Check for updates

OPEN ACCESS

EDITED BY Chris Davis, University of Glasgow, United Kingdom

REVIEWED BY Masaru Yokoyama, National Institute of Infectious Diseases (NIID), Japan Walter Randazzo, Spanish National Research Council (CSIC), Spain

*CORRESPONDENCE Baihai Jiao ⊠ Bjiao@uchc.edu Tiantian Liu Iliutian_98@163.com Boyan Jiao ⊠ j198319831983@126.com

[†]These authors have contributed equally to this work

RECEIVED 03 February 2025 ACCEPTED 11 April 2025 PUBLISHED 30 April 2025

CITATION

Fu Z, Wu C, Yu Y, Zhao J, Dong Y, Liu W, Dou H, Shi X, Cai C, Jiao B, Liu T and Jiao B (2025) Genomic characterization of norovirus and Sapovirus from a diarrhea outbreak in a school linked to heavy rainfall. *Front. Microbiol.* 16:1570161. doi: 10.3389/fmicb.2025.1570161

COPYRIGHT

© 2025 Fu, Wu, Yu, Zhao, Dong, Liu, Dou, Shi, Cai, Jiao, Liu and Jiao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication

in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Genomic characterization of norovirus and Sapovirus from a diarrhea outbreak in a school linked to heavy rainfall

Zhongyan Fu^{1,2†}, Changjing Wu^{3,4†}, Yinghui Yu⁵, Jian Zhao^{3,4}, Yan Dong⁶, Wei Liu^{2,3}, Huixin Dou^{2,3}, Xuezhen Shi⁷, Chengzhi Cai⁸, Baihai Jiao⁹*, Tiantian Liu¹⁰* and Boyan Jiao^{3,4}*

¹Department of Infectious Disease Control, Shandong Center for Disease Control and Prevention, Jinan, China, ²Shandong Provincial Key Laboratory of Intelligent Monitoring, Early Warning, Prevention and Control for Infectious Diseases, Jinan, China, ³Department of Laboratory, Jining Center for Disease Control and Prevention, Jining, China, ⁴Jining Key Laboratory of Infectious Disease Control and Prevention, Jining, China, ⁵Department of Infectious Disease Control, Jining Center for Disease Control and Prevention, Jining, China, ⁶Department of Infectious Disease Control, Sishui Center for Disease Control and Prevention, Sishui, China, ⁷Gynecology Department, Baoding First Central Hospital, Baoding, China, ⁸Computer Science Department, Central Connecticut State University, New Britain, CT, United States, ⁹Department of Medicine, School of Medicine, University of Connecticut Health Center, Farmington, CT, United States, ¹⁰Department of School Health, Jining Center for Disease Control and Prevention, Jining, China

Background: Diarrhea poses a serious threat to human health, and rainfall is known to increase the incidence of diarrheal diseases. On July 7, 2024, a diarrhea outbreak occurred in a school in Sishui County, Jining City, Shandong Province, China, following heavy rainfall. This study aimed to identify the pathogens responsible for the outbreak and characterize their whole genomes.

Methods: On July 8, 2024, a total of 21 stool samples from diarrhea cases, 2 water samples from private wells, and 1 drinking water sample from the school cafeteria were collected. Real-time quantitative PCR was used to detect Rotavirus A (RV-A), Norovirus genogroup I (NV GI), Norovirus genogroup II (NV GI), Sapovirus (SaV), Human Astrovirus (HAstV), and Human Adenovirus (HAdV). Whole-genome sequencing was performed for NV GI and SaV-positive samples to determine their genotypes, construct phylogenetic trees, and analyze amino acid variation sites in encoded proteins.

