Human calicivirus diversity in wastewater in South Africa

T.Y. Murray[1*], J. Mans[1*] and M.B. Taylor[1,2]

1 Department of Medical Virology, Faculty of Health Sciences, University of Pretoria, Private Bag X323, Arcadia, Pretoria, 0007, South Africa

2 National Health Laboratory Service, Tshwane Academic Division, Pretoria, South Africa

*Both authors contributed equally to this manuscript

Running title

HuCVs in South African wastewater

Correspondence

Tanya Y. Murray

Department of Medical Virology

University of Pretoria

Private Bag X323

Arcadia

Pretoria 0007

South Africa

Keywords

Genotypes, norovirus, sapovirus, sewage, South Africa, wastewater

Abstract

Aim: To investigate the diversity of human caliciviruses (HuCVs) in wastewater from small- to medium-sized communities in five provinces of South Africa (SA).

Methods and Results: Wastewater samples (51) were screened for norovirus (NoV) GI, GII, GIV and sapovirus (SaV) using real-time reverse transcription (RT)-PCR. Partial capsid nucleotide sequences were analysed for genotyping. At least one HuCV was detected in 42 samples (82%) with NoV GI being detected in 15 (29%), NoV GII in 32 (63%) and SaV in 37 (73%) samples. Norovirus GIV was not detected. Five NoV GI genotypes (GI.1, GI.3 GI.4, GI.8, GI.unassigned), eight NoV GII genotypes (GII.2, GII.3, GII.4, GII.6, GII.7, GII.12, GII.13, GII.17) and six SaV genotypes (GI.2, GI.3, GI.6, GI.7, GII.1, GII.2) were characterised.

Conclusions: Many NoV and SaV genotypes were detected in wastewater, demonstrating a high genetic diversity of HuCVs in the surrounding communities. Caliciviruses were characterised from several provinces in SA, indicating widespread occurrence in the country.

Significance and Impact of the Study: This study provides valuable new data on CVs circulating in SA, including the first data on SaV strains from wastewater in Africa. Environmental surveillance is especially important in countries like SA where outbreak reporting systems or routine HuCV surveillance is lacking.

Introduction

Caliciviruses (CVs) are small, single-stranded RNA viruses that are responsible for a significant proportion of non-bacterial gastroenteritis worldwide (Green 2007). Human CVs belong to the norovirus (NoV) and sapovirus (SaV) genera, which are characterised by high genetic diversity and are each classified into at least five genogroups (G) (Farkas *et al.* 2004; Zheng *et al.* 2006). Norovirus GI, GII and GIV contain strains that infect humans and based on the capsid gene sequence are divided into eight, twenty-one and two genotypes, respectively. Norovirus GII.4 is the predominant NoV genotype causing gastroenteritis outbreaks worldwide, with periodic emergence of new epidemic GII.4 variants (Siebenga *et*

al. 2007). Human SaV strains are found in genogroups GI, GII, GIV and GV, with GI and GII being divided into seven genotypes each based on the complete capsid gene sequence (Oka *et al.* 2012). Human CVs typically cause acute, self-limiting gastroenteritis, are shed in high concentrations and are transmitted via the faecal-oral route. Studies on NoV have shown the virus to have a very low infectious dose (10-100 infectious virions) (Teunis *et al.* 2008) and to be highly stable in the environment, including resistance to wastewater treatment processes (van den Berg *et al*. 2005).

Characterisation of enteric viruses present in treated and untreated wastewater provides valuable information on the viruses circulating in both symptomatic and asymptomatic individuals of a population. Studies on viruses in wastewater are especially useful when routine surveillance is lacking and could contribute more comprehensive, less biased epidemiological data than clinical-based studies. The presence and diversity of NoV GI and GII in wastewater have been described on several continents: Europe (Lodder and de Roda Husman 2005; La Rosa *et al.* 2010a), Asia (He *et al.* 2011; Kitajima *et al.* 2012), South America (Victoria *et al.* 2010; Blanco Fernandez *et al.* 2012) and New Zealand (Hewitt *et al.* 2011), however, data from Africa is limited. To date, NoV strains from wastewater in Egypt (Kamel *et al.* 2010) and Tunisia (Sdiri-Loulizi *et al.* 2010) have been characterised. Norovirus strains have also been genotyped from sewage-polluted river water in South Africa (SA) (Mans *et al.* 2013). In contrast to GI and GII, NoV GIV has only been detected in sewage in a few countries in Europe [Italy (La Rosa *et al.* 2010b), the Netherlands (van den Berg *et al.* 2005), Luxembourg (Skraber *et al.* 2011)] and in Japan (Kitajima *et al.* 2009). When compared to data on NoVs, there is less information available on SaVs in the environment. Sapoviruses have been characterised from treated and untreated wastewater in Japan (Hansman *et al.* 2007; Kitajima *et al.* 2011) and Spain (Sano *et al.* 2011) as well as from one sewage sample in the United States (US) (Ng *et al.* 2012). While CVs have been

characterised from several countries, these studies most often focus on one specific site and therefore do not necessarily represent the country-wide diversity.

In SA, NoVs and SaVs have been detected in hospitalised paediatric patients in the Gauteng province (GP). Noroviruses were more prevalent than SaVs and were most frequently detected after rotavirus (Mans *et al.* 2010). On-going routine enteric virus surveillance of sewage-polluted surface water in GP indicates frequent occurrence of human CVs, especially SaV, in this province (Mans *et al.* 2013; unpublished data). This implies higher prevalence in the surrounding communities than suggested by hospital-based studies. Characterisation of CV strains from wastewater could provide insight into their diversity in the general population. The currently available data on CVs in SA concentrates on one densely-populated province (GP) and focuses on urban populations (Mans *et al*. 2013).

Since NoVs are highly diverse and constantly evolving, it is important to establish which genotypes are circulating in a population and to monitor for the emergence and spread of new genotypes and variants within genotypes. With less information available for SaVs, data is needed to ascertain the predominant genotypes worldwide. This study investigates the diversity of both NoVs and SaVs in wastewater from several provinces in one country (SA), with a focus on small- to medium-sized communities.

Materials and Methods

Wastewater sample collection

From August 2010 to December 2011, sludge and effluent samples were collected from wastewater treatment plants (21 sites) in several (5/9) provinces of South Africa. These treatment plants service small- to medium-sized towns with populations ranging from several

hundred to 50 000 individuals. South Africa has a temperate climate, with a winter season from June to August and summer season from November to February. Fifty-one wastewater samples (47 treated effluent and four untreated wastewater), with thermotolerant (faecal) coli counts of greater than 10^6 colony forming units (cfu) 100 ml^{-1}, were selected for this study. Twenty samples originated from Mpumalanga province (MP), 16 from North West province (NW) and nine from GP. Two samples each were selected from KwaZulu-Natal province (KZN) and the Western Cape province (WC) and two samples had unknown provincial origin (Fig. 1).

