Fout, G. S. (2009, April). New
623 Electropositive Filter for Concentrating Enteroviruses and Noroviruses from Large
624 Volumes of Water. *Applied and Environmental Microbiology, 75*(8), 2393-2399.
625 doi:10.1128/AEM.00922-08
626 Kitajima, M., Iker, B. C., Pepper, I. L., & Gerba, C. P. (2014, August 1). Relative abundance and
627 treatment reduction of viruses during wastewater treatment processes--identification of

potential viral indicators. *Science of the Total Environment, 488-489*, 290-296. doi:10.1016/j.scitotenv.2014.04.087

La Rosa, G., Bonadonna, L., Lucentini, L., Kenmoe, S., & Suffredini, E. (2020, July 15). Coronavirus in water environments: Occurrence, persistence and concentration methods - A scoping review. *Water Research, 179*. doi:10.1016/j.watres.2020.115899

Li, D., Gu, A. Z., Zeng, S.-Y., Yang, W., He, M., & Shi, H.-C. (2011, May). Monitoring and evaluation of infectious rotaviruses in various wastewater effluents and receiving waters revealed correlation and seasonal pattern of occurrences. *Journal of Applied Microbiology, 110*(5), 1129-37. doi:10.1111/j.1365-2672.2011.04954.x

Li, X., Cheng, Z., Dang, C., Zhang, M., Zheng, Y., & Xia, Y. (2021, July). Metagenomic and viromic data mining reveals viral threats in biologically treated domestic wastewater. *Environmental Science and Ecotechnology, 7*. doi:10.1016/j.ese.2021.100105

Lizasoain, A., Tort, L., García, M., Gillman, L., Alberti, A., Leite, J., . . . Colina, R. (2017). Human enteric viruses in a wastewater treatment plant: evaluation of activated sludge combined with UV disinfection process reveals different removal performances for viruses with different features. *Letters in Applied Microbiology, 66*(3), 215-221. doi:DOI: 10.1111/lam.12839

Makowski, D., Ben-Shachar, M., Patil, I., & Lüdecke, D. (2020). Automated Results Reporting as a Practical Tool to Improve Reproducibility and Methodological Best Practices Adoption. Retrieved from https://github.com/easystats/report.

Mbanga, J., Abia, A. L., Amoako, D. G., & Essack, S. Y. (2020). Quantitative microbial risk assessment for waterborne pathogens in a wastewater treatment plant and its receiving surface water body. *BMC Microbiology*. doi:10.1186/s12866-020-02036-7

651 Mezzanotte, V., Antonelli, M., Citterio, S., & Nurizzo, C. (2007). Wastewater disinfection
652 alternatives: chlorine, ozone, peracetic acid, and UV light. *Water Environment Research,*
653 *79*(12), 2373-2379. doi:10.2175/106143007x183763.
654 Miura, T., Okabe, S., Nakahara, Y., & Sano, D. (2015, May 15). Removal properties of human
655 enteric viruses in a pilot-scale membrane bioreactor (MBR) process. *Water Research, 75*,
656 282-291. doi:10.1016/j.watres.2015.02.046
657 Molecular Microbiology & Genomics Team, British Columbia Centre for Disease Control.
658 (2017). Detecting Norovirus by Fast Real-Time RT-PCR. British Columbia, Canada.
659 Molecular Microbiology & Genomics Team, British Columbia Centre for Disease Control.
660 (2017). Performing the GI Virus Panel by Real-Time PCR Procedure. British Columbia,
661 Canada.
662 Murphy, H. (2017). Persistence of Pathogens in Sewage and Other Water Types. In J. Rose, & B.
663 Jiménez-Cisneros (Eds.), *Global Water Pathogen Project* (Vol. 4). E. Lansing, MI,
664 UNESCO. doi:10.14321/waterpathogens.51
665 Nour, I., Hanif, A., Zakri, A. M., Al-Ashkar, I., Alhetheel, A., & Eifan, S. (2021, April 29).
666 Human Adenovirus Molecular Characterization in Various Water Environments and
667 Seasonal Impacts in Riyadh, Saudi Arabia. *International Journal of Environmental*
668 *Research and Public Health, 18*(9), 4773. doi:10.3390/ijerph18094773
669 Oka, T., Katayama, K., Hansman, G. S., Kageyama, T., Ogawa, S., Wu, F.-T., . . . Takeda, N.
670 (2006). Detection of human sapovirus by real-time reverse transcription-polymerase
671 chain reaction. *Journal of Medical Virology, 78*(10), 1347-1353. doi:10.1002/jmv.20699