Results: Among the 21 case stool samples, 7 tested positive for both NV GI and SaV, 10 were positive for NV GI only, and 1 was positive for SaV only. Of the 2 private well water samples, one was positive for NV GI and the other for SaV. Whole-genome sequences were obtained for 11 NV GI strains and 2 SaV strains. The 11 NV GI sequences from the outbreak exhibited high homology, with whole-genome similarity ranging from 99.96% to 100%, and were all identified as the NV GI.6 [P11] genotype. Phylogenetic analysis showed that these 11 sequences clustered within the same evolutionary branch. Similarly, the 2 SaV sequences were highly homologous, with 99.97% similarity, and were identified as the SaV GI.6 genotype, clustering within the same phylogenetic branch.

Conclusions: This diarrhea outbreak was caused by the combined presence of NV GI and SaV following heavy rainfall. These findings provide valuable reference data for the prevention and control of diarrhea outbreaks caused by heavy rainfall or multiple pathogens.

KEYWORDS

heavy rainfall, diarrhea outbreak, Norovirus, Sapovirus, whole genome

1 Introduction

Diarrhea is a significant global public health issue, causing approximately 1.17 million deaths annually worldwide and ranking as one of the leading causes of child mortality (GBD 2021 Diarrhoeal Diseases Collaborators, 2024). The causes of diarrhea are diverse, with viral infections, such as Norovirus (NV) and Sapovirus (SaV), being major contributors (Cohen et al., 2022). NV accounts for an estimated 680 million diarrhea cases each year and approximately 200,000 deaths annually (Winder et al., 2022). SaV is responsible for about 10% of diarrhea cases and is estimated to cause 23,000 deaths annually among children under the age of five (Cohen et al., 2022; Pang et al., 2000).

Both NV and SaV belong to the Caliciviridae family. The NV genome contains approximately 7,500 nucleotides (Lin et al., 2023), comprising three open reading frames (ORFs). ORF1 encodes a polyprotein that is cleaved into six non-structural proteins, including p48, NTPase, p22, VPg, Pro, and RdRp. ORF2 encodes the structural protein VP1, while ORF3 encodes the structural protein VP2 (Lin et al., 2023). Based on the sequences of the VP1 and RdRp genes, NV is classified into G (genogroup) and P (polymerase) genotypes (Chhabra et al., 2019), with GI and GII being the primary genogroups infecting humans (Winder et al., 2022; Chigor et al., 2024; Wu et al., 2024).

The SaV genome consists of 7,100–7,700 nucleotides and generally contains two open reading frames (ORFs). ORF1 encodes a polyprotein that is cleaved into six non-structural proteins (NS1, NS2, NS3, NS4, NS5, and NS6-7) and the structural protein VP1, while ORF2 encodes the structural protein VP2 (Oka et al., 2015; Yinda et al., 2017). Based on the VP1 coding region sequence, SaV can be classified into different genotypes (Oka et al., 2012), with GI, GII, GIV, and GV being the primary genotypes infecting humans (Zhao et al., 2024).

The occurrence of diarrhea is closely related to environmental and climatic factors, as heavy rainfall can lead to the contamination of water sources with pathogens like NV (de Man et al., 2014; Joosten et al., 2017), thereby increasing the incidence of diarrheal diseases (Levy et al., 2016; Dhimal et al., 2022; Azage et al., 2017). On July 7, 2024, a heavy rainfall event occurred in Sishui County, China, with precipitation exceeding 240 mm. Between July 7 and July 9, a diarrhea outbreak involving 26 cases was reported in a local school in Sishui County. In this study, stool samples from diarrhea cases and water samples were tested for RV-A, NV GI, NV GII, SaV, HAstV, and HAdV. The results identified NV GI and SaV as the causative agents. Whole-genome sequencing and analysis of NV GI and SaV were performed, providing insights to inform the monitoring, prevention, and control of diarrheal diseases following heavy rainfall.

2 Materials and methods

2.1 Sample collection

From July 7 to July 9, 2024, a diarrhea outbreak characterized by nausea and vomiting as the primary clinical symptoms was reported at a school in Sishui County, Jining City, Shandong Province, China. A total of 26 cases were identified, with 12 cases on July 7, 9 cases on July 8, and 5 cases on July 9.