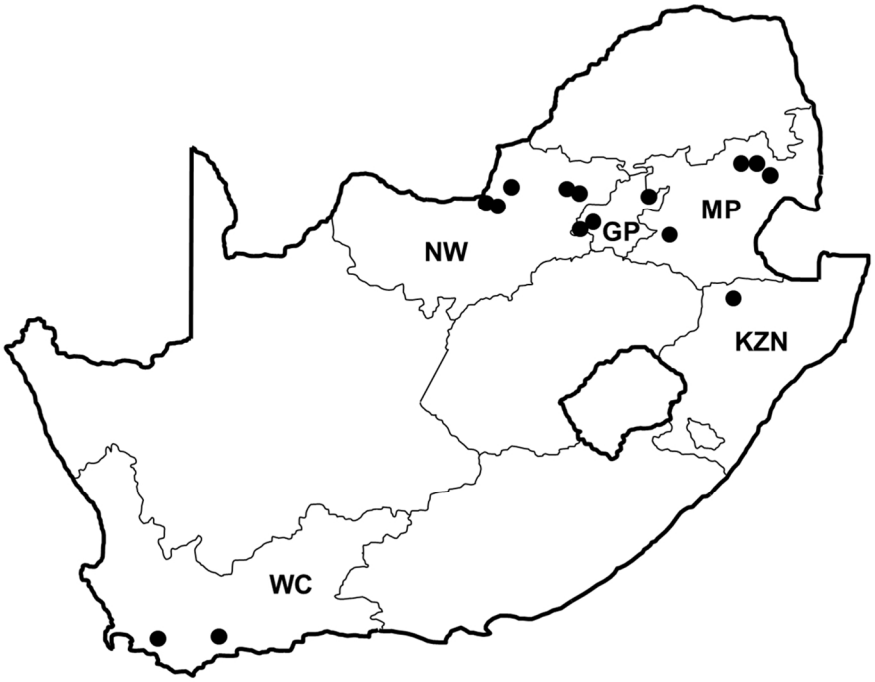

Figure 1 Distribution of wastewater sampling sites in five provinces of South Africa. Treatment plants located within 20 km of each other are represented by the same circle. GP = Gauteng province, KZN = KwaZulu-Natal province, MP = Mpumalanga province, NW = North West province, WC = Western Cape

Virus concentration and nucleic acid extraction

Viruses were concentrated from 100 ml (2010) or 75 ml (2011) wastewater samples to a volume of 2 ml in phosphate-buffered saline (pH 7.4) (PBS; Sigma-Aldrich Co., St. Louis, MO) by polyethylene glycol/sodium chloride precipitation (Minor 1985; European Committee for Standardization (CEN) Technical Committee 275 2010).

Total nucleic acid was extracted from 1 ml virus concentrate using the MagNA Pure LC Total Nucleic Acid Isolation kit (large volume) on the automated MagNA Pure LC instrument (Roche Diagnostics, Mannheim, Germany). An extraction negative control (nuclease-free water) was included for every 15 samples. Extracted nucleic acid was eluted in 100 µl and stored at -70°C in 5 µl or 10 µl aliquots until use.

Detection of caliciviruses

NoV GI and GII were detected with individual one-step real-time RT-PCRs using the QuantiTect Probe RT-PCR kit (Qiagen Inc., Valencia, CA) according to the manufacturer's instructions. The following primer sets and Taqman probes were used: QNIF4, NV1LCR (da Silva *et al.* 2007) and NVGGIp (Svraka *et al.* 2007) for NoV GI and QNIF2 (Loisy *et al.* 2005), COG2R (Kageyama *et al.* 2003) and QNIFSp (Loisy *et al.* 2005) for NoV GII. All samples that tested negative for NoV GI and GII on initial screening, were repeated at a 1:10 dilution of the RNA.

A published one-step real-time RT-PCR assay to detect NoV GIV (Trujillo *et al.* 2006) was modified to a two-step assay in this study. Briefly, cDNA was synthesised from the extracted RNA using RevertAid^{TM} Premium Reverse Transcriptase (Thermo Scientific, Waltham, MA). The 20 µl reaction contained 10 µl RNA, 30 µM random hexamer primers

and 50 U RevertAid^{TM} Premium Reverse Transcriptase. Five µl of cDNA was added to a 20 µl reaction for real-time PCR detection as described by Trujillo *et al.* (2006). Sapoviruses were detected as previously described (Murray *et al.* 2012).

Genotyping of caliciviruses

The 5'-end (320 bp, region C) of the capsid gene of NoV GI and GII strains was amplified as described previously (Mans *et al.* 2013) with minor modifications. Either AmpliTaq Gold (Applied Biosystems, Foster City, CA) or KAPA Taq DNA (Kapa Biosystems (Pty) Ltd., Cape Town, South Africa) polymerases were used for PCR amplification and the first-round PCR was performed with primers QNIF4/G1SKR or QNIF2/G2SKR.

Selected SaVs were characterised by amplifying and sequencing approximately 300 bp of a partial 5'-region of the capsid gene. The region was amplified by nested PCR using one of the following primer sets as described by Kitajima *et al.* (2010) (SaV124F, SaV1F, SaV5F, SV-R14 and SV-R14 for the first round of PCR and SaV1245Rfwd and SV-R2 for the second round) and Sano *et al.* (2011) (SV-F13, SV-F14, SV-DS3 and SV-DS4 for the first round of PCR and SaV1245Rfwd and SV-DS5 and SV-DS6 for the second round). In the first round of PCR, two µl cDNA was added to a 50 µl reaction with 0.4 µM of each primer and 1.25 U AmpliTaq Gold DNA polymerase (Applied Biosystems). The following cycling parameters were used: enzyme activation and initial denaturation at 95°C for 15 min, followed by 40 cycles of 94°C for 30 s, 50°C for 30 s and 72°C for 2 min and a final extension at 72°C for 10 min. The second round of PCR was performed in a 50 µl reaction containing 5 µl PCR1 product, 0.4 µM of each primer and 1.25 U AmpliTaq Gold DNA polymerase. The following cycling parameters were used: enzyme activation and initial

denaturation at 95°C for 15 min, followed by 45 cycles of 94°C for 30 s, 52°C for 30 s and 72°C for 1 min and a final extension at 72°C for 10 min.

Sequencing and phylogenetic analysis

All amplicons were directly sequenced in both directions using the ABI PRISM BigDye® Terminator v. 3.1 Cycle sequencing kit on an ABI 3130 automated analyser (Applied Biosystems). M13(-21) and M13-Rev primer sequences were added to the 5' end of primers to facilitate sequencing. If mixed sequences were detected in a sample, the amplicons were cloned using the CloneJET™ PCR cloning kit and 5 randomly selected clones were sequenced using pJET1.2/blunt specific primers (Thermo Scientific, Waltham, MA).

