

Complete genome sequence of a sapovirus from a child in Zhejiang, China

Xiaohong Zhou¹ · Yi Sun² · Xiaochun Shang¹ · Jian Gao² · Xueqin Zhao¹ · Huiqun Shuai¹ · Rui Zhang¹ · Yanjun Zhang²

Received: 17 January 2016 / Accepted: 16 April 2016 / Published online: 28 April 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Although *Sapovirus* (*Caliciviridae*) has been accepted as one of the causes of acute gastroenteritis worldwide, little is known about the genetic characteristics of the whole genome of sapoviruses in China, especially those that infect humans. Here we report the complete genome sequence of a sapovirus strain, Human/Zhejiang1/2015/China, obtained from a child with acute gastroenteritis in Hangzhou, Zhejiang Province, China. Samples were collected and delivered to the CDC laboratories and were detected by RT-PCR. Sanger sequencing was used to obtain the full genome and molecular characterization of the genome was determined. A phylogenetic analysis of the genome was also performed. The results indicated that Human/Zhejiang1/2015/China belongs to Genogroup I. No recombination events were detected. This is the first complete sequence from a child to be reported in China. The sequence information is important for surveillance of this emerging gastrointestinal infection.

Keywords Sapovirus · Genome · Phylogeny

Edited by Hartmut Hengel.

Xiaohong Zhou, Yi Sun, and Xiaochun Shang contributed equally.

Electronic supplementary material The online version of this article (doi:10.1007/s11262-016-1343-9) contains supplementary material, which is available to authorized users.

✉ Yanjun Zhang
yjjzhang@cdc.zj.cn

¹ Xiacheng District Center for Disease Control and Prevention, Hangzhou, Zhejiang, China

² Zhejiang Provincial Center for Disease Control and Prevention, 3399 Binsheng Road, Hangzhou 310051, Zhejiang, China

Introduction

Sapovirus (SAV; in the family *Caliciviridae*), causes acute gastroenteritis in both humans and animals. Major clinical symptoms include diarrhea and vomiting with constitutional symptoms such as nausea, chills, and myalgia. Compared to Norovirus, SAV infections are less common and are known to cause disease primarily in children, usually those under the age of 5 years [1]. SAV has received wide attention in public health studies because SAV has recently been a causative agent of gastroenteritis in people of all ages in both confined outbreaks and in sporadic cases [1, 2]. It has also been reported in outbreaks in long-term care facilities for the elderly [3]. Mortality is rare and the symptoms are generally mild. No licensed vaccines or antivirals are available for SAV infections.

Many outbreaks caused by foodborne transmission of SAV have been reported since 1976 in England, where SAV particles were first detected in human diarrheic stool samples [4]. The largest foodborne SAV outbreak recorded, which involved 665 people, was in Japan in 2010 [5]. Many worldwide studies have reported that SAV can be divided into several genogroups, ranging from GI to GV [6–8]. Recent studies found that SAV genogroup VI in isolated fecal samples from diarrheic pigs [9]. Different genogroups result in the infection of different hosts; GI, GII, GIV, and GV infect humans and other primates, whereas GIII infects swine. Furthermore, there are at least eight genotypes of GI, five of GII, and one each of GIII, GIV, and GV. Most SAV detected in humans belong to GI and GII [10, 11].

Although SAV has been accepted as one of the causes of acute gastroenteritis worldwide, little is known about the genetic characteristics of sapoviruses in China based on whole genome analysis, especially about genotypes

infecting humans, due to infrequent detection and difficulties in culturing and sequencing, especially whole genome sequencing. Here we report the complete genome sequence of SAV strain Human/Zhejiang1/2015/China obtained in late 2014 from a child with acute gastroenteritis in Hangzhou, Zhejiang Province, China. Phylogenetic analysis was performed for comparison with other genogroups/genotypes reported previously. The purpose of this study was to characterize and present the first full-length genome of SAV as fundamental information in mainland China.

Materials and methods

Samples were collected from a patient who had acute gastroenteritis in a kindergarten in the Xiacheng district of Hangzhou on December 25 and 26, 2014. Samples were delivered to the CDC laboratories and were made into 10 % (w/v) in phosphate-buffered saline suspensions (PBS) and centrifuged at 2000×g for 10 min. Viral RNA was extracted from 200 µl of the supernatant using a Qiagen RNeasy Mini kit (Qiagen) according to the manufacturer’s instructions. Eluted RNA was stored at –80 °C until use. RNA was detected by reverse transcription PCR with SAV universal primers (SV-F13, SV-F14, SV-R13, and SV-R14) using the following thermal cycling profile: 42 °C for 15 min; 95 °C for 3 min; 35 cycles of 95 °C for 30 s, 48 °C for 30 s, and 74 °C for 45 s according to the previous study of Okada et al. [14]. Full genome Sanger sequencing was performed according to the study by Liu et al. [15].