672 Pérez, M. V., Guerrero, L. D., Orellana, E., Figuerola, E. L., & Erijman, L. (2019, July 2). Time
673 Series Genome-Centric Analysis Unveils Bacterial Response to Operational Disturbance
674 in Activated Sludge. *mSystems, 4*(4). doi:10.1128/mSystems.00169-19
675 Prevost, B., Lucas, F. S., Ambert-Balay, K., Pothier, P., Moulin, L., & Wurtzer, S. (2015,
676 October). Deciphering the Diversities of Astroviruses and Noroviruses in Wastewater
677 Treatment Plant Effluents by a High-Throughput Sequencing Method. *Applied and*
678 *Environmental Microbiology, 81*(20), 7215-7222. doi:10.1128/AEM.02076-15.
679 R Core Team. (2021). R: A language and environment for statistical computing. R Foundation
680 for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/
681 Racz, L., T. Datta, and R. Goel. 2010. Effect of organic carbon on ammonia oxidizing bacteria in
682 a mixed culture. Bioresource Technology, 101 (16), 6454-60.
683 Raeiszadeh, M., & Adeli, B. (2020, October 14). A Critical Review on Ultraviolet Disinfection
684 Systems against COVID-19 Outbreak: Applicability, Validation, and Safety
685 Considerations. *ACS Photonics*. doi:10.1021/acsphotonics.0c01245
686 Ritalahti, K. M., Amos, B. K., Sung, Y., Wu, Q., Koenigsberg, S. S., & Löffler, F. E. (2006).
687 Quantitative PCR targeting 16S rRNA and reductive dehalogenase genes simultaneously
688 monitors multiple Dehalococcoides strains. *Applied and Environmental Microbiology,*
689 *72*(4), 2765-74. doi:10.1128/AEM.72.4.2765-2774.2006
690 Rosario, K., Symonds, E. M., Sinigalliano, C., Stewart, J., & Breitbart, M. (2009). Pepper Mild
691 Mottle Virus as an Indicator of Fecal Pollution. *Applied and Environmental*
692 *Microbiology, 75*(22), 7261-7267. doi:10.1128/AEM.00410-09
693 Rosman, N. H., Anuar, A. N., Chelliapan, S., Din, M. F., & Ujang, Z. (2014, June).
694 Characteristics and performance of aerobic granular sludge treating rubber wastewater at

different hydraulic retention time. *Bioresource Technology, 161*, 155-61. doi:10.1016/j.biortech.2014.03.047

RStudio Team. (2021). RStudio: Integrated Development Environment for R. Boston, MA: RStudio, PBC. Retrieved from http://www.rstudio.com/

Ruggeri, F. M., Bonomo, P., Ianiro, G., Battistone, A., Delogu, R., Germinario, C., . . . Fiore, L. (2015, January). Rotavirus Genotypes in Sewage Treatment Plants and in Children Hospitalized with Acute Diarrhea in Italy in 2010 and 2011. *Applied and Environmental Microbiology, 81*(1), 241-249. doi:10.1128/AEM.02695-14

Sarkar, D. (2008). *Lattice: Multivariate Data Visualization with R.* New York: Springer.

Scholz, M. Chapter 15-Activated Sludge Processes, Editor(s): Miklas Scholz, Wetlands for Water Pollution Control (Second Edition), Elsevier, Pages 91-105.

Schmitz, B. W., Kitajima, M., Campillo, M. E., Gerba, C. P., & Pepper, I. L. (2016). Virus Reduction during Advanced Bardenpho and Conventional Wastewater Treatment Processes. *Environmental Science & Technology, 50*(17), 9524-9532. doi:10.1021/acs.est.6b01384.

Sima, L. C., Schaeffer, J., Saux, J.-C. L., Parnaudeau, S., Elimelech, M., & Guyader, F. S. (2011, August). Calicivirus Removal in a Membrane Bioreactor Wastewater Treatment Plant. *Applied and Environmental Microbiology, 77*(15), 5170-5177. doi:10.1128/AEM.00583-11.

Sizirici, B., & Yildiz, I. (2020). Organic matter removal via activated sludge immobilized gravel in fixed bed reactor. E3S Web of Conferences, 191, 03006, 1-5.

