First detection of human sapoviruses in river water in South Africa
T. Y. Murray, J. Mans and M. B. Taylor

ABSTRACT

Over a 2-year period, from January 2009 to December 2010, water samples were collected from three rivers (Klip, Rietspruit and Suikerbosrand) in the Vaal River System, South Africa. Enteric viruses were recovered by a glass wool adsorption–elution method and concentrated using polyethylene glycol/sodium chloride precipitation. Sapoviruses (SaVs) were detected using published sapovirus (SaV)-specific primers and Taqman probes in a two-step real-time reverse transcription-polymerase chain reaction assay. Based on sequence analysis of the 5'-end of the capsid gene, SaVs were genotyped. In 2009, SaVs were detected in 39% (15/38) of samples from the Klip river, 83% (5/6) from the Rietspruit and 14% (1/7) of samples from the Suikerbosrand river. In 2010, SaVs were detected in 54% (14/26) of Klip river samples, 92% (11/12) from the Rietspruit and 20% (2/10) of samples from the Suikerbosrand river. SaV strains identified in the water samples were characterised into several GI and GII genotypes. The presence of SaVs in these rivers indicates human faecal contamination which may pose a potential health risk to persons exposed to these water sources during domestic or recreational activities.

Key words | calicivirus, environment, genotypes, sapovirus, South Africa

INTRODUCTION

Sapovirus (SaV) is a genus of the Caliciviridae (CV) family and contains strains which infect humans, causing gastroenteritis (Green 2007). They are small, single-stranded RNA viruses and are non-enveloped which makes them stable in the environment and consequently well-suited to water contamination (Rzezutka & Cook 2004). The SaV genus is further divided into five genogroups (GI – GV), of which four (GI, GII, GIV and GV) infect humans (Farkas et al. 2004). Sapoviruses (SaVs) were originally known as ‘Sapporo-like’ viruses and the prototype, Sapporo virus, was detected in 1977 following an outbreak of gastroenteritis at a children’s home in Sapporo, Japan (Chiba et al. 1979). Subsequently, SaVs have been implicated in several outbreaks of gastroenteritis (Iizuka et al. 2010; Yamashita et al. 2010), including waterborne outbreaks (Bon et al. 2005). The investigation of SaVs in water sources has been largely overshadowed by that of NoVs, another genus within the CV family. To date, the presence and genetic diversity of SaVs in river water has only been reported in Japan and Spain (Hansman et al. 2007; Kitajima et al. 2010; Sano et al. 2011). Several countries have reported GI as the most prevalent genogroup detected in clinical specimens (Malasao et al. 2008; Lyman et al. 2009; Chan-it et al. 2010) and environmental samples (Kitajima et al. 2010; Sano et al. 2011).

The occurrence and diversity of SaVs in South Africa (SA) is largely unknown. SaVs have been detected in stool specimens from paediatric patients with gastroenteritis in the Pretoria region of Gauteng. SaVs were found in 4% (10) of the 250 specimens tested (Mans et al. 2010), but the circulating genotypes were not established. The presence and diversity of SaVs in river water in SA has not yet been documented.

In SA, rivers are used by many communities for domestic purposes and recreational activities. It is, therefore, important to monitor the presence of SaVs in river water, as water contaminated with SaVs may pose a potential health risk to these individuals. The characterisation of SaVs from rivers also provides information on the strains circulating in the surrounding communities. This is essential in countries like SA, where routine surveillance is lacking. The aim of this study was to monitor the presence of...
human SaVs in three rivers in the Gauteng province in order to ascertain which viruses are circulating in the surrounding populations. Characterisation of the detected SaV strains provides new data on which SaV strains are circulating in Gauteng.

**METHODS**

**Sampling sites**

The three rivers from which samples were collected are tributaries of the Vaal river, and form part of the Vaal River System (VRS). High demands for water are placed on the VRS due to extensive urbanisation, mining, industrial and agricultural activities (Department of Water Affairs 2004). The three tributaries are the Klip river, the Rietspruit and the Suikerbosrand river and they join the Vaal river downstream of the Vaal dam (Figure 1). The Klip river flows through Soweto, a densely populated region of southern Johannesburg where many informal settlements are located. The Rietspruit flows through a region with smaller towns and the Suikerbosrand river originates in and flows predominantly though a nature reserve, but also through human settlement areas.

**Water sample collection**

From January 2009 to December 2010, water samples were collected at the same sampling sites for these three rivers. Samples (10 L) were collected at a depth of 25 cm below the surface of the river. Water samples were transported to the laboratory in cooler bags with cold packs and the temperature and pH were recorded upon arrival. Samples were stored at 4 °C until processing.

