Detection and characterization of human pathogenic viruses circulating in community wastewater using multi target microarrays and polymerase chain reaction

Mark V. M. Wong, Syed A. Hashsham, Erdogan Gulari, Jean-Marie Rouillard, Tiong Gim Aw and Joan B. Rose

ABSTRACT

Sewage pollution remains the most significant source of human waterborne pathogens. This study describes the detection and characterization of human enteric viruses in community wastewaters using cell culture coupled with multiple target microarrays (with a total of 780 unique probes targeting 27 different groups of both DNA and RNA viruses) and polymerase chain reaction (PCR) assays. Over a 13-month sampling period, RNA viruses (astroviruses and enteroviruses) were more frequently detected compared to DNA viruses (adenoviruses, particularly type 41 and BK polyomavirus). Overall, many more viruses were shed during the winter months (December–February) compared to the summer months. Exploration of the multiple types of enteric viruses particularly in winter months identified much more significant prevalence of key viral pathogens associated with sewage pollution of the water environment than previously realized and seasonal disinfection used in some parts of the world may lead to a seeding of ambient waters. Molecular characterization of pathogenic viruses in community wastewater will improve the understanding of the potential risk of waterborne disease transmission of viral pathogens.

Key words | enteric viruses, fecal contamination, microarray, PCR, sewage

Mark V. M. Wong
Department of Crops and Soil Sciences,
A570 Plant and Soil Sciences,
Michigan State University,
East Lansing, MI 48824,
USA

Tiong Gim Aw
Joan B. Rose (corresponding author)
Department of Fisheries and Wildlife,
3 Natural Resources, Michigan State University,
East Lansing, MI 48824,
USA
E-mail: *rosejo@msu.edu*

Syed A. Hashsham
Department of Civil and Environmental Engineering, A126 Engineering Research Complex,
Michigan State University,
East Lansing, MI 48824,
USA

Erdogan Gulari
Department of Chemical Engineering,
2300 Hayward St, University of Michigan,
Ann Arbor, MI 48109,
USA

Jean-Marie Rouillard
Department of Chemical Engineering,
3146 H.H. Dow, University of Michigan,
Ann Arbor, MI 48109,
USA

INTRODUCTION

Enteric viruses are important biological hazards in wastewater and frequently impair water quality due to high level of excretion by host, stability in water environments and resistance to wastewater treatment (Fong & Lipp 2005). These human pathogens are a major concern for the water industry including drinking water, wastewater and reclaimed water as well as for those involved in watershed protection and safety of recreational waters. While vaccines are available for some enteric viral diseases such as severe diarrhea caused by rotavirus, the resurgence of poliovirus in the developing world and the rapid spread of more virulent types of noroviruses have demonstrated that characterization and detection remain important for defining public health strategies in the future (Tambyah 2004; CDC 2006; Wong *et al*. 2007b).

Currently, it is estimated that over 150 types of viruses may be excreted in human and animal waste (Wong *et al*. 2007b). However, a true catalog of the numbers and types of viruses remains elusive. This is partly due to the lack of high throughput screening tools for human pathogenic viruses. There are many groups of enteric viruses in the *Caliciviridae, Picornaviridae, Reoviridae* and *Adenoviridae*

doi: 10.2166/wh.2013.322

families which have been shown to cause diarrhea, neurological and respiratory diseases worldwide. They have also been associated with other chronic conditions (e.g., hepatitis and myocarditis) (Leclerc *et al*. 2002; Oberste & Pallansch 2003).

At present, cell culture is the gold standard for detecting infectious viruses in water, but inherent problems exist, as the methods are time consuming, labor intensive and cell culture alone cannot provide information about different types of viruses present in the sample. To circumvent such drawbacks, cell culture integrated with polymerase chain reaction (PCR) has been used for detection of infectious pathogenic viruses in water although it is highly dependent on the virus type. PCR has improved the sensitivity and specificity of viral analysis of environmental samples. However, the limited capacity of PCR for multiplexing hinders its practical application as a primary high throughput screening tool for pathogenic viruses in water. Moreover PCR is incapable of targeting new and emerging infectious viruses for which no known primers exist.

Microarrays are arrays of spots on specially prepared glass or silicon surfaces containing high density immobilized DNA oligonucleotide probes that enable the simultaneous interrogation of thousands of genes in a single assay via nucleic acid hybridization (Aw & Rose 2012). Current limitations of microarrays include the lack of sufficient sensitivity to detect the low concentration of the target pathogens in relation to non-targets in the environment and the presence of inhibitory substances in the environment that compromise sample labeling and hybridization efficiency. Multi target microarrays have been developed to detect and characterize viral pathogens from clinical samples allowing for broad screening and pathogen discovery (Wang *et al*. 2003; Korimbocus *et al*. 2005; Lin *et al*. 2007; Wong *et al*. 2007a). However, these methods have yet to be applied to real-world water and wastewater samples for the detection of waterborne viral pathogens.

It has been suggested that rapid increases in population and the subsequent increase in sewage discharges contribute to surface water quality impairment worldwide (Rose & Molloy 2007). However, the diversity of human viral populations in sewage is not well understood since current detection approaches have been focused on a few specific, well-known virus groups (mainly single-target approach) (Girones *et al*. 2010). The aim of this study was to detect and characterize pathogenic viruses in municipal sewage using cell culture coupled with multi target microarrays. Samples were also assayed by PCR to compare to the microarray results. The characterization of circulating viruses within a community was addressed broadly as an approach for biomonitoring and the screening of both community health and water pollution-associated disease potential.