Nucleotide sequences were aligned using Sequencher™ 4.9 (Gene Codes Corporation, Ann Arbor, MI), BioEdit Sequence Alignment Editor (V.7.0.9.0) and BLAST-n (Altschul *et al.* 1997). Sequences were aligned with reference NoV and SaV strains using MAFFT version 6 (http://mafft.cbrc.jp/alignment/software/). Reference strains for SaV genotypes were selected according to Oka *et al.* (2012) and for NoV according to Green *et al.* (2007) and Zheng *et al.* (2006). Phylogenetic analysis was performed in MEGA5 using the neighbour-joining method, validated by 1000 bootstrap replicates. Genotypes were assigned based on clustering with reference strains in the phylogenetic tree with >70% bootstrap support. Norovirus genotypes were confirmed with the Norovirus Genotyping Tool (Kroneman *et al.* 2011). Nucleotide sequences determined in this study were submitted to GenBank under accession numbers: KC495619-KC495687 (NoV) and KC484983-KC484998 (SaV).

Results

Detection of caliciviruses

At least one CV was detected in 42/51 (82.4%) of wastewater samples. Of these, 35/51 (68.6%) were positive for NoV, 37/51 (72.5%) were positive for SaV and 30/51 (58.8%) were positive for both viruses (Table 1). With regards to NoV, GII predominated overall (32/51 samples), GI was detected in 15/51 samples and GIV was not detected in any samples. Sapoviruses were detected in all five provinces, whereas NoVs were not detected in the WC.

Table 1 Provincial distribution of caliciviruses in wastewater in South Africa

Province*	Total no. samples	No. (%) positive samples					
		NoV+SaV	NoV				SaV†
			GI	GII	GI+GII	GIV	
GP	9	5(55.6)	1(11.1)	2(22.2)	3(33.3)	0	6 (66.7)
KZN	2	1(50)	0	1(50)	1(50)	0	1 (50)
MP	20	15(75)	2 (10)	11(55)	2(10)	0	16 (80)
NW	16	7(43.8)	0	5(31.3)	5(31.3)	0	11 (68.8)
WC	2	0	0	0	0	0	1 (50)
Unknown	2	2(100)	0	1(50)	1(50)	0	2 (100)
Total	51	30(58.8)	3(5.9)	20(39.2)	12(23.6)	0	37 (72.5)

* GP = Gauteng province; KZN = KwaZulu-Natal province; MP = Mpumalanga province; NW = North West province; WC = Western Cape province and Unknown = unknown provincial location
†Detection data for sapoviruses was previously published (Murray *et al.* 2012)

Genotyping of noroviruses

Overall 13 NoV genotypes (5 GI, 8 GII) were characterised from the SA wastewater samples (Figs 2 and 3). Up to four different genotypes were detected in a single treated effluent sample. Genogroup I strains (17 in total) were detected in GP, MP and NW provinces and comprised of GI.1 GI.3, GI.4, GI.8 and one unassigned genotype (Fig. 2). The predominant genotypes were GI.1 (6/17) and the unassigned genotype (6/17). They exhibit high identity (98 % identity over 285 nucleotides) with NoV strains isolated from paediatric patients in

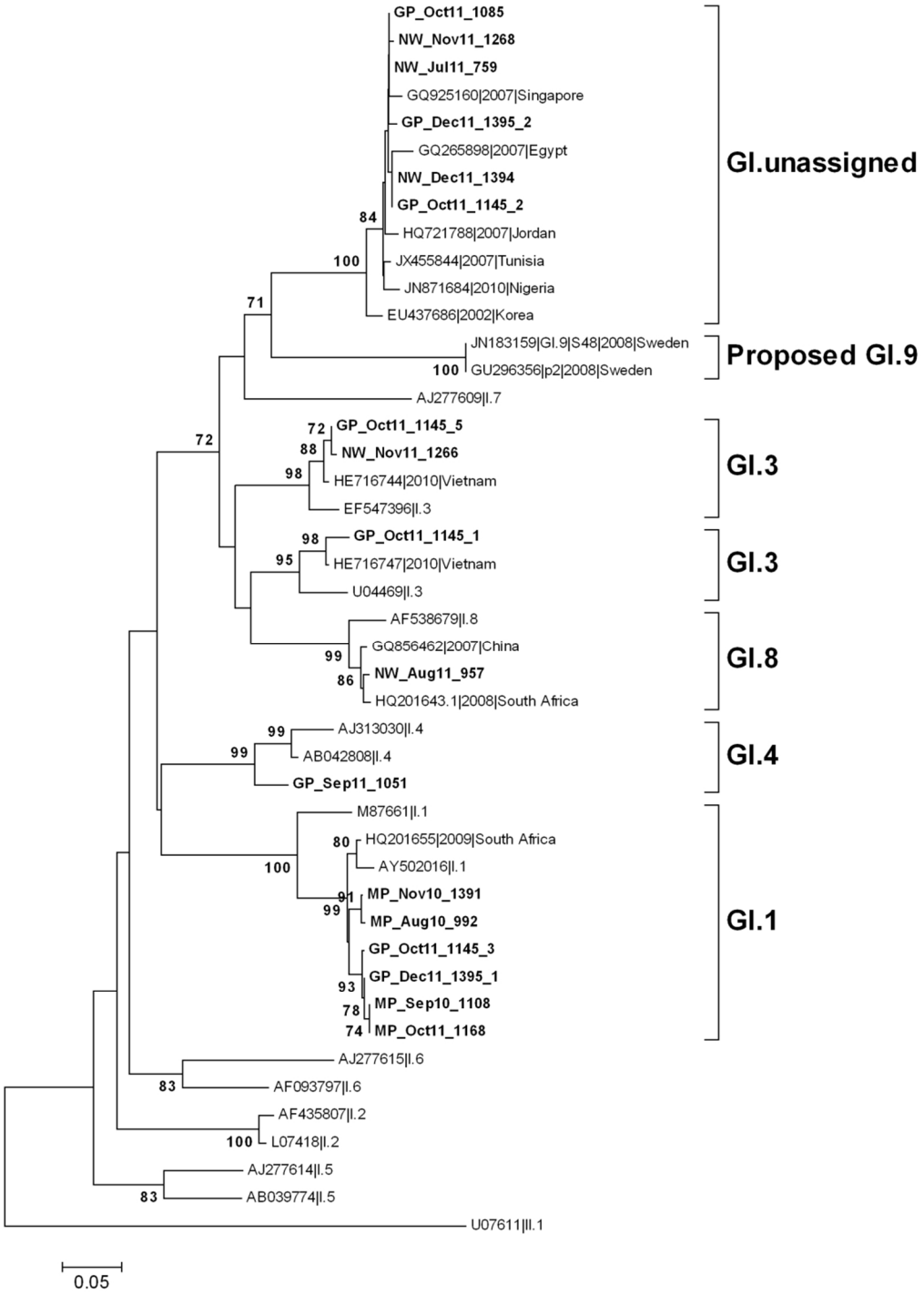

Figure 2 Phylogenetic analysis of partial capsid sequences (288 bp, region C) of 17 norovirus GI strains detected in wastewater in South Africa. The tree was generated using the neighbour-joining method and 1000 bootstrap replicates. Groupings with bootstrap support of > 70% were considered significant. Reference sequences are indicated by GenBank accession numbers and strains from this study are shown in bold. The province, date of collection (month and year), sample number and clone number are included in each sample name. The scale represents nucleotide substitutions per site.