716 Song, K., Lin, X., Liu, Y., Ji, F., Zhang, L., Chen, P., . . . Xu, A. (2021). Detection of Human
717 Sapoviruses in Sewage in China by Next Generation Sequencing. *Food and*
718 *Environmental Virology, 13*, 270-280. doi:10.1007/s12560-021-09469-x
719 Stachler, E., Akyon, B., Carvalho, N. A., Ference, C., & Bibby, K. (2018). Correlation of
720 crAssphage qPCR Markers with Culturable and Molecular Indicators of Human Fecal
721 Pollution in an Impacted Urban Watershed. *Environmental Science & Technology,*
722 *52*(13), 7505-7512. doi:10.1021/acs.est.8b00638
723 Stachler, E., Kelty, C., Sivaganesan, M., Li, X., Bibby, K., & Shanks, O. C. (2017). Quantitative
724 CrAssphage PCR Assays for Human Fecal Pollution Measurement. *Environmental*
725 *Science & Technology, 51*(16), 9146–9154. doi:10.1021/acs.est.7b02703
726 Tandukar, S., Sherchan, S. P., & Haramoto, E. (2020). Applicability of crAssphage, pepper mild
727 mottle virus, and tobacco mosaic virus as indicators of reduction of enteric viruses during
728 wastewater treatment. *Scientific Reports, 10*(3616). doi:10.1038/s41598-020-60547-9
729 Therneau, T. (2021). A Package for Survival Analysis in R. Retrieved from https://CRAN.R-
730 project.org/package=survival
731 Ushey, K., Allaire, J., Wickham, H., & Ritchie, G. (2020). rstudioapi: Safely Access the RStudio
732 API. Retrieved from https://CRAN.R-project.org/package=rstudioapi
733 Uyaguari-Diaz, M. I., Chan, M., Chaban, B. L., Croxen, M. A., Finke, J. F., Hill, J. E., . . . Tang,
734 P. (2016). A comprehensive method for amplicon-based and metagenomic
735 characterization of viruses, bacteria, and eukaryotes in freshwater samples. *Microbiome,*
736 *4*(20). doi:10.1186/s40168-016-0166-1
737 Varela, M. F., Ouardani, I., Kato, T., Kadoya, S., Aouni, M., Sano, D., & Romalde, J. L. (2018,
738 March 1). Sapovirus in Wastewater Treatment Plants in Tunisia: Prevalence, Removal,

and Genetic Characterization. *Applied and Environmental Microbiology, 84*(6). doi:10.1128/AEM.02093-17

Vu, V. Q. (2011). ggbiplot: A ggplot2 based biplot. Retrieved from http://github.com/vqv/ggbiplot

Wang, X. M., Wei, Z. M., Guo, J. M., Cai, J. M., Chang, H. M., Ge, Y. M., & Zeng, M. M. (2019, November). Norovirus Activity and Genotypes in Sporadic Acute Diarrhea in Children in Shanghai During 2014–2018. *The Pediatric Infectious Disease Journal, 38*(11), 1085-1089. doi:10.1097/INF.0000000000002456

Wickham, H. (2011). The Split-Apply-Combine Strategy for Data Analysis. *Journal of Statistical Software, 40*(1), 1-29. doi:10.18637/jss.v040.i01

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer-Verlag.

Wickham, H. (2020, April 9). reshape2: Flexibly Reshape Data: A Reboot of the Reshape Package. Retrieved from https://cran.r-project.org/package=reshape2

Wickham, H., & Bryan, J. (2019). readxl: Read Excel Files. Retrieved from https://CRAN.R-project.org/package=readxl

Wickham, H., & Bryan, J. (2021). usethis: Automate Package and Project Setup. Retrieved from https://CRAN.R-project.org/package=usethis

Wickham, H., & Seidel, D. (2020). scales: Scale Functions for Visualization. Retrieved from https://CRAN.R-project.org/package=scales

Wickham, H., François, R., Henry, L., & Müller, K. (2021). dplyr: A Grammar of Data Manipulation. Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., Hester, J., & Chang, W. (2021). devtools: Tools to Make Developing R Packages Easier. Retrieved from https://CRAN.R-project.org/package=devtools

