Norovirus GII.17 predominates in selected surface water sources in Kenya

Kiulia NM[1, 2], Mans J[2], Mwenda JM [1], Taylor MB[2]

1. Enteric Viruses Research Group, Institute of Primate Research, P.O Box 24481, 00502 Karen, Nairobi, Kenya
2. Department of Medical Virology, School of Medicine, Faculty of Health Sciences, University of Pretoria, Private Bag X323 Arcadia, 0007, Pretoria, South Africa

Corresponding author:

Nicholas M. Kiulia

Enteric Viruses Research Group

Institute of Primate Research

P. O. Box 24481

00502, Karen, Nairobi,

Kenya

Tel: +254-020-3882571-4

Fax: +254-020-3882546

E-mail: nick.kiulia@gmail.com

Shortened title: Norovirus GII.17 in water sources in Kenya

Abstract

In this study the prevalence and genotypes of NoVs in selected water sources from rural, urban and refugee settings in Kenya was investigated. Ten liters each of river, household and borehole water were collected in rural (Mboone river), urban (Nairobi and Mutoine river), and refugee (Dadaab refugee camp) settings. Noroviruses were recovered from the water samples by a glass wool adsorption-elution technique and/or PEG/NaCl precipitation. Nucleic acid was extracted using the automated MagNA Pure platform. Noroviruses were detected with singleplex real-time reverse transcription-polymerase chain reaction assays and characterised by nucleotide sequence analysis. Noroviruses were detected in 63% (25/40) of the selected water samples comprising GII (42.5%), GI (2.5%) and mixed GI/GII (17.5%) positive samples. The prevalence of NoVs in the Mutoine river (urban area) was higher than in the Mboone river (rural area) (P=0.0013). Norovirus GI.1, GI.3, GI.9, GII.4, GII.6, GII.12, GII.16 and GII.17 were identified, with GII.17 accounting for 76% (16/21) of the typed strains. The NoV GII.17 predominance differs to other studies in Africa and further surveillance of NoVs in clinical and environmental settings is required to clarify/elucidate this observation. As information regarding NoVs in Kenyan water sources is limited this report provides valuable new data on NoV genotypes circulating in environmental water sources and the surrounding communities in Kenya.

Key words: norovirus, genotypes, water samples, real-time RT-PCR, Kenya

Introduction

Noroviruses (NoVs) are the leading cause of acute human viral gastroenteritis worldwide, affecting people of all ages (Koo et al. 2010). After rotavirus, NoVs are the second most important cause of severe paediatric gastroenteritis (Glass et al. 2009) and in the elderly NoV infection can be particularly severe, even resulting in death (Harris et al. 2008; Tian et al. 2014). Noroviruses are transmitted primarily via the faecal-oral route and contaminated food and water have been implicated in NoV-associated outbreaks of gastroenteritis (Mathijs et al. 2011; Matthews et al. 2012).

Noroviruses are classified in the genus *Norovirus* of the family *Caliciviridae* (Clarke et al. 2012). They have a non-enveloped capsid with a single-stranded RNA genome 7.5-7.7 kb in length encoding three open reading frames (ORF) (Green 2013). Based on the nucleotide sequence diversity of the ORF2, human NoVs are currently classified into five genogroups (GI-GV) (Clarke et al. 2012; Green 2013). Human infection has been associated with genogroups I, II and IV while genogroups III and V infect cows and mice, respectively (Zheng et al. 2006; Scipioni et al. 2008; Clarke et al. 2012). Porcine strains and a strain from lions are grouped in GII (Martella et al. 2007) while a strain from a dog was identified in GIV (Martella et al. 2008). Norovirus GII is detected more frequently than GI (Patel et al. 2009; Vega et al. 2013), though GI has been associated with outbreaks of waterborne gastroenteritis in Sweden (Nenonen et al. 2012). On the basis of >85% sequence similarity in the complete VP1 gene, NoVs can further be classified into genotypes, with at least 9 genotypes within GI and 22 genotypes in GII (Green 2013; Kroneman et al. 2013). Recently a new nomenclature system for the classification of NoV has been proposed by using both polymerase (ORF1) and capsid (VP1) sequences, as recombination is common (Kroneman et al. 2013). Within the genotypes, GII.4 is the most prevalent worldwide (Zakikhany et al. 2012; Kim et al. 2013;

Maritschnik et al. 2013). The GII.4 strains evolve through serial changes in the VP1 sequence, which allows evasion of immunity in the human population (Siebenga et al. 2009) and the observed increase in the number of outbreaks has coincided with the emergence of novel variants (Desai et al. 2012; Centers for Disease Control and Prevention (CDC) 2013).

Over the last ten years many surveillance studies have shown that NoVs are a global problem (Hoa Tran et al. 2013). In Africa, the prevalence of NoVs has been investigated in both clinical (Mans et al. 2010, 2014; Hassine-Zaafrane et al. 2013; Huynen et al. 2013; Trainor et al. 2013) and environmental settings (Kamel et al. 2010; Kiulia et al. 2010; Sdiri-Loulizi et al. 2010; Gibson 2011; Mans et al. 2013; Murray et al. 2013). Noroviruses have also been detected in shellfish collected along the Mediterranean Sea and Atlantic Coast of Morocco (Benabbes et al. 2013). In a pilot study in Kenya (2007-2008) NoVs were detected in >90% of selected urban water sources and in 8-25% of rural water sources (Kiulia et al. 2010), but the circulating NoVs genotypes were not identified. Therefore, in this study the molecular epidemiology of NoVs detected in selected Kenyan water sources was investigated.

Materials and methods

Ethical approval

This study was approved by the Institutional Review Committee (IRC) of the Institute of Primate Research (IPR) (Protocol number: IRC/24/11) and the Student Ethics Committee of the Faculty of Health Sciences, University of Pretoria (Protocol number S172/2011).

Water samples

From February 2012 to January 2013, 40 water samples (10 ℓ) were collected from rural (Maua), urban (Nairobi) and refugee settings in Kenya. Samples included surface water from

the Mboone river (12) (Maua-rural setting), the Mutoine river (12) (Kibera, informal settlement, Nairobi-urban setting) and the Nairobi river (12) (Nairobi-urban setting). Sampling was done at set intervals on a monthly basis for a full calendar year to cover both the wet and dry seasons. Borehole water (2) and household container water (2) samples were collected on two occasions (May and August 2012) from the Dadaab refugee camp. All water samples were transported in cooler bags from the sites to the Enteric Viruses Laboratory, IPR, Nairobi and stored at 4°C until processing.