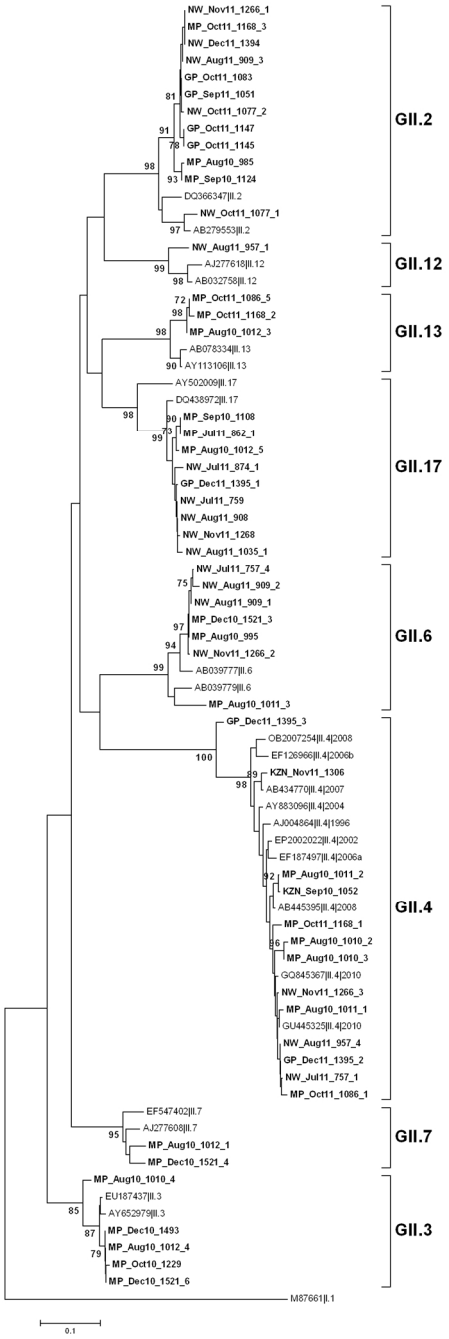

Figure 3 Phylogenetic analysis of partial capsid sequences (285 bp, region C) of 52 norovirus GII strains detected in wastewater in South Africa. The tree was generated using the neighbour-joining method and 1000 bootstrap replicates. Groupings with bootstrap support of > 70% were considered significant. Reference sequences are indicated by GenBank accession numbers and strains from this study are shown in bold. The province, date (month and year) of collection, sample number and clone number are included in each sample name. The scale represents nucleotide substitutions per site.

Nigeria (JN871684) and Jordan (HQ721788), as well as strains detected in sewage in Egypt (GQ265898), Tunisia (JX455844) and Singapore (GQ925160) and in water in Korea (EU437686). GI.1 and GI.8 strains from the current study displayed high nucleotide sequence identity (≥97%) with strains previously detected in river water in SA (Mans *et al.* 2013).

A total of 52 NoV GII strains could be classified in eight different genotypes, with GII.2, GII.4 and GII.17 being identified most frequently (Fig. 3). Norovirus GII.4 was the most prevalent genotype (13/52) with the widest distribution (four provinces). The majority of GII.4 strains clustered with recent 2008 and 2010 variants. Strains from GII.2 (12/52) and GII.17 (9/52) were detected in samples from GP, MP and NW but not in KZN. GII.6 was detected in three samples from MP in 2010 and in four samples from NW in 2011. Three genotypes GII.3 (5/52), GII.7 (2/52) and GII.13 (3/52) were only detected in MP. Genotype GII.12 was detected in a single sample from NW.

Genotyping of sapoviruses

The SaV strains identified in wastewater were characterised into 6 genotypes (4 GI, 2 GII) based on phylogenetic analysis of approximately 300 bp of the 5'-end of the capsid gene (Fig. 4). Multiple SaV genotypes were not detected in any of these samples. The genogroup I strains clustered with GI.2, GI.3, GI.6 and GI.7 and the genogroup II strains with GII.1 and GII.2. Half (8/16) of the SaV strains grouped in GI.2 and these included SaVs detected in wastewater from MP, GP and NW. Genotype I.3 was the second most prevalent at 18.75% (3/16) and comprised of two SaVs from MP and one from NW. Genotype II.1 contained two strains detected in MP and the other genotypes (GI.6, GI.7 and GII.2) were each represented by one SaV strain from MP. The strains from SA shared high nucleotide identities with SaV