On July 8, 2024, 21 stool samples (each weighing more than 5 g) were collected from affected individuals. Additionally, 2 private well water samples and 1 cafeteria drinking water sample (each 500 mL) were collected. All samples were sent to the Microbiology Laboratory at the Jining Center for Disease Control and Prevention for testing.

2.2 Nucleic acid detection

The nucleic acid detection protocol was based on our previously described method (Wu et al., 2024). Briefly, 0.1 g of stool sample was mixed with 0.9 mL PBS and vortexed three times. The mixture was centrifuged at 8,000 g for 5 min. After filtering 500 mL of water using medium-speed qualitative filter paper, MgCl₂ was added to achieve a final concentration of 0.05 mol/L, adjusting the pH to 3.0–3.5. The suspension was filtered through a 0.45 μ m nitrocellulose membrane using a large-volume pathogen concentrator (Beijing ZC Bio, Model: PurLVS), and the membrane was shaken in 1 mL PBS, and after vortexing and allowing it to stand, the supernatant was extracted. A 200 μ L aliquot of the supernatant was used for nucleic acid extraction with a fully automated nucleic acid extraction instrument (GeneRotex 96, Xi'an Tianlong Technology Co., Ltd.) and a viral nucleic acid extraction kit (Cat. No. T138, Xi'an Tianlong Technology Co., Ltd.).

Real-time quantitative PCR was performed using triplex nucleic acid detection kits for RV-A, NV GI, and NV GII (Cat. No. A2593YH, Beijing ZC BioScience Co., Ltd.) and for SaV, HAstV, and HAdV (Cat. No. A2723YH, Beijing ZC BioScience Co., Ltd.).

2.3 Genome sequencing

Nucleic acid samples positive for NV GI and SaV were sent to Shanghai BioGerm Technology Co., Ltd. for whole-genome sequencing. NV-specific target gene amplification was conducted using the NV whole-genome targeted capture kit (Shanghai BioGerm Technology Co., Ltd.), followed by library construction using a whole-genome DNA library construction kit (Shanghai BioGerm Technology Co., Ltd.). Sequencing was performed using the DNBSEQ-T7 PE150 platform (Shenzhen BGI Technology Co., Ltd.).

2.4 Genetic sequence analysis

Genotyping of NV GI sequences was conducted using the Norovirus Typing Tool Version 2.0 (https://www.rivm.nl/mpf/ typingtool/norovirus/), while SaV sequences were genotyped using the Human Calicivirus Typing Tool (https://calicivirustypingtool. cdc.gov/). BLAST analyses were performed against the NCBI database. Full-genome sequences of NV GI.1 [P1], GI.2 [P2], GI.3 [P3], GI.3 [P13], GI.4 [P4], GI.5 [P5], GI.5 [P4], GI.6 [P6], GI.6 [P11], and GI.7 [P7] genotypes and SaV GI.1, GI.2, GI.3, GI.5, GI.6, and GI.7 genotypes were downloaded from the NCBI database.



Epidemiological characteristics of the diarrheal outbreak. (A) Distribution of onset times; a total of 12 cases were reported on July 7, 2024, 9 cases on July 8, 2024, and 5 cases on July 9, 2024. (B) Distribution of affected populations; there were 24 students and 2 restaurant staff members. (C) Gender distribution; there were 16 male cases and 10 female cases. (D) Pathogen detection results indicated that among 19 student cases, the positivity rate for both NVGI and SaV viruses was 36.84% (7/19), the positivity rate for NVGI only was 52.63% (10/19), and the positivity rate for SaV only was 5.26% (1/19). Both samples from restaurant staff and the cafeteria's drinking water tested negative. However, self-supplied well A tested positive for SaV, and self-supplied well B tested positive for NVGI.

Sequence alignment was conducted using MEGA 7.0.14, and phylogenetic trees for the NV whole genome, ORF1, ORF2, ORF3, and SaV whole genome, ORF1, and ORF2 were constructed using the maximum-likelihood method with 1,000 bootstrap replicates. Phosphorylation sites in proteins were analyzed using NetPhos-3.1 (https://services.healthtech.dtu.dk/services/NetPhos-3.1/), and N-linked glycosylation sites were analyzed using NetNGlyc-1.0 (https://services.healthtech.dtu.dk/services/NetNGlyc-1.0/).

