Diverse norovirus genotypes identified in sewage-polluted river water in South Africa

J. MANS[1]*, R. NETSHIKWETA[1], M. MAGWALIVHA[1], W. B. VAN ZYL[1]
AND M. B. TAYLOR[1,2]

[1] *Department of Medical Virology, Faculty of Health Sciences, University of Pretoria, South Africa*
[2] *National Health Laboratory Services, Tshwane Academic Division, Pretoria, South Africa*

SUMMARY

This study aimed to assess norovirus (NoV) contamination and genotype diversity in surface water in Gauteng, South Africa. Between January 2008 and December 2010, three rivers, namely Klip, Suikerbosrant, and Rietspruit were monitored for NoV genogroup (G)I and GII. Viruses were recovered using the glass wool adsorption-elution technique and detected by real-time reverse transcription–polymerase chain reaction. From 2008 to 2010, NoVs were detected in 66% (70/106) of Klip river samples. The Rietspruit and Suikerbosrant rivers were contaminated with NoV in 95% (20/21) and 21% (5/24) of samples, respectively. NoV-positive samples comprised of 33% GI, 29% GII and 38% of both GI and GII strains. Based on partial capsid gene analysis (region C), 16 NoV genotypes (6 GI, 10 GII) were identified. The major genotypes detected were GI.4, GI.5 and GII.4. These rivers could be a potential source of NoV infection for communities using the water for domestic or recreational purposes.

Key words: Genotypes, norovirus, river water, South Africa.

INTRODUCTION

Noroviruses (NoVs), characterized by high genetic and antigenic diversity, are important human pathogens that cause acute gastroenteritis [1]. These small, non-enveloped viruses have a positive-sense, single-stranded RNA genome, organized into three open reading frames (ORFs). ORF1 encodes a polyprotein that is processed in infected cells to produce several non-structural proteins, e.g. RNA polymerase. ORF2 encodes the major capsid protein (VP1) and a minor structural protein (VP2) is encoded by ORF3 [1]. NoVs are divided into five genogroups (G) based on phylogenetic analysis of the capsid protein [2]. NoVs GI, GII and GIV infect humans and GI and GII have been associated with epidemic [3] as well as sporadic gastroenteritis in all age groups [4]. In the last decade GII.4 strains have caused the majority of NoV outbreaks. Novel variants of GII.4 viruses appear to emerge almost annually, and some variants spread globally and replace previously dominant GII.4 strains [5].

NoVs are characterized by high environmental stability and a very low infectious dose [6]. Accordingly, NoVs are frequently implicated in foodborne and waterborne gastroenteritis outbreaks [7]. Outbreaks of gastrointestinal disease caused by NoV in South Africa (SA) were first described in 1993 [8]. A seroprevalence survey indicated widespread exposure to NoVs in a large portion of the population

* Author for correspondence: Dr J. Mans, Department of Medical Virology, Faculty of Health Sciences, University of Pretoria, Private Bag X323, Arcadia, 0007, South Africa.
(Email: janet.mans@tuks.co.za)

Fig. 1. Diagrammatic map of the Vaal river and selected tributaries in Gauteng province, South Africa. The sampling sites in the Klip (1), Suikerbosrant (2) and Rietspruit (3) rivers are shown in the enlargement. (Map of South Africa obtained from http://d-maps.com/carte.php?num_car = 11688&lang = en.)

[9]. More recently a study in Gauteng province, SA indicated that after rotavirus (24 % prevalence), NoV was the most frequently detected viral pathogen in paediatric patients (14 % prevalence) hospitalized with viral gastroenteritis [10]. In spite of the evident importance of NoV infections, NoVs are not routinely tested for in diagnostic laboratories and NoV infections are probably underreported. The rapid evolution and global spread of NoV strains has been the subject of intense study in the past decade. However, very limited sequence data is available on the NoV strains currently circulating in SA and the African continent; therefore, it is important to characterize NoVs from this region.

In addition to the analysis of clinical specimens, the detection and characterization of viruses in the environment is an effective approach to gain insight into the enteric viruses circulating in a given population [11]. The presence and diversity of NoVs in surface water has been described in Asia [12, 13], Europe [14] and South America [15, 16]. However, no investigations have addressed the occurrence and molecular epidemiology of NoVs in surface water, or other water sources, in SA.

The Klip, Suikerbosrant and Rietspruit rivers in Gauteng province, are tributaries of the Vaal river which forms part of one of the largest river systems in SA. The aim of this study was to monitor these three rivers for the presence of NoV GI and GII, to genotype the detected NoVs and to establish which strains are circulating in the Gauteng population. These data will provide new sequence information for NoV strains occurring in SA, as well as information regarding the infectious potential of the river water for communities using the water for domestic, agricultural or recreational purposes.

METHODS

Water sampling and virus concentration

From January 2008 to December 2010, water samples were collected from the same sampling sites from three rivers in Gauteng, SA (Fig. 1). Samples (10 l) from the Klip river (106) were collected weekly in 2008 and 2009 and twice a month in 2010, except when technical difficulties arose. The Suikerbosrant river (24 samples) and the Rietspruit river (21 samples), were sampled less frequently over the 3-year period. The Klip river flows through the southern part of greater Johannesburg, specifically the Soweto area which has an estimated population of >2 million and includes informal settlements. The Rietspruit river is situated in a region with smaller towns and the Suikerbosrant river originates in a nature reserve and then flows through human settlements. Samples were collected at a depth of 25 cm below the surface of the river and transported to the laboratory in cooler bags with cold packs. The temperature and pH were recorded upon arrival. Viruses were recovered from the 10-l samples using a modification of the glass wool adsorption-elution method developed by Vilaginès and co-workers [17]. The preparation of the glass wool column was modified as follows: 15 g glass wool was used per column and a steel gauze grid (pore size = 1 mm², 30 mm diameter) was inserted between each of the three 5-g portions of glass wool. The viruses in the eluate (100 ml) were concentrated to a final volume of 20 ml in phosphate-buffered saline (pH 7·4) (PBS; Sigma-Aldrich Co., USA) by polyethylene glycol/sodium chloride precipitation [18]. The average recovery efficiency of NoVs from turbid water using this method was <1 % (V. Ruhanya, unpublished data).