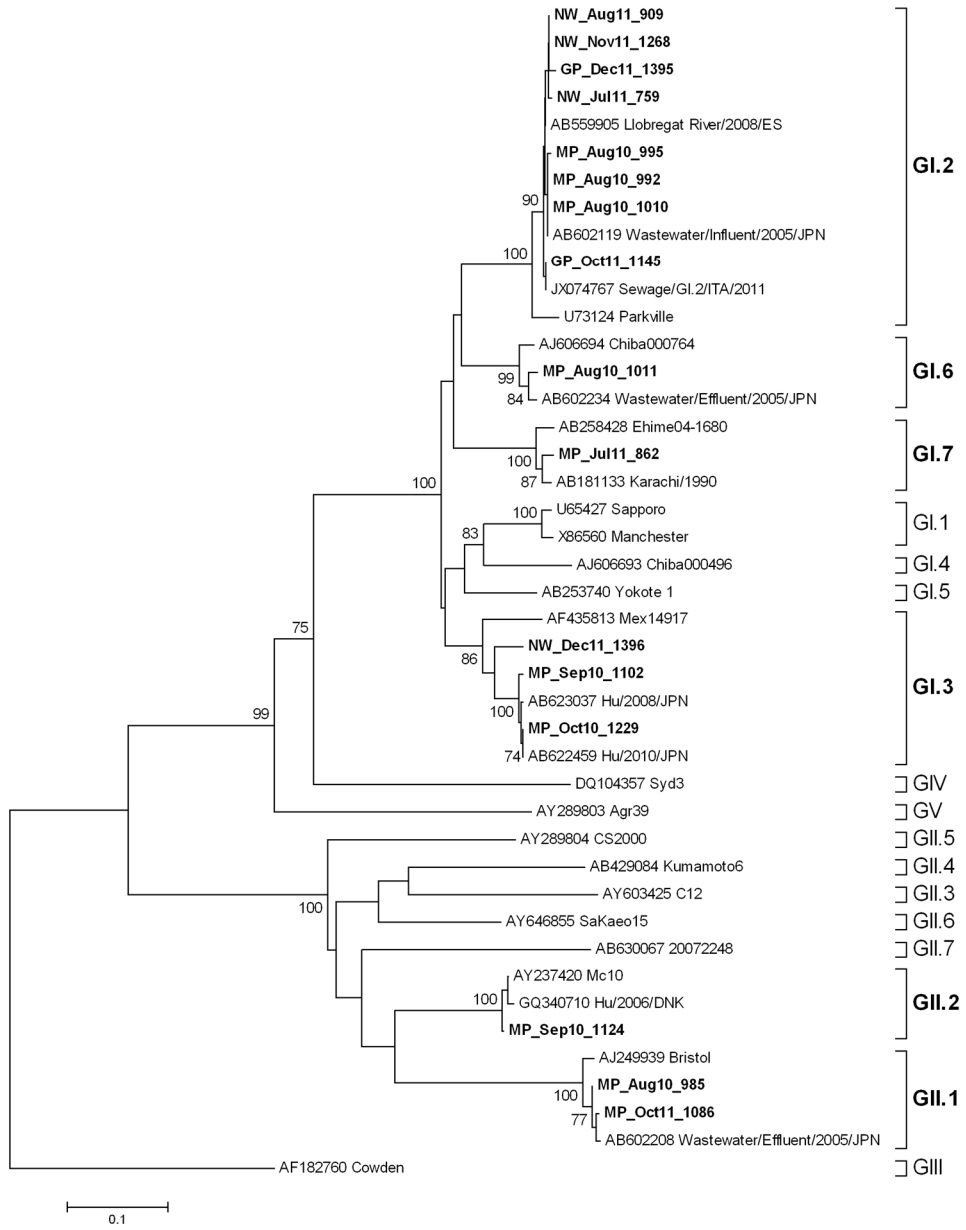

Figure 4 Phylogenetic analysis of partial capsid sequences (304 bp) of 16 sapovirus strains characterised from wastewater in South Africa. The tree was generated using the neighbour-joining method and 1000 bootstrap replicates. Groupings with bootstrap support of > 70% were considered significant. Reference sequences are indicated by GenBank accession numbers and strains from this study are shown in bold. The province, date (month and year) of collection, sample number and clone number are included in each sample name. The scale represents nucleotide substitutions per site.

strains detected in water sources and clinical specimens from various countries in Europe and Asia.

Discussion

Diverse CVs were characterised from wastewater collected from several provinces in SA. These included 13 different genotypes for NoV and six for SaV. A recent study in Argentina found limited NoV diversity (two genotypes) in rivers impacted by medium-sized populations when compared to larger cities (Blanco Fernandez *et al.* 2012). In contrast, many NoV and SaV genotypes were found in wastewater from treatment plants servicing small- and medium-sized communities in SA. The diversity observed was comparable to studies on effluent wastewater from larger urban communities in Japan (Kitajima *et al.* 2011; 2012). It is noteworthy that this high diversity was found even though only 75-100 ml of effluent was analysed compared to 1 L in the Japanese studies. Considerable NoV diversity was observed in each province of SA, with up to four strains detected in one sample and up to seven strains detected in one month. Interestingly, no mixed sequences were detected from direct sequencing of SaV amplicons. In a study based in Tunisia, the absence of mixed virus genotypes in sewage samples was also noted (Sdiri-Loulizi *et al.* 2010) and in Spain, few SaV-positive wastewater samples included multiple genotypes (Sano *et al.* 2011). In contrast, the majority of both influent and effluent samples from Japan contained multiple genotypes (Kitajima *et al.* 2011). While this lack of mixed sequences suggests a single SaV genotype in each wastewater sample, it is possible that the characterised genotype rather predominated in the specific sample or that other genotypes were present in the influent, but less resistant to the wastewater treatment process and thus absent from the effluent samples. Analysis of influent wastewater may yield more genotypes in a single sample.

Out of the five NoV GI genotypes that were identified, GI.1 and an unidentified GI strain predominated in this study. GI.1 had not been detected in clinical specimens in SA recently, but has been found in surface water in GP (Mans *et al.* 2013). The unassigned GI strain has not been detected in SA before. Interestingly, very similar viruses (97%-99% nucleotide identity over 285-311 residues) have been detected in North and West-Africa, the Middle East and Asia in wastewater (Aw and Gin 2010; Kamel *et al.* 2010) or clinical specimens (Kaplan *et al.* 2011; Oluwatoyin *et al.* 2012). On-going studies on NoVs in children with diarrhoea in various areas of SA should reveal whether this strain is clinically relevant in the SA population. The closely matched strain from Nigeria, detected in 2010, was proposed as a new GI.9 genotype (Oluwatoyin *et al.* 2012). However, a different strain detected in Sweden in 2008 had already been designated as GI.9 (Nenonen *et al.* 2012). Therefore, the strains clustering with the Nigerian strain should be designated GI.10 or remain unassigned until updated NoV nomenclature guidelines are available. So far only a few GI strains (GI.2, GI.7, GI.8) have been detected in clinical specimens in SA (Mans *et al.* 2010) of which GI.8 was detected in wastewater. In contrast, the prevalence of GI strains in surface water in GP (Mans *et al.* 2013) and in wastewater (present study) suggests that NoV GI infection is widespread in the SA community.

The two most prevalent NoV genotypes detected in treated wastewater in SA were GII.2 (12/52) and GII.4 (13/52). Eleven of the GII.2 strains formed a cluster separate from the GII.2 reference strains and one sewage sample NW_Oct11_1077_1. These GII.2 strains are most closely related to NoV strains detected in river water in SA in 2010 (Mans *et al.* 2013). They exhibit 97% identity over 282 bp at the 5'-end of the capsid gene. Other NoV strains with the highest similarity (95% identity over 305 bp) were detected in the US in 2002 (GenBank accession number: JQ320072) and in Russia in 2009 (JX524567). The other NoV GII.2 strain closely matched viruses detected in Korea in 2006 (EU249130).

Consistent with several other studies (Bucardo *et al.* 2011; Kremer *et al.* 2011; Blanco Fernandez *et al.* 2012) NoV GII.4 was one of the most prevalent genotypes found in wastewater in SA. The NoV GII.4 2010 variant (designated New_Orleans_2009 by the NoV Genotyping Tool) predominated (9/13) but other variants were also identified. A single sludge sample from KZN (2011) contained a NoV GII.4 2007 variant. The most closely matched strain (98.6% nucleotide identity over 287 residues) on GenBank was detected in 2007 in India (EU921388). The 2007 variant is considered a minor variant and has been detected at low frequency between 2007 and 2009 on several continents (Siebenga *et al.* 2009). Two other types of GII.4 strains were observed, neither of which could be assigned to a variant type based on phylogenetic analysis of the region C sequence with MEGA5 or by using the NoV Genotyping Tool. The first type clustered with one GII.4 2008 reference strain, without bootstrap support. Similar strains (with up to 99% identity over 303 bp) were circulating in Asia (HM635196, JX459638) and Europe (AB445395, GQ246793) between 2008 and 2010. The second type was represented by a single strain which clustered with 99% bootstrap support with the GII.4 genotype, but on a separate branch from all other GII.4 strains. Based on nucleotide sequence similarity this strain is 95% identical over 294 bp to NoV strains detected in human stool in October 2010 in China and January 2011 in Singapore. BLAST-p analysis revealed a strain detected in 2005 in China with 94% identity over 101 amino acids as another close match. All NoV GII genotypes detected in wastewater in SA, except GII.2 and GII.17, have been detected in clinical specimens in GP (Mans *et al.* 2010; unpublished data) confirming that wastewater effluent in several provinces of SA are contaminated with clinically relevant NoVs.