METHODS

Wastewater sampling and virus cell culture

Thirteen 6-liter samples of untreated sewage were collected monthly between August 2006 and September 2007. The viruses in the sewage were concentrated by organic flocculation and centrifugation as described by the USEPA information collection rule (USEPA 1995). Concentrated wastewater samples were frozen until processed on cell culture. Poliovirus 1 LSc strain, Adenovirus type 40 and type 41 were obtained from the American Type Culture Collection (ATCC # VR-59, VR-931 and VR-930, respectively) and used as controls. 12 ml of virus concentrate was inoculated onto African Green Monkey Kidney cell line (BGM) (USEPA 1995). Viral infections were allowed to proceed until the onset of cytopathic effects (typically within 24 hr for poliovirus and two days for sewage samples). To recover the viruses, the cells were freeze-thawed three times to disrupt the cell integrity and release the viral particles into the supernatant, additional cell-attached viruses were lifted off the flask surface by mechanical scraping. Virus particles were recovered and concentrated from the media by centrifugation through an Amicon Ultra 100 K centrifugation column (Millipore, Billerica, MA). Virus nucleic acid was extracted using the QIAamp viral RNA mini kit (Qiagen Inc, Valencia, CA) for both DNA and RNA viruses following the manufacturer's instructions.

Probe design and construction

Viral sequence data were obtained from GenBank. Seventeen RNA virus groups and 10 DNA virus groups were identified as targets (Table 1). Genome sequences were

parsed through the probe design software Oligoarray 2.1 (Rouillard *et al*. 2003). The software analyzes the complete virus genome and selects the optimal probes with the least amount of secondary structure and lowest likelihood to cross hybridize with other sequences provided. The probes were designed to conform to the following specifications: Max $T_m = 75\ ^\circ C$ (except for torovirus $= 80\ ^\circ C$), Min $T_m = 70\ ^\circ C$ (except for torovirus $= 65\ ^\circ C$), Max GC = 60%, Min GC = 40%, Max temp for secondary structure $= 45\ ^\circ C$, Max temperature for cross hybridization $= 45\ ^\circ C$. Probes were

Table 1 | List of viral targets, type of genomes and Genbank accession numbers

Virus	Type of genome	Accession no.	Seq. length (bp)
Hepatitis A virus[a]	+ ssRNA no DNA state	NC_001489	7,478
Hepatitis E virus[a]	+ ssRNA no DNA state	NC_001434	7,176
Human adenovirus A[a]	dsDNA	NC_001460	34,125
Human adenovirus B[a]	dsDNA	NC_004001	34,794
Human adenovirus C[a]	dsDNA	NC_001405	35,937
Human adenovirus D[a]	dsDNA	NC_002067	35,100
Human adenovirus E[a]	dsDNA	NC_003266	35,994
Human adenovirus type 40	dsDNA	NC_001454	34,214
Human adenovirus type 41[a]	dsDNA	DQ315364	34,189
Norwalk virus[a]	+ ssRNA no DNA state	NC_001959	7,654
Sapovirus[a]	+ ssRNA no DNA state	NC_010624	7,458
Human enterovirus A[a]	+ ssRNA no DNA state	NC_001612	7,413
Human enterovirus B[a]	+ ssRNA no DNA state	NC_001472	7,389
Human enterovirus C[a]	+ ssRNA no DNA state	NC_001428	7,401
Human enterovirus D[a]	+ ssRNA no DNA state	NC_001430	7,390
Human enterovirus E[a]	+ ssRNA no DNA state	NC_003988	7,374
Poliovirus[a]	+ ssRNA no DNA state	NC_002058	7,440
Rotavirus A[b]	dsRNA virus	AB077766	2,359
		AB071404	1,062
Rotavirus B[b]	dsRNA virus	AY539857	2,306
		AY539856	814
Rotavirus C[b]	dsRNA virus	AB008670	2,283
		AB008671	1,063
Coronavirus[a]	+ ssRNA no DNA state	NC_002645	27,317
Cytomegalovirus (HH5)[a]	dsDNA virus	NC_006273	235,645
Torovirus[c]	+ ssRNA no DNA state	AF159585	1,251
		AF024539	219
Picobirnavirus[d]	dsRNA virus	AF246940	1,674
		AF246941	1,572
Astroviruses[a]	−ssRNA	NC_001943	6,813
JC polyomavirus[a]	ds DNA	NC_001699	5,130
BK polyomavirus[a]	ds DNA	NC_001538	5,153

[a]Complete genome.
[b]VP4 and VP7.
[c]Hemagglutinin esterase and nucleocapsid protein mRNA.
[d]RNA dependent RNA polymerase and segment 1 gene.
Abbreviations: +/− ss: positive/negative single stranded, ds: double stranded.

designed from the positive strand of the genetic sequence. A total of 780 specific probes were designed (approximately 30 probes per viral family target).

The microarrays used in this study were synthesized using the in situ synthesis technology developed by the University of Michigan as described previously (LeProust *et al.* 2000; Gao *et al.* 2001; Komolpis *et al.* 2002). Its format consisted of seven blocks of 1,000 probes which allowed for a maximum of 7,000 probes. Five copies for each of the 780 virus probes were randomly assigned to be synthesized on the chip. The remaining features on the microarray were either left blank to accommodate future probes or used for the control probe sequences which served as a quality check during fabrication (Ctl_1: 5′-TATGTGTTAGCTGACCATCCAT-3′ and Ctl_2: 5′-AATATCGACGAAGCATTATTTC-3′). Multiple copies of probes were used to provide technical replication of the signals.

Preparation of samples for hybridization

Viral RNA was labeled by a semi-random primed labeling and Sensiscript III reverse transcriptase as described by Wang *et al.* (2003). Briefly, 0.5 μg of viral nucleic acid was used as a template for the generation of cDNA with a discrete 5′ terminal consisting of the sequence (5′-GTTTCCCAGTCACGATC-3′) using the semi-random primer A (5′-GTTTCCCAGTCACGATCNNNNNNNNN-3′). Primer B (5′-GTTTCCCAGTCACGATC-3′) and Qiagen Hotstart Taq polymerase were then used to amplify the generated cDNA and label it with amino-allyl dUTP for 40 cycles using the following profile: 30 s at 94 °C, 30 s at 40 °C, 30 s at 50 °C, 60 s at 72 °C. To label viral DNA, 0.5 μg of viral DNA was first digested with 2U of DPNII restriction endonuclease at 37 °C for 1 hr. The digested DNA was incubated at 37 °C for 2 hr with Klenow enzyme and Primer A to generate complementary DNA with a discrete 5′ terminal. Primer B and Qiagen Hotstart Taq polymerase were then used to amplify the DNA and label it with amino-allyl dUTP for 40 cycles using the following profile: 30 s at 94 °C, 30 s at 40 °C, 30 s at 50 °C, 60 s at 72 °C.