Sample processing and virus recovery

Viruses were recovered from the river water samples using a glass wool adsorption-elution technique based on the method Vilaginès et al. (1993) and as described by Mans et al. (2013). Briefly the glass wool columns, 20 centimetre (cm) in length and with an internal diameter of about 30 millimetre (mm) were packed with 15 grams (g) oiled sodocalcic glass wool (Glass wool Bourre 725 QN, Ouest Isol, Alizay, France). During packing three 5 g portions of glass wool were compressed into the columns at different angles to each other with a steel sieve grid with pore sizes of about 1 mm^2 inserted between each 5 g portion of compressed glass wool. The packed columns were then treated to positively charge the glass wool by soaking in sterile distilled water and then pre-treating with 40 millilitre (mℓ) of 1 M hydrochloric acid (HCl) (Merck, Darmstadt, Germany), 100 mℓ sterile distilled water and then 40 mℓ of 1 M NaOH (Merck), with a final rinse using 100 mℓ sterile distilled water to adjust the pH to 7.0. To recover the viruses, 10 ℓ of water were filtered through the positively charged glass wool columns using negative pressure at a flow rate of 10 ℓ/h. The negatively charged viruses, which adsorbed to the glass wool, were eluted twice with 50 mℓ glycine-beef extract buffer pH 9.5 (GBEB; 0.05 M glycine (Merck); 0.5% beef extract [BBL™ Becton Dickinson and Co., Sparks MD]). Immediately after elution the pH of the eluate was neutralised to pH 7

using 1 M HCl. The viruses in 100 mℓ eluate were further concentrated by polyethylene glycol/sodium chloride precipitation to a final volume of 10 mℓ in phosphate-buffered saline pH 7.2 (Sigma-Aldrich Co., St Louis, MO). The recovered virus suspensions were stored at -20°C until further analysis.

Nucleic acid extraction

The recovered virus suspensions were clarified by the addition of 200 μℓ of chloroform (Merck) to 1.5 mℓ of the recovered viral suspension followed by rigorous vortexing for 30 s. The mixture was then centrifuged at 3000 x g (Eppendorf 5402D Microcentrifuge, Hamburg, Germany) for 30 s. Prior to nucleic acid extraction, 10 μℓ of mengovirus (5 x 10^4 copies/10 μℓ) was added to 1 mℓ of each clarified sample as a process control to monitor the efficiency of nucleic acid extraction (Bosch et al. 2011). Genomic viral nucleic acid was extracted directly from the seeded recovered virus concentrate (1 mℓ) using the MagNA Pure LC Total Nucleic Acid Isolation Kit (large volume) (Roche Diagnostics GmbH, Mannheim, Germany) in a MagNA Pure LC Robotic instrument (Roche Diagnostics), following the manufacturer's instructions. The extracted nucleic acid was eluted in 100 μℓ and stored in 10 μℓ aliquots at -70°C until use.

Virus detection

Mengovirus, NoV GI and NoV GII were detected by real-time reverse transcriptase-polymerase chain reactions (rt RT-PCR) based on TaqMan technology using virus-specific environmental CeeramTools™ kits (Ceeram S.A.S, La Chappelle-Sur-Erdre, France). These NoV GI and NoV GII assays have an internal control (IC) to monitor for amplification inhibition. The reactions were carried out in a total volume of 20 μℓ in sealed glass capillaries in a LightCycler® v2.0 Instrument (Roche Diagnostics). The amplification was

done using 5 µℓ RNA and the cycling conditions were as follows: 45°C for 10 min, 95°C for 10 min and 45 cycles of 95°C for 15 s, 60°C for 45 s.

Viral characterisation

The NoV strains with a cycle threshold (Ct) value <38 were genotyped by nucleotide sequence and phylogenetic analysis of the 5' end of the NoV capsid gene (Region C; 320 bp). Briefly, the capsid gene of NoV was amplified using a semi-nested RT-PCR as described by Mans et al. (2013). Primer details are shown in Table 1. The cDNA was synthesised at 42°C for 60 min in a 20 µℓ RT reaction as follows; 10 µℓ of template RNA, 1.5 mM $MgCl_2$, 10 mM concentrations of each dNTPs, 200 U of RevertAid^{TM} Premium Reverse Transcriptase (Thermo Scientific, Waltham, MA), 20 U of RNase inhibitor (Promega Corp., Madison, WI), and 100 pmol/ µℓ of random hexamer primers (Roche Diagnostics). Ten microliters cDNA was mixed with 5 µℓ PCR buffer with 1.5 mM $MgCl_2$, 200 µM of each dNTP, 0.2 µM of primers QNIF4 and G1SKR (GI) or primers QNIF2 and G2SKR(GII), and 0.25 U/ µℓ of AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA) or KAPATaq™ HotStartDNA polymerase (Kapa Biosystems, Cape Town, SA). Initial denaturation at 95°C for 10 min was followed by 45 cycles of denaturation at 94°C for 30 s; annealing at 50°C for 30 s; and extension at 72°C for 40 s, with a final extension step at 72°C for 10 min. At the end of the amplification, 1 µℓ of the first-round PCR product was added to a semi-nested PCR mix containing the same reagents as the first round of PCR but with the primers G1SKF and G1SKR (GI) or primers G2SKF and G2SKR (GII) (Table 1). The amplification products were analysed on 1.5% agarose gels containing 0.5 ng/ mℓ of ethidium bromide and visualised under UV illumination.

Table 1. Primers used for genotyping of noroviruses

Virus	Primer	Sequence (5'-3')[c]	Nucleotide position	Reference
Norovirus GI				
1st round	QNIF4	CGCTGGATGCGNTTCCAT	5291-5308[a]	Da Silva et al. 2007
	G1SKR	CCAACCCARCCATTRTACA	5653-5671[a]	Kojima et al. 2002
2nd round	G1SKF	CTGCCCGAATTYGTAAATGA	5342-5361[a]	Kojima et al. 2002
	G1SKR	CCAACCCARCCATTRTACA	5653-5671[a]	Kojima et al. 2002
Norovirus GII				
1st round	QNIF2	ATGTTCAGRTGGATGAGRTTCTCWGA	5012-5037[b]	Loisy et al. 2005
	G2SKR	CCRCCNGCATRHCCRTTRTACAT	5367-5389[b]	Kojima et al. 2002
2nd round	G2SKF	CNTGGGAGGGCGATCGCAA	5046-5064[b]	Kojima et al, 2002
	G2SKR	CCRCCNGCATRHCCRTTRTACAT	5367-5389[b]	Kojima et al. 2002