3 Results

3.1 Epidemiological characteristics and nucleic acid testing results

The school affected by the infectious diarrhea outbreak includes 26 classes across the 1st and 2nd years of senior high school, with approximately 1,500 students and 120 staff members. A total of 26 cases of diarrhea were reported during this outbreak, all of which presented symptoms of nausea and vomiting, with some also experiencing diarrhea and fever. The first case was reported at 16:00 on July 7, 2024, with a total of 12 cases occurring on that day, 9 cases on July 8, and 5 cases on July 9. Among the patients, 24 were students aged 16 to 18, distributed across 9 classes; 2 cafeteria staff members were aged 39 and 41, respectively. The gender distribution was 16 male cases and 10 female cases (Figures 1A–C).

The school has two self-supplied wells: Well A provides water for the public restrooms in the school office building, while Well B supplies water for the student dormitories and public restrooms in the teaching building. The daily water supply for the school cafeteria and drinking water for all students and staff is sourced from the city water supply of Shou County. On July 7, 2024, Shou County experienced heavy rainfall exceeding 240 mm, which led to the backflow or leakage of sewage in both self-supplied wells.

A total of 21 fecal samples were collected from the cases involved, including 19 samples from students and 2 samples from cafeteria staff. Additionally, one sample each was collected from Well A, Well B, and the cafeteria drinking water for testing of NV, rotavirus, SaV, astrovirus, and adenovirus nucleic acids. Among the 19 student fecal samples, 7 tested positive for both NV GI and SaV, 10 tested positive only for NV GI, 1 tested positive only for SaV, and 1 tested negative. Both of the cafeteria staff fecal samples were negative. SaV was detected positive in Well A, NV GI was detected positive in Well B, while the cafeteria drinking water tested negative (Figure 1D).

3.2 NV whole genome sequencing analysis

3.2.1 Similarity analysis

In this diarrheal outbreak, 11 complete NV GI genome sequences were obtained from the fecal samples of patients, with sequence lengths ranging from 7,682 bp to 7,702 bp. Unfortunately, Whole-genome sequencing of norovirus was conducted on the self-sourced well water, but only a 75-nucleotide fragment (positions 3,287–3,361) of the ORF1 region was obtained. The nucleotide sequence of this fragment from the self-sourced water was identical

Gene		Similarity among the 11 NV GI sequences from the current diarrhea outbreak		Similarity of the 11 NV sequences from the current diarrhea outbreak with the 2019 isolate MW243609.1 from Shanxi, China	
Gene name	Protein name	Nucleotide	Amino acid	Nucleotide	Nucleotide
Complete genome sequence		99.96%-100%	-	99.00%-99.02%	-
ORF1	p48	99.94%-100%	100%	98.92%-98.96%	99.75%
	NTPase		100%		99.72%
	p22		100%		99.50%
	VPg		100%		99.27%
	Pro		100%		100%
	RdRp		99.80%-100%		99.80%-100%
ORF2	VP1	99.94%-100%	99.81%-100%	99.07%-99.14%	99.81%-99.63%
ORF3	VP2	100%	100%	99.52%	100%

TABLE 1 Similarity analysis of NV GI genes.

to the corresponding sequence from 11 NV GI strains involved in the diarrhea outbreak. A BLAST analysis of this 3,287–3,361 nucleotide sequence from the self-sourced water in the NCBI database did not reveal any exact matches. The closest match was with 18 norovirus GI.6 [P11] strains, showing a similarity of 98.67%, with a single nucleotide difference. Among the 11 NV GI sequences, 6 were found to be completely identical, and the sequences exhibited 7 nucleotide variations, with similarities ranging from 99.96% to 100%. The only differences in amino acid sequences among the 11 NV GI whole genome sequences were observed in the RdRp and VP1 regions, with amino acid similarities of 99.80% to 100% and 99.81% to 100%, respectively, while the amino acid similarities for P48, NTPase, p22, VPg, Pro, and VP2 were 100%, indicating complete identity.