Labeled DNA and RNA were separately coupled with Cyanine dye 3 and Cyanine dye 5, respectively. This reaction was carried out in the absence of light and using 1 M sodium bicarbonate (pH 9.5) as a coupling buffer with a 1 hr incubation. Following coupling, uncoupled dye was removed using the QIAgen PCR purification kit and the labeled viral DNA and RNA were dessicated in a DNA 120 SpeedVac (Thermo Fisher Scientific, MA, USA) for 1.5 hr. Labeled samples were kept dry at −20 °C until hybridization.

Microarray hybridization

Microarray hybridization was performed as described previously (Wick *et al.* 2006). The microarrays were hybridized and washed in a M-2 microfluidic station (Invitrogen, Carlsbad, CA, formerly Xeotron Corporation, Houston, TX) using a flow rate of 400 μl/min. After hybridization, the arrays were washed with a series of buffers to remove unbound and non-specifically hybridized DNA. In addition, a series of washes and scans were performed at 1 °C incremental temperatures (25 to 35 °C; 1.4 min each; 500 μl/min flowrate; 10 mM Na_2HPO_4, 5 mM EDTA, wash buffer, pH 6.6) to generate a 'melting' profile allowing the identification of non-specific hybridization through their atypical melting profile. Scanning was done with a GenePix 4000B laser scanner (Axon Instruments, Union City, CA). All solutions were filtered through a 0.22 μm filter to prevent clogging of the microfluidic channels. The high stringency wash buffer was degassed under vacuum. Scanned images were analyzed using the GenePix Pro 5.0 software supplied with the scanner.

Data analysis using DetectiV

The DetectiV software package was used to visualize, normalize and test the significance of the hybridization data (Watson *et al.* 2007) and to generate bar plots of the hybridization signals against the median signal values for each sample. A combination of a statistically significant p values arising as a result of the pairwise *t*-test, and a high $\log_2$ mean signal intensity was proposed by Watson *et al.* (2007) as the basis by which the software can be used to identify the presence of a target virus. DetectiV was also used to carry out statistical *t*-test comparisons of the hybridization signal values between different virus groups in order to determine which groups of virus targets had statistically significant probe signals. The software has previously been used to detect and genotype lyssaviruses as well as to correctly predict the presence of a

target virus in 67 out of 68 public microarray datasets (Watson *et al*. 2007; Gurrala *et al*. 2009). Using the software package DetectiV, bar plots, illustrating the hybridization intensities for the various samples were derived and served as a visual indication of the probable viruses as well as the extent of signal noise present in a given sample.

Steps taken to reduce cross hybridization and non-target hybridization for water samples

One of the challenges in use of a microarray has always been the validation of the signals, particularly when applied to environmental samples. It is impractical to evaluate all targets through seeded experiments, thus the following approaches were adopted in this study. Firstly, the probes used in these experiments were between 25 to 28 base pairs in length, which made them more specific than the 40 base pair probes commonly used in conventional clinical pathogen detection arrays (Koltai & Weingarten-Baror 2008). Secondly, each probe was filtered using the OligoArray 2.1 software to have a low propensity to cross hybridize with the genomes of non-target viruses. Also, multiple probes (in most cases, 30) derived from the entire virus genome and not just a small segment of the genome, were designed to reduce the chance that a significant number of probes for a particular virus group would all show false positive signals. In addition, each probe was represented on the array by 5 copies, randomly dispersed throughout the surface of the chip to ensure that the results could not be affected by localized scanning errors. Furthermore, to ensure a high degree of specificity in the hybridization, stringent hybridization and wash protocols for the microarray were employed as described.

Group-specific PCR detection of viral targets and sequencing confirmation of targets

Cell culture extracts were also evaluated with group-specific PCR for key viruses. PCR primers used are reported in Table 2. PCR was performed using the QIAGEN OneStep RT-PCR kit for RNA viruses and the QIAGEN Hotstar Taq kit for DNA viruses. Reverse transcription was carried out as follows: 2 μl of OneStep RT-PCR enzyme mix, 1.5 mM $MgCl_2$, 1× PCR buffer, 1× Q solution, 0.5 μM of each primer, 0.5 mM of each dNTP and 0.5 μg template. Thermocycler settings for the reverse transcription PCR were as follows: 50 °C, 30 min first strand synthesis, 95 °C, 15 min initial denaturation to activate the Hotstart Taq polymerase (Qiagen, Valencia, CA), 95 °C, 0.5 min denaturation, 57 °C, 0.5 min annealing and 72 °C, 0.5 min elongation for 35 cycles followed by a final elongation step of 72 °C for 5 min. The PCR reaction was carried out as follows: 1 unit of Hotstart Taq polymerase, 1.5 mM $MgCl_2$, 1× PCR buffer, 1× Q solution, 0.5 μM of each primer, 0.5 mM of each dNTP and 0.5 μg of template. Thermocycler settings for the PCR reaction were as follows: 95 °C, 15 min initial denaturation to activate the Hotstart Taq polymerase (Qiagen, Valencia, CA), 95 °C, 0.5 min denaturation, 57 °C, 0.5 min annealing and 72 °C, 0.5 min elongation for 35 cycles followed by a final elongation step of 72 °C for 5 min.

