Epidemiological profile and genetic diversity of sapoviruses (SaVs) identified in children suffering from acute gastroenteritis in Pune, Maharashtra, Western India, 2007–2011

N. LASURE AND V. GOPALKRISHNA*

Enteric Viruses Department, National Institute of Virology, Pune, India

Received 6 January 2016; Final revision 22 July 2016; Accepted 8 August 2016; first published online 9 September 2016

SUMMARY

Sapoviruses (SaVs) are responsible for sporadic cases and outbreaks of acute gastroenteritis. Despite this, few studies in India have focused on the epidemiological investigation of SaV in cases of acute gastroenteritis. The aim of this study was to understand the molecular epidemiology, genetic diversity and clinical impact of SaV in diarrhoeic children from Pune, Western India. Between 2007 and 2011, a total of 985 faecal samples from diarrhoeic cases and non-diarrhoeic controls were collected and examined for the presence of SaV by nested RT–PCR. SaV was detected in 2.7% (21/778) of the cases and 1.9% (4/207) of the controls. We observed that the majority of SaV mono-infections caused severe gastroenteritis (67%) with clinical manifestations of diarrhoea (100%), vomiting (73%) and dehydration (80%). All known human SaV genogroups were detected in the study. At least eight genotypes were identified from cases and controls. Genogroups GIV and GV, along with genotypes GI.5, GII.4 and GII.6, were discovered for the first time in India. Two GII.4 study strains were found to be 98.5–99% identical, having a novel intra-genogroup recombinant (GII.1/GII.4) recently reported from the Philippines, suggesting probable evidence of recombination. The circulation pattern of SaV genotypes varied during the study period, with GII.1 being predominant in 2007 and 2009, GIV.1 in 2008, and GV.1 in 2011.

Key words: Acute gastroenteritis, genetic diversity, sapovirus, severe dehydration, Western India.

INTRODUCTION

Acute gastroenteritis (AG) is one of the leading causes of morbidity and mortality worldwide, with ~70% of the cases attributed to viruses. Rotavirus (RV) is the foremost causative agent of viral gastroenteritis [1]. Sapovirus (SaV), a member of the family Caliciviridae, is known to cause diarrhoea in humans and animals. The family Caliciviridae is divided into five genera: Norovirus, Sapovirus, Vesivirus, Lagovirus and Nebovirus [2]. Of these, certain genogroups of norovirus (NoV) and SaV are known to infect humans. SaV infections are generally milder than NoV infections; however, severe cases and outbreaks of SaV have also been reported [3]. The non-enveloped SaV capsid encloses a single-stranded positive-sense RNA genome, flanked by Vpg at the 5′-end and a poly A tail at the 3′-end. It is ~7.1–7.7 kb in size. SaV strains are classified into five genogroups (GI–GV) on the basis of nucleotide sequence variation in the capsid region [4]. In these genogroups, GI, GII, GIV and GV are responsible for infection in humans, while GIII infects porcine hosts. In addition to these, nine new groups are proposed (GVI–GXIV), which cause infections in pigs.
bats, minks and sea lions [5]. The number of open reading frames (ORFs) in the SaV genome varies between genogroups. However, irrespective of the genogroup, ORF1 codes for at least six of the non-structural proteins and the major structural capsid protein (VP1). The genome of GII and GIII strains consists of two ORFs [2]. A third ORF, the function of which is still unknown, has been predicted in GI, GIV and GV. The GI and GII strains are further classified into seven genotypes each (GI.1–GI.7 and GII.1–GII.7), while the GV strains include two genotypes (GV.1 and GV.2), and GIV strains have a single genotype (GIV.1) [4, 6]. In addition, inter- and intra-genogroup (inter- and intra-genotype) SaV recombinants have also been reported [7, 8].

