



# Detection of Human Sapoviruses in Sewage in China by Next Generation Sequencing

Ke Song<sup>1</sup> · Xiaojuan Lin<sup>2</sup> · Yao Liu<sup>2</sup> · Feng Ji<sup>2</sup> · Li Zhang<sup>2</sup> · Peng Chen<sup>3</sup> · Chenxu Zhao<sup>4</sup> · Yanyan Song<sup>1</sup> · Zexin Tao<sup>2</sup> · Aiqiang Xu<sup>2,4</sup>

Received: 28 November 2020 / Accepted: 25 February 2021 / Published online: 23 March 2021  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

Human sapovirus (SaV) is an important causative agent of nonbacterial gastroenteritis in humans. However, little is known about its circulation in China. To study the prevalence and diversity of human SaV genotypes circulating in eastern China, a 3-year environmental surveillance combined with next generation sequencing (NGS) technology was conducted. A total of 36 raw sewage samples were collected from January 2017 to December 2019 in Jinan and processed. Thirty-five (97.22%) samples were positive for human SaV genome in quantitative RT-PCR assay; 33 (91.67%) samples were positive in nested RT-PCR assay on partial capsid VP1 sequence and all amplicons were further analyzed separately by NGS. Among those, ten genotypes belonging to the genogroups of GI, GII, GIV, and GV were identified by NGS, including 4 major genotypes (GI.2, GI.1, GV.1 and GI.3) and 6 uncommon genotypes (GII.5, GII.1, GII.NA1, GII.3, GI.6 and GIV.1). A temporal switch of predominant genotype was observed from GI.2 to GI.1 around June 2019. Local and foreign sequences clustered together in some branches according to phylogenetic analysis, indicating frequent transmission of various lineages in different regions of the world. Environmental surveillance provides a comprehensive picture of human SaV in China. NGS-based environmental surveillance improves our knowledge on human SaV circulating in communities greatly and should be encouraged as a sensitive surveillance tool.

**Keywords** Human sapovirus · Next generation sequencing · Phylogenetics · Environmental surveillance

## Introduction

Sapovirus (SaV) is a small nonenveloped virus belonging to the family *Caliciviridae*. The SaV genome has a positive-sense, single-stranded RNA, which is approximately 7.1–7.7 kb in size organized into two or three open reading frames (ORFs). Human SaV is classified into 4 genogroups (GI, GII, GIV and GV) based on the complete VP1 nucleotide sequences, which are further subdivided into 18 genotypes (GI.1–GI.7, GII.1–GII.8, GIV.1, GV.1 and GV.2) (Oka et al. 2015; Kagning Tsinda et al. 2017).

Human SaV is an important causative agent of nonbacterial gastroenteritis among population (Platts-Mills et al. 2018). According to recent data, human SaV resulted in about 2.2–22.6% of the gastroenteritis worldwide (Mancini et al. 2019). All age groups especially infants are susceptible to human SaV (Xiao Li L. Pang et al. 2014; Rockx et al. 2002; de Wit et al. 2001). Although the severity of SaV-associated gastroenteritis is generally milder than norovirus and rotavirus-associated gastroenteritis (Page et al. 2016; Sakai et al. 2001),

✉ Yanyan Song  
yysong@sdu.edu.cn

✉ Zexin Tao  
zexin.tao@163.com

<sup>1</sup> Department of Microbiology, School of Public Health, Cheeloo College of Medicine, Shandong University, No. 44 Wenhuxi Road, Jinan 250012, People's Republic of China

<sup>2</sup> Shandong Provincial Key Laboratory of Infectious Disease Control and Prevention, Shandong Center for Disease Control and Prevention, No. 16992 Jingshi Road, Jinan 250014, People's Republic of China

<sup>3</sup> School and Hospital of Stomatology, Cheeloo College of Medicine, Shandong University, 44-1 Wenhuxi Road, Jinan 250012, People's Republic of China

<sup>4</sup> Department of Epidemiology, School of Public Health, Cheeloo College of Medicine, Shandong University, No. 44 Wenhuxi Road, Jinan 250012, People's Republic of China

human SaV can result in hospitalization (Lee et al. 2012; Medici et al. 2012). Human SaV has been identified in both sporadic and outbreak cases of the gastroenteritis (Oka et al. 2015). A Meta-analysis reported that outbreaks were mainly caused by GI and GIV worldwide (Yu et al. 2019), although there were outbreaks associated with other genogroups (Yu et al. 2019; Hergens et al. 2017; Oka et al. 2017). Outbreaks caused by human SaV always occurred in closed and semi-closed settings, such as kindergartens, hospitals, ships, long-time care facilities, and schools (Yamashita et al. 2010; Pang et al. 2009; Usuku et al. 2008; Yan et al. 2005). In China, 2 gastroenteritis outbreaks associated with human SaV occurred in Shenzhen during 2015–2016 (Wang et al. 2018). Human SaV poses significant disease burden, which highlights the emerging role as a public health issue (Liu et al. 2016).

Due to the relatively low positive rate of human SaV in gastroenteritis patients (compared to norovirus and rotavirus), information on its genotype diversity in China is limited. Human SaV can be detected in sewage. Environmental surveillance on human SaV have been conducted in Thailand, Japan, Italy, Brazil, Tunisia, etc (Khamrin et al. 2020; Mancini et al. 2019; Ibrahim et al. 2019; Thongprachum et al. 2018; Fioretti et al. 2016; Murray et al. 2013; Kitajima et al. 2010) to study its molecular epidemiology, whereas, to the best of our knowledge, no studies on human SaV in sewage in China have been reported yet.

Generally, PCR amplicons of human SaV genomes from sewage contain multiple genotypes and variants. Cloning and Sanger sequencing offers a labor-consuming and inefficient approach in previous studies (Kumthip et al. 2020; Ibrahim et al. 2019; Thongprachum et al. 2018; Fioretti et al. 2016; Murray et al. 2013; Kitajima et al. 2010). Recently, next generation sequencing (NGS)-based amplicon sequencing has been carried out successfully for detection of viruses in sewage, such as SARS-CoV-2 (Ahmed et al. 2020), enterovirus (Majumdar and Martin 2018; Montmayeur et al. 2017), norovirus (Fumian et al. 2019; Suffredini et al. 2018), adenovirus (Iaconelli et al. 2017), and human SaV (Mancini et al. 2019). NGS has the advantages of high sensitivity and high throughput for detecting viruses in mixed samples and it can detect less prevalent genotypes undetectable comparing to Sanger sequencing (Mancini et al. 2019). Here, we collected sewage samples monthly during 2017–2019 in Jinan, China and analyzed human SaV by quantitative PCR and NGS-based amplicon sequencing to study its genotypes and genetic diversity.

## Materials and Methods

### Sampling

Between 2017 to 2019, 36 raw sewage samples were collected monthly by using grab sampling method from the

influent of a wastewater treatment plant (WWTP) in Jinan, the capital city of Shandong Province, China. The WWTP collects the domestic sewage from approximately one million inhabitants. A total of 1 liter of sewage sample was collected into sterile containers each time and stored at a low temperature (4 °C) before processing.

