



# Viral metagenomics reveals sapoviruses of different genogroups in stool samples from children with acute gastroenteritis in Jiangsu, China

Wang Li<sup>1,2</sup> · Surong Dong<sup>1,2</sup> · Juan Xu<sup>1,2</sup> · Xiaobin Zhou<sup>1,2</sup> · Junling Han<sup>1,2</sup> · Zhaqing Xie<sup>1,2</sup> · Qin Gong<sup>1,2</sup> · Hailin Peng<sup>1,2</sup> · Chenglin Zhou<sup>1,2</sup>  · Mei Lin<sup>1,2</sup>

Received: 10 October 2019 / Accepted: 13 January 2020 / Published online: 11 February 2020

© The Author(s) 2020

## Abstract

Sapoviruses (SAVs), including several genogroups (GI to GV), are one of the causes of acute gastroenteritis (AGE). In this study, viral metagenomics revealed the presence of sapoviruses of different genogroups in stool from children with AGE. Eight different complete SAV genomes were determined, of which five belonged to GI and the other three belonged to GII, GIV and GV, respectively. Although they were highly similar to published sequences, the GIV and GV were the first complete genome sequences of these SAVs found in China. In a prevalence investigation, 19% of subjects with AGE were positive for SAVs, while none of the control group was positive.

Sapoviruses (SAVs) belong to the family *Caliciviridae* and cause both sporadic cases and occasional outbreaks of acute gastroenteritis (AGE) in humans and animals worldwide. Although all ages are affected, children younger than five years have the highest burden of disease. Recently, studies have shown that infections with SAVs can result in hospitalization and severe dehydration [1–3]. SAVs have increasingly been detected in low- and middle-income countries [4], but unfortunately, no licensed vaccines or antivirals are available for SAV infections. SAVs are genetically diverse viruses that can be classified into several genogroups (GI to GV) [5], four of which (I, II, IV, and V) infect humans. The existence of several additional SAV genogroups infecting different animal species has been shown [6]. Although SAV is known to cause acute gastroenteritis worldwide, the genetic characteristics of human SAVs in China are not known.

In 2016, 100 fecal samples were collected from children ( $\leq 6$  years old) with suspected viral AGE from the cities of Taizhou ( $n = 50$ ), Zhenjiang ( $n = 30$ ) and Xuzhou ( $n = 20$ ) in Jiangsu province. Antigen testing at the local hospitals for viral pathogens, including norovirus, astrovirus, and rotavirus, gave negative results. As controls, 40 fecal samples (20 from Taizhou, 10 from Zhenjiang, and 10 from Xuzhou) were collected in 2016 from healthy children ( $\leq 6$  years old) who had shown no symptoms of diarrhea within at least eight weeks before or after sample collection. Ethical approval was given by the Ethics Committee of Taizhou Hospital with the reference no. TZRMYY2015023.

To detect viral nucleic acids, 10% suspensions of fecal samples were prepared in phosphate-buffered saline and mixed thoroughly by vortexing. Fecal suspensions were then centrifuged for 10 min at  $15,000 \times g$ . One hundred  $\mu\text{L}$  of supernatant from each of the 140 samples was collected and combined into 14 pools based on sampling region and health status, each including 10 samples (Table 1). For viral metagenomic analysis, sample treatment, library construction, and data processing were done as described previously [7–9]. To investigate the sequence read distribution of different types of viruses detected in these pooled libraries, BLASTx results based on clean unassembled sequence reads from each pooled library were further analyzed in the MEGAN program, which yielded the viral read distribution in the 14 libraries.

Handling Editor: Reimar Johnne.