BLAST analysis in the NCBI database revealed that the NV GI sequences from this diarrheal outbreak showed the highest similarity, ranging from 99.00% to 99.02%, to the Norovirus GI isolate 2019 (GenBank: MW243609.1) derived from human feces in Shanxi, China. Specifically, the ORF1 similarity was 98.92% to 98.96%, ORF2 similarity was 99.07% to 99.14%, and ORF3 similarity was 99.52%. Compared to MW243609.1, the amino acid sequences of Pro and VP2 exhibited 100% similarity, while differences were observed in the amino acid sequences of P48, NTPase, p22, VPg, RdRp, and VP1 (Table 1).

3.2.2 Gene evolution analysis

The Norovirus Typing Tool Version 2.0 was utilized to analyze the genotypes of the 11 NV GI sequences from this diarrheal outbreak, and all 11 NV GI sequences were confirmed to be of the GI.6 [P11] genotype. Reference sequences for various genotypes, including GI.6P (Oka et al., 2012), GI.1P (GBD 2021 Diarrhoeal Diseases Collaborators, 2024), GI.2P (Cohen et al., 2022), GI.3P (Winder et al., 2022), GI.3P (de Man et al., 2014), GI.4P (Pang et al., 2000), GI.5P (Pang et al., 2000), GI.5P (Chhabra et al., 2019), GI.6P (Chhabra et al., 2019), and GI.7P (Chigor et al., 2024), were obtained from the GenBank database to construct an evolutionary tree. In the NV GI whole genome phylogenetic tree, the 11 NV GI sequences from this diarrheal outbreak formed an evolutionary cluster positioned within the GI.6 [P11] evolutionary branch (Figure 2).

In the ORF1 phylogenetic tree, the 11 NV GI sequences from this diarrheal outbreak formed an evolutionary cluster positioned within the GI.P11 evolutionary branch, as shown in Figure 3. In the ORF2 phylogenetic tree, the 11 NV GI sequences from this diarrheal outbreak also formed an evolutionary cluster within the GI.6 evolutionary branch, depicted in Figure 4. In the ORF3 phylogenetic tree, the 11 NV GI sequences from this diarrheal outbreak formed an evolutionary cluster that is located on the same evolutionary branch as the NV GI.6 [P11] sequences, as illustrated in Figure 5.

3.2.3 Gene variation analysis

The ORF1 of NV encodes a polyprotein that can be cleaved into p48, NTPase, p22, VPg, Pro, and RdRp, while ORF2 encodes VP1 and ORF3 encodes VP2, with VP1 and VP2 being structural proteins of NV. Compared to the isolated strain MW243609.1 from Shanxi, China, the p48 protein exhibited a Q137L mutation, where Q137 is located in the H-box/NC domain of p48. The H-box/NC domain is involved in regulating lipid synthesis and metabolism by influencing the formation of enlarged lipid droplets (Hung et al., 2023).

NTPase possesses RNA helicase activity, with motifs A (aa162–169), B (aa206–213), and C (aa245–259) representing the helicase core domains of NTPase (Li et al., 2018). In the 11 NV GI strains from this diarrheal outbreak, no mutations were found at the amino acid sites in the helicase core domains. NTPase contains two lipid droplet targeting motifs (LTM), with LTM-2 located at amino acids 51–90, which plays a role in NTPase's ER localization and the induction of cellular apoptosis (Yen et al., 2021). In the NV GI strains from this outbreak, a mutation E67D was observed in the LTM-2 of NTPase.

The amino acids at positions 65–72 of the p22 protein form the YXWESDG motif, which is essential for p22/s entry into COPII vesicles (Sharp et al., 2012, 2010). In this outbreak, the MERES motif of p22 did not exhibit any mutations. The amino acids at positions 112–127 of p22 comprise the membrane association domain (MAD), a critical domain for the localization of p22 to





Phylogenetic tree of the ORF1 region of NV GI. The tree was constructed using the maximum likelihood method with 1,000 bootstrap replicat MEGA 7.0.14 software. The yellow branches represent the NV GI sequences associated with the current diarrheal outbreak.