### Sewage Concentration and RNA Extraction

The sewage samples were concentrated 100-fold by mixed cellulose ester (MCE) membrane adsorption and ultrasonication elution method as described previously (Matsuura et al. 1984; Berg et al. 1971). Briefly, 1 liter of sewage was centrifuged at 3200×g for 30 min at 4 °C. The supernatant was adjusted to a final Mg<sup>2+</sup> concentration of 0.05M and a pH value of 3.5 by using MgCl<sub>2</sub> and hydrochloric acid. After the solution was filtered through an MCE membrane, the virus absorbed on the membrane was eluted with 10 ml of 3.0% beef extract solution (pH 8.5, adjusted by NaOH) by 3-min ultrasonication. The eluent was centrifuged again at 3000×g for 30 min, filtered through a 0.22 μm filter, and was adjusted to the pH value of 7 by hydrochloric acid.

Total viral RNA was extracted from 420 μl of concentration solution to a final volume of 50 μl by using QIAamp Viral RNA Mini Kit (QIAGEN, USA), according to the manufacturer's instructions.

### qRT-PCR

The qRT-PCR assay was carried out using SaV124F, SaV1F, SaV5F, and SaV1245R primers and SaV5TP and SaV124TP probes, which targeted polymerase-capsid junction region (Oka et al. 2006). Five microliters of RNA extract were subjected to amplification by using AgPath-ID One-Step RT-PCR reagents (ABI) with a final volume of 25 μl. Each sample was tested in duplicates. The amplification conditions were reverse transcription at 45 °C for 10 min, denaturation at 95 °C for 10 min, and followed by 40 cycles of 95 °C for 15 s and 60 °C for 45 s.

Quantification of Pepper mild mottle virus (PMMoV) RNA in sewage was performed via qRT-PCR using AgPath-ID One-Step RT-PCR reagents as internal control (Kitamura et al. 2020). The primers and probes in the qRT-PCR assay was accordant with published literatures (Haramoto et al. 2013; Zhang et al. 2006). Five microliters of RNA extract were amplified in a final volume of 25 μl with the cycling conditions of reverse transcription at 50 °C for 30 min and denaturation at 95 °C for 30 s, followed by 45 cycles of 95 °C for 5 s and 60 °C for 60 s.

## Nested RT-PCR and NGS

A nested RT-PCR-targeting polymerase-capsid junction region of all the human SaV was performed according to previous reports (Kitajima et al. 2010). The first round PCR was performed by using SuperScript™ IV One-Step RT-PCR System with a final volume of 25 µl. The forward primers were SaV124F, SaV1F, and SaV5F, while the reverse primers were SV-R13 and SV-R14. The amplification conditions were 45 °C for 30 min and 98 °C for 2 min and followed by 40 cycles of 98 °C for 10 s, 50 °C for 10 s, and 72 °C for 1 min with a final extension step of 72 °C for 5 min. The second round was performed by using Platinum Taq DNA Polymerase in a final volume of 100 µl. The forward primer and the reverse primer were SV-1245Rfwd and SV-R2, respectively. The amplification conditions were 94 °C for 2 min and followed by 35 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s with a final extension step of 72 °C for 5 min. PCR products were analyzed by agarose (1.5%) gels electrophoresis. The lengths of products in the first and second round PCR were 800 bp and 430 bp, respectively. The positive products were forward to NGS analysis.

NGS library preparation and Miseq sequencing using 2×150 bp paired-end reads method were performed by Shanghai BioGerm Medical Biotechnology Company. Then clean data were assembled de novo to form contigs using CLC Genomics Workbench 12.0 (QIAGEN, USA) with default parameters. Following trimming, contigs length less than 200 bp were removed. Contigs with the average coverage > 30 were exported and classified into different genotypes using BLAST against a local SaV database. Sequences with E value less than E-100 were forwarded for further analysis.

## Nucleotide Diversity and Phylogenetic Analysis

The Simpson's diversity index based on the numbers of NGS reads was calculated to describe the diversity of human SaV. The nucleotide sequences were aligned by Bioedit 7.0.9.0. The nucleotide substitution model that fitted our data best was identified by MegaX. Kimura 2-parameter model with gamma-distributed rates (K2+G) was the best-fit

nucleotide substitution model. Phylogenetic tree including the sequences obtained in this study and those from GenBank was constructed based on partial VP1 nucleotide sequences (nt 5179–5571 corresponding to strain Hu/SaV/Manchester/1993/UK with accession number X86560) via Neighbor-Joining method with K2+G model by MegaX. Bootstrap test with 1000 replicates was used to measure of the robustness each node (Kumar et al. 2016).

## Results

### Human SaV Prevalence in Sewage

In the present study, 36 sewage samples were collected monthly during January 2017–December 2019. Using qRT-PCR assay, 35 out of 36 (97.22%) sewage samples were positive for human SaV nucleic acid and the sample collected in February 2019 was negative. Using nested RT-PCR assay, 33 (91.67%) samples were positive and samples collected in September 2018, February 2019, and September 2019 were negative.

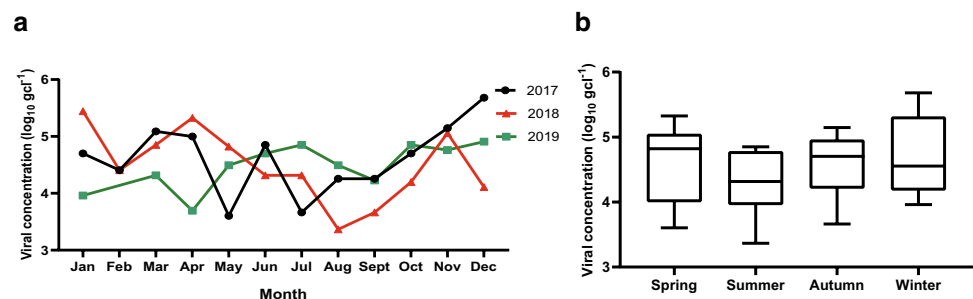
### Quantitative RT-PCR (qRT-PCR)

According to the qRT-PCR, the peak of human SaV concentration in sewage ( $4.8 \times 10^5$  genome copies per liter) was observed in December 2017, whereas the sample from August 2018 had the lowest concentration of human SaV ( $2.3 \times 10^3$  genome copies per liter) (Fig. 1a). The human SaV copies in spring (March to May), summer (June to August), autumn (September to November), and winter (December to February) were compared, and no statistically significant difference in viral copies concentration among four seasons was observed (Kruskal–Wallis test,  $P > 0.05$ ) (Fig. 1b).