✉ Chenglin Zhou  
18762340015@126.com

✉ Mei Lin  
l\_mei@163.com

<sup>1</sup> Taizhou People's Hospital, 366 Taihu Road, Taizhou 225300, Jiangsu, China

<sup>2</sup> The Fifth Affiliated Hospital of Nantong University, Taizhou 225300, Jiangsu, China

**Table 1** Information about the sample pools included in the viral metagenomic analysis

Pool ID	Sample size	Group	Sampling site
Dia01	10	AGE	Taizhou
Dia02	10	AGE	Taizhou
Dia03	10	AGE	Taizhou
Dia04	10	AGE	Taizhou
Dia05	10	AGE	Taizhou
Dia06	10	AGE	Zhenjiang
Dia07	10	AGE	Zhenjiang
Dia08	10	AGE	Zhenjiang
Dia09	10	AGE	Xuzhou
Dia10	10	AGE	Xuzhou
Norm1	10	Normal	Taizhou
Norm2	10	Normal	Taizhou
Norm3	10	Normal	Zhenjiang
Norm4	10	Normal	Xuzhou

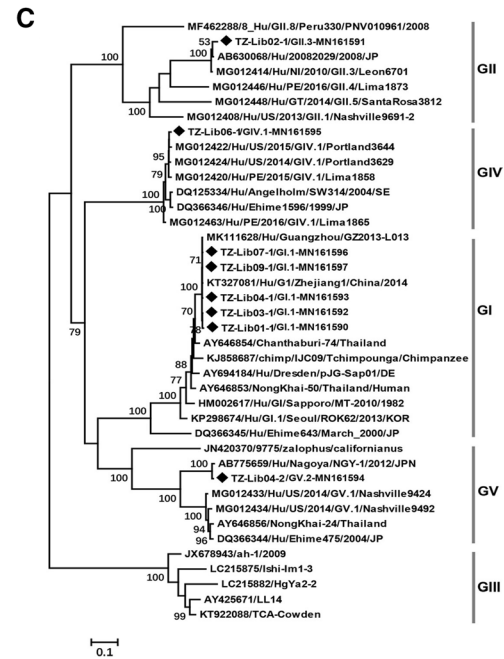
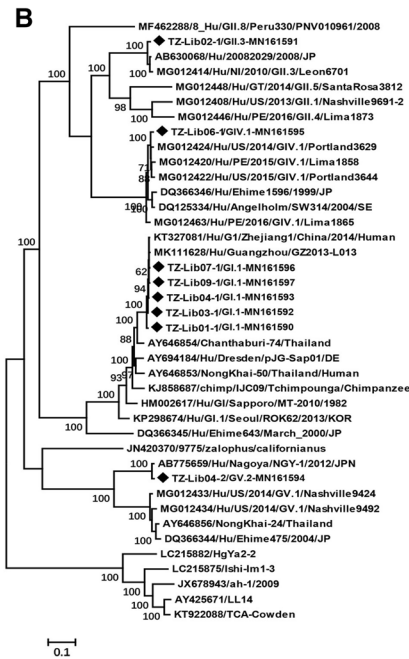
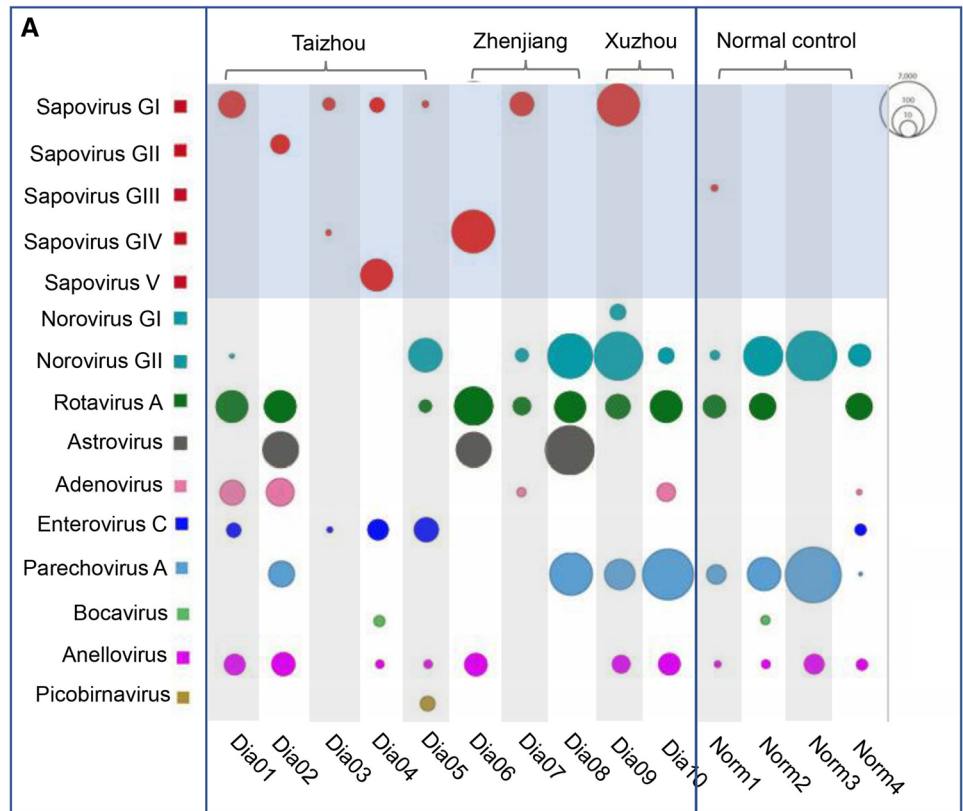
As shown in Fig. 1A, the viral sequences detected in the 14 libraries were compared to those of members of the families *Caliciviridae*, *Picornaviridae*, *Anelloviridae*, *Picornaviridae*, *Parvoviridae*, *Reoviridae*, *Astroviridae*, and *Adenoviridae*. Viruses of the families *Caliciviridae*, *Reoviridae*, *Astroviridae* and *Adenoviridae* are generally associated with AGE. Figure 1A also shows the sequence reads related to SAVs, which have rarely been reported to be associated with AGE in China. We detected SAV sequences in eight out of the ten pooled libraries from the AGE group, belonging to genogroups GI, GII, GIV and GV, while only a few reads of GIII, which is non-pathogenic for humans, were detected in the control group. From the eight SAV-positive pooled libraries, eight different complete SAV genome sequences were obtained by de novo assembly of the SAV reads in the individual pooled library, combined conventional PCR, which was used for bridging sequence gaps. Of the eight SAV-positive pooled libraries, one (Dia03) yielded too few SAV sequence reads to allow a complete genome sequence to be obtained, another (Dia04) yielded two different complete genome sequences, belonging to GI and GV, and the remaining six yielded one complete genome sequence. Of the eight complete SAV genomes identified, five belonged to GI and the other three belonged to GII, GIV, and GV, respectively.