**Virus recovery and concentration**

Viruses were recovered from 10 L water samples using a previously described glass wool adsorption–elution method (Mans et al. 2013). Viruses in the eluate (100 mL) were concentrated to a final volume of 10 mL in phosphate-buffered saline pH 7.4 (PBS; Sigma-Aldrich Co., St Louis, MO) by polyethylene glycol/sodium chloride (PEG/NaCl) precipitation (Minor 1985).

**RNA extraction**

Total nucleic acid was extracted from 1 mL virus concentrate using the MagNA Pure LC Total Nucleic Acid Isolation Kit (large volume) (Roche Diagnostics, Mannheim, Germany) in the automated MagNA Pure LC instrument (Roche Diagnostics). Extracted nucleic acid was eluted in 100 μL and stored at −70 °C until use.

**Positive control**

RNA extracted from a SaV-positive stool specimen from a patient with gastroenteritis was used as a positive control.

**Detection of SaVs**

**Reverse transcription**

Reverse transcription (RT) was performed using the Transcriptor First Strand cDNA synthesis kit (Roche Diagnostics) and random hexamer primers, according to...
the manufacturer’s instructions. Extracted RNA (10 μL) was added to 10 μL RT mixture and the reaction was incubated at 25 °C for 10 min, 50 °C for 60 min and 85 °C for 5 min.

Real-time polymerase chain reaction

TaqMan-based real-time polymerase chain reaction (PCR) for human SaVs was performed in a 20 μL reaction using the method described by Chan and co-workers (2006) with several modifications. Briefly, the reaction volume contained 5 μL of cDNA, 900 nmol/L of each primer (CU-SV-F1, CU-SV-F2 and CU-SV-R) and 200 nmol/L of TaqMan probe (CU-SV-probe) (Table 1) in the LightCycler TaqMan Master mix (Roche Diagnostics). The following cycling parameters were used in the Roche LightCycler 2.0 (Roche Diagnostics): 95 °C for 15 min and 45 cycles of 95 °C for 15 sec, 56 °C for 1 min and 65 °C for 60 min and 85 °C for 5 min.

Conventional PCR

Nested PCR amplification was performed using two different published primer sets targeting the 5′-end of the SaV capsid gene as described by Kitajima et al. (2010) (SaV124F, SaV1F, SaV5F, SV-R13 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) (Table 1). The first round of PCR was performed in a 50 μL reaction containing 5 μL of cDNA, 1.25 U AmpliTaq Gold polymerase (Applied Biosystems, Foster City, CA) and 0.4 μmol/L of each primer. Nested PCR was performed in a 50 μL reaction containing 2 μL of first PCR product, 1.25 U AmpliTaq Gold polymerase and 0.4 μmol/L of each primer. Thermal cycling, conditions as described by Kitajima et al. (2010), was conducted on the BIOER Little Genius TC-25/H (BIOER Technology Co., Ltd, China). Results were analysed by electrophoresis through a 1.5% agarose gel and amplicons were visualised by staining with ethidium bromide. Negative controls were included in the first and second rounds of PCR in order to avoid false-positive results due to cross-contamination.

Genotyping and phylogenetic analysis

Cloning and sequencing

PCR products of the expected size were excised from the gel and purified using the Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA). The purified amplicon was directly sequenced in both directions using the ABI PRISM BigDye® Terminator v3.1 Cycle Sequencing Kit on

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Primers and probes used in this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR</td>
<td>Primer/probe</td>
</tr>
<tr>
<td>Real-time</td>
<td>CU-SV-F1</td>
</tr>
<tr>
<td></td>
<td>CU-SV-F2</td>
</tr>
<tr>
<td></td>
<td>CU-SV-R</td>
</tr>
<tr>
<td>Nested: first round</td>
<td>SaV124F</td>
</tr>
<tr>
<td></td>
<td>SaV1F</td>
</tr>
<tr>
<td></td>
<td>SaV5F</td>
</tr>
<tr>
<td></td>
<td>SV-F13</td>
</tr>
<tr>
<td></td>
<td>SV-F14</td>
</tr>
<tr>
<td></td>
<td>SV-R13</td>
</tr>
<tr>
<td></td>
<td>SV-R14</td>
</tr>
<tr>
<td></td>
<td>SVR-DS5</td>
</tr>
<tr>
<td></td>
<td>SVR-DS3</td>
</tr>
<tr>
<td></td>
<td>SVR-DS4</td>
</tr>
</tbody>
</table>

*UPAC codes used to indicate degenerate positions.
1AY237422.
2U56464.
3AY564656.
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. Where mixed sequences were detected in a sample, amplicons were cloned using the cloneJET™ PCR cloning kit (Fermentas, Burlington, Ontario). Ten clones were randomly selected for sequencing using pJET1.2/blunt specific primers.