### Quality Control

In this study, PMMoV RNA in all 36 samples was examined via qPCR as internal quality control. As shown in Supplementary Table S1, PMMoV was tested positive in all

**Fig. 1** Sapovirus concentration (copies per liter) in sewage monthly from January 2017 to December 2019, by month (a) and by season (b). The sample collected in February 2019 was negative in SaV qRT-PCR assay, and is not included in b

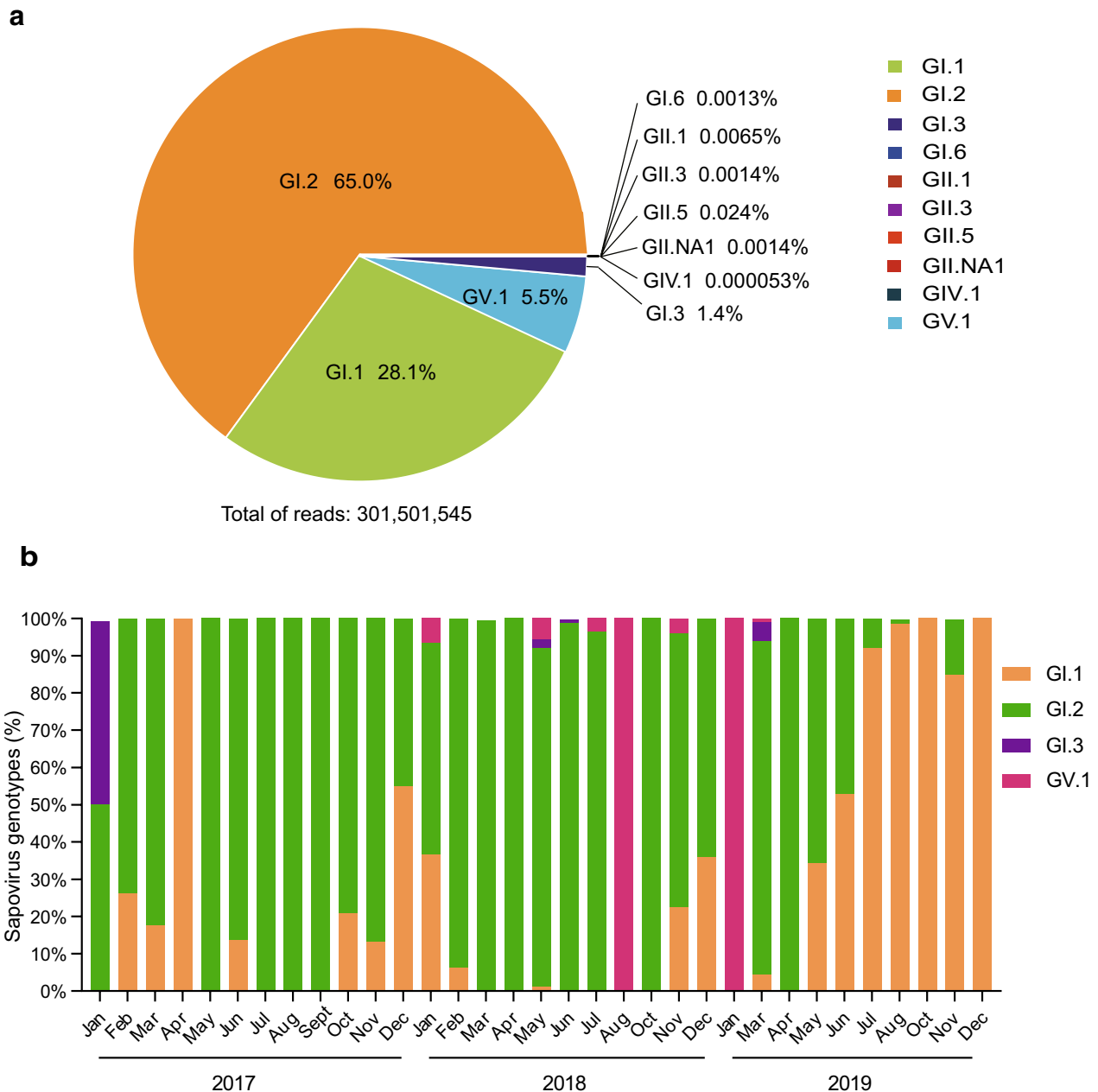


samples and the concentration of PMMoV in sewage samples ranged from  $1.26 \times 10^7$  to  $7.44 \times 10^8$  genome copies per liter, which was relatively stable, suggesting the reliability of genome quantification in this study.

### NGS-Based Amplicon Sequencing and Genotypes

Nested RT-PCR and NGS analysis showed that human SaV sequences in sewage were classified into 10 genotypes

belonging to 4 genogroups (GI, GII, GIV and GV). Of the total 301,501,545 reads, 195,852,867 reads were aligned to GI.2 reference sequences (65.0%), followed by 84,714,998 reads to GI.1 (28.1%), 16,624,161 reads to GV.1 (5.5%), and 4,205,026 reads to GI.3 (1.4%) (Fig. 2b). In addition, some rare genotypes (<0.05% of monthly identified sequences) were detected in this study, including GII.5 (72,332 reads, 0.024%), GII.1 (19,668 reads, 0.0065%), GII.NA1 (4,239 reads, 0.0014%), GII.3 (4,196 reads, 0.0014%), GI.6 (3,897



**Fig. 2** Sapovirus genotype distribution in sewage samples from NGS in three years (a). Monthly distribution of major genotypes (>1% of monthly identified sequences) in raw sewage in Jinan, China from 2017 to 2019 (b)

reads, 0.0013%), and GIV.1 (161 reads, 0.000053%). (Table 1, Fig. 2a).

## Major and Minor Genotypes

It was found that GI.2 and GI.1 were the most predominant genotypes in sewage in Jinan city. Among the 33 samples which were forwarded to NGS, GI.2 nucleic acid and GI.1 nucleic acid were detected in 30 (90.91%) and 20 samples (60.61%), respectively. A switch of predominant human SaV genotype in sewage was observed during the study period. Before May 2019, GI.2 was the predominant genotype of most months accounting for 76.28 % of total reads. After July 2019, however, the main genotype had been changed to GI.1 which accounted for 94.67% of total reads (Fig. 2b). Some genotypes appeared only in several months. For example, 10 out of 33 samples were positive for both GII.3 and GII.5 nucleic acid (30.30%). GII.NA1 nucleic acid was found only in 4 samples (12.12%). GI.5 and GIV.1 nucleic acid were only found in the samples collected in May 2019 and June 2019, respectively (3.03%) (Table 1). Compared to the common genotypes, the reads of these rare genotypes were not only detected in fewer months, but also with smaller number.

## Diversity

Generally, multiple genotypes were coexisting in sewage in Jinan during the study period. The Simpson's diversity index was calculated to analyze the richness and evenness of human SaV in China. It ranged from 0 to 0.539, which implied the variable diversity of genotypes in the population during three years. The Simpson's diversity index was generally consistent with the number of genotypes except for some samples in several months (Fig. 3).

## Homology and Phylogeny

The phylogenetic tree based on SaV partial VP1 sequences was constructed to investigate the relationship between strains obtained in this study and those detected from human feces, wastewater, and shellfishes throughout the world (Fig. 4). The nucleotide sequences of strains in the present study were close to reference strains around the world. The phylogenetic tree was grouped into 10 main clusters. The homology of the nucleotide sequences in this study ranged from 87.5% to 100% for GI.1 and from 87.9 to 100% for GI.2. As to GII.5 nucleotide sequences obtained in this study, they had high homology (93–98.5%) with the reference strain from Guatemala and Germany. For the newly detected genotype, GII.NA1, identity ranged from 87.9 to 94.5% compared with that isolated from Kenya and Cameroon. Moreover, we detected only one GIV.1 sequence in

this study, and it shared high homology (99.4%) with the reference strain from China in 2008 and Japan in 2011.