[a]GenBank accession number M87661,

[b]GenBank accession number X86557,

[c]IUPAC codes used to indicate degenerate positions. R = A/G, S = G/C, Y = C/T

The PCR products were cloned using the ClonJET™ PCR cloning kit (Thermo Scientific). Briefly, 5 µℓ of the purified PCR product was used for the ligation reaction according to the manufacturer's instructions. The ligation mixture (4 µℓ) was then transformed in competent *E. cloni*® cells (Lucigen® Corp., Middleleton, WI). Fifteen discrete colonies were randomly selected and colony PCR was performed in a 20 µℓ reaction consisting of 4 µℓ of 1X PCR buffer, 1.2 µℓ 1.5 mM $MgCl_2$, 0.4 µℓ 200 µM dNTPs, 0.4 µℓ of 0.2 µM pJET1.2 forward and reverse primers and 0.1 GoTaq® Flexi DNA polymerase (Promega Corp.) and nuclease free water (Promega Corp.). The amplification conditions were as follows; initial denaturation at 95°C for 3 min followed by 30 cycles of denaturation at 94°C for 30 s, annealing at 60°C for 30 s, extension at 72°C for 40 s and a final extension at 72°C for 10 min. The amplified product (448 bp) of 10 clones was selected for sequencing using pJet1.2/blunt specific primers (Thermo Scientific). Briefly the PCR products were purified with the DNA Clean and Concentrator kit (Zymo Research, Irvine, CA) and directly sequenced with the ABI PRISM BigDye® Terminator v. 3.1 Cycle Sequencing kit (Applied Biosystems) on an ABI

3130 automated analyser (Applied Biosystems). The sequencing reaction (20 μℓ) containing 3 μℓ 5 X sequencing buffer, 1 μℓ reaction terminator mix, 1 μℓ of 3.2 pmol of the forward or the reverse primers and 13 μℓ nuclease-free water (Promega Corp.) was used. The amplification conditions for the sequencing reaction was an initial denaturation step at 95°C for 3 min followed by 25 cycles of denaturation at 94°C for 30 s, annealing at 50°C for 10 s and extension at 60°C for 4 min.

Phylogenetic analysis

Nucleotide sequences for NoVs were edited and analysed using Sequencher™ 4.9 (Gene Codes Corporation, Ann Arbor, MI) and BioEdit Sequence Alignment Editor (V.7.0.9.0) [Hall, 1999]. A Basic Local Alignment Search Tool (BLAST) search (http://blast.ncbi.nlm.nih.gov) was performed to compare the identity of the sequenced clones with the reference strains from GenBank by pairwise comparison. Where the nucleotide sequence of all 10 clones from a single sample were identical, 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 NoV GI and NoV GII reference sequences from GenBank using MAFFT version 7 (http://mafft.cbrc.jp/alignment/server/). After manual adjustment of the alignment, phylogenetic analysis was performed with MEGA5 using the neighbour-joining methods (Tamura et al. 2011). All the genotypes assigned were confirmed by analysis of the sequences with the NoV Genotyping Tool as described by Kroneman et al. (2011) (http://www.rivm.nl/mpf/norovirus/typingtool).

The nucleotide sequences determined in this investigation have been deposited in GenBank under the following accession numbers NoV GI (KF793788 to KF793798) and NoV GII (KF808211 to KF808254, KF916584 and KF916585).

Statistical analysis

Analysis of the viral prevalence rates in the river samples was done using Fisher's exact test using StatView software (version 5.0, SAS Institute Inc, Cary, NC). P values <0.05 were considered significant at 95% confidence interval (Cl).

Results

Viral prevalence in water samples

The Ct values for mengovirus detection were within three cycles (1 log) for all samples in each water source indicating that the processing and nucleic acid extraction procedures were adequate and valid. The Ct values of the ICs of the NoV GI and NoV GII RT-PCR assays were within the value range specified for the particular kit lot number used thereby indicating that there was no amplification inhibition. Noroviruses were detected in 62.5% (25/40) of the selected water samples comprising GII (42.5%), GI (2.5%) and mixed GI/GII (17.5%) positive samples (Table 2). Norovirus GI was detected in 8.3% (1/12) and NoV GII in 4/12 (33.3%) of the Mboone River samples (Maua-rural setting) (Table 2). In the Nairobi-urban setting, NoV GI and NoV GII were detected in 3/12 (25%) and 12/12 (100%) of the Mutoine river samples and in 4/12 (33.3%) and 7/12 (58%) of the Nairobi river samples, respectively (Table 2). Norovirus GII was detected in 25% (1/4) and no NoV GI was detected in the water samples from the Dadaab refugee camp. The overall prevalence of NoVs detected in the Mboone river (rural setting) was low (33.3%) compared to the 100% in the Mutoine river (urban setting) (P=0.0013) (Table 2). The difference in the NoV GI prevalence rate in the

rural (Mboone river) and urban (Mutoine river) settings, i.e. 8.3% vs 25% (Table 2), was not significant (*P*=0.0590). There was a significant difference in NoV GII detection rates between the rural (Mboone river; 33.3%) and urban (Mutoine river; 100%) settings (*P*=0.0013). When the detection of NoVs in wet and dry periods was compared, no differences were noted as Ct values were within the same range for both seasons. In the rural setting however, NoVs were only detected from April 2012 to July 2012 which included wet and dry periods while in the urban setting NoVs were detected throughout the year.

Table 2. Prevalence of process control and noroviruses in water sources from rural, urban and refugee settings from February 2012 to January 2013

Site / Virus type	Rural setting (Maua)	Urban setting (Nairobi)		Refugee setting	Total
	Mboone	Mutoine	Nairobi	Dadaab	All sites
	n=12 (%)	*n=12* (%)	*n=12* (%)	*n=4* (%)	*n=40* (%)
Mengovirus	12 (100%)	12 (100%)	12 (100%)	4 (100%)	40 (100%)
NoV GI	0	0	1 (8.3%)	0	1 (2.5%)
NoV GII	3 (25%)	9 (75%)	4 (33.3%)	1 (25%)	17 (42.5%)
NoV GI+GII	1 (8.3%)	3 (25%)	3 (25%)	0	7 (17.5%)
Total NoV	4 (33.3%)	12 (100%)	8 (67%)	1 (25%)	25 (62.5%)