Thermotolerant coliform determination

The thermotolerant (faecal) coliform count [colony-forming units (c.f.u.)/100 ml] in the water samples was determined using the membrane filtration technique and m-FC agar (Difco™, Becton Dickinson and Company, USA) with incubation at 44·5 °C for 18–24 h [19].

Virus isolation

Using conventional cell culture techniques, monolayers of the human hepatoma cell line PLC/PRF/5 (ECACC 85061113) and the BGM African Green Monkey cell line (ECACC 90092601) in 25-cm² cell culture flasks were infected with 1 ml of antibiotic- and antimycotic-treated virus concentrates [20]. After a 14-day incubation period, which included a blind passage at day 7 post-infection, the infected cells were harvested for molecular analysis for human adenoviruses [21] and enteroviruses [20], as well as for blind passage onto monolayers of Vero African Green Monkey cell line (ECACC 84113001) in cell culture tubes with flying coverslips. After a further incubation of 7 days at 37 °C, the infected cell cultures were examined for cytopathic effects, stained with haematoxylin and eosin and examined for virus-specific inclusion bodies [22].

Nucleic acid extraction and detection

Total nucleic acid was extracted from the virus concentrate (1 ml) with the MagNA Pure LC Total Nucleic Acid Isolation kit (large volume) (Roche Diagnostics, Germany) in a robotic MagNA Pure LC instrument (Roche Diagnostics). Nucleic acids were eluted in 100 μl and stored at −70 °C until use.

Published primer sets and Taqman probes (Table 1) were used for the detection of NoV GI [23, 24] and NoV GII [25, 26]. For NoV GI, cDNA (20 μl) was prepared from 10 μl total RNA with random primers using the Transcriptor First Strand cDNA Synthesis kit (Roche Diagnostics). Five microlitres of cDNA was subsequently used for virus-specific amplification and detection with the LightCycler TaqMan Master mix (Roche Diagnostics). NoV GII was detected with a one-step real-time reverse transcription–polymerase chain reaction (RT–PCR) assay (5 μl RNA, QuantiTect Probe RT–PCR kit; Qiagen Inc., USA). The two-step real-time RT–PCR was used for NoV GI as this was found to be more sensitive than a one-step assay. A positive control (viral nucleic acid extracted from a characterized clinical strain) and negative control (nuclease-free water: Promega Corp., USA) were included for all real-time RT–PCR reactions. Samples with cycle threshold (Ct) values >40 were considered negative. The detection limit for the GI and GII detection assays was determined using plasmid DNA standards (CeeramTools®, Ceeram S.A.S, France) and was found to be 170 and 150 cDNA copies/reaction, respectively.

RT–PCR inhibition control

All NoV-negative samples were re-tested at a 1/10 dilution of the RNA in nuclease-free water to ensure that the assay was not inhibited by compounds present in the water samples. Samples that still tested negative with diluted RNA were spiked with ∼5 × 10³ copies of NoV GI or NoV GII RNA transcripts. The Ct value obtained in the RT–PCR on nuclease-free water spiked with the same number of gene copies was compared to that obtained in the spiked water samples. Since intra-assay variation was <1 Ct value, any reaction exhibiting a shift of >1 Ct value was considered to show levels of inhibition.

Genotyping and phylogenetic analysis

For characterization, a highly conserved region at the 5′-end of the NoV capsid gene (region C) was amplified and sequenced. cDNA was prepared as described for the detection of NoV GI. Samples collected in 2008 and 2009 were typed as follows: conventional PCR was performed with published primers G1SKF, G1SKR, G2SKF and G2SKR (Table 1) [27] and AmpliTaq Gold (Applied Biosystems, USA) according to the manufacturer's recommendations. If no PCR product was obtained after one PCR reaction, a second round PCR was performed using the same primers with 1 μl of the first PCR reaction as template. In order to reduce non-specific amplification, a semi-nested PCR was introduced to type samples collected in 2010. First-round PCR was performed with primers QNIF4/G1SKR or QNIF2/G2SKR and second-round PCR with primers G1SKF/G1SKR or G2SKF/G2SKR (Table 1). PCR products were cloned using the ClonJET™ PCR cloning kit (Fermentas, Canada) and 10 randomly selected clones were sequenced using pJet1.2/blunt specific primers and the ABI PRISM BigDye® Terminator v. 3.1 Cycle Sequencing kit on an ABI 3130 automated analyser (Applied Biosystems, USA).