## Discussion

With the development of molecular techniques, human SaV could be detected with highly diagnostic efficacy to investigate its prevalence among population. Sporadic and outbreaks related to human SaV have been reported in Asia (Thongprachum et al. 2018; Kitajima et al. 2010), Europe (Mancini et al. 2019), North America (Kitajima et al. 2018), and South Africa (Ibrahim et al. 2019; Murray et al. 2013). However, only a few studies in China have investigated human SaV in clinical samples, providing limited information on genetic and genotype diversities of human SaV in local population (Wang et al. 2014). The results of this study are useful in understanding human SaV circulation in China.

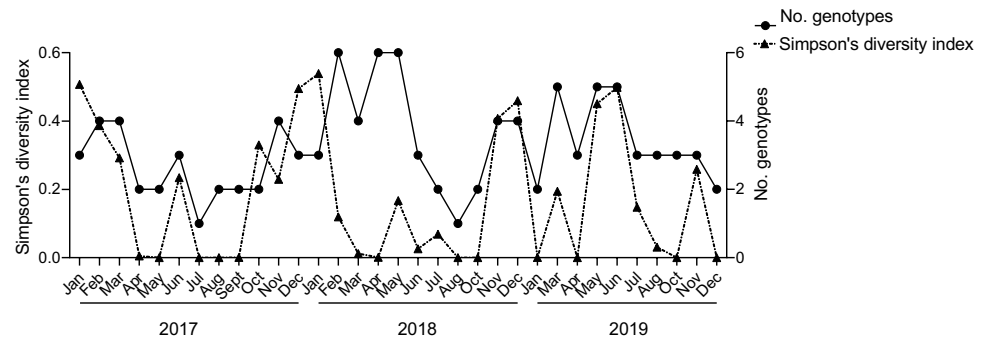
Human SaV can be discharged into the environment via sewage effluents, where it can remain infective persistently (Sinclair et al. 2008). It can be transmitted by fecal-oral route, especially feces-contaminated water. There have been studies proving that asymptomatic patient shed viruses at levels comparable to those shed by gastroenteritis patients (Kobayashi et al. 2012; Yoshida et al. 2009). Environmental surveillance has the advantages of sensitivity, wide representative scopes, and good correlation with the presence of viruses in population, which can be used as a strong supplement to clinical surveillance (Iwai et al. 2009; Ozawa et al. 2019). Our findings show a high detection rate of 97.22% with detection of multiple genotypes, reflect continuous circulation of human SaV among local population, and reveal high sensitivity and importance of environmental surveillance in monitoring enteric viruses.

Traditionally, environmental surveillance is performed using Sanger sequencing technology. However, it has some defects in that it can only detect major genotypes in the mixed pool. On the contrary, NGS enhances the understanding of genetic diversity for it can recognize rare genotypes which were concealed by major ones when using Sanger sequencing. For example, a study from Italy showed that NGS revealed 3 additional genotypes (GI.6, GII.6 and GV.1) beyond the 4 (GI.1 GI.2 GI.3 and GII.1) detected by Sanger sequencing (Mancini et al. 2019). In this study, NGS was performed with 33 sewage samples which were positive for nested RT-PCR during the three-year period. Ten genotypes were identified, including 4 major genotypes (> 1% of monthly identified sequences) and 6 rare genotypes (< 0.05% of monthly identified sequences), demonstrating NGS-based amplicon sequencing is an effective approach in analyzing complicated samples. The existence of extremely rare genotypes was also observed in similar studies on enterovirus and norovirus (Tao et al. 2020; Fumian et al. 2019).

**Table 1** The number of sapovirus reads detected in raw sewage, by month and by genotype

Geno- type	2018												2019												Viral detection frequency (%)	
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Oct	Nov	Dec			
GI.1	2.9E+06	1.9E+06	1.9E+06	9.7E+06	1.3E+06	7.9E+06	9.2E+06	7.9E+06	8.5E+06	4.4E+05	1.2E+06	4.7E+06	3.7E+06	6.7E+06	8.6E+06	9.7E+06	9.7E+06	7.9E+06	7.9E+06	7.2E+06	5.1E+06	8.4E+06	9.2E+06	1.2E+06	20/33	
GI.2	3.2E+06	8.1E+06	9.1E+06	2.6E+06	8.3E+06	8.5E+06	9.2E+06	8.5E+06	1.7E+06	7.8E+06	3.8E+06	5.8E+06	1.0E+06	6.3E+06	7.7E+06	9.7E+06	9.5E+06	9.0E+06	8.8E+06	7.1E+06	4.2E+06	1.2E+06	1.1E+06	7.9E+06	4.4E+06	30/33
GI.3	3.2E+06	1.9E+06	1.9E+06	7.3E+06	7.3E+06	7.1E+06	9.7E+06	1.9E+06	1.1E+06	7.1E+06	9.7E+06	1.9E+06	1.1E+06	9.7E+06	1.9E+06	1.1E+06	9.7E+06	8.2E+06	9.3E+06	9.3E+06	6.9E+06	6.9E+06	4.8E+06	3.4E+06	12/33	
GI.6																										(36.36)
GI.1			1.7E+02																							1/33
GI.3																										(3.03)
GI.5																										8/33
GI.1																										(24.24)
GI.3																										(30.30)
GI.5																										(30.30)
GI.1																										(12.12)
NA1																										4/33
GV1																										1/33
GV1																										(3.03)
GV1																										(33.33)

**Fig. 3** The Simpson's diversity index and the number of genotypes monthly in three years



During the process from nested PCR to NGS, there may be some factors influencing the proportion of nucleic acids from different genotypes. Thus, it is reasonable to conclude that constitution of the number of NGS reads cannot fully reflect the nucleic acid composition in the original sample.