762 Worley-Morse, T., Mann, M., Khunjar, W., Olabode, L., & Gonzalez, R. (2019, September).
763 Evaluating the fate of bacterial indicators, viral indicators, and viruses in water resource
764 recovery facilities. *Water Environment Research, 91*(9), 830-842. doi:10.1002/wer.1096
765 Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012, June 18).
766 Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction.
767 *BMC Bioinformatics, 13*(134). doi:10.1186/1471-2105-13-134
768 Zeileis, A., & Croissant, Y. (2010). Extended Model Formulas in R: Multiple Parts and Multiple
769 Responses. *Journal of Statistical Software, 34*(1), 1-13. doi:10.18637/jss.v034.i01
770 Zeng, S. Q., Halkosalo, A., Salminen, M., Szakal, E. D., Puustinen, L., & Vesikari, T. (2008).
771 One-step quantitative RT-PCR for the detection of rotavirus in acute gastroenteritis.
772 *Journal of Virological Methods, 153*(2), 238-40. doi:10.1016/j.jviromet.2008.08.004
773 Zhang, T., Breitbart, M., Lee, W. H., Run, J.-Q., Wei, C. L., Soh, S. W., Hibberd, M. L., Liu, E.
774 T., Rohwer, F., & Ruan, Y. (2006, January). RNA Viral Community in Human Feces:
775 Prevalence of Plant Pathogenic Viruses. *PLOS Biology, 4*(1).
776 doi:10.1371/journal.pbio.0040003
777 Zhao, Y., Liu, D., Huang, W., Yang, Y., Ji, M., Nghiem, L. D., Trinh, Q. T., Tran, N. H. (2018).
778 Insights into biofilm carriers for biological wastewater treatment processes: Current state-
779 of-the-art, challenges, and opportunities. Bioresource Technology, 288, 121619.
780 doi.org/10.1016/j.biortech.2019.121619.

Table 1(on next page)

Table 1. Primers and probes used in the present study.

Target	DNA or RNA	Primer/ Probe	Sequence (5'-3')	Genomic Target	References
Adenovirus 40/41	DNA	AdV-F	GCC TGG GGA ACA AGT TCA G	Hexon	Molecular Microbiology & Genomics Team, British Columbia Centre for Disease Control, 2017a
		AdV-R	ACG GCC AGC GTA AAG CG		
		AdV-P (Probe)	NED-ACC CAC GAT GTA ACC AC-MGB-NFQ		
Astrovirus	RNA	Ast-F	AAG CAG CTT CGT GAR TCT GG	Junction of polymerase and capsid	Molecular Microbiology & Genomics Team, British Columbia Centre for Disease Control, 2017a
		Ast-R	GCC ATC RCA CTT CTT TGG TCC		
		Ast-P (Probe)	Cy5-CAC AGA AGA GCA ACT CCA TCG CAT TTG-Tao-IBDRQ		
crAssphage	DNA	056F1	CAG AAG TAC AAA CTC CTA AAA AAC GTA GAG	Genomic base pair region: 14731 bp-14856 bp	Stachler, Akyon, Carvalho, Ference, & Bibby, 2018
		056R1	GAT GAC CAA TAA ACA AGC CAT TAG C		
		056P1	FAM-AAT AAC GAT TTA CGT GAT GTA AC		
Escherichia coli	DNA	784F	GTG TGA TAT CTA CCC GCT TCG C	*uidA*	Frahm and Obst, 2003
		866R	AGA ACG GTT TGT GGT TAA TCA GGA		
		EC807 (Probe)	FAM-TCG GCA TCC GGT CAG TGG CAG T-BHQ1		
GI	RNA	COG1F-flap	AATAAATCATAACGYTGGATGCGNTT YCATGA	Norovirus GI	Kageyama, et al., 2003; Wang, et al., 2019
		COG1R-flap	AATAAATCATAACTTAGACGCCATCA TCATTYAC		
		Ring1a.2 (Probe)	6-FAM- AGATYGCGR/ZEN/ TCYCCTGTCCA -IBFQ		Molecular Microbiology & Genomics Team, British Columbia Centre for Disease Control, 2017b
GII	RNA	COG2F-flap	AATAAATCATAACARGARBCNATGTT YAGRTGGAT GAG	Norovirus GII	Kageyama, et al., 2003; Wang, et al., 2019
		COG2R-flap	AATAAATCATAATCGACGCCATCTTC ATTCACA		
		Ring 2.2 (Probe)	JOE - TGGGAGGGY/ZEN/ GATCGCAATCT - IBFQ		Molecular Microbiology & Genomics Team, British Columbia Centre for Disease Control, 2017b
Pepper Mild Mottle Virus	RNA	PMMV-FP	GAG TGG TTT GAC CTT AAC GTT TGA	1878 bp-1901 bp[a] and 1945 bp-1926 bp[a]	Rosario, Symonds, Sinigalliano, Stewart, & Breitbart, 2009
		PMMV-RP	TTG TCG GTT GCA ATG CAA GT		
		PMMV-Probe	FAM-CCT ACC GAA GCA AAT G-MGB-NFQ		
Rotavirus Type A	RNA	Tampere NSP3-F	ACC ATC TWC ACR TRA CCC TCT ATG AG	Non-structural Protein 3	Zeng, et al., 2008
		Tampere NSP3-R	GGT CAC ATA ACG CCC CTA TAG C		
		Tampere NSP3 (Probe)	VIC-AGT TAA AAG CTA ACA CTG TCA AA-MGB-NFQ		
Sapovirus	RNA	SaV1F	TTG GCC CTC GCC ACC TAC	Junction of polymerase and capsid	Oka, et al., 2006
		SaV5F	TTT GAA CAA GCT GTG GCA TGC TAC		
		SaV124F	GAY CAS GCT CTC GCY ACC TAC		
		SaV1245R	CCC TCC ATY TCA AAC ACT A		
		SaV124TP (Probe)	FAM-CCR CCT ATR AAC CA-MGB-NFQ		
		SaV5TP (Probe)	FAM-TGC CAC CAA TGT ACC A-MGB-NFQ		