06





the Golgi apparatus, and this domain also remained unchanged. Notably, a mutation A162S was identified in p22, and analysis using NetPhos-3.1 indicated that S162 acts as a phosphorylation site, leading to the introduction of an additional phosphorylation site S162 in the p22 of the 11 NV GI strains from this outbreak.

During the NV replication process, the conserved tyrosine (Y26 in Murine Norovirus VPg and Y30 in Human noroviruses) in VPg binds to the 5/ terminus of the genomic RNAs and serves as a protein primer for the first base of RNA polymerase extension (Eden et al., 2011; Subba-Reddy et al., 2011; Leen et al., 2013). A stretch of 20 amino acids at the C terminus of NV VPg can initiate RNA translation by interacting with translation initiation factors (Leen et al., 2016; Daughenbaugh et al., 2003, 2006). In the 11 NV GI strains from this outbreak, neither Y30 nor the C terminus of VPg exhibited mutations. However, an S26N mutation occurred in VPg, with S26 identified as a phosphorylation site, resulting in the loss of the phosphorylation site S26 in the NV GI strains from this outbreak.

RdRp acts as the polymerase of NV, and the residues R182, D242, S300, N309, D343, and D344 are crucial for its enzymatic activity (Zamyatkin et al., 2009, 2008; Deval et al., 2017). No

TABLE 2 Analysis of variant sites in 11 NV GI protein strains associated with the current diarrheal outbreak.

Protein name	Variant sites of the 11 NV GI protein strains from the current diarrheal outbreak compared to the Chinese isolate MW243609.1 from Shanxi		
p48	Q137L (11)		
NTPase	E67D (11)		
p22	A162S (11)		
VPg	S26N (11)		
RdRp	T133I (1)		
VP1	S58N (11), R293G (1)		

The numbers in parentheses indicate the count of mutations occurring in the 11 NV strains associated with the current diarrheal outbreak.

TABLE 3 Similarity analysis of SaV genes.

mutations were detected at these polymerase active sites in the 11 NV GI strains from this outbreak. However, a T133I mutation was found in the RdRp of one NV GI strain, changing residues 131–133 from NGT to NGI. Analysis using NetNGlyc-1.0 showed that the NV GI with the T133I mutation lost the N131 glycosylation site.

VP1 is the major structural protein of NV, with amino acids 337–346, 351–360, 366–377, 386–394, and 431–442 serving as binding sites for human monoclonal antibodies (Kimura-Someya et al., 2024). The histo-blood group antigens (HBGAs) function as receptors for NV, with receptor binding sites located at D331, H333, Q347, W384, S386, and S389 on NV GI (Tan and Jiang, 2010). In this diarrheal outbreak, NV GI exhibited S58N and R293G mutations; however, neither the binding sites for human monoclonal antibodies nor the receptor binding sites were affected (Table 2).

3.3 SaV whole genome sequencing analysis

3.3.1 Similarity analysis

Two strains of SaV sequences were obtained from the stool samples of patients involved in the diarrhea outbreak; one strain has a sequence length of 7,389 bp, while the other has a sequence length of 6,714 bp (lacking the *ORF2* gene sequence). Unfortunately, this study did not obtain the SaV genomic sequences from the water samples. Between the two SaV sequences from the diarrhea outbreak, there are two nucleotide site differences in the *ORF1* gene, with a similarity of 99.97%. The only differences in the encoded proteins of the two SaV sequences are found in NS1 and NS6–7, with similarities of 97.00% and 99.53%, respectively, while the amino acid sequences of NS2, NS3, NS4, NS5, and VP1 show 100% similarity, indicating they are completely identical.