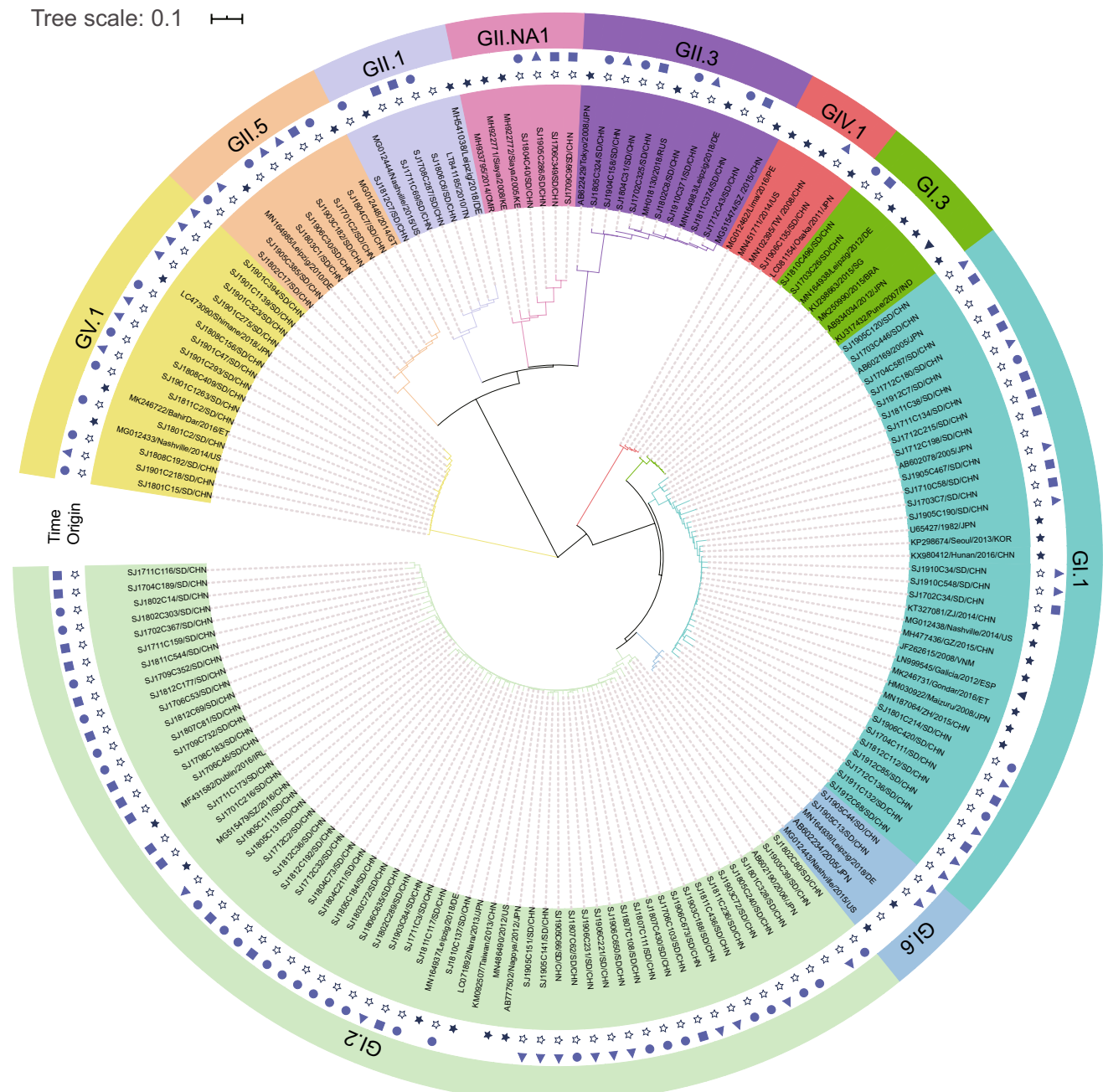
Previously, several studies in China have investigated human SaV genotypes in clinical samples. Wang et al. detected human SaV in 42/1,125 (3.73%) samples collected from adult outpatients with acute gastroenteritis in Shanghai, China from April 2011 to March 2013 (Wang et al. 2014), and GI.2 was the most predominant genotype (78.5%; 33/42). Subsequently, Xue et al. detected SaV in 11/569 (1.93%) fecal samples from acute diarrhea patients in south China from 2013 to 2017, and GI was positive in 9 samples (Xue et al. 2019). In keeping with previous studies (Makhaola et al. 2020), GI was the most prevalent genotype, detected in 32 out of 33 samples, and GI.2 was the most prevalent genotype, reflecting its high activity in local population during the study period.

Previous studies from the USA (Kitajima et al. 2018) and Japan (Harada et al. 2013; Harada et al. 2012; Harada et al. 2009) had observed dynamic changes of human SaV genotypes. Similarly, a switch of major genotype from GI.2 to GI.1 was observed around June 2019 in this study. This phenomenon might result from the changes of population immunity levels or the changes of infectivity, as needs further investigation. In addition, GV.1 was a rarely detected genotype in China. However, it was the most prevalent genotype in two months during the study period (Fig. 2b), suggesting high activity at that time. Moreover, we identified two genotypes, GII.NA1 and GII.5, to the best of our knowledge, these two genotypes had not been reported before in China, demonstrating the high sensitivity of NGS. GII.5 has been described in a food-borne gastroenteritis outbreak among adults in Japan (Oka et al. 2017), in pediatric patients with acute gastroenteritis in Thailand (Kumthip et al. 2020), and in children younger than 5 years of age in Guatemala (M. Diez-Valcarce et al. 2019a, b). According to the phylogenetic analysis, GII.NA1 strains we obtained in this study were closely related to the strains detected in human stool samples in Kenya in 2005 and 2008 (Marta Diez-Valcarce et al.

2019a, b) and in Cameroon in 2014 (Yinda et al. 2019). Consecutive clinical and environmental surveillance is needed to provide meaningful information for understanding the prevalence and pathogenicity of these two genotypes in China in the future.

Although PMMoV belongs to plant viruses, it is one of the most abundant virus types in a metagenomic survey of RNA viruses from human feces (Zhang et al. 2006). The abundance of PMMoV in stool samples does not depend on its infection status in human and will not change seasonally (Haramoto et al. 2013). Therefore, it is a potential indicator of pollution degree of water by human feces (Kuroda et al. 2015; Malla et al. 2019). PMMoV was used as an indicator to evaluate the accuracy of quantified values of human SaV genome in sewage. Significant levels of PMMoV were detected in all sewage samples, indicating that the observation of samples with low or no human SaV detection were not due to the presence of PCR inhibitors in the quantification process. No statistically significant seasonality of viral concentration was observed in the present study, similar with a study from the USA which detected no clear seasonality pattern over one-year period (Kitajima et al. 2014). However, studies in Brazil showed the existence of seasonal differences in virus concentration. Among four seasons, higher rate of human SaV from wastewater was observed in rainy seasons (summer and autumn) (Fioretti et al. 2016). The difference of dynamics of the human SaV may occur due to different continental dimensions. The viral concentration in this study ranges from  $10^3$  to  $10^5$  genome copies per liter, which is in accordance with that in Brazil, Japan, and the USA (Fioretti et al. 2016; Kitajima et al. 2014; Haramoto et al. 2007), suggesting Jinan is also an endemic area of human SaV. Among GI, GI.1 was likely to be more prevalent during cold seasons, in line with studies reporting a higher positive proportion of clinical samples in winter (Varela et al. 2019).

There may be some limitations in this study. We only detected human SaV in sewage supernate, ignoring the sewage sludge, which is needed to be investigated in the future. Also, all samples tested in this study were collected from raw urban sewage, lacking stool samples from gastroenteritis patients. The actual prevalence of human SaV infection



**Fig. 4** Phylogenetic tree of sapovirus strains based on partial VP1 nucleotide sequences (nt position 5179–5571 corresponding to strain Hu/SaV/Manchester/1993/UK with accession number X86560). The tree was generated using the Neighbor-Joining method with Kimura 2-parameter model and gamma-distributed rates in MegaX, with representative strains derived from sewage in Jinan, China, and reference

strains from GenBank. The time of sequences detected in present study is characterized as different shapes: (filled square) for 2017; (filled circle) for 2018; and (filled triangle) for 2019. The origin of sequences in the tree was presented in other shapes: (filled star) from human stool; (open star) from sewage; and (filled inverted triangle) from shellfishes

among people could not be estimated exactly. Further studies are needed to test human SaV from both clinical and environmental specimens to acquire a comprehensive understanding of human SaV.