Maximum-likelihood phylogenetic trees were constructed based on the nucleotide sequences of the complete genome and the VP1 region using MEGA 7.0 with 1,000 bootstrap resamplings of the alignment data. The analysis included the eight complete genome sequences determined in this study, their best matches found in a BLASTn search, and complete genome sequences of representatives of each genogroup. Both trees indicated that the five SAV GI strains clustered closely with SAVs belonging to the GI.1 genotype. The remaining three strains clustered within GII.3, GIV.1, and GV.2, respectively, and were closely related to their best BLASTn matches (Fig. 1B and C). Based on their complete genome sequences, the five GI strains shared 98.2–98.9% sequence identity to two previous Chinese SAV strains, KT327081 and MK111628, which were isolated from patients in Zhejiang province and Guangdong province, respectively. The GII strain showed 97.9% sequence identity to a Japanese SAV strain (AB630068). The GIV strain shared 98.1% sequence identity with an SAV isolate from the USA (MG012424). The remaining GV SAV genome sequences were 98.5% identical to another Japanese strain (AB775659) isolated from a foodborne gastroenteritis outbreak [10]. The GIV and GV sequences from this study were the first complete genome sequences of members of these genotypes obtained in China.

To investigate the prevalence of SAVs in the 140 individual fecal samples, universal primers capable of detecting human SVA genogroups I, II, IV and V as well as genogroup-specific primers were used for conventional PCR screening [11]. showed that 19% (19/100) Sanger sequencing of 420-bp PCR products of the screening assay of the samples from diarrheic children were positive for SAVs, 12 of which belonged to GI, two of which belonged to GII, three of which belonged to GIV, and two of which belonged to GV. None of the 40 samples from healthy children were positive for any of the four genogroups of SAV. The prevalence results were in accordance with the results from metagenomics.