Table 1. *Primer and probe sequences used for real-time RT-PCR detection and genotyping of norovirus GI and GII*

Virus	Primer/probe	Sequence (5′–3′)*	Polarity	Position
Detection				
Norovirus GI	QNIF4	CGCTGGATGCGNTTCCAT [23]	+	5291–5308†
	NV1LCR	CCTTAGACGCCATCATCATTTAC [24]	−	5354–5376†
	NVGG1 probe	FAM-TGGACAGGAGAYCGCRATCT-TAMRA [24]		5321–5340†
Norovirus GII	QNIF2	ATGTTCAGRTGGATGAGRTTCTCWGA [26]	+	5012–5037‡
	COG2R	TCGACGCCATCTTCATTCACA [25]	−	5080–5100‡
	QNIFS probe	FAM-AGCACGTGGGAGGGCGATCG-TAMRA [26]		5042–5061‡
Genotyping				
Norovirus GI	G1SKF	CTGCCCGAATTYGTAAATGA [27]	+	5342–5361†
	G1SKR	CCAACCCARCCATTRTACA [27]	−	5653–5671†
or	QNIF4	CGCTGGATGCGNTTCCAT [23]	+	5291–5308†
	G1SKR	CCAACCCARCCATTRTACA [27]	−	5653–5671†
Norovirus GII	G2SKF	CNTGGGAGGGCGATCGCAA [27]	+	5046–5064‡
	G2SKR	CCRCCNGCATRHCCRTTRTACAT [27]	−	5367–5389‡
or	QNIF2	ATGTTCAGRTGGATGAGRTTCTCWGA [26]	+	5012–5037‡
	G2SKR	CCRCCNGCATRHCCRTTRTACAT [27]	−	5367–5389‡

* IUPAC codes used to indicate degenerate positions.
† GenBank accession number M87661.
‡ GenBank accession number X86557.

Nucleotide sequences were analysed as described previously [10]. A representative sequence of each set of clones was used for phylogenetic analyses. The 5′-end of the capsid gene (285 bp for GI and 273 bp for GII) of the NoV strains was aligned with reference sequences from genogroups I and II using MAFFT version 6 (https://align.bmr.kyushu-u.ac.jp/mafft/online/server/). After manual adjustment of the alignment, phylogenetic analysis was performed with MEGA4 using the neighbour-joining method [28]. The evolutionary distances were computed using the maximum composite likelihood method [29]. Genotypes were assigned based on the clustering in the phylogenetic tree ($>70\%$ bootstrap support). Genotype assignment was confirmed by analysis of all sequences with the NoV Genotyping Tool [30] (http://www.rivm.nl/mpf/norovirus/typingtool).

Nucleotide sequence accession numbers

The nucleotide sequences determined in this investigation have been deposited in GenBank under accession numbers HQ201641–HQ201679 and JN191355–JN191381.

Statistical analyses

Statistical analyses were performed using simple interactive statistical analysis (SISA) [31], and associations were considered significant at the 95% confidence level. The t test was used to determine the significance of differences between genogroup proportions. Two-by-two table analysis was used to determine association between the presence of NoVs and thermotolerant coliform counts.

RESULTS

Occurrence of NoVs in three rivers in Gauteng

Over the 3-year period NoVs were detected in 66% [81% (2008), 53% (2009) and 61·5% (2010)] of the Klip river samples. Overall NoV GI (64%) and GII (67%) were detected at similar frequencies; however, the GI/GII ratio changed from predominant GII detection in 2008 (16/34 GI *vs.* 28/34 GII, $P = 0.002$) to predominant GI detection in 2009 (16/20 GI *vs.* 7/20 GII, $P = 0.002$) with similar detection rates in 2010 (13/16 GI *vs.* 12/16 GII) (Table 2). The proportions were compared using the t test and samples that contained both genogroups were included. NoVs, detected in the Klip river in at least 7 months of each year, showed no seasonal pattern over the 3-year period (Fig. 2). Thermotolerant coliform counts frequently exceeded recommended levels for water used for recreational purposes (>2000 c.f.u./100 ml) [32], showed sharp increases at sporadic intervals during the year and reached levels of 84 000 c.f.u./100 ml in

Table 2. *Norovirus (NoV) genogroups I and II detected at three sampling sites during 2008–2010 in Gauteng, South Africa. The NoV genotypes that could be determined are indicated. Samples positive for cytopathogenic viruses (adenovirus, enterovirus and reovirus) are also shown*

Year	River	Real-time RT–PCR detection (% positive)			Total (%) NoV +/total	Typing efficiency Typed/NoV +	NoV genotypes Based on partial capsid sequence (region C)	Virus isolation Cytopathogenic virus/NoV +
		NoV GI	NoV GII	NoV GI and GII				
2008	Klip (n=42)	6 (14·3%)	18 (42·9%)	10 (23·8%)	34/42 (81%)	12/34	I.1, I.4, II.2, II.4, II.6, II.9, II.13, II.17	14/34
	Rietspruit (n=3)	0	0	3 (100%)	3/3 (100%)	3/3	I.5, I.8, II.9	3/3
	Suikerbosrant (n=8)	0	1 (12·5%)	0	1/8 (12·5%)	0/1	n.t.	0/1
2009	Klip (n=38)	13 (34·2%)	4 (10·5%)	3 (10·5%)	20/38 (53%)	6/20	I.1, I.2, I.4, I.5, II.12	4/20
	Rietspruit (n=6)	1 (16·6%)	1 (16·6%)	3 (50%)	5/6 (83%)	4/5	I.5, II.3, II.17	4/5
	Suikerbosrant (n=6)	1 (16·6%)	1 (16·6%)	0	2/6 (33%)	0/2	n.t.	1/2
2010	Klip (n=26)	4 (15·4%)	3 (11·5%)	9 (34·6%)	16/26 (61·5%)	9/16	I.1, I.3, I.4, I.5, II.1, II.2, II.3, II.4	10/16
	Rietspruit (n=12)	4 (33·3%)	0	8 (66·7%)	12/12 (100%)	6/12	I.3, I.4, II.2, II.4, II.7, II.13	12/12
	Suikerbosrant (n=10)	2 (20%)	0	0	2/10 (20%)	0/2	n.t.	0/2