In conclusion, this study provided a comprehensive picture of genotypes and genetic characterization of human

SaV in sewage in Jinan, China by NGS-based environmental surveillance, which greatly improves our understanding on human SaV circulation in communities. NGS should be encouraged as a sensitive surveillance tool in the future.



**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12560-021-09469-x>.

**Acknowledgements** This study was supported by a Grant from the National Natural Science Foundation of China (81573209), a Grant from Taishan Scholar Program (ts201511105), and a Grant from Young Experts of Taishan Scholar Program.

**Author Contributions** Conceptualization: ZT, YS, and AX. Formal analysis: KS. Funding acquisition: AX, ZT, and YS. Investigation: XL, FJ, and CZ. Methodology: AX, ZT, and YS. Project administration: YS and ZT. Resources: YL, FJ, CZ, and PC. Supervision: AX, ZT, PC, and YS. Writing-original draft: KS. Writing-review & editing: ZT and KS.

**Date Availability** Sequences determined in this study were submitted to GenBank database under the Accession Number MW245378 to MW245652. The individual SRA codes are SRX9772850-SRX9772882 and are deposited in the NCBI SRA archive as BioProject PRJNA689353.

## Compliance with Ethical Standards

**Conflict of interest** We declare that we have no known conflict of financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J. W., et al. (2020). First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: A proof of concept for the wastewater surveillance of COVID-19 in the community. *Science of the Total Environment*, 728, 138764. <https://doi.org/10.1016/j.scitotenv.2020.138764>
- Berg, G., Dahling, D. R., & Berman, D. (1971). Recovery of small quantities of viruses from clean waters on cellulose nitrate membrane filters. *Applied Microbiology*, 22(4), 608–614.
- de Wit, M. A., Koopmans, M. P., Kortbeek, L. M., Wannet, W. J., Vinjé, J., van Leusden, F., et al. (2001). Sensor, a population-based cohort study on gastroenteritis in the Netherlands: Incidence and etiology. *American Journal of Epidemiology*, 154(7), 666–674. <https://doi.org/10.1093/aje/k154.7.666>
- Diez-Valcarce, M., Lopez, M. R., Lopez, B., Morales, O., Sagastume, M., Cadena, L., et al. (2019). Prevalence and genetic diversity of viral gastroenteritis viruses in children younger than 5 years of age in Guatemala, 2014–2015. *Journal of Clinical Virology*, 114, 6–11. <https://doi.org/10.1016/j.jcv.2019.03.006>
- Diez-Valcarce, M., Montmayeur, A., Tatusov, R., & Vinje, J. (2019). Near-complete human sapovirus genome sequences from Kenya. *Microbiology Resource Announcements*. <https://doi.org/10.1128/mra.01602-18>.
- Fioretti, J. M., Rocha, M. S., Fumian, T. M., Ginuino, A., da Silva, T. P., de Assis, M. R., et al. (2016). Occurrence of human sapoviruses in wastewater and stool samples in Rio De Janeiro, Brazil. *Journal of Applied Microbiology*, 121(3), 855–862. <https://doi.org/10.1111/jam.13205>
- Fumian, T. M., Fioretti, J. M., Lun, J. H., Dos Santos, I. A. L., White, P. A., & Miagostovich, M. P. (2019). Detection of norovirus epidemic genotypes in raw sewage using next generation sequencing. *Environment International*, 123, 282–291. <https://doi.org/10.1016/j.envint.2018.11.054>
- Harada, S., Oka, T., Tokuoka, E., Kiyota, N., Nishimura, K., Shimada, Y., et al. (2012). A confirmation of sapovirus re-infection gastroenteritis cases with different genogroups and genetic shifts in the evolving sapovirus genotypes, 2002–2011. *Archives of Virology*, 157(10), 1999–2003. <https://doi.org/10.1007/s00705-012-1387-7>
- Harada, S., Okada, M., Yahiro, S., Nishimura, K., Matsuo, S., Miyasaka, J., et al. (2009). Surveillance of pathogens in outpatients with gastroenteritis and characterization of sapovirus strains between 2002 and 2007 in Kumamoto Prefecture, Japan. *Journal of Medical Virology*, 81(6), 1117–1127. <https://doi.org/10.1002/jmv.21454>
- Harada, S., Tokuoka, E., Kiyota, N., Katayama, K., & Oka, T. (2013). Phylogenetic analysis of the nonstructural and structural protein encoding region sequences, indicating successive appearance of genomically diverse sapovirus strains from gastroenteritis patients. *Japanese Journal of Infectious Diseases*, 66(5), 454–457. <https://doi.org/10.7883/yoken.66.454>
- Haramoto, E., Katayama, H., Oguma, K., & Ohgaki, S. (2007). Quantitative analysis of human enteric adenoviruses in aquatic environments. *Journal of Applied Microbiology*, 103(6), 2153–2159. <https://doi.org/10.1111/j.1365-2672.2007.03453.x>
- Haramoto, E., Kitajima, M., Kishida, N., Konno, Y., Katayama, H., Asami, M., et al. (2013). Occurrence of pepper mild mottle virus in drinking water sources in Japan. *Applied and Environmental Microbiology*, 79(23), 7413–7418. <https://doi.org/10.1128/aem.02354-13>
- Hergens, M. P., Nederby Öhd, J., Alm, E., Askling, H. H., Helgesson, S., Insulander, M., et al. (2017). Investigation of a food-borne outbreak of gastroenteritis in a school canteen revealed a variant of sapovirus genogroup V not detected by standard PCR, Sollentuna, Sweden, 2016. *Eurosurveillance*. <https://doi.org/10.2807/1560-7917.Es.2017.22.22.30543>.
- Iaconelli, M., Valdazo-Gonzalez, B., Equestre, M., Ciccaglione, A. R., Marcantonio, C., Della Libera, S., et al. (2017). Molecular characterization of human adenoviruses in urban wastewaters using next generation and Sanger sequencing. *Water Research*, 121, 240–247. <https://doi.org/10.1016/j.watres.2017.05.039>
- Ibrahim, C., Hammami, S., Cherif, N., Mejri, S., Pothier, P., & Hasen, A. (2019). Detection of Sapoviruses in two biological lines of Tunisian hospital wastewater treatment. *International Journal of Environmental Health Research*, 29(4), 400–413. <https://doi.org/10.1080/09603123.2018.1546835>
- Iwai, M., Hasegawa, S., Obara, M., Nakamura, K., Horimoto, E., Takizawa, T., et al. (2009). Continuous presence of noroviruses and sapoviruses in raw sewage reflects infections among inhabitants of Toyama, Japan (2006 to 2008). *Applied and Environmental Microbiology*, 75(5), 1264–1270. <https://doi.org/10.1128/aem.01166-08>
- Kagning Tsinda, E., Malasao, R., Furuse, Y., Gilman, R. H., Liu, X., Apaza, S., et al. (2017). Complete coding genome sequences of uncommon GII.8 sapovirus strains identified in diarrhea samples collected from Peruvian children. *Genome Announc*. <https://doi.org/10.1128/genomeA.01137-17>.
- Khamrin, P., Kumthip, K., Thongprachum, A., Sirilert, S., Malasao, R., Okitsu, S., et al. (2020). Genetic diversity of norovirus genogroup I, II, IV and sapovirus in environmental water in Thailand. *Journal of Infection and Public Health*. <https://doi.org/10.1016/j.jiph.2020.05.004>.
- Kitajima, M., Iker, B. C., Pepper, I. L., & Gerba, C. P. (2014). Relative abundance and treatment reduction of viruses during wastewater treatment processes—identification of potential viral indicators. *Science of the Total Environment*, 488–489, 290–296. <https://doi.org/10.1016/j.scitotenv.2014.04.087>
- Kitajima, M., Oka, T., Haramoto, E., Katayama, H., Takeda, N., Katayama, K., et al. (2010). Detection and genetic analysis of human sapoviruses in river water in Japan. *Applied and Environmental*