The prevalence of NoV GI and GII in the environment likely depends on complex factors such as possible differences in severity of disease caused by the two genogroups (Friesema *et al.* 2009), higher levels of virus shedding in individuals infected with GII (Chan

et al. 2006) and differences in survival of GI and GII viruses during wastewater treatment (da Silva *et al.* 2007). Contrasting reports on the predominance of a particular NoV genogroup in wastewater exist in the literature. Some studies found NoV GI more frequently than GII in treated wastewater samples (da Silva *et al.* 2007; Nordgren *et al.* 2009) or detected both genogroups at the same frequency (Kitajima *et al.* 2012), whereas the majority of investigations described NoV GII as the predominant genogroup detected in wastewater (Lodder and de Roda Husman 2005; Bucardo *et al.* 2011; Kremer *et al.* 2011; Blanco Fernandez *et al.* 2012; Perez-Sautu *et al.* 2012). This corresponds with the observation that NoV GII accounts for the majority of clinical gastroenteritis cases worldwide (Atmar 2010). However, the detection rates of NoV GI in the environment imply much more widespread circulation in the human population than suggested by its prevalence in the clinical setting (Kitajima *et al.* 2012).

Norovirus GIV was not detected in any of the wastewater samples in SA and while there is little information on NoV GIV worldwide, studies have documented the presence of NoV GIV in stool specimens from patients with gastroenteritis (Vinjé and Koopmans 2000; Fankhauser *et al.* 2002). Norovirus GIV was first reported in wastewater in the Netherlands in 2001, where 28% of treated and untreated samples were positive for GIV (van den Berg *et al.* 2005). In Italy 8% of raw sewage samples were positive for NoV GIV (La Rosa *et al.* 2010b) and in Luxembourg and France, NoV GIV has been detected in influent wastewater (Sima *et al.* 2011; Skraber *et al.* 2011). In Japan, NoV GIV was reported in both influent and effluent wastewater (Kitajima *et al.* 2009). In this study, NoV GIV was not detected in wastewater samples which comprised mostly of effluent samples. Either NoV GIV was not present or will only be detected in more concentrated samples such as sewage.

Several different SaVs within genogroups I and II were identified from wastewater in SA. Genogroup I SaVs predominated and this has also been noted in wastewater in other

countries. In Japan, a study on SaVs in raw sewage indicated that GI.1 predominated (Iwai *et al.* 2009) and another study on influent and effluent wastewater showed GI.2 to be most prevalent (Kitajima *et al.* 2011). In Spain, SaV GI.2 was also identified most often in wastewater (Sano *et al.* 2011).

The SA GI.2 cluster shared high (99-100%) nucleotide identity with SaVs detected in treated wastewater in Spain (2008) (Sano *et al.* 2011), influent wastewater in Japan (2005) (Kitajima *et al.* 2011) and sewage in Italy (2011). Two additional SaV strains from SA, which clustered in different genotypes (GI.6 and GII.1) grouped with strains detected in effluent wastewater in Japan (2005) (Kitajima *et al.* 2011). Overall, there was high nucleotide identity (98-100%) between the SaVs found in wastewater in SA and strains found in wastewater in other countries. The other SA strains (GI.3, GI.7 and GII.2) shared high nucleotide identity (95-100%) with SaVs identified in stools specimens from individuals with gastroenteritis in Japan (Oka *et al.* 2012), Pakistan (Phan *et al.* 2004) and Denmark (Johnsen *et al.* 2009). In a clinical setting, GI.2 has been associated with outbreaks of gastroenteritis in Europe from 2007 to 2009 (Svraka *et al.* 2010) and Japan in 2009 (Miyoshi *et al.* 2010). Sapovirus GIV, also frequently detected in stool from individuals with gastroenteritis and reported in outbreaks of gastroenteritis (Svraka *et al.* 2007; Harada *et al.* 2012; Lee *et al.* 2012), was not detected in wastewater in SA.

This study provides comprehensive data on the presence of CVs in wastewater in the northern region of SA. Analysis of wastewater from small- to medium-sized communities reveals remarkable CV diversity and the frequent occurrence of NoV GI and SaV indicates more extensive circulation in the population than is suggested by clinical studies (Mans *et al.* 2010). Since SA lacks a gastroenteritis outbreak reporting system as well as routine CV surveillance, characterisation of viruses in wastewater is a good approach to monitor for emerging genotypes and variants. This investigation determined the dominant CV genotypes

circulating in the country and led to a better understanding of CV epidemiology by contributing new data from the African continent where information is lacking.

Acknowledgements

The authors would like to thank Roney Hoffmann Consultants cc for providing the wastewater samples and Ms Mmapula Sima for technical assistance during the study.

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402.

Atmar, R.L. (2010) Noroviruses - State of the art. *Food Environ Virol* **2**, 117-126.

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

Blanco Fernandez, M.D., Torres, C., Poma, H.R., Riviello-Lopez, G., Martinez, L.C., Cisterna, D.M., Rajal, V.B., Nates, S.V. and Mbayed, V.A. (2012) Environmental surveillance of norovirus in Argentina revealed distinct viral diversity patterns, seasonality and spatio-temporal diffusion processes. *Sci Total Environ* **437**, 262-269.

Bucardo, F., Lindgren, P.E., Svensson, L. and Nordgren, J. (2011) Low prevalence of rotavirus and high prevalence of norovirus in hospital and community wastewater after introduction of rotavirus vaccine in Nicaragua. *PloS one* **6**, e25962.

Chan, M.C., Sung, J.J., Lam, R.K., Chan, P.K., Lee, N.L., Lai, R.W. and Leung, W.K. (2006) Fecal viral load and norovirus-associated gastroenteritis. *Emerg Infect Dis* **12**, 1278-1280.

da Silva, A.K., Le Saux, J.C., Parnaudeau, S., Pommepuy, M., Elimelech, M. and Le Guyader, F.S. (2007) Evaluation of removal of noroviruses during wastewater treatment, using real-time reverse transcription-PCR: different behaviors of genogroups I and II. *Appl Environ Microbiol* **73**, 7891-7897.

European Committee for Standardization (CEN) Technical Committee 275 (2010) CEN/TC 275/WG6/TAG4 - Detection of viruses in food and water.

Fankhauser, R.L., Monroe, S.S., Noel, J.S., Humphrey, C.D., Bresee, J.S., Parashar, U.D., Ando, T. and Glass, R.I. (2002) Epidemiologic and molecular trends of "Norwalk-like Viruses" associated with outbreaks of gastroenteritis in the United States. *J Infect Dis* **186**, 1-7.

Farkas, T., Zhong, W.M., Jing, Y., Huang, P.W., Espinosa, S.M., Martinez, N., Morrow, A.L., Ruiz-Palacios, G.M., Pickering, L.K. and Jiang, X. (2004) Genetic diversity among sapoviruses. *Arch Virol* **149**, 1309-1323.