[a]Corresponding nucleotide position of GenBank accession number M81413 (PMMoV strain S)

1

Figure 1

Figure 1. Graphical abstract of workflow.

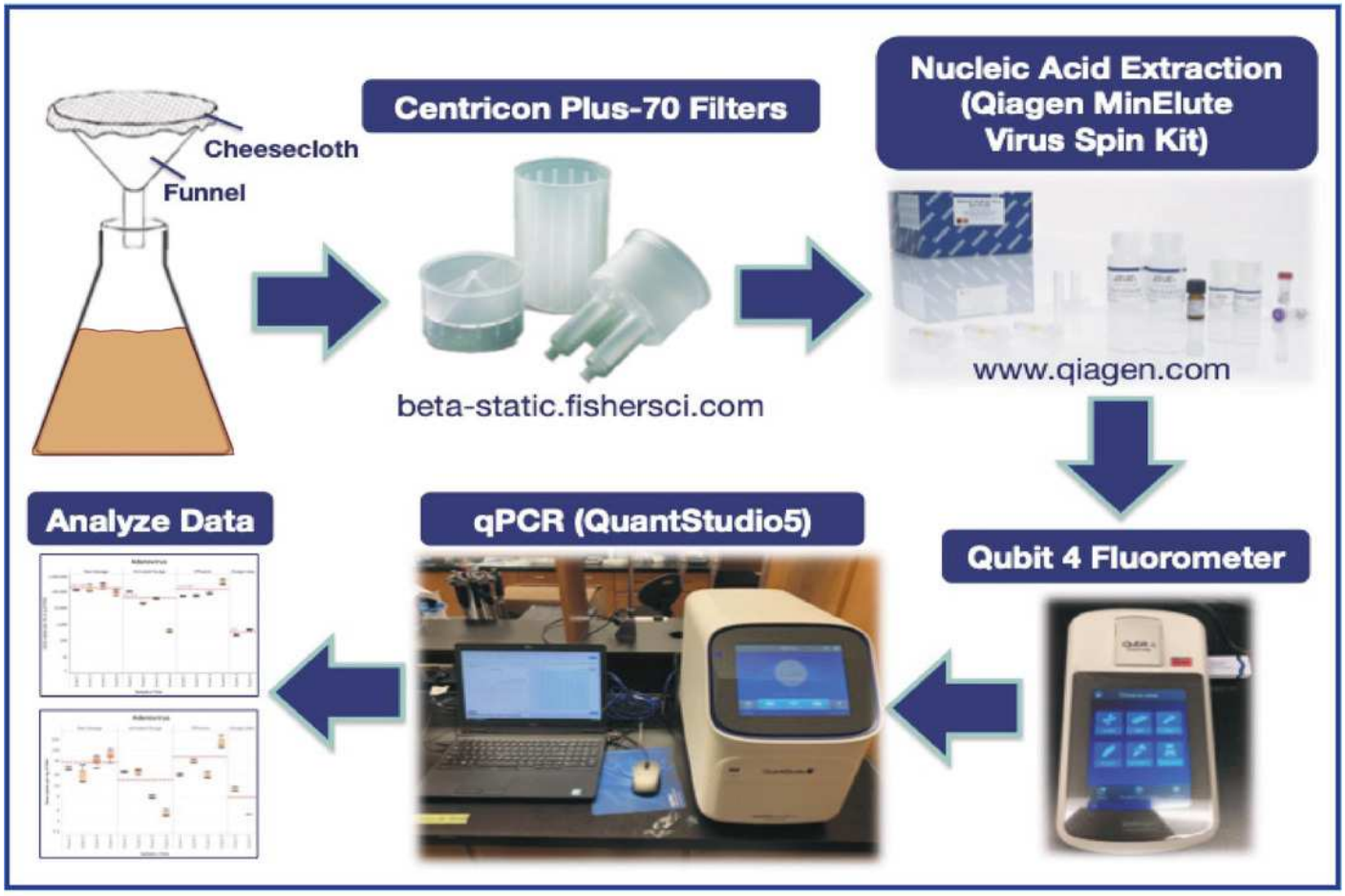

Figure 2

Figure 2. Box plots of the number of gene copies of DNA enteric viruses across each wastewater stage throughout Events 1-4.

Figures 2A and 2C visualize the number of gene copies per mL or g of sample, while *Fig. 2B* and *Fig. 2D* visualize the number of gene copies per ng of DNA. In *Fig. 2C*, this quantity was $\log_{10}$-transformed for aesthetic purposes. The unit for the SC in *Fig. 2A* and *Fig. 2C* is gene copies per g of sample.