Geneious 4.8.3 (www.geneious.com) was used to assemble and check the sequences of this SAV, Human/Zhejiang1/2015/China. For the VP1 region and the whole genome, we performed multiple alignments with data matrixes of other sequences downloaded from GenBank for each segment, respectively. We blasted the homology sequences in NCBI for both the VP1 region and the genome, then checked for variant positions in amino acid sequences using MEGA 6.0 (www.mega.software.informer.com) for the whole genomic range. Dataset-specific models that were selected using the Akaike Information Criterion in Modeltest 3.7 were analyzed for each matrix [16]. Maximum likelihood (ML) analysis was processed in RAXML v7.2.8 using the VP1 region and genome, respectively (<http://sco.h-its.org/exelixis/software.html>). Optimal ML trees and bootstrap percentages (BP) were estimated in the same run. The ML BP values were obtained from 1000 BP replicates using the rapid BP algorithm. Both Simplot v3.5.1 and RDP 4.26 analysis with default set were used for detecting recombination event in this strain [17, 18].

Table 1 Unique amino acid changes observed in Human/Zhejiang1/2015/China compared to Chanthaburi-74/2004/Thailand (AY646854)

Polypeptide position	Nonstructural proteins																												
	VP1							VP2																					
Human/Zhejiang1/2015/China	V	V	Y	V	K	K	A	P	G	T	K	S	T	V	I	N	H	H	T	I	V	V	I	S	E	S	S	L	
Chanthaburi-74/2004/Thailand	A	I	F	M	E	R	T	N	V	A	R	P	A	I	V	S	Q	Q	K	V	V	N	D	P	D	P	P	P	S