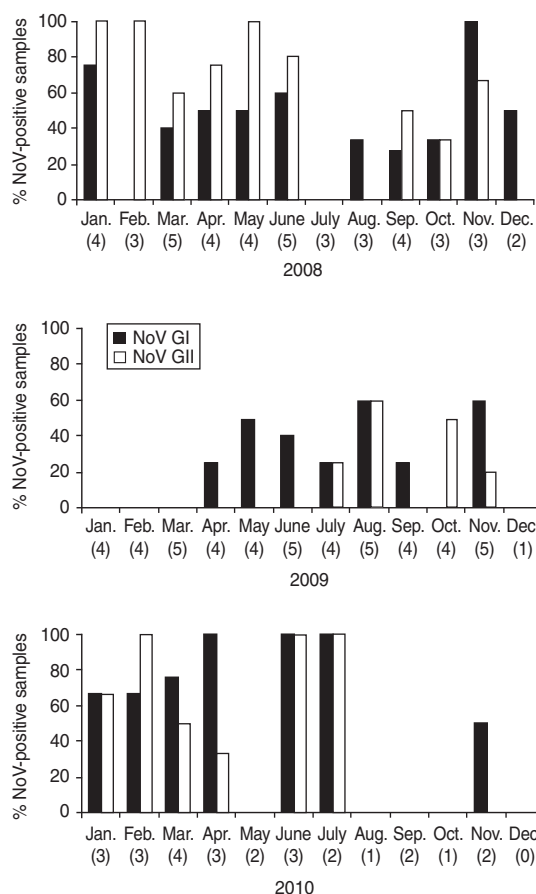n.t., Not typed as conventional RT–PCR amplification was unsuccessful.

Fig. 2. Percentage of norovirus (NoV) GI-positive (■) and NoV GII-positive (□) Klip river water samples detected in 2008, 2009 and 2010. The number in parentheses below each month indicates the number of samples analysed.

2008, 19 500 c.f.u./100 ml in 2009 and 26 000 c.f.u./100 ml in 2010. Cytopathogenic viruses (adenovirus, enterovirus, reovirus) were detected in 40% (28/70) of the NoV-positive Klip river samples, suggesting that potentially infectious viruses were present in the water samples (Table 2).

The Rietspruit river was sampled only three times in 2008 and all samples were positive for NoV GI and GII. Increased sampling of Rietspruit river during 2009 (six samples) and 2010 (12 samples) confirmed the high levels of enteric virus contamination, with 83% and 100% NoV-positive samples identified in 2009 and 2010, respectively. Thermotolerant coliform counts ranged from 880 to 40 000 c.f.u./100 ml, highlighting the extent of faecal pollution in this river. Both NoV GI and GII were detected in the majority (14/21) of the Rietspruit samples. The third river, Suikerbosrant, was sampled 6–10 times per year and NoV contamination was detected in 12·5–33% of the samples. Thermotolerant coliform count levels ranged

from 37 to 9700 c.f.u./100 ml. The total number of NoV-positive water samples with acceptable levels of thermotolerant coliforms (<2000, limit for recreational water) was compared to NoV-positive water samples with levels >2000. Two-by-two table analysis showed no association between the levels of bacterial indicators and NoV presence (χ^2, $P = 0.08$).

The real-time RT–PCR Ct values were used as an indication of relative concentration of NoVs in the water samples. The Ct values observed for the three rivers ranged between 33·7 and 40 (Klip river), 33·1–39·4 (Rietspruit river) and 35·3–37·5 (Suikerbosrant river). The mean Ct value observed for Rietspruit samples (35·76) was lower than the mean Ct values observed in Klip (36·93) and Suikerbosrant (36·87) river samples suggesting a higher viral load. Although the difference between the means was not statistically significant, genotyping viruses detected in the Rietspruit river was more successful.

Overall, 62·9% (95/151) of the water samples screened in this study were positive for NoV. Fifteen of the 95 NoV-positive samples (3 GI, 12 GII) were detected only when diluted target nucleic acid was tested. To confirm whether the remaining NoV-negative samples were true negatives, the samples were spiked with NoV GI or GII RNA transcripts and the Ct values obtained were compared to that of nuclease-free water spiked with the same number of RNA copies. Partial inhibition (0·5–5 Ct value shift) was observed in 29% and 72% of the NoV GI- and Nov GII-negative samples, respectively, and total RT–PCR inhibition was observed in 3·6% of the samples.

NoV genotypes detected in the Klip and Rietspruit rivers

Out of 95 NoV-positive samples, 40 could be amplified successfully for nucleotide sequencing and genotyping based on region C of the capsid gene. The typing efficiency varied from site to site, 38·6% of NoVs from the Klip river and 65% of NoV strains from the Rietspruit river could be genotyped whereas no NoVs from the Suikerbosrant river were typable (Table 2). Sixteen different NoV genotypes (6 GI, 10 GII) were identified during the 3-year sampling period in the Klip and Rietspruit rivers (Table 2, Figs 3 and 4). Based on real-time RT–PCR results GII NoVs (32/38 vs. 19/38 GI, $P = 0.001$) predominated in 2008, whereas GI NoVs predominated in 2009 (GI 21/27 vs. GII 12/27, $P = 0.01$) and 2010 (GI 27/30,

GII 20/30, $P = 0.027$). Overall, 38% of samples contained both GI and GII viruses and up to three different genotypes were detected simultaneously in single river water samples.