Overall, SaVs have been detected in faecal samples of sporadic cases of AG at a frequency of 2.2–12.7%, worldwide [2]. They are also known to cause outbreaks of gastroenteritis, primarily in semi-closed settings such as day-care centres and schools [3, 9]. Several studies from Asia have documented the presence of SaV strains and genetic variations within both occurrences, namely sporadic cases and outbreaks of AG from Western India, between January 2007 and December 2011. Faecal samples collected from non-diarrhoeic controls (n = 207) during the study period were also included. A sporadic case of AG was defined as the passage of ≥3 episodes of loose (watery) stools per day, with or without associated symptoms like vomiting, fever or abdominal pain. Informed consent permitting the collection and utilization of faecal samples was obtained from the parents or legal guardians of the children. Detailed case-reporting forms (CRFs) covering demographic and clinical information were completed for each registered case. The clinical severity of cases was determined from CRFs using the 20-point Vesikari scoring system [12]. This study was approved by the Institutional Human Ethical Committee of the National Institute of Virology, Pune. All faecal samples were analysed for the presence of enteric viruses such as RV, NoV, adenovirus (AdV), astrovirus (AstV) and enterovirus (EV). Detection of group A rotavirus was carried out by ELISA (Dako Cytomation, UK), while NoV, AdV, AstV and EV were detected by conventional polymerase chain reaction (PCR) using specific primers targeting the RdRp (126 bp), Hexon (300 bp), ORF1a (289 bp) and 5′-UTR region (537/400 bp), respectively.

Sample processing and nucleic acid extraction

Clarified 30% (w/v or v/v) faecal suspensions were prepared in 0.1 m PBS (pH 7.4). Viral RNA extraction was carried out using QIAamp viral RNA mini kit (Qiagen, Germany) according to the manufacturer’s protocol.

SaV detection in clinical samples

Reverse transcription (RT)–PCR was performed in a single 25 μl reaction mixture containing 5 μl of extracted RNA, 5 μl of 5× RT–PCR buffer, and 1 μl RT–PCR enzyme mix (Qiagen One-Step RT–PCR kit, Germany), using a nested set of primers described previously [13]. RT was performed at 47 °C for 30 min, according to the manufacturer’s protocol. The thermal cycling conditions used for the first-round PCR were as follows: initial denaturation at 94 °C for 10 min, followed by 35 cycles at 94 °C for 1 min, 47 °C for 1 min, and 72 °C for 1 min, followed by final extension for 10 min at 72 °C. Nested PCR was carried out under identical conditions by changing the annealing temperature to 42 °C. The PCR products were resolved on 2% agarose gel, and positive amplicons (~420 bp) were visualized under a UV transilluminator using a gel documentation system.

Nucleotide sequencing and phylogenetic analysis

All positive amplicons were subjected to cycle sequencing using BigDye Terminator v. 3.1 Cycle Sequencing kit (Applied Biosystems, USA), and were further sequenced
using the ABI 3130XL Genetic Analyzer (Applied Biosystems). Alignment of study sequences, along with prototype strains and reference strains deposited in GenBank and phylogenetic analysis, were performed using MEGA 6 software [14]. Nucleotide sequences identified in the study have been submitted to GenBank under accession numbers KU317432–KU317451.

**Statistical analysis**

The positivity between groups was compared using Fisher’s exact test, while the difference in median values between mono- and mixed infection groups was compared using the Mann–Whitney U test. A P value of ≤0.05 was considered as significant.

**RESULTS**

**Detection of SaV in faecal samples from AG cases and controls**

During the 5-year study period, SaV was detected at a frequency of 2.7% (21/778) in children hospitalized with AG, and 1.9% (4/207) in asymptomatic controls. Although the frequency of detection in diarrhoeic cases was more than that in non-diarrhoeic controls, this finding was not statistically significant (P = 0.54, Fisher’s exact test). Co-infection with other enteric viruses was observed in 28.6% (6/21) of the positive cases and 25.0% (1/4) of the controls. Dual infection with RV (33.3%), AdV (33.3%) and AstV (16.7%) was observed in 83.3% cases (5/6), while triple infection with RV and EV was observed in 16.7% cases (1/6). In the SaV-positive control sample, dual infection with EV was observed.

**Age distribution of SaV-positive cases**

SaV was the sole infection in children aged ≤3 years, in gastroenteritis cases as well as controls. In SaV-positive AG, 12/21 cases (57.1%) belonged to children aged <12 months, while 8/21 (30.1%) and 1/21 (4.8%) cases were observed in children aged between 12–24 and 24–36 months, respectively (Fig. 1).

**Seasonal distribution of SaV-positive cases**

In AG cases, SaV infections were reported throughout the year, with peak infections recorded in summer: 9/21 (42.8%) cases occurred in summer, 6/21 (28.6%) during monsoon, 3/21 (14.3%) each in post-monsoon and winter seasons, respectively (Fig. 2). In controls, SaV was detected only during monsoon (50%) and winter (50%).