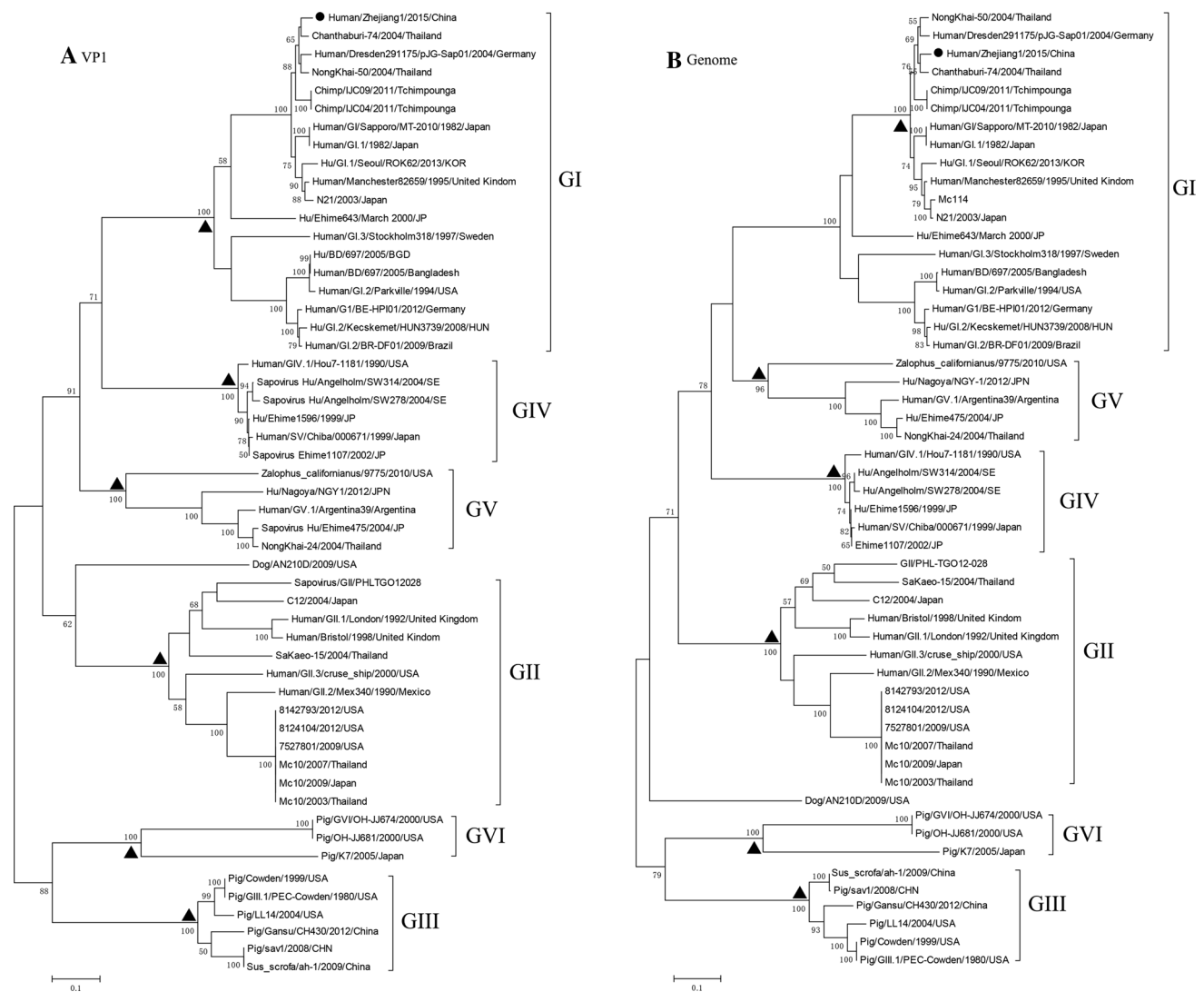


Fig. 1 Maximum likelihood trees for sapovirus based on VP1 region and whole genome, respectively. BP values based on 1000 replicates are indicated above *branches*. *Solid circles* in each tree indicate

sequences obtained in our study. *Triangles* at nodes indicate monophyletic clades of different genogroups. **a** VP1 tree, **b** genome tree

Results and discussion

This has been the first genome obtained from a child in China, in contrast to 47 complete human/porcine sapovirus sequences available in GenBank till the data published. The full genome sequence of the SAV from the patient was deposited in GenBank with accession number KT327081. The SAV obtained was a positive-sense, single-stranded RNA 7441 bp in size with a 3'-end poly (A) tail. Its 5'-UTR was 9 bp long; 3'-UTR was 94 bp long. The genome contained three open reading frames (ORFs). ORF1 encodes the nonstructural proteins (NS) from 10 to 6853 bp. ORF2 is a minor structural protein VP2 (from 6848 to 7347 bp). ORF3 is a hypothetical protein CDS from 5176 to 5662 bp, which has not been clearly defined

[12]. The major capsid protein VP1, from 5164 to 6853 bp, was embraced in ORF1, which followed the seven non-structural proteins from NS1 to NS7 in the 5' to 3' direction (supplementary materials). The NS pattern of ORF1 in our sequence was similar to the GI.1 Manchester strain (GenBank accession: X86560) except for one putative cleavage site between NS5 (VPg) and NS6–NS7 (Protease-RdRp), which was E/A rather than E/G, as mentioned in Oka's review [13]. Typical amino acid motifs were also detected in our sequence, for example: GAPGIGKT from 480 to 487 amino acid (aa) in NS3, KGKTK, and DDEYDE from 941 to 945 aa and 962–967 aa in NS5, respectively. There were three PPG motifs in the VP1 region, which differed from the reference strain of GI.1 Manchester. The conserved GWS motif in the VP1 region was also detected in our

sequence, which proved strictly characteristic of caliciviruses [13]. Chanthaburi-74/2004/Thailand (GenBank accession: AY646854) showed the highest query cover and percent nucleotide identity with our sequence based on the whole genome. We therefore compared amino acid substitutions in the coding regions (mainly including ORF1 and ORF2) in the two sequences (Table 1). Twenty-two variants were found in the nonstructural region while one and two substitutions were discovered in the VP1 and VP2 areas, respectively.

To better classify the sapovirus, we searched the NCBI database and found Hu/GI.1/8743/Maizuru/2008/JPN (HM030922) to be a close match in the VP1 region, while Chanthaburi-74/2004/Thailand (AY646854) was close to the whole genome. In these two maximum likelihood trees with Bootstrap percentages assigned, different genogroups, such as GI, GII, GIII, GIV, GV, and GVI, were reciprocally monophyletic with clades receiving 100 % BP support, with the exception of the GV clade in the genomic ML tree which had a BP value of 96 % (Fig. 1a, b). Our sequence, Human/Zhejiang1/2015/China, was in the GI clade and shared a relative high homology with Chanthaburi-74/2004/Thailand in both trees. The results of Blast from NCBI indicated sapovirus Human/Zhejiang1/2015/China to be in genogroup I/genotype 1 (GI.1). The phylogenetic structure in both ML trees was basically similar, although there was an inconsistency in that either GIV or GV could be the closest sister group of GI according to the VP1 region and the whole genome (Fig. 1a, b). Several possible scenarios may have led to this phenomenon, such as incomplete lineage sorting between genes and genome. We are further analyzing this hypothesis. This study, focused on the interior of the GI clade, indicated that Human/Zhejiang1/2015/China is not a recombinant, since both trees showed the strain we obtained to be clustered in a similar GI clade. This finding was subsequently substantiated by RDP analysis and Simplot (supplementary materials). Further studies are underway with additional viruses in the family *Caliciviridae*, including norovirus, based on surveillance programs of foodborne infections established by the national health and family planning commission of China.