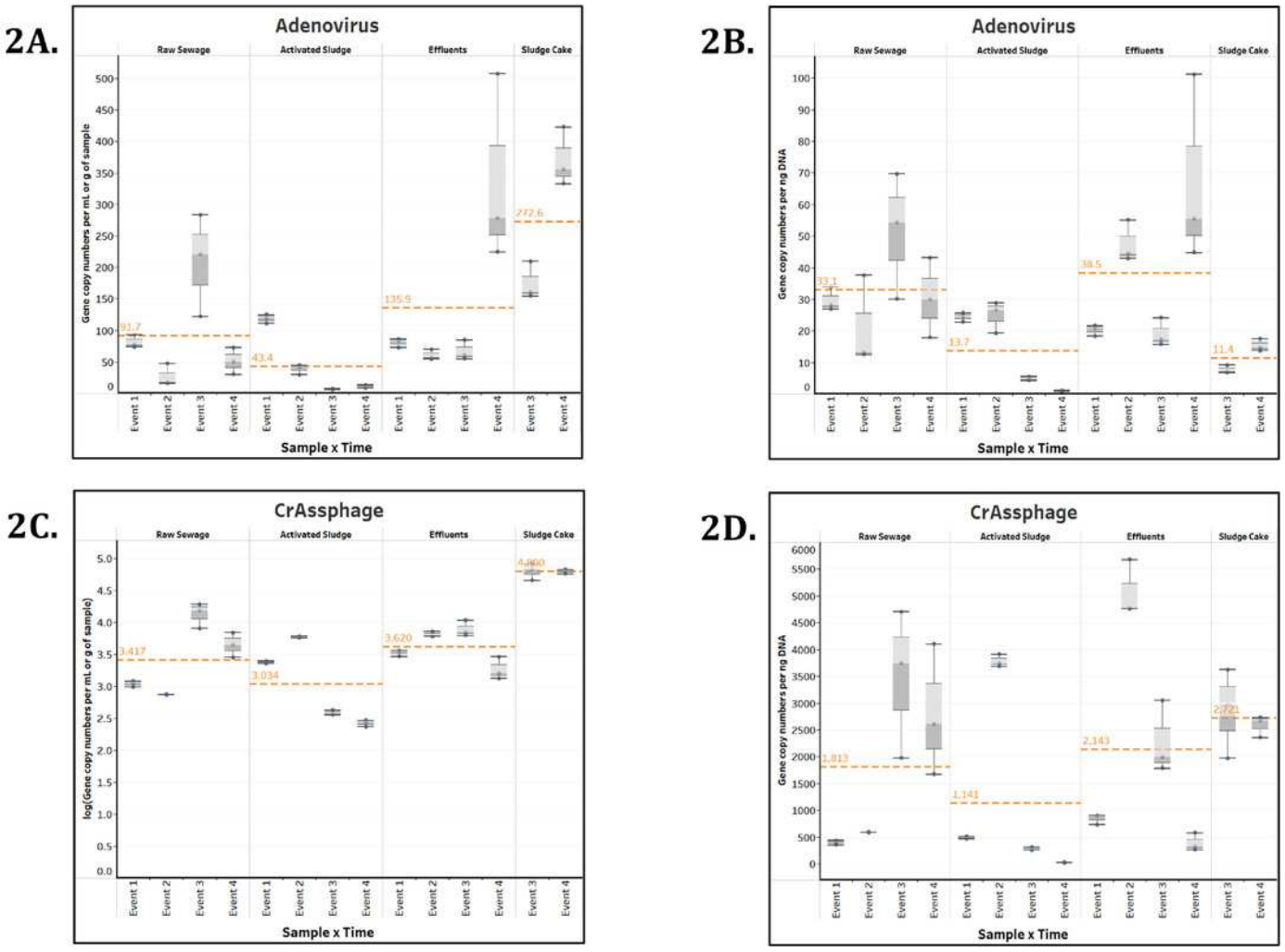

Figure 3

Figure 3. Box plots of the number of genes copies of PMMV across each wastewater stage throughout Events 1-4.

Figure 3A visualizes the number of gene copies per mL or g of sample, while *Fig. 3B* visualizes the number of gene copies per ng of DNA. Both quantities were $\log_{10}$-transformed for aesthetic purposes. The unit for the SC in *Fig. 3A* is gene copies per g of sample.

