water

MDPI

Supplementary Material

Membrane Bioreactor-Based Wastewater Treatment Plant in Saudi Arabia: Reduction of Viral Diversity, Load, and Infectious Capacity

Muhammad Raihan Jumat [1], Nur A. Hasan [2,3], Poorani Subramanian [2], Colin Heberling [2], Rita R. Colwell [2,3,4] and Pei-Ying Hong [1,*]

[1] Biological and Environmental Science & Engineering Division (BESE), Water Desalination and Reuse Center (WDRC), King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia; raihan.jumat@kaust.edu.sa

[2] CosmosID, Inc., Rockville, MD 20850, USA; nur.hasan@cosmosid.com (N.A.H.); poorani.subramanian@cosmosid.com (P.S.); colin.heberling@cosmosid.com (C.H.); rita.colwell@cosmosid.com (R.R.C.)

[3] Center for Bioinformatics and Computational Biology, University of Maryland Institute of Advanced Computer Studies, University of Maryland, College Park, MD 20742, USA

[4] Bloomberg School of Public Health, The Johns Hopkins University, Baltimore, MD 21205, USA

* Correspondence: peiying.hong@kaust.edu.sa; Tel.: +966-0-808-2218

1. Supplementary Text

Text S1: Measurement of Nutrient Content and COD

The chemical oxygen demand (COD) of each wastewater sample was measured by HACH TNT reagents (HACH, Germany) LCK 314 for influent and TNT 820 for effluent and chlorinated effluent. Briefly, each cuvette was shaken to dissolve any precipitated material before adding 2 mL of wastewater. The mixtures were then inverted several times to prior to digestion at 148 °C for 2 h. The cuvettes were then inverted several times before cooling down to room temperature. The COD was measured by a DR 2800 spectrophotometer (HACH, Germany). Each of the wastewater samples were characterized for the total nitrogen (TN) and non-purgable organic carbon (NPOC). Each wastewater sample was filtered through 0.22 μm WhatmanTM Puradisc 23-mm syringe filters (GE Healthcare, Little Chalfort, Buckinhamshire, UK) prior to measurement on the Organic Carbon Analyzer, TOC-VCPH (Shimadzu Scientific Instruments, Japan). A standard of 2.5 mg/L of TN and NPOC was used as a positive control and deionized water was used as negative control. Deionized water had an average reading of 0.229 ± 0.072 mg/L of NPOC and 0.0385 ± 0.0176 mg/L of TN. Each of the wastewater samples had values higher than deionized water. Ammonia in influent, as well as nitrite and nitrate in effluent streams were measured and provided by the plant operator.

Nutrient content decreased throughout the WWTP. Each of the wastewater samples were analysed for their carbon oxygen demand (COD), non-purgable organic carbon (NPOC) and total nitrogen (TN) content as an indirect indicator of organic matter present (Table S3). The COD values of the wastewater samples decreased significantly from the influent (mean: 110.8 ± 68.4 mg/L), to the effluent (mean: 12.3 ± 5.1 mg/L) and to the chlorinated effluent (mean: 14.1 ± 4.7 mg/L) as validated by a 1-tailed, paired T-test ($p < 0.05$). COD values were significantly lower in the influent of the "cold" months (68.3 ± 60.1 mg/L) than the "hot" months (164.0 ± 29.5 mg/L) ($p < 0.05$) (Table S3).The reduction of COD by the MBR process was significant in the months having a maximum temperature of 38 °C ("hot") (influent: 164.0 ± 29.5 mg/L, effluent: 12.8 ± 3.2 mg/L) than in the "cold" months (influent: 68.3 ± 60.1 mg/L, effluent: 11.9 ± 6.4 mg/L) ($p < 0.05$).

The only major difference in the NPOC was observed between the influent (6.1 ± 2.5 mg/L) and the effluent (2.5 ± 0.8 mg/L) ($p < 0.05$), suggesting that the main organic carbon reduction took place within the MBR. Chlorination did not yield any significant reduction in NPOC ($p > 0.05$). Unlike the significant reduction of COD, NPOC was significantly reduced through the MBR process in the cold

months (influent: 6.2 ± 2.7 mg/L, effluent: 2.4 ± 0.6 mg/L) than the hot months (influent: 6.3 ± 2.7 mg/L, effluent: 2.7 ± 1.2 mg/L) ($p < 0.05$) (Table S3).