Conclusions

This first genome sequence of SAV from a child in Zhejiang, China, provides genetic characteristics of SAV infections in humans in the eastern China region. Based on complete capsid gene sequences from the phylogenetic results we determined that this SAV, Human/Zhejiang1/

2015/China, belongs to GI, which is consistent with previous reports that GI is known to infect humans. Recombination was not detected. Our SAV sequence may help in both surveillance and in understanding the characterization of emerging gastroenteritis infections in China.

Acknowledgments This work was supported by the Major Science and Technology Special Projects of Zhejiang, China (2013C03045-1), Provincial Medical Research Fund of Zhejiang (2015RCB010), Zhejiang Provincial Program for the Cultivation of High-level Innovative Health Talents and Key Disciplinary of Health and Family Planning Commission of Zhejiang Province (CX-9). Dr. David E Boufford is gratefully acknowledged for editing the manuscript.

Author contributions XHZ and YJZ conceived and designed the study. XCS and JG carried out the sequencing. YS analyzed the data and drafted the manuscript. XQS and RZ collected the sample and provided the reagents. All authors read and approved the final manuscript.

Compliance with ethical standards

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

Ethics statement This study was approved by the ethics committee of the Zhejiang Provincial Center for Disease Control and Prevention (ZJCDC), China. This study was done in accordance with the Declaration of Helsinki, Good Clinical Practice, and Chinese regulatory requirements.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. S. Svraka, H. Vennema, B. van der Veer, K.O. Hedlund, M. Thorhagen, J. Siebenga, E. Duizer, M. Koopmans, J. Clin. Virol. **48**, 2191–2198 (2010)
2. T. Oka, Q. Wang, K. Katayama, L.J. Saif, Clin. Microbiol. Rev. **28**, 32–53 (2015)
3. L.E. Lee, E.A. Cebelinski, C. Fuller, W.E. Keene, K. Smith, J. Vinje, J.M. Besser, Emerg. Infect. Dis. **18**, 873–876 (2012)
4. C.R. Madeley, B.P. Cosgrove, Lancet **307**, 199–200 (1976)
5. S. Kobayashi, N. Fujiwara, Y. Yasui, T. Yamashita, R. Hiramatsu, H. Minagawa, Arch. Virol. **157**, 1995–1997 (2012)
6. T. Farkas, W.M. Zhong, Y. Jing, P.W. Huang, S.M. Espinosa, N. Martinez, A.L. Morrow, G.M. Ruiz-Palacios, L.K. Pickering, X. Jiang, Arch. Virol. **149**, 1309–1323 (2004)
7. G.S. Hansman, T. Oka, N. Takeda, Virus Res. **137**, 261–265 (2008)
8. L. Dufkova, I. Scigalkova, R. Moutelikova, H. Malenovska, J. Prodelalova, Arch. Virol. **158**, 549–558 (2013)
9. T. Oka, L.F. Saif, Z.Z. Yang, K.A. Scheuer, G.T. Stoltzfus, Q.H. Wang, Genome Announc. **2**, e0027-14 (2014)

10. S. Akihara, T.G. Phan, T.A. Nguyen, F. Yagyu, S. Okitsu, W.E. Muller, H. Ushijima, *J. Med. Virol.* **77**, 595–601 (2005)
11. T. Oka, K. Mori, N. Iritani, S. Harada, Y. Ueki, S. Iizuka, K. Mise, K. Murakami, T. Wakita, K. Katayama, *Arch. Virol.* **157**, 349–352 (2012)
12. G.S. Hansman, N. Takeda, T. Oka, M. Oseto, K.O. Hedund, K. Katayama, *Emerg. Infect. Dis.* **11**, 1916–1920 (2005)
13. R. Chen, J.D. Neill, J.S. Noel, A.M. Hutson, R.I. Glass, M.K. Estes, B.V. Prasad, *J. Virol.* **78**, 6469–6479 (2004)
14. M. Okada, Y. Yamashita, M. Oseto, K. Shinozaki, *Arch. Virol.* **151**, 2503–2509 (2006)
15. X. Liu, D. Yamamoto, M. Saito, T. Imagawa, *J. Clin. Virol.* **68**, 83–88 (2015)
16. D. Posada, K.A. Crandall, *Bioinformatics* **14**, 817–818 (1998)
17. M. Smith, *J. Mol. Evol.* **34**, 126–129 (1992)
18. D. Martin, E. Rybicki, *Bioinformatics* **16**, 562–563 (2000)