**Clinical features of SaV infection in AG cases**

Clinical features of SaV mono-infections (15/21) and mixed infections (6/21) are presented in Table 1. While diarrhoea was primarily observed in all positive cases, occurrence of fever and vomiting varied between the two groups. Children with mono-infections experienced vomiting more frequently than those with mixed infections. In contrast, fever was observed more frequently in mixed infections. However, these differences were not statistically significant (P = 0.29 and P = 0.55, respectively, using Fisher’s exact test). Dehydration was assessed on the basis of presence or absence of clinical features such as lethargy, restless/irritable behaviour, sunken eyes, skin pinch, and feeding habits of the child. Some degree of dehydration was observed in both groups (80% in SaV mono-infections and 83% in mixed infections). A high number of children with mono-infection had severe dehydration (60%). The median duration and frequency of diarrhoea in children with mono-infections were slightly lower than that observed in children with mixed infections, whereas the duration and frequency of vomiting was higher for mono-infections. The median duration of hospitalization for mono-infection cases was significantly longer than mixed infection cases (P = 0.039, Mann–Whitney U test). Overall, the severity assessment of positive cases revealed that 67% of SaV mono-infections resulted in severe gastroenteritis.

**Nucleotide sequence and phylogenetic analysis**

Nucleotide sequence analysis of the SaV capsid region (~420 bp) showed a high degree of variation in study
strains. All four human SaV genogroups were detected, and at least eight different genotypes could be identified from AG cases: GI.1 (n = 2), GI.3 (n = 1), GI.5 (n = 1), GII.1 (n = 9), GII.4 (n = 2), GIV.1 (n = 3), GV.1 (n = 1), and in controls: GI.1 (n = 1), GI.2 (n = 1), GII.1 (n = 1) and GII.6 (n = 1). Phylogenetic analysis of the SaV strains from AG cases indicated that these strains were distributed in seven genoclusters within GI, II, IV and V (Fig. 3). Out of these 21 strains, 11 (52.3%) clustered with GII9 GII.1, and two GII.4 strains forming a separate cluster with a recombinant strain TG012-28, four (19%) with GI in three clusters (two GI.1, one GI.3, one GI.5), three (14.3%) with GV, and three (14.3%) with GIV. One GIV strain could not be incorporated in the tree due to sub-optimal amplification (∼280 bp).

GI and GII strains displayed more diversity than GIV and GV strains, which appeared to be almost identical (Table 2). Nucleotide sequence identity for study strains between the same genotype was >95%, while genotypes of the same genogroup varied between 20% and 30%. GII.1 strains showed the highest level of variation (96.1–99.3%) among all the genotypes detected in the study. GI.1, GII.1, GIV.1 and GV.1 strains most closely resembled their prototype strains, with nucleotide identity >95%, while GI.3, GI.5 and GII.4 strains showed greater divergence to their respective prototypes. The SaV genotypes and their variants identified in this study were most closely related to strains from Japan, the Philippines, Russia, Thailand, UK and the United States.

The circulation pattern of SaV strains varied throughout the study period and showed a changing trend (Table 3). While GII strains were detected throughout the study period, the occurrence of GI, IV and V strains varied greatly. Although genotype GII.1 was consistently detected, it was observed to be predominant only in 2007 and 2009. GIV.1 emerged in 2008 and was the predominant genotype that year whereas GV.1 emerged in 2010 and became the predominant genotype in 2011. In 2010, the frequency of detection of GII.1 was similar to that of GV.1 (Table 3).