PCR products were purified using the Qiagen QIAquick PCR purification kit and eluted with molecular grade water. 10 ng of purified PCR product and 30 picomoles of each primer were provided to the Michigan State University Research Technology Support Facility for custom sequencing. Sequencing results were reassembled from the forward and reverse sequencing runs by hand and the sequences were used to perform a Basic Local Alignment and Search Tool query using the GenBank database to determine the closest matches based on sequence homology.

RESULTS AND DISCUSSION

Performance of the multi target microarray with seeded viruses

The microarray was first examined in the laboratory using deionized water seeded with poliovirus 1 LSc, adenovirus type 40 and type 41. The viral array signals were unambiguous when tested with poliovirus virus and no cross hybridization above a signal to noise ratio cutoff of 2 was observed between labeled poliovirus cDNA with non-poliovirus probes (Figure 1(a)). Limited cross reaction was observed with two probes from human enterovirus A, two probes from human enterovirus B and three probes from

Table 2 | List of primers used for the PCR detection and sequencing for each viral target

Virus	Primers (5′-3′)	Amplicon size (bp)	Ref.
Hepatitis E virus	HE364: CTGGGMYTGGTCDCGCCAAG HE361: GCRGTGGTTTCTGGGGTGAC	164	Inoue *et al.* (2006)
Adenovirus	hexAA1885: GCCGCAGTGGTCTTACATGCACATC hexAA1913: CAGCACGCCGCGGATGTCAAAGT	301	Chapron *et al.* (2000)
	hexAA1893: GCCACCGAGACGTACTTCAGCCTG hexAA1905: TTGTACGAGTACGCGGTATCCTCGCGGTC	143	
Norovirus	GIF: TGTCACGATCTCATCATCACC GIR: GTGAACAGC/TATAAAT/CCACT/CGG	123	Borchardt *et al.* (2003)
	GIIF: TGTCACGATCTCATCATCACC GIIR: TGGAATTCCATCGCCCACTGG	123	
Sapovirus	F22: SMWAWTAGTGTTTGARATG R2: GWGGGRTCAACMCCWGGTGG	420	Okada *et al.* (2006)
Enteroviruses	Ent1: CGGTACCTTTGTACGCCTGT Ent2: ATTGTCACCATAAGCAGCCA	534	Chapron *et al.* (2000)
	neEnt1: TCCGGCCCCTGAATGCGGCTA neEnt2: GAAACACGGACACCCAAAGTA	138	
Rotavirus	TTGCCACCAATTCAGAATAC ATTTCGGACCATTTATAACC	211	Borchardt *et al.* (2003)
Coronavirus	Cor-FW: ACWCARHTVAAYYTNAARTAYGC Cor-RV: TCRCAYTTDGGRTARTCCCA	251	Moes *et al.* (2005)
Cytomegalovirus (HH5)	IE: GCTGCGGCATAGAATCAAGGAGCA IE: GGTTGGTGGTCTTAGGGAAGGCTGAG	393	Caballero *et al.* (1997)
Torovirus	TRVF: CACCACGTAATCAGTATAG TRVR: GTACGACACACAACATC	160	This study
Astrovirus	RT: GTAAGATTCCCAGATTGGT PCR: CCTGCCCCGAGAACAACCAAG	243	Chapron *et al.* (2000)
	A1: CCTTGCCCCGAGCCAGAA A2: TATTCACAAACTTATGGCAA	143	
JC polyomavirus	JCSR: TGATTACAGCATTTTTGTCTGCAAC JCSL: GGAAGTCCTTCTGTTAATTAAATCAG	170	Bofill-Mas *et al.* (2001)
BK polyomavirus	BK3: ACTGTAACACCTGCTCTT BK4: AGTAGATTTCCACAGGTTAG	350	Bofill-Mas *et al.* (2001)

human enterovirus C, but these were clearly distinguishable from positive reactions observed in the probes targeting poliovirus by their numbers (two or three probes versus 30 for poliovirus) and lower signal to noise ratio. More cross hybridizations were observed using Cy3 labeled adenovirus type 40 and Cy5 labeled adenovirus type 41 with probes targeting other virus groups. Specifically, probes targeting human adenovirus A, human adenovirus C and human enterovirus D showed some cross reactivity with adenovirus type 40 and to a lesser extent adenovirus type 41. In addition, cross-hybridization was observed between the two closely related virus species adenovirus type 40 and 41, both members of enteric adenovirus subgenus F (Figure 1(b)). But in all cases the non-specific signals were distinguished from the specific target virus signals by their lower average normalized intensity (1.03 versus 2.78 when Cy3 labeled adenovirus type 40 was hybridized to the microarray and 0.94 versus 2.69 when Cy5 labeled adenovirus type 41 was hybridized to the microarray) and based on the full analysis of the probes, these two DNA viruses could be distinguished. All hybridization signal values were deposited in Gene Expression Omnibus (GEO) accession #GSE10566 and GEO accession #GSE10569, for the poliovirus and adenoviruses experiments, respectively.