Phylogenetic analysis of the NoV GI reference strains and virus strains detected in the Klip and Rietspruit rivers from 2008 to 2010 are shown in Figure 3. The following strains were detected: GI.1 (Norwalk), GI.2, GI.3, GI.4, GI.5 and GI.8. No single GI genotype dominated during 2008. In 2009, more than 60% of the typed viruses were GI.5 (Fig. 3) and genotypes GI.1, GI.2, and GI.4 were also identified. In 2010, the genotype distribution changed again with GI.4 (50%) and GI.3 (33%) detected most commonly. The comparison of the GII virus strains with reference strain sequences is depicted in Figure 4. In 2008, GII.6 (29%), GII.4 (24%) and GII.13 (18%) viruses were detected most frequently. BLAST analysis of the GII.6 strains showed that they are most similar to strains identified in Mexico in 2007 and in Japan in 2008. These GII.6 viruses were also 98–100% identical (over 273 nucleotides) to GII.6 strains identified in clinical specimens from paediatric patients in Gauteng [10]. Several other genotypes were also found (Fig. 4). In 2009, GII.3, GII.7, GII.12 and GII.17 were identified in single samples. In 2010 the majority of strains that could be typed belonged to GII.4 (7/15). GII.2 was detected in three samples of which two were closely related to a GII.2 strain detected in water in SA in 2008 (present study) and a strain identified in Japan in 2008. The other GII.2 virus (2010-09-13-R2) grouped with a more recently (2009) identified strain from Russia. Of the GII.4 strains, only one could be assigned a variant type based on this phylogenetic tree. Strain 2010-03-29-R was classified as GII.4 2010 as it was identical to the reference GII.4 2010 strain (GenBank ID: GU445325) over 273 nucleotides. The other GII.4 strains all clustered with the 2006a lineage of GII.4 variants, but could not be assigned variant types due to insignificant bootstrap support.

DISCUSSION

This investigation demonstrated a high diversity of NoVs in three rivers in Gauteng, SA. These data provided new insight into the variety of NoV genotypes co-circulating in this community and the high degree of environmental NoV pollution. In these three rivers the thermotolerant coliform counts frequently exceeded quality limits for drinking water
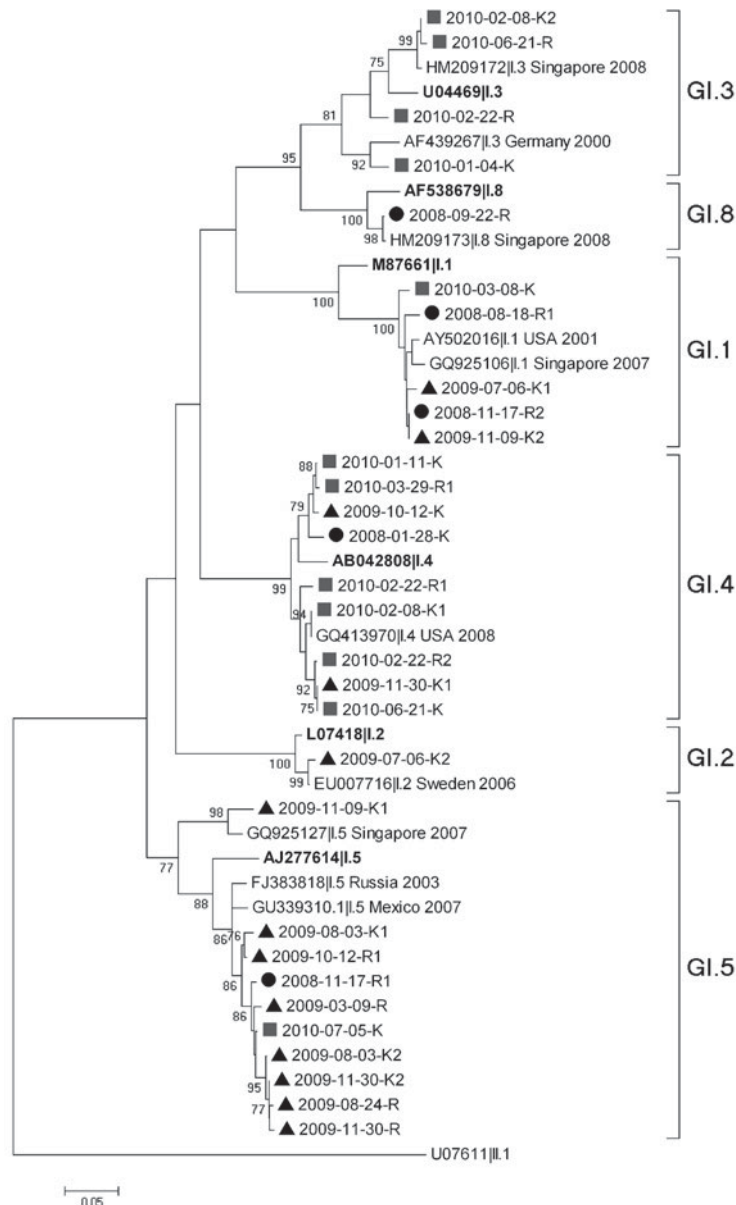
Fig. 3. Neighbour-joining phylogenetic tree of norovirus (NoV) GI reference strains (indicated by their GenBank accession numbers in bold), NoV strains detected in the Klip (K) and the Rietspruit (R) rivers, during 2008 (●), 2009 (▲) and 2010 (■) and their most closely related sequences in GenBank. The bootstrap consensus tree inferred from 1000 replicates is shown with bootstrap support >70% indicated. All positions containing alignment gaps and missing data were eliminated in pairwise sequence comparisons.

(0 c.f.u./100 ml) [33], irrigation water (<1000 c.f.u./ 100 ml) [34] and water used for recreational purposes (<2000 c.f.u./100 ml) [32]. NoVs were also detected in water samples that conformed to quality guidelines for thermotolerant coliforms for water used for irrigation (63%) and recreational (64%) purposes. These findings are of potential public health concern, since the river water is used for domestic purposes in informal settlements and for irrigation of fresh produce which is often eaten raw or minimally processed [35].