DISCUSSION

Worldwide, SaV is detected at a frequency ranging between 0.9% and 12.7% by conventional RT-PCR [4, 15]. Higher prevalence (17–18%) has been observed in recent studies using the real-time PCR approach [16, 17]. Previous reports available from southern (Vellore) and northern (New Delhi) regions of India, along with Japan, have reported SaV infections in
Fig. 3. Phylogenetic tree based on partial nucleotide sequence analysis of the capsid region of sapovirus (SaV) (~420 bp). The tree was generated using neighbour-joining method with 1000 bootstrap replicates, in MEGA 6. Scale bar indicates nucleotide substitutions per site. SaV strains detected in the study are indicated by bold font.
5·1–12·7% of cases hospitalized with AG [11, 18, 19]. In the present study, SaV was detected in 2·7% of children with AG and in 1·9% of asymptomatic children. Our data is similar to other studies from eastern (Kolkata) and northern (Lucknow) India, wherein the frequency of detection of SaV in children with AG was reported between 0·9% and 3·2% [15, 20]. SaV prevalence in asymptomatic controls was reported to be lower than in AG cases. In a study from Denmark, SaV infections were observed in 3·2% of the diarrhoeal cases and in 1·2% of the non-diarrhoeic controls [21]. In an earlier study from Vellore, India, SaV prevalence was reported to be higher in AG cases than in controls (5·1% vs 3·5%) [19]. Similarly to our finding, the difference was not statistically significant ($P = 0·4$). SaVs are often reported to cause mixed infections with other enteric pathogens. High rates of co-infections were reported in earlier studies from India (61%) and Burkina Faso (56%) [17, 19]. The lower rate of co-infection (28·6%) in our study was comparable to similar studies from France (26·2%) and the Philippines (34%) [8, 22]. The reason for these differences could be explained by the fact that only a few other enteric pathogens were examined in our study. In the present study, SaV infections were observed only in children aged <3 years. The occurrence of infections was observed to decrease with increasing age (1 > 2 > 3 years), with 95·2% of the infections observed in children aged <2 years. Comparable data on age susceptibility to SaV infections was previously reported from the United States and Nicaragua, where ~75% and 93% of infections were observed by age 2 years, respectively [17, 23]. Globally, SaV infections have mainly been reported in winter months [10, 23–25]. However, in our study, the infections were observed throughout the year, with peak activity during the summer months. Although human calicivirus infections are detected more frequently during winter, sporadic SaV cases and outbreaks have been reported in summer months as well. Our data on seasonality is in agreement with studies from Finland and Japan, where highest SaV prevalence was reported during the warmer months [26, 27]. Although SaVs are reported to cause milder infections than RVs and NoVs, they are frequently known to cause hospitalization [28]. In our study, the most common clinical manifestations in children with SaV mono-infections were diarrhoea (100%), dehydration (80%) and vomiting (73%). In similar studies from Bangladesh and Burkina Faso, cases with SaV infection experienced diarrhoea more frequently than vomiting [10, 16]. Variable levels of dehydration in SaV-positive gastroenteritis cases, have been reported from countries like Bangladesh (88%) and Nicaragua (40%) [10, 17]. However, the status of dehydration in mono- and mixed infections was not

| Table 2. Nucleotide sequence identities of sapovirus (SaV) strains detected in acute gastroenteritis cases |
|--------------------------------------------|---------------------|---------------------|---------------------|
| SaV genotype $(n)$ | Prototype strain | Within study strains | With prototype strain | Maximum identity with reference strain |
| GI.1 (2) | Sapporo/1982/JP | 99·3 | 99·3 | FJ214044 (Russia) |
| GI.3 (1) | Stockholm318/97/SE | – | 90·5 | LC071859 (Japan), |
| GI.5 (1) | Yokote1/06/JP | – | 96·5 | AY538716 (Russia) |
| GII.1 (9) | London/1992/UK | 96·1–99·3 | 95·5–96·3 | AJ249939 (UK) |
| GII.4 (2) | Kumamoto/6/JP | 99·5 | 91·6 | AY237419 (Thailand) |
| GIV.1 (2) | Hou7-1181/US/1990 | 100 | 95·4 | KM067444 (Philippines) |
| GV.1 (2) | Arg39/Argentina/1995 | 100 | 96·4 | HM800919 (USA) |

| Table 3. Year-wise distribution of sapovirus genotypes in acute gastroenteritis cases and their changing trend, 2007–2011 |
|--------------------------|--------------------------|
| Year | Genotypes in circulation |
| 2007 $(n = 6)$ | GII.1 (4), GI.3 (1), GI.4 (1) |
| 2008 $(n = 6)$ | GIV.1 (2), GII.1 (1), GI.4 (1), GII.1 (1), GI.5 (1) |
| 2009 $(n = 4)$ | GII.1 (2), GI.1 (1) |
| 2010 $(n = 2)$ | GII.1 (1), GV.1 (1) |
| 2011 $(n = 4)$ | GV.1 (2), GI.1 (1), GIV.1 (1) |

Predominant strains are indicated by bold font.