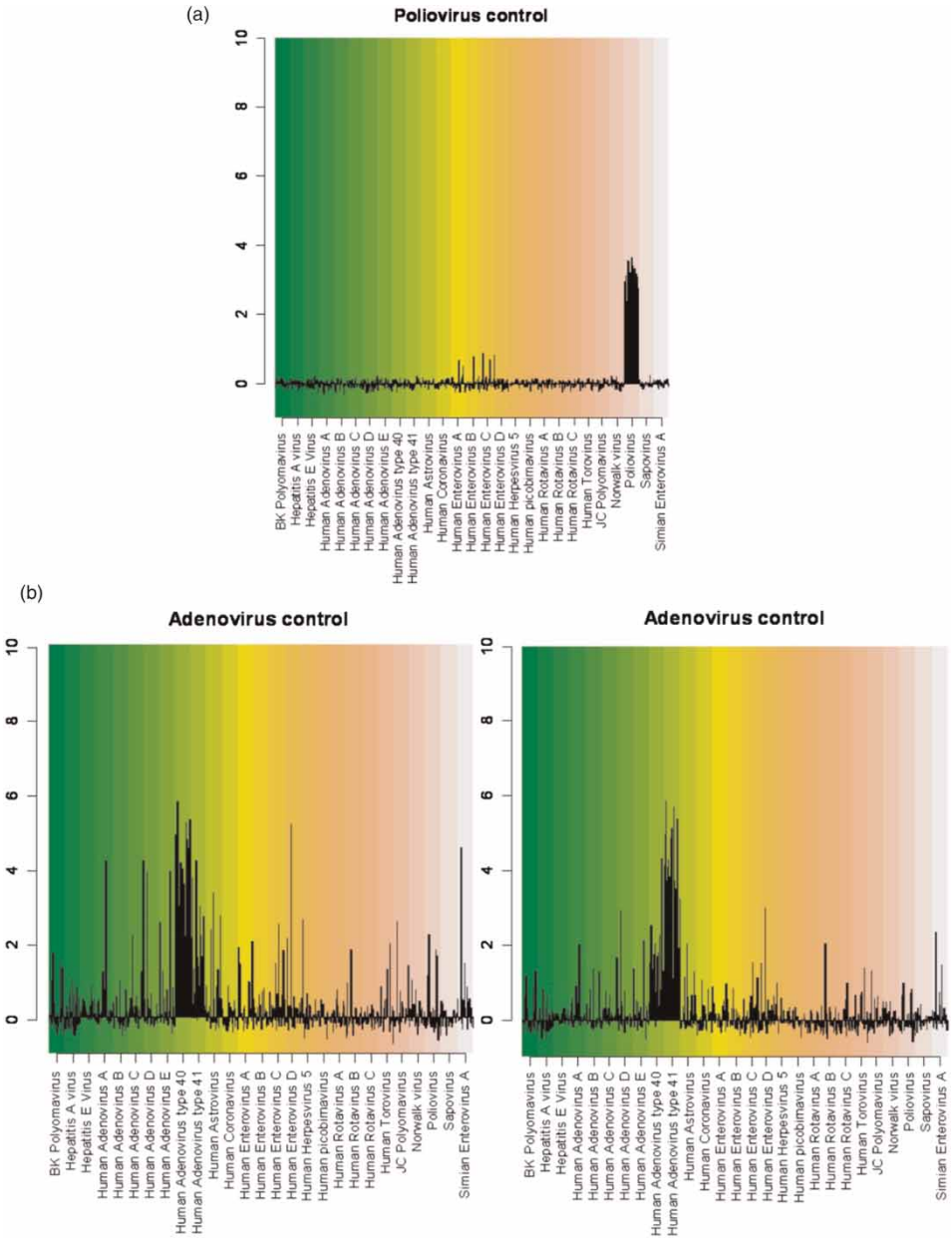

Figure 1 | Bar plots showing the $\log_2$-normalized hybridization signals for various control experiments. (a) Hybridization of poliovirus 1 LSc RNA extracts labeled with Cy3. (b) Hybridization of Cy3-labeled adenovirus type 40 (right) and Cy5-labeled adenovirus type 41 (left) co-hybridized unto the viral array.

Characterization of the viral hazard in sewage using multi target microarrays and PCR

The viral hazards were assessed in sewage over a 13-month period. Both RNA and DNA viruses were examined with the Cy5-labeled (RNA) and Cy3 (DNA) nucleic acids, respectively derived from sewage samples collected between August 2006 to August 2007. RNA viruses were labeled with Cy5 dye and DNA viruses were labeled with Cy3 dye to allow quick visual addressing of the probe locations. Based on the *t*-test results, the most probable virus present in each sample using microarrays, as determined by the DetectiV

Table 3 | Detection and characterization of human pathogenic viruses in community wastewaters by multi target microarrays (M) and group-specific PCR (P)

Virus type[a]		August 2006	September 2006	October 2006	November 2006	December 2006	January 2007	February 2007	March 2007	April 2007	May 2007	June 2007	July 2007	August 2007
DNA virus														
Adenoviridae	Adenovirus A													
	Adenovirus B					P, M	M	M						
	Adenovirus C						M							
	Adenovirus D			P			M							
	Adenovirus E													
	Adenovirus type 40													
	Adenovirus type 41	P				M	P, M	P, M				P		
Herpesviridae	Cytomegalovirus (HH5)													
Polyomaviridae	JC polyomavirus													
	BK polyomavirus					P	P, M	P						
RNA virus														
Astroviridae	Astroviruses	P	P		P	P, M	P, M	P, M				P		
Caliciviridae	Norwalk virus					P, M	P, M							
	Sapovirus						M							
Coronaviridae	Coronavirus							M						
	Torovirus					P	P							
Hepeviridae	Hepatitis E virus					P	P, M							
Picobirnaviridae	Picobirnavirus													
Picornaviridae	Human enterovirus A			P		P, M	P, M	P, M	P	P				P
	Human enterovirus B						M	M						
	Human enterovirus C													
	Human enterovirus D						M							
	Human enterovirus E					M	M							
	Poliovirus													
	Hepatitis A virus													
Reoviridae	Rotavirus A													
	Rotavirus B						P							
	Rotavirus C					P								

[a]All viruses are human specific.

software package and utilizing a p value cut off of less than 0.02 and a $\log_2$ mean intensity value of 0.6, are listed in Table 3. All processed sewage sample hybridization signals were deposited at the GEO as accession #GSE11195. Samples were also processed with 12 different PCR methods followed by sequencing. PCR detected viruses in 11 out of 13 months compared to 3 out of 13 months for the microarray. In total, PCR detected 31 virus positives versus 25 virus positives for microarray using DetectiV analysis (Table 3). The enhanced sensitivity when using PCR was evident in samples collected during the non-winter months (March to October) where PCR detected seven virus-positive samples versus none for the microarray and was able to identify viruses in the family *Adenoviridae*, *Picornaviridae* and *Astroviridae* throughout the year.