Phylogenetic analysis

Nucleotide sequences were analysed using Sequencher™ 4.9 (Gene Codes Corporation, Ann Arbor, MI) and BioEdit Sequence Alignment Editor (V.7.0.9.0). BLAST-n was used to determine the most closely related nucleotide sequences in GenBank. The 5′-end of the capsid gene (approximately 300 nucleotides) was aligned with reference SaV strains, selected according to Oka et al. (2012), and closely matched sequences from GenBank using MAFFT Version 6 (https://align.bmr.kyushuu.ac.jp/mafft/online/server/). Phylogenetic analysis was performed in MEGA5 using the neighbour-joining method with 1,000 replicates for bootstrap analysis. Genotypes were assigned based on clustering with reference strains in the phylogenetic tree. SaV genogroup III (accession number AF182760) was used as an outgroup. SaV sequences were submitted to GenBank under the following accession numbers: KC511558-KC511581.

RESULTS AND DISCUSSION

Detection of SaVs in river water in Gauteng

From January 2009 to December 2010, SaVs were detected in 22 out of the 23 months in which water samples were drawn. SaVs were detected in the majority of samples received in May, August and November 2009 and in January and April through to July 2010 (Figure 2). No distinct seasonal distribution was seen over the 2 years; however, a seasonal pattern may become apparent when SaV concentrations are determined. An overall increase in SaV prevalence from 41% in 2009 to 58% in 2010 was noted.

The number of water samples varied between the three river sites over the 2 years (Table 2). The Klip river was sampled most frequently, with 64 samples collected in total. In 2009, 39% (15/38) of the samples tested positive for SaV and in 2010, 54% (14/26) were positive. Fewer samples were collected from the Rietspruit and the Suikerbosrand rivers, with 18 and 17 samples respectively. Samples from the Rietspruit frequently tested positive for SaV; 83% (5/6) of samples were positive in 2009 and 92% (11/12) in 2010. The Suikerbosrand had the lowest occurrence of SaV, with 14% (1/7) of samples testing positive in 2009 and 20% (2/10) in 2010. Each of the three rivers showed an increase in SaVs detected from 2009 to 2010.

Genetic diversity of SaVs in the Klip and Rietspruit rivers

SaVs were genotyped from 21 water samples, 10 of which were collected from the Klip river and 11 from the Rietspruit. Phylogenetic analysis clustered the 24 strains typed
from 21 water samples into two genogroups (GI and GII) and nine genotypes (GI.1, GI.2, GI.3, GI.5, GI.6, GI.7 GII.3, GII.5 and GII.7) (Figure 3). Genotype I.2 strains were detected most frequently in 43% (9/21) of the water samples, suggesting that it is a prevalent genotype circulating in these rivers. The GI.2 cluster contained strains from both the Klip river (5) and the Rietspruit (4), mostly from 2010 and two samples from 2009. The other genotypes were represented by three (GI.5, GI.6), two (GI.3, GI.5, GII.7) or one (GI.1, GI.7, GII.3) strain each. Overall, there is high nucleotide identity (94–100%) over the typed region between the SaV strains found in SA and those found in stool and water from several other countries.

Multiple genotypes are often detected in a single water sample and out of the 21 samples, three contained multiple genotypes. They were all Klip river samples from 2010, two from March and one from June. In the first sample (2010-03-01_K) the strains clustered with GI.1 and GI.7 and in the second (2010-03-08_K), with GI.2 and GI.3. The third sample (2010-06-21_K) contained strains that grouped with GI.2 and GI.5.

SaVs are present in these three tributaries of the Vaal river and the prevalence has increased from 2009 to 2010 in all three rivers. The occurrence of SaV suggests that the rivers are contaminated with human faecal pollution, which has also been confirmed by faecal indicators in the water exceeding recommended levels (>2,000 colony forming units/100 mL) for recreational purposes (data not shown). This increase in human faecal pollution may be a result of poor sanitation conditions and malfunctioning sewage treatment plants in the areas surrounding the rivers. Notably, the Klip river flows through Soweto which is densely populated with informal settlements where sanitation conditions are not adequate. A sewage treatment plant which is frequently non-functional is situated near, and flows into, the Rietspruit. The Suikerbosrand river, which is the least contaminated, flows mainly through a nature reserve and is the least impacted on by urbanisation.