In addition, the rivers feed into the Vaal Barrage which is used for many recreational activities [35].

Treated sewage effluent is discharged into the Klip river system from sewage treatment plants servicing the city of Johannesburg. The effluent is generally of good quality; however, occasionally the performance of the treatment works can be affected by excessive storm water loads and problems with ageing infrastructure [36]. Run-off from informal settlements with inadequate or no sanitation also contributes to
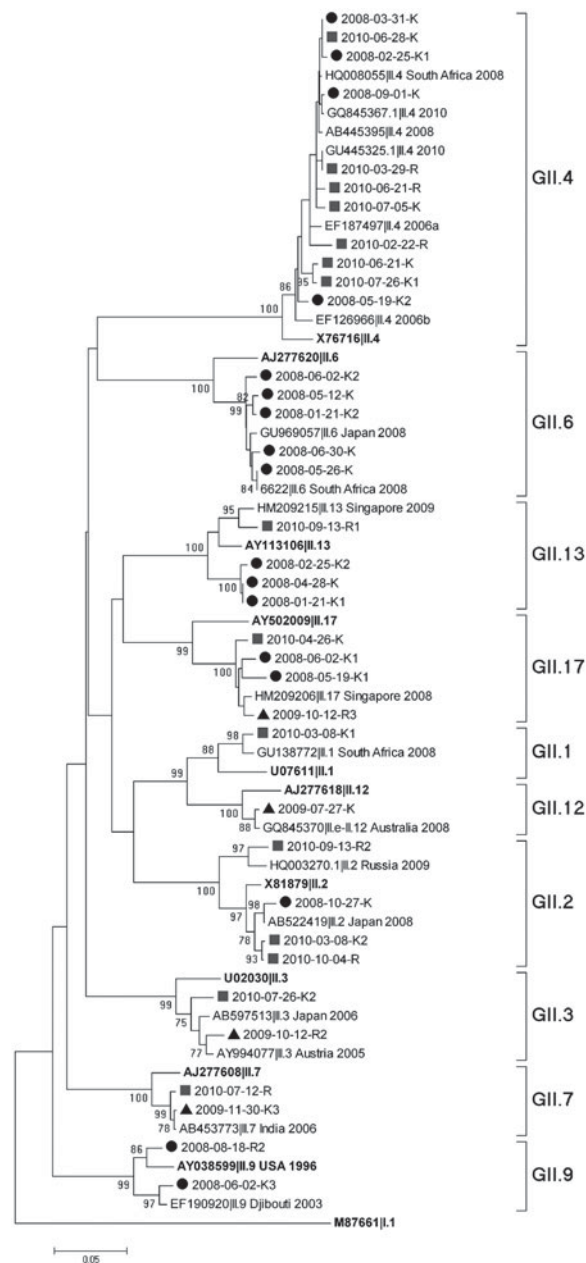
Fig. 4. Neighbour-joining phylogenetic tree of norovirus (NoV) GII reference strains (indicated by their GenBank accession numbers in bold), NoV strains detected in the Klip (K) and the Rietspruit (R) rivers, during 2008 (●), 2009 (▲) and 2010 (■) and their most closely related sequences in GenBank. The bootstrap consensus tree inferred from 1000 replicates is shown with bootstrap support >70% indicated. All positions containing alignment gaps and missing data were eliminated in pairwise sequence comparisons.

the pollution of the rivers. The most contaminated site in the study, the Rietspruit river, is most likely impacted by upstream dysfunctional municipal waste-water treatment works.

No seasonal pattern was observed over the 3-year period which is in contrast with environmental NoV surveys in Japan, where a higher occurrence of NoVs was observed in the winter/spring season [12]. There is evidence that NoV epidemics tend to occur in the colder months in the northern hemisphere and in the warmer months in the southern hemisphere [3]. There is little information on the seasonality of NoV outbreaks in the clinical setting in SA. In a 1-year study on sporadic paediatric gastroenteritis in the hospital setting in the northern region of Gauteng, NoV infections were observed throughout the year with an increase in the summer months [10]. Factors that may have masked a seasonal pattern include the relatively low efficiency of the recovery procedure, the presence of inhibitors in water samples and bias introduced by interruptions in the sampling schedule. The difference in sampling schedules between the three rivers is a shortcoming of this study, as more targeted and consistent sampling might have revealed seasonal fluctuations. The efficiency of recovery (<1%) is low compared to previously reported values of between 7% (well water) and 60% (tap water) [37]. The high levels of turbidity in the water sources in this investigation could have resulted in a lower efficiency of recovery. Taking into account possible RT–PCR inhibition and the low recovery efficiency, it is likely that the overall NoV positive rate of 62·9% is an underestimation.

The majority of NoV infections in the clinical setting are caused by NoV GII, particularly GII.4 strains [38]. In this and other studies [13, 23], GI strains were detected at similar frequencies to GII strains in the environment. This indicates that GI viruses might be as prevalent in the population as GII strains but may cause less severe disease or asymptomatic infection. In agreement with a study in Singapore [39], the GII.4 strain (11/66) was the dominant GII strain found in this investigation. The GII.4 2010 variant (2010-03-29-R) identified in this study was 99–100% identical, over 282 nucleotides, to NoV strains identified in Australia (GQ845367), France (GQ246793), Japan (AB541322), Korea (GU390902) and Singapore (HM209210) during 2008, as well as strains detected in Hong Kong (HM191767), India (AB539162) and the USA (GU445325) during 2009 and Japan (AB555587) and Russia (HQ003284) in 2010. This shows the global distribution of this GII.4 variant. In addition, strain 2010-03-29-R was 99% identical, in 282 bp of the 5′-end of the capsid gene, to GII.4 strains (e.g. HQ08055) detected

in clinical specimens from paediatric patients in the Gauteng province [10]. The other ten GII.4 strains identified in the water samples could not be assigned to a variant group with the nucleotide sequence data available. Sequence analysis of the full-length capsid gene would increase the phylogenetic signal and would allow more accurate variant assignment [30]; however, it is often problematical to amplify the entire capsid gene from environmental NoVs.