Virus characterisation

Norovirus GII strains from 87.5% (21/24) of the samples with Ct values <38 and four NoV GI strains were successfully sequenced (Table 3). A neighbour-joining tree constructed from an alignment of a 285 bp region of the capsid gene for NoV GI and a 273 bp region for NoV GII and NoV reference strains revealed a diversity of genotypes in the three rivers during different sampling periods (Table 3). Three GI genotypes were identified, namely, GI.1, GI.3 and GI.9 (Figure 1). The predominant GI genotype, GI.3, was detected in the Nairobi river in June 2012 and August 2012 (Table 3). Norovirus GI.1 was detected in the Mutoine river in May 2012 and GI.9 was detected in the Nairobi river in June 2012 (Table 3). The NoV GII

strains clustered into five genotypes, i.e. GII.4, GII.6, GII.12, GII.16 and GII.17 (Figure 2). Norovirus GII.17 was the predominant genotype and accounted for 76% of the NoV GII strains identified. Norovirus GII.17 was detected in both rural and urban settings being identified in the Mboone river in May 2012 to July 2012. Genotype GII.17 was detected throughout the year in both urban rivers, namely the Mutoine (February 2012, March 2012, August 2012 and October 2012 to January 2013) and Nairobi (February 2012 to April 2012, June 2012, July 2012 and October 2012) rivers (Table 3). Strains of NoV GII.4 were only detected in the urban setting (Table 3), i.e. in the Mutoine river in March 2012 and the Nairobi river in August 2012 (Table 3). Norovirus GII.6 was detected in both urban and refugee settings where it was detected in Mutoine river in 6/12 months (March 2012, April 2012, June 2012, July 2012, November 2012, December 2012) and in the Dabaab household water in May 2012 (Table 3). Norovirus GII.16 was detected in both the rural and urban setting, namely in the Mboone river in May 2012 and the Mutoine river in March 2012 while a GII.12 strain was only identified in the Mutoine river in August 2012 (Table 3).

Table 3. Norovirus genotypes identified in water samples from February 2012 to January 2013

Site / Month	Rural setting	Urban setting		Refugee setting
	Mboone river	Mutoine river	Nairobi river	Dadaab
Feb 2012	-	GII.17	GII.17	
Mar 2012	-	GII.4;GII.6;GII.16;GII.17	GII.17	
Apr 2012	-	GII.6	GII.17	
May 2012	GII.16;GII.17	GI.1	-	GII.6
Jun 2012	GII.17	GII.6	GI.3;GI.9;GI1.17	
Jul 2012	GII.17	GII.6	GII.17	
Aug 2012	-	GII.12;GII.17	GI.3;GII.4	-
Sept2012	-	-	-	
Oct 2012	-	GII.17	GII.17	
Nov 2012	-	GII.6;GII.17	-	
Dec 2012	-	GII.6;GII.17	-	
Jan 2013	-	GII.17	-	

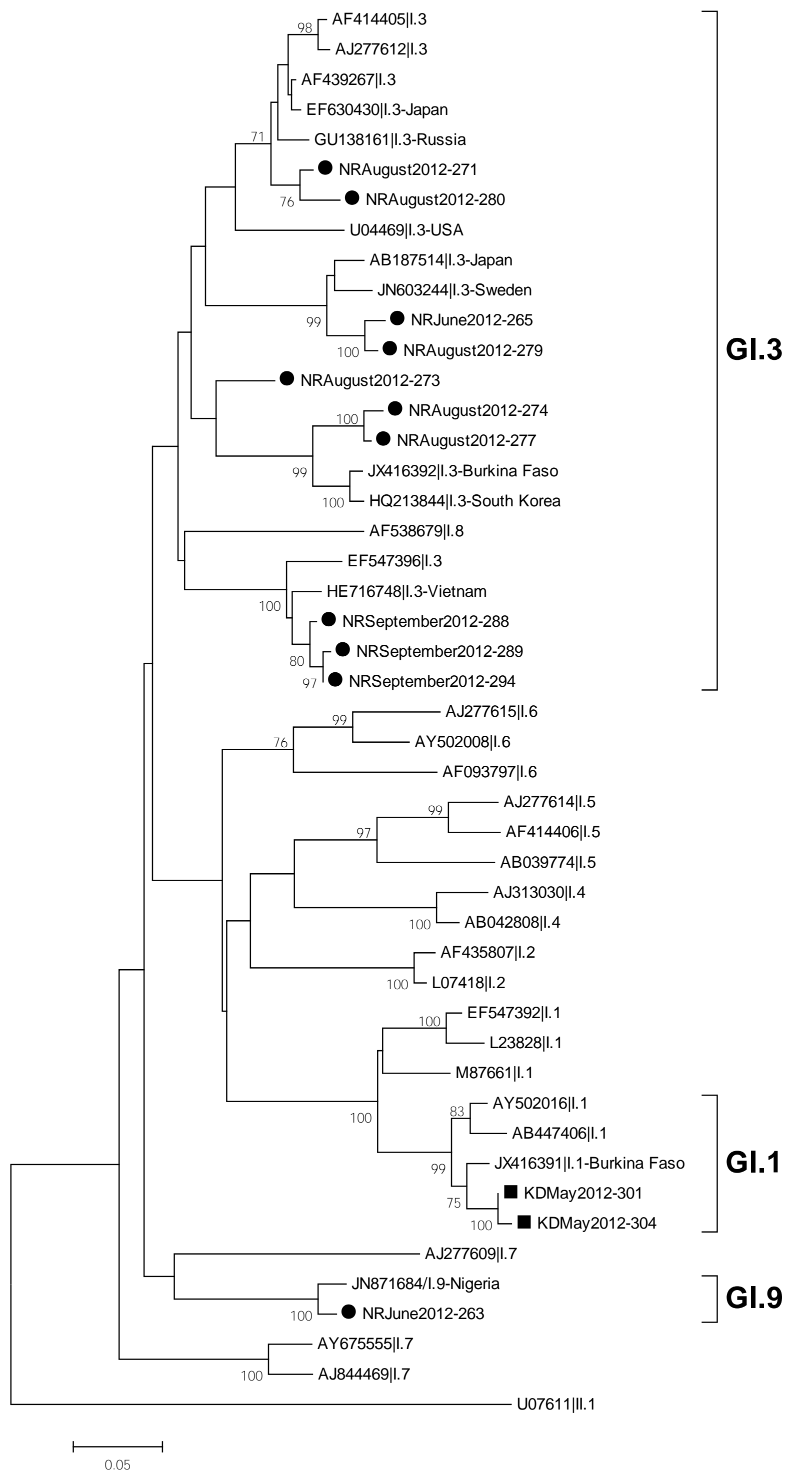

Figure 1: Phylogenetic tree based on partial nucleotide sequences (285 bp) of the capsid gene of NoV GI strains detected in Kenyan river water samples [● Nairobi river (NR) ■ Mutoine river (KD)]. Collection month/year and clone number for different genotypes are indicated. Numbers at each branch indicate bootstrap values for the clusters supported by that branch. Bootstrap values >70 are shown.