Detection using the microarray appeared to be more specific compared to PCR, particularly for enteroviruses and adenoviruses. In the samples collected, PCR was unable to detect any of the group B, C and D human adenoviruses (confirmed by sequencing) that were detected by the microarray and was more likely to pick up group F human adenoviruses (particularly human adenoviruses type 41). PCR also failed to detect any of the group B, D and E human enteroviruses which were detected by the microarray in the samples from December to February. The ability of the microarray method to differentiate between viral species could be due to the use of multiple probes derived from the entire viral genome rather than a small segment of the genome as usually used in the PCR.

In this study, the most frequently detected viruses year round from sewage using cell culture coupled with microarrays and PCR were human enterovirus A and human astroviruses (seven out of 13 samples) (Table 3). The next most frequently detected virus was human adenovirus type 41 (five out of 13 samples). Overall, a large number of groups of viruses were detected during the colder months from November to February (17 viral types) compared to the warmer months of March to October (four viral types). This demonstrated that a great diversity of human enteric viruses circulating among a given community even though no outbreaks of these diseases were reported during the year. High prevalence of enteric viruses in the colder months could be attributed either to a higher concentration of viruses being discharged to the sewers during the winter months or to a higher degree of survival by the viruses due to the lower temperatures. Astroviruses in particular may be associated with illness that has yet to be fully recognized and only a few studies describe the prevalence of astroviruses in wastewater (Meleg *et al.* 2006; Aw & Gin 2010). Our microarray and PCR data also revealed the prevalence of human enterovirus A in community wastewater samples. This distribution pattern was not in agreement with the majority of previous studies in temperate regions which showed the prevalence of human enterovirus B (Costan-Longares *et al.* 2008; Petrinca *et al.* 2009). This suggested that the prevalence of human enterovirus A infections could be underestimated as most strains belonging to this enterovirus species (particularly group A coxsackieviruses) grow poorly or do not produce a discernible cytopathic effect in cell culture (Lipson *et al.* 1988).

This study demonstrates the application of multi target microarrays to characterize diverse groups of enteric viruses in the environment after cell culture which primarily focuses on infectious enteroviruses and adenoviruses. It is important to note that many viral targets in this study, particularly members of the family *Caliciviridae* and hepatitis E virus do not grow in BGM cells. The detection of these viruses by microarrays and PCR in BGM cell culture extracts did not indicate their replication. This could be due to the inoculation of a large volume of raw sewage virus concentrate onto the cells and the further concentration of cell culture extracts prior to viral nucleic acid extraction. Direct detection without cell culture using microarrays is hampered by the presence of inhibitory substances in environmental samples that compromise sample labeling and hybridization efficiency and the low detection levels (Straub 2011). To make this practical, multi target microarrays would need to be further coupled with sample concentration and purification techniques as well as other amplification techniques, such as recent studies which have shown that whole genome amplification using universal oligonucleotide primers offers a promising approach to increase the amount of nucleic acid in the environmental sample prior to microarray detection (Binga *et al.* 2008; Parker *et al.* 2011). Yet the infectivity with these approaches will remain a question mark. While direct hybridization on a microarray is less sensitive than target specific PCR detection, the primer sets that need to be established and the many reactions that are necessary in order to describe the wide diversity of pathogenic viruses

that might be present in sewage using PCR makes it difficult to do comparative assessment of the various hazards. Moreover, the use of multiplex PCR is often limited by the number of targets due to primer–primer interactions.

Most recently, next generation sequencing approach (mainly Roche 454 sequencing platform) has been used in metagenomic studies of viral communities in wastewater, which has the potential to replace microarrays in the examination of environmental microbial communities. A general finding from these studies using next generation sequencing is that the majority of the viral metagenomes have no similarities to known genes in the database and low abundance of established human viral pathogen sequences was detected by a metagenomic approach (Rosario *et al*. 2009; Cantalupo *et al*. 2011). Moreover, the most critical challenge for the application of next generation sequencing technologies in viral pathogen detection and characterization in environmental samples is the development of improved user-friendly computational methods to handle large amounts of sequencing data and facilitate rapid and robust analysis of sequencing data. Our study suggested that multi target microarrays could be a more feasible method for public health laboratories and wastewater industry to directly and rapidly characterize multiple selected enteric viruses of public health concern in environmental samples. Microarray is a more mature technique and the analysis is relatively straightforward. Nevertheless, next generation sequencing has the potential to provide a detailed understanding of viral gene diversity and expand viral genome databases to improve future sequence annotation, which can be used in the design of multi target microarray probe sets.

CONCLUSIONS

This study demonstrates the feasibility of the use of multi target microarrays to screen community wastewater for a broad number of pathogenic viruses. This could serve as an example for improved environmental health surveillance of wastewater pollution and allows an improved understanding of the potential risk for waterborne transmission. While the level of sensitivity offered by microarrays may be inadequate for direct pathogen detection in environmental samples, microarrays may prove valuable as a viral fingerprinting tool for environmental samples when coupled with random amplification. Further research into multi target microarray platform optimization is warranted.

Using both microarray and PCR assays, we showed that sewage discharges are contributing significant array of viral hazards to waters and many communities do not disinfect or only disinfect during the swimming season. This means that large loading of potentially infectious viruses to receiving rivers and lakes is occurring. A recent study has found that even limited-contact water recreation activities such as boating on non-disinfected effluent-dominated waters were associated with an elevated risk of gastroenteritis illness (Dorevitch *et al*. 2012). Viruses remain a public health concern and should remain a priority for the water and health community. The ability of both DNA and RNA viruses to rapidly evolve means new and emerging viral pathogens will need to be addressed. Microarrays with multiple probes such as those described by Wang *et al*. (2003) and the one presented here provide a distinct advantage for unknown viruses with somewhat varying sequences to the known viruses.