The unassigned variants form a separate cluster, with significant bootstrap support, with the 2006a, 2008 and 2010 variants. BLAST analysis of strain 08-05-19-K2 indicated 98 % sequence identity with NoVs detected in 2003 in The Netherlands (AB303930), in 2005 in Russia (FJ383862) and in 2006 in Belgium (EU794859) and Sweden (EU007808), suggesting that this is a ≤2006 GII.4 variant. Another GII.4 strain (08-09-01-K) displayed 99 % nucleotide sequence identity with a virus identified in Australia (GQ845367.1) during 2008. Therefore, this is likely to be a more recent GII.4 variant.

In 2009, GI.5 was most often identified followed by GI.1 and GI.4. The majority (9/10) of GI.5 strains showed between 94 % and 99 % sequence identity with strains isolated from Russia and Mexico and one GI.5 strain showed 96 % nucleotide identity with a virus detected in Singapore. These results emphasize the wide geographical distribution of similar NoV genotypes. Interestingly, the GII.12 strain (2007-07-27-K) detected in a single Klip river sample in 2009 was 98 % identical over 273 residues to the GII.12 strain that caused 16 % of all reported NoV gastroenteritis outbreaks in the USA in the 2009–2010 season [40]. In 2010, GI.4 was the predominant genotype identified followed by GI.3. In a study in Spain from 2007 to 2009, GI.4 was one of the most abundant genotypes identified in sewage and river water [41]. Furthermore, GI.4 has been associated with foodborne NoV outbreaks [38]. These findings might reflect better survival of GI.4 in the environment compared to other NoV strains [38]. Genotypes I.2 and I.8 were detected at low frequency in environmental samples in this study; however, these strains exhibited 98 % and 99 % sequence identity, respectively, to clinical NoV isolates from Gauteng [10].

Even though NoVs could only be genotyped in 41 % of positive samples, the observed genetic diversity (16 genotypes) is comparable to that found in Japan (16 genotypes) [12], Korea (15 genotypes) [13] and Singapore (16 genotypes) [39]. Low viral titres in the water samples as well as sequence variation between the regions used for detection and genotyping might account for the inability to amplify the 5′-end of the capsid gene of some viruses. Up to five different genotypes were found to occur within the time-frame of 1 month (March 2010), underlining the diversity of strains co-circulating in the community. Overall, the results suggest that the rivers are contaminated with clinically relevant strains (GI.2, GI.8, GII.1, GII.4, GII.6, GII.12, GII.13) [10, 38].

Although the real-time RT–PCR method used to detect NoVs in this study does not give any indication of virus infectivity, it is likely that contact with, or ingestion of, contaminated river water could pose a potential health risk as other cytopathogenic and, therefore, potentially infectious viruses were detected in 38 % of NoV-positive samples. Given the high prevalence of HIV/AIDS in SA and the consequent vulnerability of a large proportion of the population, domestic or recreational exposure to contaminated river water could pose a serious public health threat [42]. The results from this study provide the first NoV sequence data from environmental samples in Sub-Saharan Africa and reveal that at least 16 NoV genotypes are actively circulating in SA. The majority of these NoV strains are also circulating on other continents in the same time-frame. A variety of recent GII.4 variants are present in the environment and the large contribution of GII.4 viruses to NoV disease is reflected in their predominance in the environment. Further investigations in the clinical setting and the environment are warranted to fully understand the epidemiology and impact of NoV in SA.

ACKNOWLEDGEMENTS

DECLARATION OF INTEREST

None.