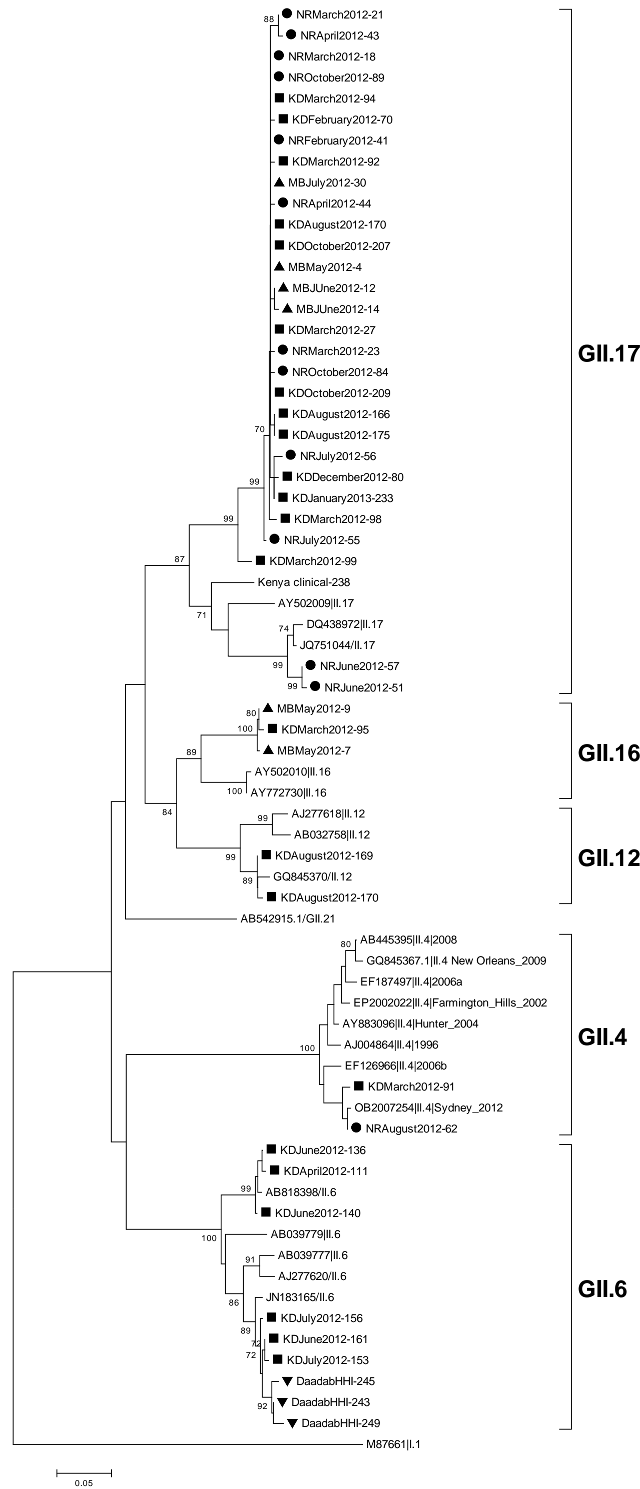

Figure 2: Phylogenetic tree based on partial nucleotide sequences (273 bp) of the capsid gene of NoV GII strains detected in Kenyan river water samples [●Nairobi river (NR) ■ Mutoine river (KD) ▲Mboone river (MB) ▼Dadaab Household water (HHI)]. Collection month/year and clone number for different genotypes are indicated. Numbers at each branch indicate bootstrap values for the clusters supported by that branch. Bootstrap values >70 are shown.

Discussion

Information with regard to the prevalence and NoVs genotypes in Kenya is limited. The analysis of water samples not only provides information as to contamination levels and potential health risks in the water sources but it serves as a passive surveillance system to identify viruses circulating in the surrounding communities. In this study NoVs were detected in 63% (25/40) of the selected water samples (Table 2). This prevalence is similar to the 66% NoV detection rate in surface water samples from South Africa (SA) (Mans et al. 2013). Comparing the results of this investigation to those of an earlier (2007-2008) Kenyan study (Kiulia et al. 2010), the prevalence of both NoV GI and NoV GII was similar for the Mboone river (NoV GI 8.3% vs 8.3%; NoV GII 33% vs 25%). However, in the Mutoine river the prevalence rates differed, with NoV GI being detected in 25% of the samples in this study while in the earlier study it was present in 90% of the samples (Kiulia et al. 2010), suggesting that NoV GII has displaced NoV GI in the surrounding community.

A number of NoV GI (GI.1, GI.3, GI.9) and NoV GII (GII.4, GII.6, GII.7, GII.12, GII.16, GII.17) strains were identified in the Kenyan surface water samples (Table 3; Figures 1 & 2). The genotype diversity is similar to that reported in SA where NoV GI.1, GI.3 and NoV GII.4, GII.12, GII.17 were identified in surface water sources (Mans et al. 2013). The characterised NoV GI strains were detected in the rivers from the urban setting. The two NoV GI.1 strains (KDMarch2012-301 and KDMarch2012-304), detected in the urban Mutoine river, showed a high nucleotide sequence identity (97%) to a clinical NoV GI.1 strain (JX416391) from Burkina Faso. Except for studies in South Africa (Mans et al. 2013; Murray 2013), Botswana (Mattison et al. 2010) and Burkino Faso (Nordgren et al. 2013), the occurrence of NoV GI.1 does not appear to be widely reported in Africa as no GI.1 strains were detected in diarrhoeal stool samples from children in Nairobi, Kenya (Mans et al. 2014)

nor in the clinical specimens (Silva et al. 2008; Abugalia et al. 2011; Ayukekbong et al. 2011; Trainor et al. 2013) or sewage samples (Sdiri-Loulizi et al. 2010) from other African countries**.** Norovirus GI.3 however, appears to more prevalent in African regions (Ayukekbong et al. 2011; Yassin et al. 2012; Trainor et al. 2013), including Kenya (Mans et al., 2014). The GI.3 strains identified in this study grouped into four distinct clusters (Figure 1) and showed 94% - 97% nucleotide sequence identity to strains from Russia (GU138161), Japan (AB187514), Sweden (JN603244), Burkina Faso (JX416392), South Korea (HQ213844) and Vietnam (HE716748). The single GI.9 strain showed a high (98%) nucleotide sequence identity to Nigerian (JN871684) strain.