The use of high throughput approaches such as multi target microarrays to screen community health via excretion of viral pathogens in urine and feces offers the possibility of integration between the wastewater industry and the health community as a means to improve our understanding of exposure, disease and ultimately prevention strategies. We can continue to improve the detection and genotyping ability of multi target microarrays for pathogenic viruses in complex environmental samples by incorporating concentration and amplification steps as well as reducing the cost and hybridization time by an order of magnitude. Nevertheless, this work demonstrates important first steps in taking multi target microarray technology into the field, enabling detection and characterization of multiple viral pathogens in a single array assay. Pathogen discovery and characterization, occurrence and survival in the environment, exposure pathways and health risks via environmental exposure are important issues that deserve future attention and elucidation.

ACKNOWLEDGEMENT

The authors thank the US Environmental Protection Agency for funding this research under the EPA's National

Center for Environment Research Science to Achieve Results (STAR) grant (R831628).

REFERENCES

Aw, T. G. & Gin, K. Y. H. 2010 Environmental surveillance and molecular characterization of human enteric viruses in tropical urban wastewaters. *J. Appl. Microbiol.* **109**, 716–730.

Aw, T. G. & Rose, J. B. 2012 Detection of pathogens in water: from phylochips to qPCR to pyrosequencing. *Curr. Opin. Biotechnol.* **23**, 422–430.

Borchardt, M. A., Bertz, P. D., Spencer, S. K. & Battigelli, D. A. 2003 Incidence of enteric viruses in groundwater from household wells in Wisconsin. *Appl. Environ. Microbiol.* **69**, 1172–1180.

Binga, E. K., Lasken, R. S. & Neufeld, J. D. 2008 Something from (almost) nothing: the impact of multiple displacement amplification on microbial ecology. *ISME J.* **2**, 233–241.

Bofill-Mas, S., Formiga-Cruz, M., Clemente-Casares, P., Calafell, F. & Girones, R. 2001 Potential transmission of human polyomaviruses through the gastrointestinal tract after exposure to virions or viral DNA. *J. Virol.* **75**, 10290–10299.

Caballero, O. L., Menezes, C. L., Costa, M. C., Fernandes, S. C., Anacleto, T. M., de Oliveira, R. M., Viotti, E. A., Távora, E. R., Vilaça, S. S., Sabbaga, E., de Paula, F. J., Távora, P. F., Villa, L. L. & Simpson, A. J. 1997 Highly sensitive single-step PCR protocol for diagnosis and monitoring of human cytomegalovirus infection in renal transplant recipients. *J. Clin. Microbiol.* **35**, 3192–3197.

Cantalupo, P. G., Calgua, B., Zhao, G., Hundesa, A., Wier, A. D., Katz, J. P., Grabe, M., Hendrix, R. W., Girones, R., Wang, D. & Pipas, J. M. 2011 Raw sewage harbors diverse viral populations. *mBio* **2**, e 00180-11.

Chapron, C. D., Ballester, N. A., Fontaine, J. H., Frades, C. N. & Margolin, A. B. 2000 Detection of astroviruses, enteroviruses, and adenovirus types 40 and 41 in surface waters collected and evaluated by the information collection rule and an integrated cell culture-nested PCR procedure. *Appl. Environ. Microbiol.* **66**, 2520–2525.

CDC 2006 Resurgence of wild poliovirus type 1 transmission and consequences of importation–21 countries, 2002–2005. *Morb. Mortal. Wkly. Rep.* **55**, 145–150.

Costan-Longares, A., Moce-Llivina, L., Avellon, A., Jofre, J. & Lucena, F. 2008 Occurrence and distribution of culturable enteroviruses in wastewater and surface waters of North-Eastern Spain. *J. Appl. Microbiol.* **105**, 1945–1955.

Dorevitch, S., Pratap, P., Wroblewski, M., Hryhorczuk, D. O., Li, H., Liu, L. C. & Scheff, P. A. 2012 Health risks of limited-contact water recreation. *Environ. Health Perspect.* **120**, 192–197.

Fong, T. T. & Lipp, E. K. 2005 Enteric viruses of humans and animals in aquatic environments: health risks, detection, and potential water quality assessment tools. *Microbial. Mol. Biol. Rev.* **69**, 357–371.

Gao, X., LeProust, E., Zhang, H., Srivannavit, O., Gulari, E., Yu, P., Nishiguchi, C., Xiang, Q. & Zhou, X. 2001 A flexible light-directed DNA chip synthesis gated by deprotection using solution photogenerated acids. *Nucleic Acids Res.* **29**, 4744–4750.

Girones, R., Ferrús, M. A., Alonso, J. L., Rodriguez-Manzano, J., Calgua, B., de Abreu Corrêa, A., Hundesa, A., Carratala, A. & Bofill-Mas, S. 2010 Molecular detection of pathogens in water – The pros and cons of molecular techniques. *Water Res.* **44**, 4325–4339.

Gurrala, R., Dastjerdi, A., Johnson, N., Nunez-Garcia, J., Grierson, S., Steinbach, F. & Banks, M. 2009 Development of a DNA microarray for simultaneous detection and genotyping of lyssaviruses. *Virus Res.* **144**, 202–208.

Inoue, J., Takahashi, M., Yazaki, Y., Tsuda, F. & Okamoto, H. 2006 Development and validation of an improved RT-PCR assay with nested universal primers for detection of hepatitis E virus strains with significant sequence divergence. *J. Virol. Methods* **137**, 325–333.

Koltai, H. & Weingarten-Baror, C. 2008 Specificity of DNA microarray hybridization: characterization, effectors and approaches for data correction. *Nucleic Acids Res.* **36**, 2395–2405.

Komolpis, K., Srivannavit, O. & Gulari, E. 2002 Light-directed simultaneous synthesis of oligopeptides on microarray substrate using a photogenerated acid. *Biotechnol. Prog.* **18**, 641–646.

Korimbocus, J., Scaramozzino, N., Lacroix, B., Crance, J. M., Garin, D. & Vernet, G. 2005 DNA probe array for the simultaneous identification of herpesviruses, enteroviruses, and flaviviruses. *J. Clin. Microbiol.* **43**, 3779–3787.