Friesema, I.H.M., Vennema, H., Heijne, J.C.M., de Jager, C.M., Teunis, P.F.M., van der Linde, R., Duizer, E. and van Duynhoven, Y.T.H.P. (2009) Differences in clinical presentation between norovirus genotypes in nursing homes. *J Clin Virol* **46**, 341-344.

Green, K. (2007) Caliciviridae: The Noroviruses. In *Field's Virology* ed. Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B and SE, S. pp. 949-979. Philadelphia: Lippincott Williams & Wilkins.

Hansman, G.S., Sano, D., Ueki, Y., Imai, T., Oka, T., Katayama, K., Takeda, N. and Omura, T. (2007) Sapovirus in water, Japan. *Emerg Infect Dis* **13**, 133-135.

Harada, S., Oka, T., Tokuoka, E., Kiyota, N., Nishimura, K., Shimada, Y., Ueno, T., Ikezawa, S., Wakita, T., Wang, Q., Saif, L.J. and Katayama, K. (2012) A confirmation of sapovirus re-infection gastroenteritis cases with different genogroups and genetic shifts in the evolving sapovirus genotypes, 2002-2011. *Arch Virol* **157**, 1999-2003.

He, X.Q., Cheng, L., Zhang, D.Y., Xie, X.M., Wang, D.H. and Wang, Z. (2011) One-year monthly survey of rotavirus, astrovirus and norovirus in three sewage treatment plants (STPs) in Beijing, China and associated health risk assessment. *Water Sci Technol* **64**, 1202-1210.

Hewitt, J., Leonard, M., Greening, G.E. and Lewis, G.D. (2011) Influence of wastewater treatment process and the population size on human virus profiles in wastewater. *Water Res* **45**, 6267-6276.

Iwai, M., Hasegawa, S., Obara, M., Nakamura, K., Horimoto, E., Takizawa, T., Kurata, T., Sogen, S. and Shiraki, K. (2009) Continuous Presence of Noroviruses and Sapoviruses in Raw Sewage Reflects Infections among Inhabitants of Toyama, Japan (2006 to 2008). *Appl Environ Microbiol* **75**, 1264-1270.

Johnsen, C.K., Midgley, S. and Böttiger, B. (2009) Genetic diversity of sapovirus infections in Danish children 2005–2007. *J Clin Virol* **46**, 265-269.

Kageyama, T., Kojima, S., Shinohara, M., Uchida, K., Fukushi, S., Hoshino, F.B., Takeda, N. and Katayama, K. (2003) Broadly reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. *J Clin Microbiol* **41**, 1548-1557.

Kamel, A.H., Ali, M.A., El-Nady, H.G., Aho, S., Pothier, P. and Belliot, G. (2010) Evidence of the co-circulation of enteric viruses in sewage and in the population of Greater Cairo. *J Appl Microbiol* **108**, 1620-1629.

Kaplan, N.M., Kirby, A., Abd-Eldayem, S.A., Dove, W., Nakagomi, T., Nakagomi, O. and Cunliffe, N.A. (2011) Detection and molecular characterisation of rotavirus and norovirus infections in Jordanian children with acute gastroenteritis. *Arch Virol* **156**, 1477-1480.

Kitajima, M., Haramoto, E., Phanuwan, C. and Katayama, H. (2011) Genotype distribution of human sapoviruses in wastewater in Japan. *Appl Environ Microbiol* **77**, 4226-4229.

Kitajima, M., Haramoto, E., Phanuwan, C., Katayama, H. and Furumai, H. (2012) Molecular detection and genotyping of human noroviruses in influent and effluent water at a wastewater treatment plant in Japan. *J Appl Microbiol* **112**, 605-613.

Kitajima, M., Haramoto, E., Phanuwan, C., Katayama, H. and Ohgaki, S. (2009) Detection of genogroup IV norovirus in wastewater and river water in Japan. *Lett Appl Microbiol* **49**, 655-658.

Kitajima, M., Oka, T., Haramoto, E., Katayama, H., Takeda, N., Katayama, K. and Ohgaki, S. (2010) Detection and genetic analysis of human sapoviruses in river water in Japan. *Appl Environ Microbiol* **76**, 2461-2467.

Kremer, J.R., Langlet, J., Skraber, S., Weicherding, P., Weber, B., Cauchie, H.M., De Landtsheer, S., Even, J., Muller, C.P., Hoffmann, L. and Mossong, J. (2011) Genetic diversity of noroviruses from outbreaks, sporadic cases and wastewater in Luxembourg 2008-2009. *Clin Microbiol Infect* **17**, 1173-1176.

Kroneman, A., Vennema, H., Deforche, K., Avoort, H.V., Penaranda, S., Oberste, M.S., Vinje, J. and Koopmans, M. (2011) An automated genotyping tool for enteroviruses and noroviruses. *J Clin Virol* **51**, 121-125.

La Rosa, G., Iaconelli, M., Pourshaban, M. and Muscillo, M. (2010a) Detection and molecular characterization of noroviruses from five sewage treatment plants in central Italy. *Water Res* **44**, 1777-1784.

La Rosa, G., Iaconelli, M., Pourshaban, M., Fratini, M. and Muscillo, M. (2010b) Molecular detection and genetic diversity of norovirus genogroup IV: a yearlong monitoring of sewage throughout Italy. *Arch Virol* **155**, 589-593.

Lee, L.E., Cebelinski, E.A., Fuller, C., Keene, W.E., Smith, K., Vinjé, J. and Besser, J.M. (2012) Sapovirus outbreaks in long-term care facilities, Oregon and Minnesota, USA, 2002–2009. *Emerg Infect Dis* **18**, 873-876.

Lodder, W.J. and de Roda Husman, A.M. (2005) Presence of noroviruses and other enteric viruses in sewage and surface waters in The Netherlands. *Appl Environ Microbiol* **71**, 1453-1461.

Loisy, F., Atmar, R.L., Guillon, P., Le Cann, P., Pommepuy, M. and Le Guyader, F.S. (2005) Real-time RT-PCR for norovirus screening in shellfish. *J Virol Methods* **123**, 1-7.

Mans, J., de Villiers, J.C., du Plessis, N.M., Avenant, T. and Taylor, M.B. (2010) Emerging norovirus GII.4 2008 variant detected in hospitalised paediatric patients in South Africa. *J Clin Virol* **49**, 258-264.

Mans, J., Netshikweta, R., Magwalivha, M., van Zyl, W.B. and Taylor, M.B. (2013) Diverse norovirus genotypes identified in sewage-polluted river water in South Africa. *Epidemiol Infect* **141**, 303-313.

Minor, P.D. (1985) Growth, assay and purification of Picornaviruses. In *Virology: a practical approach* ed. Mahy, B.W.J. pp. 25-42. Washington: IRL Press.