- Microbiology*, 76(8), 2461–2467. <https://doi.org/10.1128/AEM.02739-09>
- Kitajima, M., Rachmadi, A. T., Iker, B. C., Haramoto, E., & Gerba, C. P. (2018). Temporal variations in genotype distribution of human sapoviruses and Aichi virus 1 in wastewater in Southern Arizona, United States. *Journal of Applied Microbiology*, 124(5), 1324–1332. <https://doi.org/10.1111/jam.13712>
- Kitamura, K., Sadamasu, K., Muramatsu, M., & Yoshida, H. (2020). Efficient detection of SARS-CoV-2 RNA in the solid fraction of wastewater. *Science of the Total Environment*, 763, 144587. <https://doi.org/10.1016/j.scitotenv.2020.144587>
- Kobayashi, S., Fujiwara, N., Yasui, Y., Yamashita, T., Hiramatsu, R., & Minagawa, H. (2012). A foodborne outbreak of sapovirus linked to catered box lunches in Japan. *Archives of Virology*, 157(10), 1995–1997. <https://doi.org/10.1007/s00705-012-1394-8>
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33(7), 1870–1874. <https://doi.org/https://doi.org/10.1093/molbev/msw054>
- Kumthip, K., Khamrin, P., Ushijima, H., Chen, L., Li, S., & Maneekarn, N. (2020). Genetic recombination and diversity of sapovirus in pediatric patients with acute gastroenteritis in Thailand, 2010–2018. *PeerJ*. <https://doi.org/10.7717/peerj.8520>
- Kuroda, K., Nakada, N., Hanamoto, S., Inaba, M., Katayama, H., Do, A. T., et al. (2015). Pepper mild mottle virus as an indicator and a tracer of fecal pollution in water environments: Comparative evaluation with wastewater-tracer pharmaceuticals in Hanoi, Vietnam. *Science of the Total Environment*, 506–507, 287–298. <https://doi.org/10.1016/j.scitotenv.2014.11.021>
- Lee, L. E., Cebelinski, E. A., Fuller, C., Keene, W. E., Smith, K., Vinjé, J., et al. (2012). Sapovirus outbreaks in long-term care facilities, Oregon and Minnesota, USA, 2002–2009. *Emerging Infectious Diseases*, 18(5), 873–876. <https://doi.org/10.3201/eid1805.111843>
- Liu, X., Jahuir, H., Gilman, R. H., Alva, A., Cabrera, L., Okamoto, M., et al. (2016). Etiological role and repeated infections of sapovirus among children aged less than 2 years in a cohort study in a Peri-urban community of Peru. *Journal of Clinical Microbiology*, 54(6), 1598–1604. <https://doi.org/10.1128/JCM.03133-15>
- Majumdar, M., & Martin, J. (2018). Detection by direct next generation sequencing analysis of emerging enterovirus D68 and C109 strains in an environmental sample from Scotland. *Frontiers in Microbiology*, 9, 1956. <https://doi.org/10.3389/fmicb.2018.01956>
- Makhaola, K., Moyo, S., & Kebaabetswe, L. P. (2020). Distribution and genetic variability of sapoviruses in Africa. *Viruses*. <https://doi.org/https://doi.org/10.3390/v12050490>
- Malla, B., Ghaju Shrestha, R., Tandukar, S., Sherchand, J. B., & Haramoto, E. (2019). Performance evaluation of human-specific viral markers and application of pepper mild mottle virus and CrAssphage to environmental water samples as fecal pollution markers in the Kathmandu Valley, Nepal. *Food and Environmental Virology*, 11(3), 274–287. <https://doi.org/10.1007/s12560-019-09389-x>
- Mancini, P., Bonanno Ferraro, G., Iaconelli, M., Suffredini, E., Valdazo-Gonzalez, B., Della Libera, S., et al. (2019). Molecular characterization of human Sapovirus in untreated sewage in Italy by amplicon-based Sanger and next-generation sequencing. *Journal of Applied Microbiology*, 126(1), 324–331. <https://doi.org/10.1111/jam.14129>
- Matsuura, K., Hasegawa, S., Nakayama, T., Morita, O., & Uetake, H. (1984). Viral pollution of the rivers in Toyama City. *Microbiology and Immunology*, 28(5), 575–588. <https://doi.org/10.1111/j.1348-0421.1984.tb00710.x>
- Medici, M. C., Tummolo, F., Albonetti, V., Abelli, L. A., Chezzi, C., & Calderaro, A. (2012). Molecular detection and epidemiology of astrovirus, bocavirus, and sapovirus in Italian children admitted to hospital with acute gastroenteritis, 2008–2009. *Journal of Medical Virology*, 84(4), 643–650. <https://doi.org/10.1002/jmv.23231>
- Montmayeur, A. M., Ng, T. F., Schmidt, A., Zhao, K., Magana, L., Iber, J., et al. (2017). High-throughput next-generation sequencing of polioviruses. *Journal of Clinical Microbiology*, 55(2), 606–615. <https://doi.org/10.1128/JCM.02121-16>
- Murray, T. Y., Mans, J., & Taylor, M. B. (2013). Human calicivirus diversity in wastewater in South Africa. *Journal of Applied Microbiology*, 114(6), 1843–1853. <https://doi.org/10.1111/jam.12167>
- Oka, T., Doan, Y. H., Haga, K., Mori, K., Ogawa, T., & Yamazaki, A. (2017). Genetic characterization of rare genotype GII.5 sapovirus strain detected from a suspected food-borne gastroenteritis outbreak among adults in Japan in 2010. *Japanese Journal of Infectious Diseases*, 70(2), 223–224. <https://doi.org/https://doi.org/10.7883/yoken.JIID.2016.468>
- Oka, T., Katayama, K., Hansman, G. S., Kageyama, T., Ogawa, S., Wu, F. T., et al. (2006). Detection of human sapovirus by real-time reverse transcription-polymerase chain reaction. *Journal of Medical Virology*, 78(10), 1347–1353. <https://doi.org/10.1002/jmv.20699>
- Oka, T., Wang, Q., Katayama, K., & Saif, L. J. (2015). Comprehensive review of human sapoviruses. *Clinical Microbiology Reviews*, 28(1), 32–53. <https://doi.org/10.1128/CMR.00011-14>
- Ozawa, H., Yoshida, H., & Usuku, S. (2019). Environmental surveillance can dynamically track ecological changes in enteroviruses. *Applied and Environmental Microbiology*. <https://doi.org/https://doi.org/10.1128/aem.01604-19>
- Page, N., Groome, M. J., Murray, T., Nadan, S., Netshikweta, R., Keddy, K. H., et al. (2016). Sapovirus prevalence in children less than five years of age hospitalised for diarrhoeal disease in South Africa, 2009–2013. *Journal of Clinical Microbiology*, 78, 82–88. <https://doi.org/10.1016/j.jcv.2016.03.013>
- Pang, X. L., Lee, B. E., Tyrrell, G. J., & Preiksaitis, J. K. (2009). Epidemiology and genotype analysis of sapovirus associated with gastroenteritis outbreaks in Alberta, Canada: 2004–2007. *The Journal of Infectious Diseases*, 199(4), 547–551. <https://doi.org/10.1086/596210>
- Pang, X. L., Preiksaitis, J. K., & Lee, B. E. (2014). Enhanced enteric virus detection in sporadic gastroenteritis using a multi-target real-time PCR panel: A one-year study. *Journal of Medical Virology*, 86(9), 1594–1601. <https://doi.org/10.1002/jmv.23851>
- Platts-Mills, J. A., Liu, J., Rogawski, E. T., Kabir, F., Lertsethtakarn, P., Sigua, M., et al. (2018). Use of quantitative molecular diagnostic methods to assess the aetiology, burden, and clinical characteristics of diarrhoea in children in low-resource settings: A reanalysis of the MAL-ED cohort study. *Lancet Global Health*, 6(12), e1309–e1318. [https://doi.org/10.1016/s2214-109x\(18\)30349-8](https://doi.org/10.1016/s2214-109x(18)30349-8)
- Rockx, B., De Wit, M., Vennema, H., Vinjé, J., De Bruin, E., Van Duynhoven, Y., et al. (2002). Natural history of human calicivirus infection: A prospective cohort study. *Clinical Infectious Diseases*, 35(3), 246–253. <https://doi.org/10.1086/341408>
- Sakai, Y., Nakata, S., Honma, S., Tatsumi, M., Numata-Kinoshita, K., & Chiba, S. (2001). Clinical severity of Norwalk virus and Sapporo virus gastroenteritis in children in Hokkaido, Japan. *The Pediatric Infectious Disease Journal*, 20(9), 849–853. <https://doi.org/10.1097/00006454-200109000-00005>
- Sinclair, R. G., Choi, C. Y., Riley, M. R., & Gerba, C. P. (2008). Pathogen surveillance through monitoring of sewer systems. *Advances in Applied Microbiology*, 65, 249–269
- Suffredini, E., Iaconelli, M., Equestre, M., Valdazo-González, B., Ciccaglione, A. R., Marcantonio, C., et al. (2018). Correction to: Genetic diversity among genogroup II noroviruses and progressive emergence of GII.17 in Wastewaters in Italy (2011–2016) revealed by next-generation and sanger sequencing. *Food Environ Virol*, 10(2), 222–223. <https://doi.org/https://doi.org/10.1007/s12560-018-9346-4>