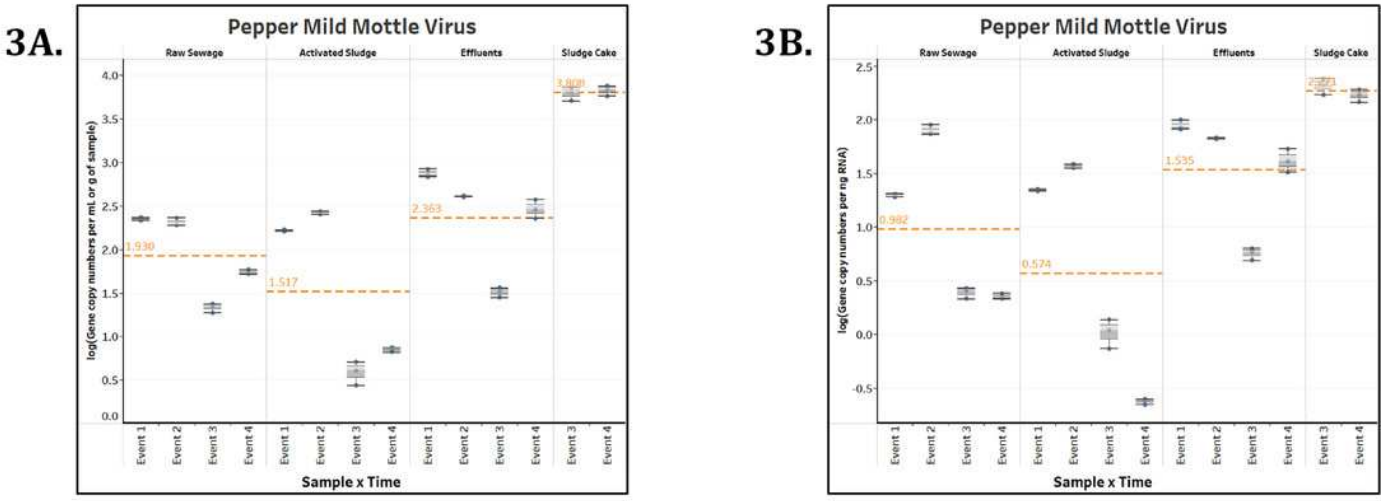

Figure 4

Figure 4. Box plots of the number of gene copies of *uidA* across each wastewater stage throughout Events 1-4.

Figure 4A visualizes the number of gene copies per mL or g of sample, while *Fig. 4B* visualizes the number of gene copies per ng of DNA. Both quantities were $\log_{10}$-transformed for aesthetic purposes. The unit for the SC in *Fig. 4A* is gene copies per g of sample.