Nucleotide sequence and phylogenetic analysis of the NoV GII strains revealed that GII.17 was the predominant (76%) genotype identified. The GII.17 strains grouped into two clusters (Figure 2) with the majority of the strains, which shared 96% - 100% nucleotide sequence identity, from urban and rural sampling sites resorting in one cluster. In the second cluster strains detected in the Nairobi River in June 2012 grouped together with strains from the United State of America (USA) (DQ438972) and China (JQ751044). Strains of NoV GII.17 were identified in clinical specimens collected during 1999-2000 from human immunodeficiency virus (HIV)-infected children at a care center in Kenya (Mans et al. 2014) suggesting that GII.17 has been in circulating in urban and rural Kenyan communities for a number of years. Nucleotide sequence information on NoV GII.17 strains from Burkina Faso (Nordgren et al. 2013), Morocco (KJ162374), Cameroon (Ayukekbong et al. 2011) and South Africa (Murray et al. 2013) are available in GenBank. Although NoV GII.17 seems to be prevalent in some parts of Africa this genotype was not identified in clinical specimens (Silva et al. 2008; Abugalia et al. 2011; Hassine-Zaafrane et al. 2013; Huynen et al. 2013; Trainor et al. 2013) and environmental samples (Sdiri-Loulizi et al. 2010) from many other African

regions and the GII.17 predominance appears to be peculiar to Kenya. Norovirus GII.17 is of clinical relevance as it has been implicated in nosocomial NoV infection (Sukhrie et al. 2011) and chronic NoV infection in a kidney transplant patient (Schorn et al. 2010). The genotype has also been widely reported in children with gastroenteritis in Central and South America (Bucardo et al. 2009; Ferreira et al. 2012; Gomes et al. 2008), Korea (Park et al. 2011) and Thailand (Kittigul et al. 2010).

The GII.4 strains, NRAugust2012-62 and KDMarch2012-91, detected in the urban setting, were closely related to the Sydney variant (OB2007254_Sydney_2012) (Figure 2), which is currently the most predominant GII.4 variant identified in NoV-associated gastroenteritis outbreaks worldwide (Kim et al. 2013; Maritschnik et al. 2013; CDC 2013). In contrast the GII.4 strain identified in the HIV-infected children in 2000-2001 was closely related to the non-epidemic GII.4 Kaiso 2003 variant (Mans et al. 2014), thus highlighting the periodic emergence of new GII.4 variants within a community. The GII.6 strains were the second most predominant strains identified and clustered together with strains from Sweden (JN183165) and Japan (AB818398). After NoV GII.4, NoV GII.6 has been shown to be the second most dominant strain in clinical studies in a number of countries namely, SA, Brazil, Japan, and Finland (Phan et al. 2005; Mans et al., 2010; Ferreira et al. 2010; Puustinen et al. 2011; Chan-it et al. 2012). Norovirus GII.6 strains were previously identified in clinical specimens from Kenya (Mans et al. 2014), indicating the presence of clinically relevant strains in the water sources. Of importance was the detection of NoV GII.6 in one of the household water samples from the Dadaab refugee camp (Table 3). The refugees collect their household water from boreholes and no NoVs were detected in the borehole water. The question therefore arises as to the source of contamination, i.e. whether the source water was contaminated or whether the household water was contaminated by an infected person or utensil. This would require further investigation as although NoVs have been detected in

groundwater in Korea (Jung et al. 2011) and Ghana (Gibson et al. 2011), water in household containers is often more contaminated than the source water, with contamination occurring post-collection (Harris et al. 2013).

The NoV GII.12 strain was detected in the Mutoine river in August 2012 (KDAugust2012-169/171). This environmental strain shares 99% nucleotide sequence identity with the GQ845370 strain, which was identified in Australia during 2008. Norovirus GII.12, described as an emerging genotype (Vega and Vinje, 2011), is currently circulating globally and the capsid gene is often identified in combination with the GII.Pg polymerase region (Giammanco et al. 2012). The GII.16 strains, identified in the Mboone (MBMay2012-7/9) and Mutoine rivers (KDMarch2012-95), were closely related to strains from Germany (AY772730) and the USA (AY502010).

From the data it is evident that selected rivers, which are often used for domestic purposes, are contaminated with clinically relevant NoV strains. In conclusion, this is the first comprehensive report into the molecular epidemiology of NoVs in Kenyan water sources thereby highlighting NoV genotypes circulating in the surrounding Kenyan communities.

Acknowledgements

This study is based on the research supported by National Research Foundation, South Africa and the National Council for Science and Technology, Kenya: South Africa/Kenya Research Programme.

Conflict of interest statement

There is no conflict of interest from all the authors.

Disclaimer: Any opinion, finding and conclusion or recommendation expressed in this article is that of the authors and the National Research Foundation/National Council for Science and Technology does not accept any liability in this regard.

References

Abugalia, M., Cuevas, L., Kirby, A., Dove, W., Nakagomi, O., Nakagomi, T., et al. (2011). Clinical features and molecular epidemiology of rotavirus and norovirus infections in Libyan children. *Journal of Medical Virology,* 83, 1849-1856.

Ayukekbong, J., Lindh, M., Nenonen, N., Tah, F., Nkuo-Akenji, T., & Bergström, T. (2011). Enteric viruses in healthy children in Cameroon: viral load and genotyping of norovirus strains. *Journal of Medical Virology,* 83, 2135-2142.

Benabbes, L., Ollivier, J., Schaeffer, J., Parnaudeau, S., Rhaissi, H., Nourlil, J., et al. (2013). Norovirus and other human enteric viruses in Moroccan shellfish. *Food and Environmental Virology,* 5, 35-40.

Bosch, A. Sánchez, G., Abbaszadegan, M., Carducci, A., Guix, S., Le Guyader, F.S., et al. (2011). Analytical methods for virus detection in water and food. *Food Analytical Methods.* 4:4-12.

Bucardo, F., Kindberg, E., Paniagua, M., Grahn, A., Larson, G., Vildevall, M., et al. (2009). Genetic susceptibility to symptomatic norovirus infection in Nicaragua. *Journal of Medical Virology,* 81,728-735.

Centers for Disease Control and Prevention (CDC). (2013). Emergence of new norovirus strain GII.4 Sydney--United States, 2012. *Morbidity Mortality Weekly Report.* 62(3), 55.

Chan-It, W., Thongprachum, A., Khamrin, P., Kobayashi, M., Okitsu, S., Mizuguchi, M., et al. (2012). Emergence of a new norovirus GII.6 variant in Japan, 2008–2009. *Journal of Medical Virology,* 84, 1089–1096.