Downloaded from https://www.cambridge.org/core. IP address: 54.70.40.11, on 26 Aug 2018 at 17:36:11, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms . doi:10.1017/S0950268816001953
clearly differentiated in these studies. In the present study, dehydration was observed in 80% and 83% of mono- and mixed infection cases, respectively. Data from Burkina Faso showed similar rates of dehydration in mono- (75%) and mixed (77%) infection cases [16]. However, while 60% of cases with mono-infection were observed to have severe dehydration in our study, only 10% of the mono-infection cases were reported to have severe dehydration in Burkina Faso. Greater severity of dehydration in our study could be explained by the fact that 73% of the SaV mono-infections were accompanied by vomiting, while only 45% from Burkina Faso were reported to have severe dehydration in Burkina Faso. Greater severity of dehydration in our study, only 10% of the mono-infection cases had severe dehydration in Burkina Faso. Greater severity of dehydration in our study could be explained by the fact that 73% of the SaV mono-infections were accompanied by vomiting, while only 45% from Burkina Faso were reported to experience vomiting. Although the median duration and frequency of diarrhoea and vomiting in our study did not vary significantly between the two groups, median duration of hospitalization was significantly longer for mono-infection cases than mixed infection cases. Our study showed that SaVs could cause severe infections in sporadic AG cases, resulting in long periods of hospitalization. Currently, similar detailed epidemiological studies on age distribution, seasonality and clinical severity of SaV infections have not been reported from India.

SaVs display a wide range of genetic diversity, which varies with geographical location. To date, human SaVs are classified into four genogroups and 17 genotypes, based on >20·1% variation in their complete capsid sequences [4, 6]. Of these, several genotypes of GI and GII were predominantly detected in AG cases worldwide, while GIV and GV genotypes are detected at lower frequencies [24, 29, 30]. Earlier reports available from Asian countries have documented the predominance of genotypes GI.1, GI.2 and GII.1 in AG cases [10, 31, 32]. Limited data with respect to genetic diversity of SaVs in AG cases are available from India. SaV GI.1 was predominantly detected in studies from Kolkata (east) and Vellore (south), while GI.1 was predominant in New Delhi (north) [15, 18, 19]. A total of six genotypes (GI.1–3 and GII.1–3) were reported to circulate in Vellore between 2001 and 2004. A recent study from Lucknow (north India) reported the presence of SaV in AG cases; however, the genetic variation among the strains was not addressed [20]. Our study reports the presence of at least eight genotypes of SaV from AG cases and asymptomatic controls. Genotypes GI.5 and GII.6, along with GIV.1 and GV.1, were reported for the first time in India. The analysis of genetic diversity within the genogroups for GIV and GV strains, was limited by their low frequency of detection in the study. The wide genetic diversity of the SaV strains detected in our study could be due to detection targeting of the RdRp-capsid junction with our assay, compared to assays targeting only the RdRp region in other studies from India. Two of the GII.4 study strains formed a separate cluster with a recombinant strain from the Philippines. As described earlier, this GII.1/GII.4 recombinant strain shared nucleotide and amino-acid identity of 94% and 99% with GII.1 in the non-structural region, and 92% and 99% similarity with GII.4 in the capsid region [8]. While the two GII.4 strains showed 9-5% divergence from the GII.4 prototype, they shared a 98-5-99% sequence identity with the Philippines strain. It is likely that these two GII.4 strains are recombinant SaV strains as suggested in the Philippines study. However, sequencing of the complete RdRp and capsid regions is mandatory to confirm this. In our study, the annual distribution of SaV genotypes varied greatly in Pune, Western India, with GI.1 being predominant in 2007 and 2009, GIV.1 in 2008, and GV.1 in 2011. Similar genotypic variation and shift was previously observed in Chiang Mai province, Thailand and the Kumamoto prefecture in Japan [33–35]. The GV.1 study strains shared an amino-acid identity of 99% with other GV.1 strains that were reported from an outbreak that occurred in Japan, indicating that SaV strains detected from sporadic cases could also be responsible for outbreaks of gastroenteritis. In view of the present data on genetic diversity and the changing trend of SaV strains identified in Pune, a constant surveillance of SaVs from other parts of India is mandatory.

Although the rate of SaV detection from Western India was slightly lower than that previously reported from other parts of India, the number of SaV mono-infection cases with severe gastroenteritis having longer periods of hospitalization, was higher in this study. The genotypic data from the study region reveals circulation of multiple genotypes, with the presence of genotypes GIV.1 and GV.1 being documented for the first time in India. Our study further emphasizes the need for continuous surveillance of SaVs from other parts of India for better management of AG, which is one of the major public health concerns in developing countries like India.

ACKNOWLEDGEMENTS

Authors are grateful to Dr D. T. Mourya, Director, National Institute of Virology, Pune for his constant support during this study.
Sapovirus gastroenteritis in Pune, India