In summary, we detected SAVs belonging to genotypes GI, GII, GIV and GV in fecal samples from children with AGE in China. Eight complete SAV genome sequences were determined, including the first SAV GIV and GV genome sequences reported in China. A prevalence investigation indicated that 19% of the children with AGE of unknown etiology were positive for SAV. Our data suggest that SAVs are important pathogens associated with AGE in children in Jiangsu, China.

**Fig. 1** Virome composition and phylogeny. **(A)** Virome composition in fecal samples from children with and without AGE. **(B)** Phylogenetic analysis based on the complete genome. **(C)** Phylogenetic analysis based on VP1 region



**Acknowledgements** C.Z., L.M., and W.L. conceived the study. W.L., S.D., J.X., X.Z., J.H., Z.X., Q.G., and H.P. performed research, and W.Z. analyzed the data. C.Z. wrote the manuscript, which was read, edited and approved by all authors.

**Funding** This work was partly supported by Taizhou Social Development Project No. TS201623, Project of Jiangsu Provincial Health and Family Planning Commission No. H2018072, Taizhou Technology Support Program No. SSF20170218, and Taizhou Technology Support Program No. TS201907.

**Data availability** The 14 sets of raw sequence reads were deposited in the Short Read Archive of the GenBank database with accession nos. SRX6651218-SRX6651231. The complete genome sequences of SAVs determined in this study were submitted to GenBank with the accession nos. MN161590- MN161597.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Svraka S, Vennema H, van der Veer B et al (2010) Epidemiology and genotype analysis of emerging sapovirus-associated infections across Europe. *J Clin Microbiol* 48:2191–2198. <https://doi.org/10.1128/JCM.02427-09>
2. Lee LE, Cebelinski EA, Fuller C et al (2012) Sapovirus outbreaks in long-term care facilities, Oregon and Minnesota, USA, 2002–2009. *Emerg Infect Dis* 18:873–876. <https://doi.org/10.3201/eid1805.111843>
3. Kobayashi S, Fujiwara N, Yasui Y et al (2012) A foodborne outbreak of sapovirus linked to catered box lunches in Japan. *Arch Virol* 157:1995–1997. <https://doi.org/10.1007/s00705-012-1394-8>
4. Magwalivha M, Kabue J-P, Traore AN, Potgieter N (2018) Prevalence of human sapovirus in low and middle income countries. *Adv Virol* 2018:1–12. <https://doi.org/10.1155/2018/5986549>
5. Diez-Valcarce M, Castro CJ, Marine RL et al (2018) Genetic diversity of human sapovirus across the Americas. *J Clin Virol* 104:65–72. <https://doi.org/10.1016/j.jcv.2018.05.003>
6. Oka T, Wang Q, Katayama K, Saif LJ (2015) Comprehensive review of human sapoviruses. *Clin Microbiol Rev* 28:32–53. <https://doi.org/10.1128/CMR.00011-14>
7. Zhang W, Li L, Deng X et al (2016) Viral nucleic acids in human plasma pools. *Transfusion* 56:2248–2255. <https://doi.org/10.1111/trf.13692>
8. Deng X, Naccache SN, Ng T et al (2015) An ensemble strategy that significantly improves de novo assembly of microbial genomes from metagenomic next-generation sequencing data. *Nucleic Acids Res*. <https://doi.org/10.1093/nar/gkv002>
9. Zhang W, Yang S, Shan T et al (2017) Virome comparisons in wild-diseased and healthy captive giant pandas. *Microbiome* 5:90. <https://doi.org/10.1186/s40168-017-0308-0>
10. Shibata S, Sekizuka T, Kodaira A, et al (2015) Complete Genome Sequence of a Novel GV.2 Sapovirus Strain, NGY-1, Detected from a Suspected Foodborne Gastroenteritis Outbreak. *Genome Announc* 3. 10.1128/genomeA.01553–14
11. Okada M, Yamashita Y, Oseto M, Shinozaki K (2006) The detection of human sapoviruses with universal and genogroup-specific primers. *Arch Virol* 151:2503–2509. <https://doi.org/10.1007/s00705-006-0820-1>

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