REFERENCES

1. **Green K.** Caliciviridae: The Noroviruses. In: Knipe DM, et al., eds. *Fields Virology*, 5th edn. Philadelphia: Lippincott Williams & Wilkins, 2007, pp. 949–979.
2. **Zheng DP, et al.** Norovirus classification and proposed strain nomenclature. *Virology* 2006; **346**: 312–323.
3. **Marshall JA, Bruggink LD.** The dynamics of norovirus outbreak epidemics: recent insights. *International Journal of Environmental Research and Public Health* 2011; **8**: 1141–1149.
4. **Patel MM, et al.** Systematic literature review of role of noroviruses in sporadic gastroenteritis. *Emerging Infectious Diseases* 2008; **14**: 1224–1231.
5. **Siebenga JJ, et al.** Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001–2007. *Journal of Infectious Diseases* 2009; **200**: 802–812.
6. **Teunis PF, et al.** Norwalk virus: how infectious is it? *Journal of Medical Virology* 2008; **80**: 1468–1476.
7. **Newell DG, et al.** Food-borne diseases – the challenges of 20 years ago still persist while new ones continue to emerge. *International Journal of Food Microbiology* 2010; **139** (Suppl. 1): S3–15.
8. **Taylor MB, et al.** Two successive outbreaks of SRSV-associated gastroenteritis in South Africa. *Journal of Medical Virology* 1993; **41**: 18–23.
9. **Taylor MB, et al.** An epidemiological investigation of Norwalk virus infection in South Africa. *Epidemiology and Infection* 1996; **116**: 203–206.
10. **Mans J, et al.** Emerging norovirus GII.4 2008 variant detected in hospitalised paediatric patients in South Africa. *Journal of Clinical Virology* 2010; **49**: 258–264.
11. **Kiulia NM, et al.** The detection of enteric viruses in selected urban and rural river water and sewage in Kenya, with special reference to rotaviruses. *Journal of Applied Microbiology* 2010; **109**: 818–828.
12. **Kitajima M, et al.** Seasonal distribution and genetic diversity of genogroups I, II, and IV noroviruses in the Tamagawa River, Japan. *Environmental Science and Technology* 2010; **44**: 7116–7122.
13. **Lee C, Kim SJ.** The genetic diversity of human noroviruses detected in river water in Korea. *Water Research* 2008; **42**: 4477–4484.
14. **Lodder WJ, de Roda Husman AM.** Presence of noroviruses and other enteric viruses in sewage and surface waters in The Netherlands. *Applied and Environmental Microbiology* 2005; **71**: 1453–1461.
15. **Rodriguez-Diaz J, et al.** Detection and characterization of waterborne gastroenteritis viruses in urban sewage and sewage-polluted river waters in Caracas, Venezuela. *Applied and Environmental Microbiology* 2009; **75**: 387–394.
16. **Victoria M, et al.** Assessment of norovirus contamination in environmental samples from Florianopolis City, Southern Brazil. *Journal of Applied Microbiology* 2010; **109**: 231–238.
17. **Vilaginès P, et al.** Glass wool for virus concentration at ambient water pH level. *Water Science and Technology* 1993; **27**: 229–306.
18. **Minor P.** Growth, assay and purification of picornaviruses. In: Mahy BWJ, ed. *Virology: A Practical Approach*. Washington: IRL Press, 1985, pp. 25–42.
19. **APHA-AWWA-WPCF.** *Standard Methods for the Examination of Water and Waste-water*, 21st edn. Washington, DC: APHA, AWWA, and WPCF, 2005.
20. **Vivier JC, Ehlers MM, Grabow WO.** Detection of enteroviruses in treated drinking water. *Water Research* 2004; **38**: 2699–2705.
21. **van Heerden J, et al.** Prevalence, quantification and typing of adenoviruses detected in river and treated drinking water in South Africa. *Journal of Applied Microbiology* 2005; **99**: 234–242.
22. **Malherbe H, Strickland-Cholmley M.** *Viral Cytopathology*. Boca Raton: CRC Press, Inc., 1980.
23. **da Silva AK, et al.** Evaluation of removal of noroviruses during wastewater treatment, using real-time reverse transcription-PCR: different behaviors of genogroups I and II. *Applied and Environmental Microbiology* 2007; **73**: 7891–7897.
24. **Svraka S, et al.** Etiological role of viruses in outbreaks of acute gastroenteritis in The Netherlands from 1994 through 2005. *Journal of Clinical Microbiology* 2007; **45**: 1389–1394.
25. **Kageyama T, et al.** Broadly reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. *Journal of Clinical Microbiology* 2003; **41**: 1548–1557.
26. **Loisy F, et al.** Real-time RT-PCR for norovirus screening in shellfish. *Journal of Virological Methods* 2005; **123**: 1–7.
27. **Kojima S, et al.** Genogroup-specific PCR primers for detection of Norwalk-like viruses. *Journal of Virological Methods* 2002; **100**: 107–114.
28. **Tamura K, et al.** MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 2007; **24**: 1596–1599.
29. **Tamura K, Nei M, Kumar S.** Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences USA* 2004; **101**: 11030–11035.
30. **Kroneman A, et al.** An automated genotyping tool for enteroviruses and noroviruses. *Journal of Clinical Virology* 2011; **51**: 121–125.
31. **Simple Interactive Statistical Analysis (SISA)** (http://www.quantitativeskills.com/sisa/). Accessed 12 October 2011.
32. **Department of Water Affairs and Forestry.** *South African Water Quality Guidelines*, 2nd edn, vol. 2: *Recreational Use*. Pretoria: Department of Water Affairs and Forestry, 1996.
33. **South African National Standard.** SANS 241: Drinking Water, Edition 6.1. Pretoria: Standards South Africa, 2006.

34. **Department of Water Affairs and Forestry.** *South African Water Quality Guidelines*, 2nd edn, vol. 4: *Agricultural Use: Irrigation*. Pretoria: Department of Water Affairs and Forestry, 1996.

35. **Department of Water Affairs and Forestry.** Report on the radioactivity monitoring programme in the Klip River catchment. Report No. N000000REQ0403. Resource Quality Services. Pretoria: Department of Water Affairs and Forestry, 2003.

36. **Department of Environmental Affairs and Tourism.** Ekurhuleni metropolitan municipality first year state of the environment report. Pretoria: Department of Environmental Affairs and Tourism, 2003.

37. **Lambertini E, et al.** Concentration of enteroviruses, adenoviruses, and noroviruses from drinking water by use of glass wool filters. *Applied and Environmental Microbiology* 2008; **74**: 2990–2996.

38. **Verhoef L, et al.** Use of norovirus genotype profiles to differentiate origins of foodborne outbreaks. *Emerging Infectious Diseases* 2010; **16**: 617–624.

39. **Aw TG, et al.** Prevalence and genotypes of human noroviruses in tropical urban surface waters and clinical samples in Singapore. *Applied and Environmental Microbiology* 2009; **75**: 4984–4992.

40. **Vega E, Vinje J.** Novel GII.12 norovirus strain, United States, 2009–2010. *Emerging Infectious Diseases* 2011; **17**: 1516–1518.

41. **Pérez-Sautu U, et al.** Human norovirus occurrence and diversity in the Llobregat river catchment, Spain. *Environmental Microbiology* 2012; **14**: 494–502.

42. **Wingfield T, et al.** Chronic norovirus infection in an HIV-positive patient with persistent diarrhoea: a novel cause. *Journal of Clinical Virology* 2010; **49**: 219–222.