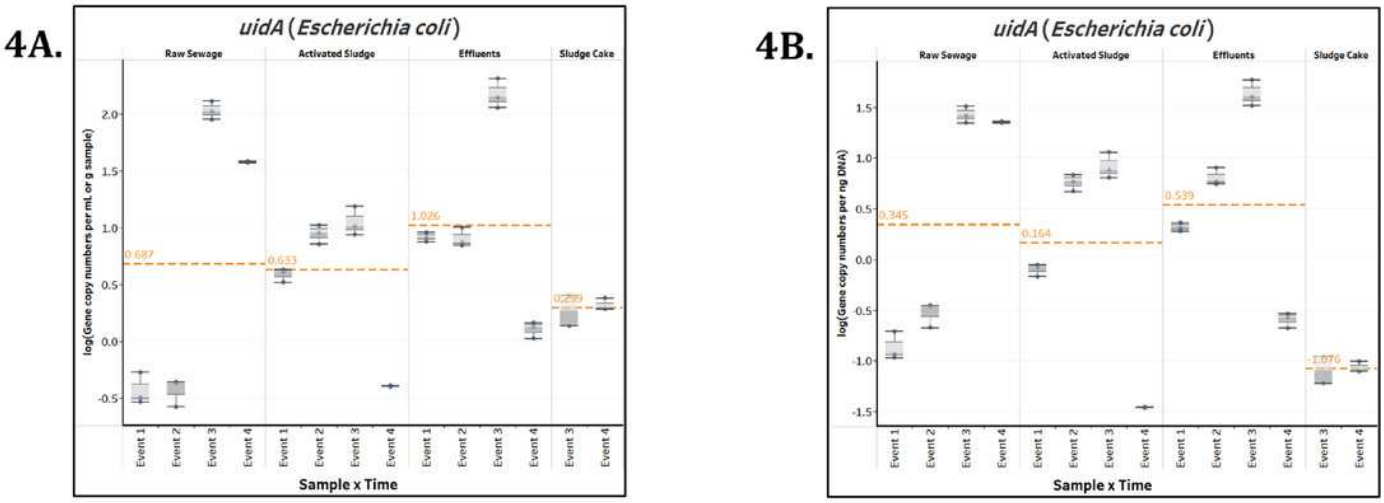

Figure 5

Figure 5. Principal Component Analysis of log_{10}-transformed EF parameters, PC1 versus PC2.

The only variable not log_{10}-transformed was precipitation due to presence of zero values.

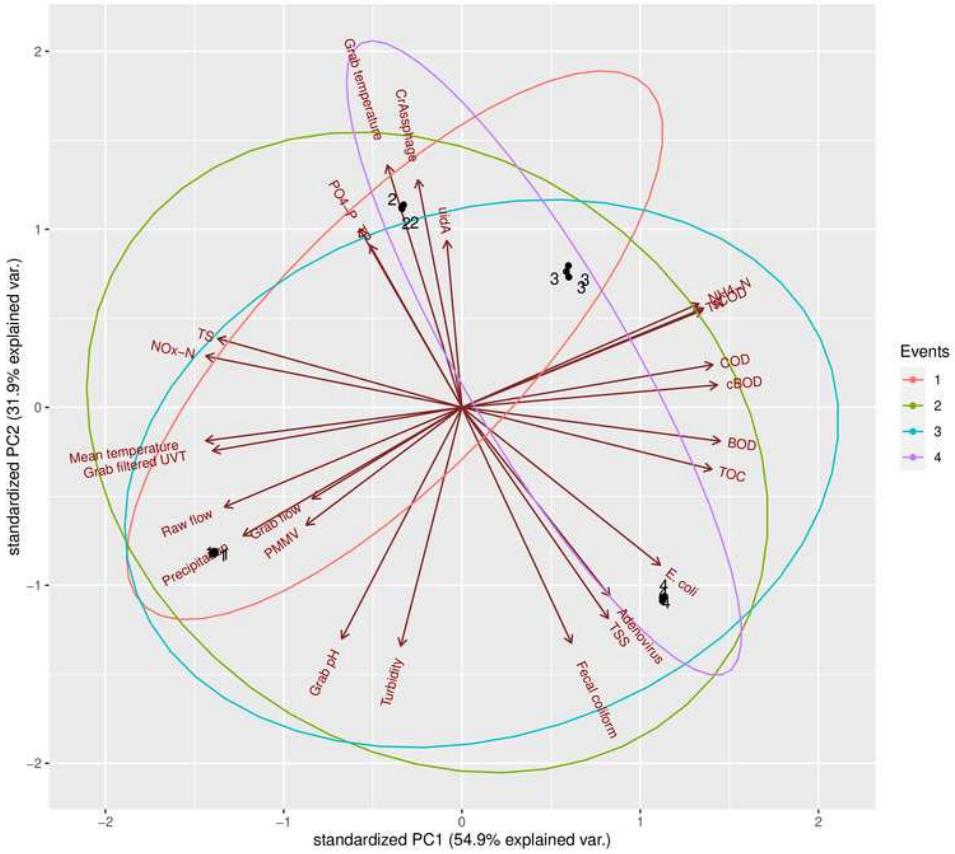

Figure 6

Figure 6. Heatmap showing Spearman's rank correlation analysis between parameters collected for EF sampling events.