Miyoshi, M., Yoshizumi, S., Kanda, N., Karino, T., Nagano, H., Kudo, S., Okano, M. and Ishida, S. (2010) Different genotypic sapoviruses detected in two simultaneous outbreaks of gastroenteritis among schoolchildren in the same school district in Hokkaido, Japan. *Jpn J Infect Dis* **63**, 75-78.

Murray, T.Y., Mans, J., van Zyl, W.B. and Taylor, M.B. (2012) Application of a competitive internal amplification control for the detection of sapoviruses in wastewater. *Food Environ Virol* DOI:10.1007/s12560-012-9101-1.

Nenonen, N.P., Hannoun, C., Larsson, C.U. and Bergström, T. (2012) Marked Genomic Diversity of Norovirus Genogroup I Strains in a Waterborne Outbreak. *Appl Environ Microbiol* **78**, 1846-1852.

Ng, T.F.F., Marine, R., Wang, C., Simmonds, P., Kapusinszky, B., Bodhidatta, L., Oderinde, B.S., Wommack, K.E. and Delwart, E. (2012) High Variety of Known and New RNA and DNA Viruses of Diverse Origins in Untreated Sewage. *J Virol* **86**, 12161-12175.

Nordgren, J., Matussek, A., Mattsson, A., Svensson, L. and Lindgren, P.E. (2009) Prevalence of norovirus and factors influencing virus concentrations during one year in a full-scale wastewater treatment plant. *Water Res* **43**, 1117-1125.

Oka, T., Mori, K., Iritani, N., Harada, S., Ueki, Y., Iizuka, S., Mise, K., Murakami, K., Wakita, T. and Katayama, K. (2012) Human sapovirusclassification based on complete capsid nucleotide sequences. *Arch Virol* **157**, 349-352.

Oluwatoyin, M.J., Adesina, O.A., Famurewa, O., Svensson, L. and Nordgren, J. (2012) Molecular epidemiology of rotavirus and norovirus in Ile-Ife, Nigeria: High prevalence of G12P[8] rotavirus strains and detection of a rare norovirus genotype. *J Med Virol* **84**, 1489-1496.

Perez-Sautu, U., Sano, D., Guix, S., Kasimir, G., Pinto, R.M. and Bosch, A. (2012) Human norovirus occurrence and diversity in the Llobregat river catchment, Spain. *Environ Microbiol* **14**, 494-502.

Phan, T.G., Okame, M., Nguyen, T.A., Maneekarn, N., Nishio, O., Okitsu, S. and Ushijima, H. (2004) Human astrovirus, norovirus (GI, GII), and sapovirus infections in Pakistani children with diarrhea. *J Med Virol* **73**, 256-261.

Sano, D., Pérez-Sautu, U., Guix, S., Pintó, R.M., Miura, T., Okabe, S. and Bosch, A. (2011) Quantification and genotyping of human sapoviruses in the Llobregat river catchment, Spain. *Appl Environ Microbiol* **77**, 1111-1114.

Sdiri-Loulizi, K., Hassine, M., Aouni, Z., Gharbi-Khelifi, H., Chouchane, S., Sakly, N., Neji-Guédiche, M., Pothier, P., Aouni, M. and Ambert-Balay, K. (2010) Detection and molecular characterization of enteric viruses in environmental samples in Monastir, Tunisia between January 2003 and April 2007. *J Appl Microbiol* **109**, 1093-1104.

Siebenga, J.J., Vennema, H., Zheng, D.P., Vinje, J., Lee, B.E., Pang, X.L., Ho, E.C., Lim, W., Choudekar, A., Broor, S., Halperin, T., Rasool, N.B., Hewitt, J., Greening, G.E., Jin, M., Duan, Z.J., Lucero, Y., O'Ryan, M., Hoehne, M., Schreier, E., Ratcliff, R.M., White, P.A., Iritani, N., Reuter, G. and Koopmans, M. (2009) Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001-2007. *J Infect Dis* **200**, 802-812.

Siebenga, J.J., Vennema, H., Renckens, B., de Bruin, E., van der Veer, B., Siezen, R.J. and Koopmans, M. (2007) Epochal Evolution of GGII.4 Norovirus Capsid Proteins from 1995 to 2006. *J Virol* **81**, 9932-9941.

Sima, L.C., Schaeffer, J., Le Saux, J.C., Parnaudeau, S., Elimelech, M. and Le Guyader, F.S. (2011) Calicivirus removal in a membrane bioreactor wastewater treatment plant. *Appl Environ Microbiol* **77**, 5170-5177.

Skraber, S., Langlet, J., Kremer, J.R., Mossong, J., De Landtsheer, S., Even, J., Muller, C.P., Hoffmann, L. and Cauchie, H.M. (2011) Concentration and diversity of noroviruses detected in Luxembourg wastewaters in 2008-2009. *Appl Environ Microbiol* **77**, 5566-5568.

Svraka, S., Duizer, E., Vennema, H., de Bruin, E., van der Veer, B., Dorresteijn, B. and Koopmans, M. (2007) Etiological role of viruses in outbreaks of acute gastroenteritis in The Netherlands from 1994 through 2005. *J Clin Microbiol* **45**, 1389-1394.

Svraka, S., Vennema, H., van der Veer, B., Hedlund, K.O., Thorhagen, M., Siebenga, J., Duizer, E. and Koopmans, M. (2010) Epidemiology and genotype analysis of emerging sapovirus-associated infections across Europe. *J Clin Microbiol* **48**, 2191-2198.

Teunis, P.F., Moe, C.L., Liu, P., Miller, S.E., Lindesmith, L., Baric, R.S., Le Pendu, J. and Calderon, R.L. (2008) Norwalk virus: how infectious is it? *J Med Virol* **80**, 1468-1476.

Trujillo, A.A., McCaustland, K.A., Zheng, D.P., Hadley, L.A., Vaughn, G., Adams, S.M., Ando, T., Glass, R.I. and Monroe, S.S. (2006) Use of TaqMan real-time reverse transcription-PCR for rapid detection, quantification, and typing of norovirus. *J Clin Microbiol* **44**, 1405-1412.

van den Berg, H., Lodder, W., van der Poel, W., Vennema, H. and de Roda Husman, A.M. (2005) Genetic diversity of noroviruses in raw and treated sewage water. *Res Microbiol* **156**, 532-540.

Victoria, M., Rigotto, C., Moresco, V., de Abreu Correa, A., Kolesnikovas, C., Leite, J.P., Miagostovich, M.P. and Barardi, C.R. (2010) Assessment of norovirus contamination in environmental samples from Florianopolis City, Southern Brazil. *J Appl Microbiol* **109**, 231-238.

Vinjé, J. and Koopmans, M.P.G. (2000) Simultaneous detection and genotyping of "Norwalk-Like Viruses" by oligonucleotide array in a reverse line blot hybridization format. *J Clin Microbiol* **38**, 2595-2601.

Zheng, D.P., Ando, T., Fankhauser, R.L., Beard, R.S., Glass, R.I. and Monroe, S.S. (2006) Norovirus classification and proposed strain nomenclature. *Virology* **346**, 312-323.