- Tao, Z., Chen, P., Cui, N., Lin, X., Ji, F., Liu, Y., et al. (2020). Detection of enteroviruses in urban sewage by next generation sequencing and its application in environmental surveillance. *Science of The Total Environment*. <https://doi.org/10.1016/j.scitotenv.2020.138818>
- Thongprachum, A., Fujimoto, T., Takanashi, S., Saito, H., Okitsu, S., Shimizu, H., et al. (2018). Detection of nineteen enteric viruses in raw sewage in Japan. *Infection Genetics and Evolution*, 63, 17–23. <https://doi.org/10.1016/j.meegid.2018.05.006>
- Usuku, S., Kumazaki, M., Kitamura, K., Tochikubo, O., & Noguchi, Y. (2008). An outbreak of food-borne gastroenteritis due to sapovirus among junior high school students. *Japanese Journal of Infectious Diseases*, 61(6), 438–441.
- Varela, M. F., Rivadulla, E., Lema, A., & Romalde, J. L. (2019). Human sapovirus among outpatients with acute gastroenteritis in Spain: A one-year study. *Viruses*. <https://doi.org/https://doi.org/10.3390/v11020144>.
- Wang, G., Shen, Z., Qian, F., Li, Y., Yuan, Z., & Zhang, J. (2014). Genetic diversity of sapovirus in non-hospitalized adults with sporadic cases of acute gastroenteritis in Shanghai, China. *Journal of Clinical Virology*, 59(4), 250–254. <https://doi.org/10.1016/j.jcv.2014.01.007>
- Wang, J., Li, Y., Kong, X., Li, H., Zhang Ba, Q., Jin, M., et al. (2018). Two gastroenteritis outbreaks caused by sapovirus in Shenzhen, China. *Journal of Medical Virology*, 90(11), 1695–1702. <https://doi.org/https://doi.org/10.1002/jmv.25236>.
- Xue, L., Cai, W., Zhang, L., Gao, J., Dong, R., Li, Y., et al. (2019). Prevalence and genetic diversity of human sapovirus associated with sporadic acute gastroenteritis in South China from 2013 to 2017. *Journal of Medical Virology*, 91(10), 1759–1764. <https://doi.org/10.1002/jmv.25511>
- Yamashita, Y., Ootsuka, Y., Kondo, R., Oseto, M., Doi, M., Miyamoto, T., et al. (2010). Molecular characterization of Sapovirus detected in a gastroenteritis outbreak at a wedding hall. *Journal of Medical Virology*, 82(4), 720–726. <https://doi.org/10.1002/jmv.21646>
- Yan, H., Abe, T., Phan, T. G., Nguyen, T. A., Iso, T., Ikezawa, Y., et al. (2005). Outbreak of acute gastroenteritis associated with group A rotavirus and genogroup I sapovirus among adults in a mental health care facility in Japan. *Journal of Medical Virology*, 75(3), 475–481. <https://doi.org/10.1002/jmv.20292>
- Yinda, C. K., Vanhulle, E., Conceição-Neto, N., Beller, L., Deboutte, W., Shi, C., et al. (2019). Gut virome analysis of Cameroonians reveals high diversity of enteric viruses, including potential interspecies transmitted viruses. *MSphere*. <https://doi.org/https://doi.org/10.1128/mSphere.00585-18>.
- Yoshida, T., Kasuo, S., Azegami, Y., Uchiyama, Y., Satsumabayashi, K., Shiraishi, T., et al. (2009). Characterization of sapoviruses detected in gastroenteritis outbreaks and identification of asymptomatic adults with high viral load. *Journal of Clinical Virology*, 45(1), 67–71. <https://doi.org/10.1016/j.jcv.2009.03.003>
- Yu, Y., Guo, X. H., Yan, H. Q., Gao, Z. Y., Li, W. H., Liu, B. W., et al. (2019). Systematic review on the characteristics of acute gastroenteritis outbreaks caused by sapovirus. *Zhonghua Liu Xing Bing Xue Za Zhi*, 40(1), 93–98. <https://doi.org/10.3760/cma.j.issn.0254-6450.2019.01.019>
- Zhang, T., Breitbart, M., Lee, W. H., Run, J. Q., Wei, C. L., Soh, S. W., et al. (2006). RNA viral community in human feces: Prevalence of plant pathogenic viruses. *PLoS Biology*, 4(1), e3. <https://doi.org/10.1371/journal.pbio.0040003>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.