Clarke, I.N., Estes, M.K., Green, K.Y., Hansman, G.S,, Knowles, N.J., Koopmans, M., et al. (2012). Caliciviridae. In King, A.M.Q., Adams, M.J., Carstens, E.B., & Lefkowitz, E.J. (eds). *Virus Taxonomy: Classification and nomenclature of viruses*. Ninth report of the International Committee on Taxonomy of Viruses. (pp. 977-985). San Diego: Elsevier Academic Press.

da Silva, A.K., Le Saux, J.C., Parnaudeau, S., Pommepuy, M., Elimelech, M., & Le Guyader, F.S. (2007). Evaluation of removal of noroviruses during wastewater treatment, using real-

time reverse transcription-PCR: different behaviours of genogroups I and II. *Applied and Environmental Microbiology,* 73, 7891-7897.

Desai, R., Hembree, C.D., Handel, A., Matthews, J.E., Dickey, B.W., McDonald, S., et al. (2012). Severe outcomes are associated with genogroup 2 genotype 4 norovirus outbreaks: a systematic literature review. *Clinical and Infectious Diseases,* 55, 189-193.
Ferreira, M.S., Victoria, M., Carvalho-Costa, F.A., Vieira, C.B., Xavier, M.P., Fioretti, J.M., et al. (2010). Surveillance of norovirus infections in the state of Rio De Janeiro, Brazil 2005–2008. *Journal of Medical Virology*, 82, 1442–1448.

Ferreira, M.S., Xavier Mda, P., Tinga, A.C., Rose, T.L., Fumian, T.M., Fialho, A.M., et al. (2012). Assessment of gastroenteric viruses frequency in a children's day care center in Rio De Janeiro, Brazil: a fifteen year study (1994-2008). *Plos One.* 7(3):e33754.

Green, K. (2013). Caliciviridae: The Noroviruses. In D.M. Knipe., P.M. Howley., J.I. Cohen., et al. (eds). *Fields Virology* 6th ed., Vol. 1. (pp. 582-608). Philadelphia: Lippincott Williams & Wilkins.

Gibson, K.E., Opryszko, M.C., Schissler, J.T., Guo, Y., & Schwab, K.J. (2011). Evaluation of human enteric viruses in surface water and drinking water resources in southern Ghana. *American Journal of Tropical Medicine and Hygiene*, 84, 20-29.

Glass, R.I., Parashar, U.D., & Estes, M.K. (2009). Norovirus gastroenteritis. *New England Journal of Medicine.* 361, 1776–1785.

Giammanco, G.M., Rotolo, V., Medici, M.C., Tummolo, F., Bonura, F., Chezzi, C., et al (2012). Recombinant norovirus GII.g/GII.12 gastroenteritis in children. *Infection Genetics and Evolution,* 12,169-174.

Gomes, K.A., Stupka, J.A., Diana, A., & Parra, G.I. (2008). Molecular characterization of calicivirus strains detected in outbreaks of gastroenteritis occurring in Argentina during 2005 and 2006. *Revista Argentina de Microbiologia*. 40:222-228.

Hall, T.A. (1999). Bio Edit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acid Symposium Series,* 41, 95-98.

Harris, J.P., Edmunds, W.J., Pebody, R., Brown, D.W., & Lopman, B.A. (2008). Deaths from norovirus among the elderly, England and Wales. *Emerging Infectious Diseases.*10,1546-1552.

Harris, A.R., Davis, J., & Boehm, A.B. (2013). Mechanisms of post-supply contamination of drinking water in Bagamoyo, Tanzania. *Journal of Water and Health,* 11, 543-554.

Hassine-Zaafrane, M., Sdiri-Loulizi, K., Kaplon, J., Salem, I.B., Pothier, P., Aouni, M., et al. (2013). Prevalence and genetic diversity of norovirus infection in Tunisian children (2007-2010). *Journal of Medical Virology,* 85, 1100-1110

Hoa Tran, T.N., Trainor, E., Nakagomi, T., Cunliffe, NA., & Nakagomi, O. (2013). Molecular epidemiology of noroviruses associated with acute sporadic gastroenteritis in children: global distribution of genogroups, genotypes and GII.4 variants. *Journal of Clinical Virology,* 56,185-193

Huynen, P., Mauroy, A., Martin, C., Savadogo, L.G., Boreux, R., Thiry, E., et al. (2013). Molecular epidemiology of norovirus infections in symptomatic and asymptomatic children from Bobo Dioulasso, Burkina Faso. *Journal of Clinical Virology*, 58,515-521

Jung, J.H., Yoo, C.H., Koo, E.S., Kim, H.M., Na, Y., Jheong, W.H., et al. (2011). Occurrence of norovirus and other enteric viruses in untreated groundwaters of Korea. *Journal of Water Health,* 9, 544-55.

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

Kim, H.S., Hyun, J., Kim, H.S., Kim, J.S., Song, W., & Lee, K.M. (2013). Emergence of GII.4 Sydney norovirus in South Korea during the winter of 2012-2013. *Journal of Microbiology and Biotechnology,* 23, 1641-1643.

Kittigul, L., Pombubpa, K., Taweekate, Y., Diraphat, P., Sujirarat, D., Khamrin, P., & Ushijima, H. (2010). Norovirus GII-4 2006b variant circulating in patients with acute gastroenteritis in Thailand during a 2006-2007 study. *Journal of Medical Virology*. 82,854-860.

Kiulia, N.M., Netshikweta, R., Page N.A., van Zyl, WB., Kiraithe, M.M., Nyachieo A., et al. (2010). 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,* 109, 818-828.

Kojima, S., Kageyama, T., Fukushi, S., Hoshino, F.B., Shinohara, M., Uchida, K., et al. (2002). Genogroup-specific PCR primers for detection of Norwalk-like viruses. *Journal of Virological Methods*, 100, 107-114.

Koo, H.L., Ajami, N., Atmar R.L., & DuPont, HL. (2010). Noroviruses: The leading cause of gastroenteritis worldwide. *Discovery Medicine,* 50,61-70.

Kroneman, A., Vega, E., Vennema, H., Vinjé, J., White, P.A., Hansman, G., et al. (2013). Proposal for a unified norovirus nomenclature and genotyping. *Archives of Virology,* 158, 2059-2068.

Kroneman, A., Vennema, H., Deforche, K., v d Avoort, H., Peñaranda, S., Oberste MS., et al. (2011). An automated genotyping tool for enteroviruses and noroviruses. *Journal of Clinical Virology,* 51, 121–125.