The high occurrence of SaVs in the environment suggests continual circulation of the viruses in surrounding communities. This differs from the clinical-based data in SA, where SaVs were only detected in 4.5% of stool specimens from hospitalised paediatric patients with gastroenteritis (Mans et al. 2010). Further studies are needed to clarify this difference between SaV presence in a clinical and environmental setting. The clinical study in SA was restricted to a hospitalised paediatric population, so possible explanations for the high presence of SaVs in the environment include that SaVs are infecting the adult population or that SaV gastroenteritis is less severe and therefore does not require hospitalisation. Recent studies from Europe and the USA have reported outbreaks of SaV, which was previously thought to mainly infect children, in adult populations (Svraka et al. 2010; Lee et al. 2012).

A diverse number of SaV strains were detected, with GI.2 being detected most often. Studies on SaVs in river water in Japan and Spain have also shown that GI genotypes predominate, with GI.2 and GI.3 being detected most frequently in Spain and Japan, respectively (Kitajima et al. 2010; Sano et al. 2011). The SA GI.2 cluster shared high nucleotide identity (98–100% over the typed region) with SaVs characterised from river water in Spain (2008), wastewater influent in Japan (2005) and from a stool specimen from a patient with gastroenteritis in Japan (2005). Two other SaV strains from SA (GI.5) also matched closely (94% nucleotide identity) with a SaV from wastewater influent in Japan (2005). The other SA strains shared high nucleotide identity (96–100%) with SaVs detected in stool specimens from patients with gastroenteritis from several countries, including Denmark (GI.6), Japan (GI.3, GI.5 and GII.3), Pakistan (GI.7), Russia (GI.5) and Thailand (GI.1). This high nucleotide identity between the strains found in water and those found in stool specimens indicates the potential for these strains to be clinically relevant.

Two SaV strains from 2010, one from the Klip river and one from the Rietspruit, clustered within GII.7 but with relatively low nucleotide identity of 80% with the reference GII.7 strain (accession number AB630067). To clarify this grouping, a larger region, preferably the full capsid would need to be sequenced from the strains in these two samples. There was, however, one match on GenBank with 96% nucleotide identity; a novel SaV that had been associated with an outbreak of gastroenteritis in Minnesota, USA, in 2010 (HM590581).

This study has shown a high percentage of SaV-positive river water samples; however this may still be an underestimation. The inclusion of an internal amplification control (IAC) for SaVs in the detection assay, such as the recently developed competitive IAC (Murray et al. 2012), may identify reactions which have failed due to molecular inhibition and therefore tested negative for SaV. The introduction of a process control prior to nucleic acid extraction would monitor extraction efficiency and quantification of SaVs would provide additional information on the concentrations of SaV circulating in the environment, potentially identifying seasonal patterns which would otherwise be overlooked.
Figure 3: Phylogenetic analysis of sapoviruses (SaVs) from the Klip river (K) and the Rietspruit (R) in Gauteng, based on partial capsid gene sequences (approximately 300 nucleotides). The South African SaVs are indicated in bold and the date that the water sample was collected is included in the strain name. Reference SaV strains and closely matched strains from GenBank are included in the analysis. This phylogenetic tree was created using the neighbour-joining method with 1,000 bootstrap replicates (bootstrap support > 70% is indicated). Sapovirus GIII (Cowden) is used as an outgroup.
CONCLUSIONS

This study provides the first data on SaV genotypes circulating in river water in SA. The detection of SaVs over the 2 years indicated an increase in presence from 2009 to 2010, suggesting an increase in human faecal pollution of the rivers. This contamination may pose a potential health risk to individuals using these waters and further studies are required to assess this risk. In addition, the characterisation of SaVs in the environment is a good indication of strains circulating in the surrounding communities and provides valuable information on the genetic diversity of SaVs in Gauteng, SA.

ACKNOWLEDGEMENTS

The authors thank Rand Water for the collection and shipment of water samples. Tanya Y Murray was supported by Poliomyelitis Research Foundation and National Research Foundation (NRF) postgraduate bursaries and her travel to the IWA Young Water Professionals conference in Budapest, 2012, was supported by Iliso Consulting. The study was supported, in part, by the NRF and Medical Research Council of South Africa.

REFERENCES


First received 22 October 2012; accepted in revised form 18 February 2013