The COD and NPOC reduction from the influent to the effluent also corresponds with the number of viral species detected by omics-based sequencing, where 81 species were detected in the influent, 27 in the effluent, and 25 in the chlorinated effluent (Figure S1).

Total nitrogen (TN) levels increased significantly through the MBR (influent: 12.7 ± 3.4 mg/L, effluent: 17.1 ± 2.6 mg/L) and while no significant change was observed post chlorination (14.7 ± 3.4 mg/L) ($p < 0.05$) (Table S3). The ammonium, nitrite and nitrate as provided by the plant operators, as well as the mixed liquor suspended solids (MLSS) concentrations were listed in Table S4.

Text S2: Recovery Efficiency of Tangential Flow Filtration (TFF)

7.0×10^7 and 3.0×10^7 copies of AV41 and EV71, respectively, were separately spiked into 50 mL of wastewater samples, which were filtered through a through a 0.22 μm WhatmanTM Puradisc filter (GE Healthcare, Little Chalfort, Buckinhamshire, UK). Each spiked sample was concentrated to 0.5 mL by TFF. To collect any viral particles adhering to the cassette, 50 mL of wash solution (PBS + 0.01% Tween 60) was concentrated by TFF to 0.5 mL. The retentate and wash was pooled prior to nucleic acid extraction. AV41 and EV71 copy numbers were determined through dPCR. Viral copy numbers of unspiked wastewater samples were enumerated to establish the background viral load. The TFF method had a recovery efficiency of 19.7–45.5% for AdV and 47.8–86.9% for EV across all the samples (Table S6). This finding suggests that the data presented in Figure 3 is an under representation of the actual copy number of both viruses that were present in the wastewater samples. Although recovery efficiencies were generally higher in the treated wastewater than untreated wastewater (Table S6), t-tests carried out on the recovered viruses from the different wastewater samples had p values > 0.1, suggesting that the wastewater matrix does not affect the recovery of viruses. Given the lack of significant difference in recovery yields among the influent, effluent and chlorinated effluent samples, the LRV as determined in this study, which is calculated as log10 (original concentration/treated concentrated), remains minimally biased by the different recovery rates across samples.

Text S3: Metagenomic Sequence Analysis via MetaGenID Software (CosmosID)

The system utilizes curated genome databases and a high performance data-mining algorithm that rapidly disambiguates hundreds of millions of metagenomic sequence reads into the discrete microorganisms engendering the particular sequences. The pipeline has two separable comparators. The first consists of a pre-computation phase and a per-sample computation. The input to the pre-computation phase is a curated reference microbial database, and its output is a whole genome phylogeny tree, together with sets of fixed length n-mer fingerprints (biomarkers) that are uniquely identified with distinct nodes of the tree. The second per-sample, computational phase searches the hundreds of millions of short sequence reads against the fingerprint sets. The resulting statistics are analyzed to give fine-grain composition and relative abundance estimates at all nodes of the tree. Overall classification precision is maintained through aggregation statistics.

Text S4: Principle behind Digital PCR

The Clarity dPCR system works with a fluorescently tagged probe to aid in template quantification. After the reaction mix is prepared, it is partitioned into 10000 sub-reactions on a high density chip using the auto-Loader and sealing enhancer. Each partitioned reaction is then subjected to thermal cycling, after which the fluorescence from each partition is detected by the reader. Ideally, each partition should contain a maximum of one DNA molecule and partitions with the target template DNA would be amplified, yielding a positive fluorescence signal. By counting the number of positive fluorescence signals, the concentration of the target template DNA can be inferred. To account for the possibility of more than one DNA molecule being included into a partition, the

following equation which takes into account Poisson distribution of the DNA molecules, is used in the algorithm for calculating the initial copies of template DNA:

$$\text{DNA Copy Number} = -\frac{1000}{\text{Vd}}\ln\left(1-\left(\frac{\text{P}}{\text{R}}\right)\right) \text{ copy number per } \mu\text{L.}$$

where Vd is the mean partition volume (nL), P is the number of partitions containing the amplified product and R is the number of partitions analysed [1].

This partitioning not only compartmentalizes the DNA template but any potential PCR inhibitors as well, making dPCR less susceptible to potential PCR inhibitors than conventional qPCR [2,3].