Leclerc, H., Schwartzbrod, L. & Dei-Cas, E. 2002 Microbial agents associated with waterborne diseases. *Crit. Rev. Microbiol.* **28**, 371–409.

LeProust, E., Pellois, J. P., Yu, P., Zhang, H., Gao, X., Srivannavit, O., Gulari, E. & Zhou, X. 2000 Digital light-directed synthesis. A microarray platform that permits rapid reaction optimization on a combinatorial basis. *J. Com. Chem.* **2**, 349–354.

Lin, B., Malanoski, A. P., Wang, Z., Blaney, K. M., Ligler, A. G., Rowley, R. K., Hanson, E. H., von Rosenvinge, E., Ligler, F. S., Kusterbeck, A. W., Metzgar, D., Barrozo, C. P., Russell, K. L., Tibbetts, C., Schnur, J. M. & Stenger, D. A. 2007 Application of broad-spectrum, sequence-based pathogen identification in an urban population. *PLoS ONE* **2**, e419.

Lipson, S. M., Walderman, R., Costello, P. & Szabo, K. 1988 Sensitivity of rhabdomyosarcoma and guinea pig embryo cell cultures to field isolates of difficult-to-cultivate group A coxsackieviruses. *J. Clin. Microbiol.* **26**, 1298–1303.

Meleg, E., Jakab, F., Kocsis, B., Banyai, K., Melegh, B. & Szucs, G. 2006 Human astroviruses in raw sewage samples in Hungary. *J. Appl. Microbiol.* **101**, 1123–1129.

Moes, E., Vijgen, L., Keyaerts, E., Zlateva, K., Li, S., Maes, P., Pyrc, K., Berkhout, B., van der Hoek, L. & van Ranst, M. 2005 A novel pancoronavirus RT-PCR assay: frequent detection of

human coronavirus NL63 in children hospitalized with respiratory tract infections in Belgium. *BMC Infect. Dis.* **5**, 6.

Oberste, M. S. & Pallansch, M. A. 2003 Establishing evidence for enterovirus infection in chronic disease. *Ann. N. Y. Acad. Sci.* **1005**, 23–31.

Okada, M., Yamashita, Y., Oseto, M. & Shinozaki, K. 2006 The detection of human sapoviruses with universal and genogroup-specific primers. *Arch. Virol.* **151**, 2503–2509.

Parker, J. K., Chang, T. Y. & Meschke, J. S. 2011 Amplification of viral RNA from drinking water using TransPlex™ whole-transcriptome amplification. *J. Appl. Microbiol.* **111**, 216–223.

Petrinca, A. R., Donia, D., Pierangeli, A., Gabrieli, R., Degener, A. M., Bonanni, E., Diaco, L., Cecchini, G., Anastasi, P. & Divizia, M. 2009 Presence and environmental circulation of enteric viruses in three different wastewater treatment plants. *J. Appl. Microbiol.* **106**, 1608–1617.

Rosario, K., Nilsson, C., Lim, Y. W., Ruan, Y. J. & Breitbart, M. 2009 Metagenomic analysis of viruses in reclaimed water. *Environ. Microbiol.* **11**, 2806–2820.

Rose, J. B. & Molloy, S. L. 2007 Globalization effects on water quality: monitoring the impact on and control of waterborne disease. In: *Globalization: Effects on Fisheries Resources* (W. W. Taylor, M. G. Schechter & L. G. Wolfson, eds). Cambridge University Press, New York, USA, pp. 92–119.

Rouillard, J. M., Zuker, M. & Gulari, E. 2003 OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. *Nucleic Acids Res.* **31**, 3057–3062.

Straub, T. M. 2011 Challenges of multiplexed detection: Detection of pathogens in water and wastewater using microarrays. In: *Environmental Microbiology: Current Technology and Water Applications* (K. Sen & N. J. Ashbolt, eds). Caister Academic Press, Norfolk, UK, pp. 127–142.

Tambyah, P. A. 2004 SARS: responding to an unknown virus. *Eur. J. Clin. Microbiol. Infect. Dis.* **23**, 589–595.

USEPA 1995 *Virus Monitoring Protocol for the Information Collection Requirements Rule*. US Environmental Protection Agency, publication EPA/814-B-95-002, Government Printing Office, Cincinnati, Ohio.

Wang, D., Urisman, A., Liu, Y. T., Springer, M., Ksiazek, T. G., Erdman, D. D., Mardis, E. R., Hickenbotham, M., Magrini, V., Eldred, J., Latreille, J. P., Wilson, R. K., Ganem, D. & DeRisi, J. L. 2003 Viral discovery and sequence recovery using DNA microarrays. *PLoS biology* **1**, E2.

Watson, M., Dukes, J., Abu-Median, A. B., King, D. P. & Britton, P. 2007 DetectiV: visualization, normalization and significance testing for pathogen-detection microarray data. *Genome Biol.* **8**, R190.

Wick, L. M., Rouillard, J. M., Whittam, T. S., Gulari, E., Tiedje, J. M. & Hashsham, S. A. 2006 On-chip non-equilibrium dissociation curves and dissociation rate constants as methods to assess specificity of oligonucleotide probes. *Nucleic Acids Res.* **34**, e26.

Wong, C. W., Heng, C. L., Wan Yee, L., Soh, S. W., Kartasasmita, C. B., Simoes, E. A., Hibberd, M. L., Sung, W. K. & Miller, L. D. 2007a Optimization and clinical validation of a pathogen detection microarray. *Genome Biol.* **8**, R93.

Wong, M., Xagoraraki, I. & Rose, J. B. 2007b Recent advances and future needs in environmental virology. In: *Human Viruses in Water* (A. Bosch, ed.). Elsevier, Oxford, UK, pp. 259–284.

First received 29 January 2013; accepted in revised form 11 July 2013. Available online 19 August 2013