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

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

Mans, J., Murray, T.Y., Kiulia, N.M., Mwenda, J.M., Musoke, R.N., & Taylor, M.B. (2014). Human caliciviruses detected in HIV-seropositive children in Kenya. *Journal of Medical Virology,* 86, 75-81

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

Maritschnik, S., Kanitz, E.E., Simons, E., Höhne, M., Neumann, H., Allerberger, F., et al. (2013). A food handler-associated, foodborne norovirus GII.4 Sydney 2012-outbreak following a wedding dinner, Austria, October 2012. *Food and Environmental Virology.* 5, 220-223

Martella, V., Campolo, M., Lorusso, E., Cavicchio, P., Camero, M., & Bellacicco, A.L. (2007). Norovirus in captive lion cub (*Panthera leo*). *Emerging Infectious Diseases*, 13, 1071-1073.

Martella, V., Lorusso, E., Decaro, N., Elia, G., Radogna, A., & D'Abramo, M. (2008). Detection and molecular characterization of a canine norovirus. *Emerging Infectious Diseases,* 14, 1306-1308.

Mathijs ,E., Denayer, S., Palmeira, L., Botteldoorn, N., Scipioni, A., Vanderplasschen, A., et al. (2011). Novel norovirus recombinants and of GII.4 sub-lineages associated with outbreaks between 2006 and 2010 in Belgium. *Virology Journal,* 8, 310.

Matthews, J.E., Dickey, B.W., Miller, RD., Felzer, JR., Dawson, B.P., Lee, A.S., et al. (2012). The epidemiology of published norovirus outbreaks: a review of risk factors associated with attack rate and genogroup. *Epidemiology and Infection,* 140, 1161-1172.

Mattison, K., Sebunya, T.K., Shukla, A., Noliwe, L.N., & Bidawid, S. (2010). Molecular detection and characterization of noroviruses from children in Botswana. *Journal of Medical Virology,* 82:321-324.

Murray, T.Y., Mans, J., & Taylor, MB. (2013). Human calicivirus diversity in wastewater in South Africa. *Journal of Applied Microbiology,* 114, 1843-1853

Nenonen, N.P., Hannoun, C., Larsson, C.U., & Bergström, T. (2012). Marked genomic diversity of norovirus genogroup I strains in a waterborne outbreak. *Applied and Environmental Microbiology*, 78, 1846-1852.

Nordgren, J., Nitiema, L.W., Ouermi, D., Simpore, J., & Svensson L. (2013). Host genetic factors affect susceptibility to norovirus infections in Burkina Faso. *PLoS One.* 2013, 8:e69557.

Park, S.H., Kim, E.J., Oh SA., Kim, CK., Choi, S.S., Cho, S.J., et al. (2011). Viral agents associated with acute gastroenteritis in Seoul, Korea. *Clinical Laboratory,* 57,59-65.

Patel, M.M., Hall, A.J., Vinjé, J., & Parashar, U.D. (2009). Noroviruses: a comprehensive review. *Journal of Clinical Virology,* 44, 1–8.

Phan, T.G., Nguyen, T.A., Kuroiwa, T., Kaneshi, K., Ueda, Y., Nakaya, S., et al. (2005). Viral diarrhea in Japanese children: results from a one-year epidemiologic study. *Clinical Laboratory,* 51, 183–191.

Puustinen L, Blazevic V, Salminen M, Hämäläinen M, Räsänen S, & Vesikari T. Noroviruses as a major cause of acute gastroenteritis in children in Finland, 2009–2010. (2011). *Scandinavia Journal of Infectious Diseases,* 43, 804–808.

Schorn, R., Höhne, M., Meerbach, A., Bossart, W., Wüthrich, RP., Schreier, E., et al. (2010) Chronic norovirus infection after kidney transplantation: molecular evidence for immune-driven viral evolution. *Clinical Infectious Diseases,* 51,307-314.

Scipioni, A., Mauroy, A., Vinjé, J., & Thiry, E. (2008). Animal noroviruses. *Veterinary Journal,* 178, 32-45.

Sdiri-Loulizi, K., Hassine, M., Aouni, Z., Gharbi-Khelifi, H., Chouchane, S., Sakly, N., et al. (2010). Detection and molecular characterization of enteric viruses in environmental samples in Monastir, Tunisia between January 2003 and April 2007. *Journal of Applied Microbiology,* 109:1093-1104.

Siebenga, J.J., Vennema, H., Zheng, D.P, Vinjé, J., Lee, B.E., Pang, X.L., et al. (2009). Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001–2007. *Journal of Infectious Diseases,* 200, 802–812.

Silva, P.A., Stark, K., Mockenhaupt, F.P., Reither, K., Weitzel, T., Ignatius, R., et al. (2008). Molecular characterization of enteric viral agents from children in northern region of Ghana. *Journal of Medical Virology*, 80, 1790-1798.

Sukhrie, F.H., Beersma, M.F., Wong, A,. van der Veer, B., Vennema, H., Bogerman, J., et al. (2011). Using molecular epidemiology to trace transmission of nosocomial norovirus infection. *Journal of Clinical Microbiology,* 49, 602-606.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar S (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution,* 28:2731-1739.

Tian, G., Jin, M., Li, H., Li, Q., Wang, J., & Duan, Z.J. (2014). Clinical characteristics and genetic diversity of noroviruses in adults with acute gastroenteritis in Beijing, China in 2008-2009. *Journal of Medical Virology*, 86,1235-1242.

Trainor, E., Lopman, B., Iturriza-Gomara, M., Dove, W., Ngwira, B., Nakagomi, O., et al. (2013). Detection and molecular characterisation of noroviruses in hospitalised children in Malawi, 1997-2007. *Journal of Medical Virology*, 85, 1299-1306

Vega, E., Barclay, L., Gregoricus, N., Shirley, S.H., Lee, D., & Vinjé, J. (2013). Genotypic and epidemiologic trends of norovirus outbreaks in the United States, 2009 to 2013. *Journal of Clinical Microbiology,52,147-55.*

Vega, E., & Vinjé, J. (2011). Novel GII.12 norovirus strain, United States, 2009-2010. *Emerging Infectious Diseases.* 17,1516-1518.

Vilaginès, Ph., Sarrette, B., Husson, G., & Vilaginès, R. (1993). Glass wool for virus concentration at ambient pH level. *Water Science and Technology,* 27, 299-306.

Yassin, M.A., Kirby, A., Mengistu, A.A., Arbide, I., Dove, W., Beyer, M., et al. (2012). norovirus and rotavirus genotypes in Ethiopia. *Paediatrics and International Child Health,* 32, 51-55.

Zakikhany, K., Allen, D.J., Brown, D., & Iturriza-Gómara, M. (2012). Molecular evolution of GII-4 norovirus strains. *PLoS One* 7(7), e41625.

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