To achieve an even partitioning of nucleic acid molecules on the high density chip, nucleic acids from the sludge was diluted 2- to 100-fold, and the influent was diluted 2- to 20-fold. Effluent and chlorinated effluent nucleic acids were undiluted, as signals from these samples were not saturating. The optimum dilution required for each sample was determined by observing for the strongest and most consistent signal over the technical replicates, to ensure that potential PCR inhibitors in the sample had minimal effect. The average values of the technical replicates were reflected in Figure 3, with error bars corresponding to the standard deviation. Each dPCR run included a Non-Template Control (NTC) with sterile water as template to determine the background fluorescence, which was then accounted for by the software.

To determine the sensitivity of the platform, DNA was extracted from 50 μL of Adenovirus 40 (AV40, VR-931, ATCC). The concentration of DNA was quantified with a Qubit 2.0 fluorometer (Thermo Fisher Scientific, Carlsbad, CA) revealing a concentration of 6.5×10^7 copies/μL. AV40 DNA was serially diluted to achieve 6.5×10^4, 6.5×10^2, 6.5×10^0 and 6.5×10^{-2} copies/μL as template for dPCR. A similar extraction was carried out for 50 μL of Enterovirus 71 (EV71, VR1775, ATCC) revealing a concentration of 3.1×10^8 copies/μL. EV71 RNA was diluted to achieve 3.1×10^5, 3.1×10^3, 3.1×10^2 and 3.1×10^1, copies/μL. The observed results from both of these dPCR runs were plotted on an expected vs observed graph. The amount of correction required for each assay was inferred by where the line intersects y = 1 (Figure S2).

To optimize the protocol and test the sensitivity of dPCR, DNA was extracted from AV40 and actual copy numbers were estimated by dividing its concentration by the molecular weight of the genomes. The nucleic acid template was serially diluted and ran through the dPCR. "Observed v.s. Expected" copy numbers were subsequently plotted (Figure S2). dPCR was able to detect template concentrations of AdV as low as 6.5 copies/μL of DNA template which translates to 13.5 copies/L of wastewater. The observed copy numbers were detected at approximately 3.8 times lower than the expected viral concentration. This value was factored into the quantification of AdV by dPCR (Figure 3A).

RNA extracted from EV 71 (EV71, ATCC VR- 1775) was quantified and the copy numbers present were calculated by dividing the concentration by the molecular weight of the genome. dPCR was able to detect EV as low as 31.1 copies per μL of RNA template, translating to 64.7 copies/L of wastewater. Serially diluted RNA were used as template for dPCR and the "Observed v.s. Expected" line was plotted. The average of the values observed were skewed by a factor of 74.5 than the expected. This factor was included in the calculation for the EV load in the wastewater samples (Figure S2).

Text S5: Cell Culture Infection with Sludge Samples

Sludge samples with diluted to 50% in DMEM supplemented with 1× Penicillin/Streptomycin and 1× Antibiotic AntiMycotic reagent (Corning). Sludge mixtures were vortexed for 10 s and treated in a water bath sonicator at 40 kHz for three pulses. Mixtures were centrifuged at 200 g for 10 minutes to remove the colloidal matter. Supernatants were diluted 2×, 10× and 100× in DMEM and 500 μL was inoculated onto confluent HEp-2 cells in flat-surfaced glass culture tubes for 1 h at 37 °C with 5% atmospheric CO_2. As a negative control, cells were inoculated with 500 μL of PBS instead of sludge. Tubes were agitated every 15 minutes to ensure even distribution of viruses. The inoculum was removed and the cells were washed with 1× PBS for twice before adding DMEM supplemented with 2% FBS, 1× Penicillin/Streptomycin and 1× Antibiotic AntiMycotic reagent (Corning). The tubes were then placed back for incubation at 37 °C with 5% atmospheric CO_2. The cells were observed under a

light microscope daily for CPE. Once CPE was observed, cells were harvested and processed for diagnostic immunofluorescence, as described in section 2.5 of main text. While the recovery efficiency of this method of viral elution is unknown, sludge samples treated in this manner showed a large increase in viral infection as compared to un-sonicated and undiluted samples which also tend to result in fungal contamination.

2. Supplementary Tables

Table S1: Average Minimum and Maximum Temperatures in Saudi Arabia during the Sampling Period

Months with maximum temperatures >38 °C are highlighted in orange. 1st October of this sampling year also experienced high temperature >38 °C (orange background not shown). Information adapted from www.weather-and-climate.com.

Month	Min Temp. (°C)	Max Temp. (°C)
Jan	8.1	20.3
Feb	10.4	22.1
Mar	14.4	27.0
Apr	18.0	32.0
May	24.3	37.8
Jun	26.1	41.0
Jul	27.0	41.9
Aug	27.0	41.9
Sep	24.3	39.6
Oct	19.4	33.8
Nov	14.4	27.0
Dec	9.5	21.2

Table S2: Coliform Counts of the Chlorinated Effluent

The coliform counts in the irrigation pump station where the chlorinated effluent was stored prior to irrigation was determined by the KAUST WWTP.

Sample	Coliform Count per 100mL (Chlorinated Effluent)	Max Temp. *
1st July 2015	nil	>38 °C
19th August 2015	nil	
8th September 2015	nil	
1st October 2015	0.28	
22nd October 2015	0.34	<38 °C
23rd November 2015	0.39	
27th December 2015	0.48	
26th January 2016	0.21	
31st March 2016	nil	

* The maximum temperature observed in the sampling months, indicating the season the sample was collected in.

Table S3: COD, NPOC and TN Measurements of the Wastewater Samples Collected in this Study

COD: Chemical Oxygen Demand; NPOC: Non-purgable organic carbon; TN: Total Nitrogen; – denotes no lab measurement taken. Collection dates were classified according to their maximum temperature with >38 °C as "hot" and <38 °C as "cold".

Sample	Date	COD	NPOC	TN	Max Temp.
		(mg/L)			
Influent	1st July 2015	170	9.4	16.0	>38 °C
	19th August 2015	124	4.3	9.75	
	8th September 2015	195	5.1	11.4	
	1st October 2015	167	-	-	
	22nd October 2015	167	5.3	8.0	<38 °C
	23rd November 2015	84.8	2.9	14.2	
	27th December 2015	28.7	9.2	18.2	
	26th January 2016	33.7	8.4	12.7	
	31st March 2016	27.3	4.4	11.0	
Effluent	1st July 2015	15.9	2.0	17.6	>38 °C
	19th August 2015	-	-	-	
	8th September 2015	9.52	4.0	18.9	
	1st October 2015	13.1	2.0	17.2	
	22nd October 2015	11.1	3.2	16.4	<38 °C
	23rd November 2015	7.61	2.7	14.6	
	27th December 2015	5.82	1.8	16.5	
	26th January 2016	22.2	2.6	22.1	
	31st March 2016	12.8	1.8	13.4	
C. Effluent	1st July 2015	15.8	2.1	14.3	>38 °C
	19th August 2015	10.8	1.8	15.7	
	8th September 2015	14.2	5.5	17.9	
	1st October 2015	20.4	5.0	12.6	
	22nd October 2015	13.1	3.7	11.1	<38 °C
	23rd November 2015	9.06	6.3	11.4	
	27th December 2015	7.66	1.9	17.3	
	26th January 2016	21.6	3.0	20.5	
	31st March 2016	14.3	1.7	11.6	

Table S4: MLSS, Ammonium and Nitrate Concentrations of Wastewater Samples

MLSS: Mixed liquor suspension solids present in the activated sludge (mg/L). Ammonium (NH_3–N) and Nitrate (NO_3–N) concentration of the influent from each of the wastewater samples (mg/L).

	MLSS (mg/L)				NH_3–N (mg/L)	NO_3–N (mg/L)
	Stream A	**Stream B**	**Stream C**	**Stream D**	**Eq. Tank**	**Eq. Tank**
1st July 2015	10,367	12,700	13,033	10,200	12.2	2
19th August 2015	10,400	12,200	13,067	10,133	9.9	1.5
8th September 2015	11,434	13,100	11,633	12,933	11.2	1.11
1st October 2015	9167	12,667	10,567	9833	9.9	2.4
22nd October 2015	9600	12,533	9400	9167	11.6	2.5
23rd November 2015	9800	13,900	108,667	9234	15	2.5
27th December 2015	-	15,367	122,667	9467	16.1	3.2
26th January 2016	-	14,067	13,433	10,267	18.5	2.8
31st March 2016	-	14,067	12,200	9000	15	1.5

Table S5: Viral Screening on HEK 293T, HeLa, HEp-2, Vero and MRC-5 cells

Listed cells were infected with Adenovirus 40 (ATCC VR-931), Adenovirus 41 (ATCC VR-930), Enterovirus 68 (ATCC-VR1825), Enterovirus 70 (VR-836) and Enterovirus 71 (VR-1775). Cells were harvested upon displaying cytopathic effect (CPE) or after 14 days post-infection. Cells were assayed with anti-adenovirus-FITC or anti-pan enterovirus antibody counterstained with secondary anti-mouse FITC antibody.

+: cells which showed CPE and a positive signal upon immunostaining; -: cells which either did not show CPE and/or did not show a positive signal on immunostaining; -/+: cells which showed significant CPE upon infection but did not exhibit any signal upon immunostaining.

Cell Line	**AV40**	**AV41**	**EV68**	**EV70**	**EV71**
HEK 293T	+	+	-	+	+
HeLa	+	+	+	+	+/-
HEp-2	+	+	+	+	+
Vero	+/-	+	-	-	+/-
MRC-5	-	-	-	+	-

Table S6. Adenovirus and Enterovirus Recovery Efficiency of Tangential Flow Filtration (TFF)

Copy numbers of spiked Adenovirus (AV41) and Enterovirus (EV71) as well as the corresponding recovered copy numbers. Recovery efficiency is expressed as a percentage of the spiked copy number.

	AV41 Copy Numbers			**EV71 Copy Numbers**		
	Spiked	**Recovered**	**Recovery Efficiency**	**Spiked**	**Recovered**	**Recovery Efficiency**
Chlorinated Effluent		2.3×10^7	33.2%		2.6×10^7	86.9%
Effluent	7.0×10^7	1.4×10^7	19.7%	3.0×10^7	1.4×10^7	47.8%
Influent		3.2×10^7	45.5%		1.7×10^7	56.6%

3. Supplementary Figures

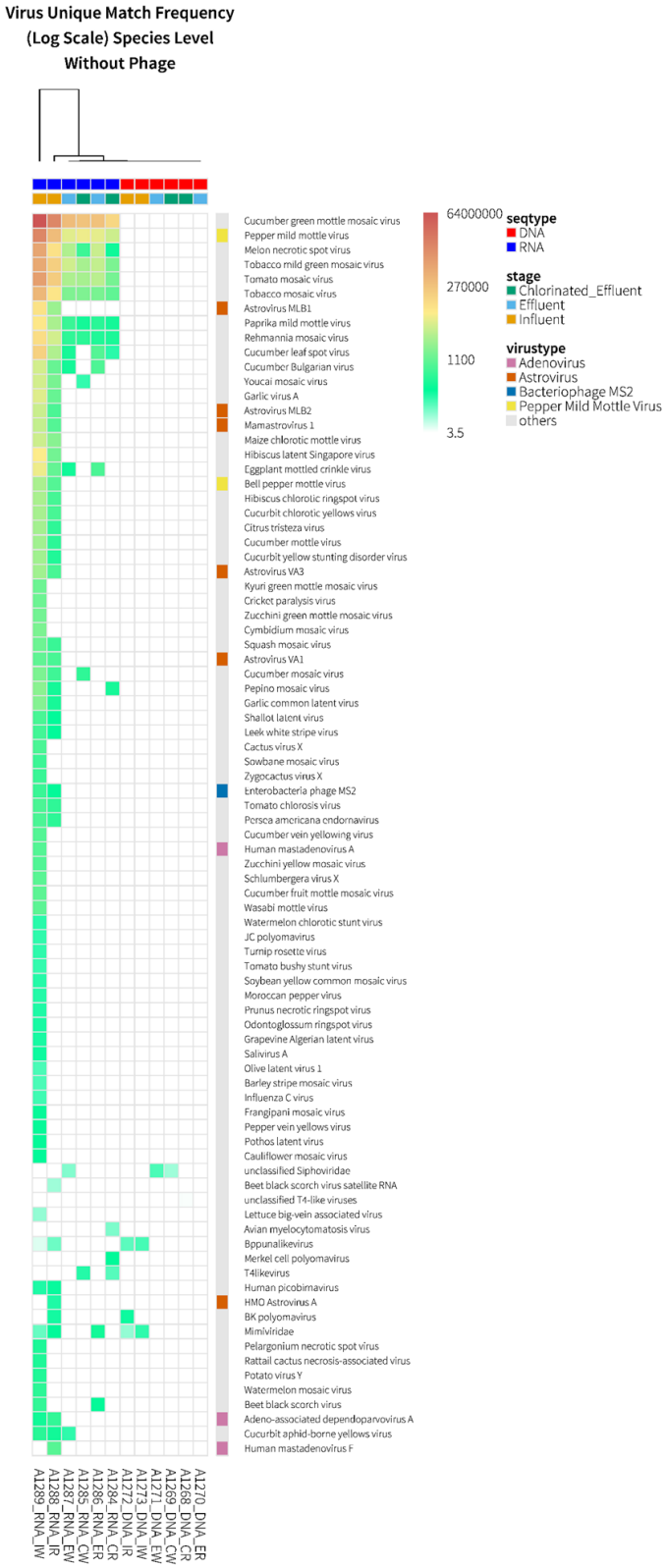

Figure S1. Heatmap of virus species detected by virome analysis of wastewater samples (excluding phages). Coloured scale bar represents unique hit frequency for each species. IR = Influent-retentate, IW = Influent wash, ER = Effluent-retentate, EW = Effluent-wash, CR = Chlorinated Effluent-retentate, CW = Chlorinated Effluent-wash.

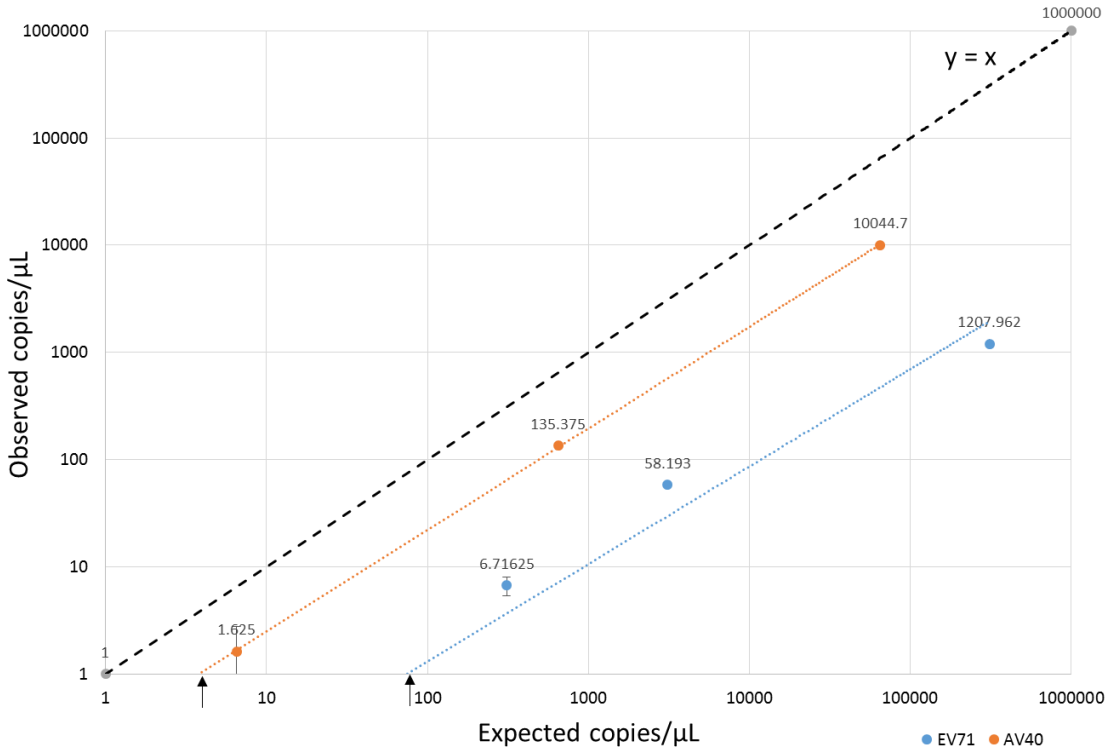

Figure S2. "Observed versus Expected" Enterovirus (EV71) (blue line) and Adenovirus (AV40) (orange line).