In the BLAST analysis from the NCBI database, the two SaV sequences from this diarrhea outbreak are most similar to the Sapovirus GI strain Hu/GI.6/CHN/2021/RC-21020, which was isolated in Jiangsu Province, China, in February 2021. Compared to the Sapovirus GI strain Hu/GI.6/CHN/2021/RC-21020, the gene sequence of the outbreak Sapovirus GI strain

Gene		Similarity of SaV sequences between two strains in the current diarrheal outbreak		Similarity of SaV sequences from the current diarrheal outbreak with Sapovirus GI strain Hu/GI.6/CHN/2021/RC-21020	
Gene name	Protein name	Nucleotide	Amino acid	Nucleotide	Amino acid
Complete Genome Sequence		-	-	98.38%	-
ORF1	NS1	99.97%	100%	98.38%-98.41%	97.00%
	NS2		100%		100%
	NS3		100%		100%
	NS4		100%		100%
	NS5		100%		100%
	NS6-7		100%		99.53%
	VP1		100%		100%
ORF2	VP2	-	-	97.99%	99.39%

"-" indicates that the sequence of one strain from this outbreak did not yield the ORF2 sequence.



Hu/GI.6/CHN/2024/JN01 has a whole genomic nucleotide sequence similarity of 98.38%, with an ORF2 gene nucleotide similarity of 97.99%. Additionally, the similarity of the SaV ORF1 between these two strains in the diarrhea outbreak is between 98.38% and 98.41% (Table 3).

3.3.2 Genetic evolution analysis

Using the Human Calicivirus Typing Tool, the genotypes of the two SaV gene sequences from this outbreak were analyzed, and both strains were found to belong to the GI.6 genotype. Phylogenetic trees for the complete SaV genome, ORF1, and ORF2 were constructed using the sequences of SaV GI.1, GI.2, GI.3, GI.5, GI.6, and GI.7 genotypes. The phylogenetic trees for the complete genome, ORF1, and ORF2 of SaV showed a high degree of similarity. In the phylogenetic trees of the complete SaV genome and ORF1, both strains from the diarrhea outbreak were located on the GI.6 evolutionary branch, and they independently formed an evolutionary cluster. This can be seen in Figures 6, 7. In the ORF2 phylogenetic tree, the sequences from this diarrhea outbreak were positioned on the same evolutionary branch as those of the SaV GI.6 genotype (Figure 8).

3.3.3 Gene variation analysis

The SaV ORF1 encodes a polyprotein that can be cleaved into NS1, NS2, NS3, NS4, NS5, NS6-7, and VP1, while ORF2 encodes VP2. VP1 and VP2 are the structural proteins of SaV. NS6-7



contains the virus's chymotrypsin-like protease domain and RNAdependent RNA polymerase domain. Among these, NS1 is known to be one of the least conserved proteins of SaV.

Compared to the Sapovirus GI strain Hu/GI.6/CHN/2021/RC-21020, the NS1 of the SaV involved in this diarrhea outbreak exhibited two amino acid mutations. Similarly, NS6-7 also showed two amino acid mutations. The H31, E52, and C116 in NS6-7 are the protease catalytic sites, while V32, T111, K113, G114, A132, G133, and T134 are part of the protease active site (Robel et al., 2008; Oka et al., 2007). However, no mutations were found at the protease catalytic sites and active sites in the two SaV strains from the diarrhea outbreak. The palm region and thumb domain of NS6-7 are associated with polymerase activity (Fullerton et al., 2007). The amino acid T554 is located in the extended loop domain of the polymerase, while E613 resides in the α -helix 14 domain of the polymerase, both of which are not included within the active domains of the enzyme. Phosphorylation analysis conducted using the NetPhos3.1 Server indicated that the two SaV strains from this diarrhea outbreak lost the phosphorylation site at T554. VP2 is known as the minor structural protein of the SaV virion, although its structure and function remain unclear (Li et al., 2022). In the SaV strain Hu/GI.6/CHN/2024/JN01 from this diarrhea outbreak, an A153T mutation was identified (Table 4).

4 Discussion

NV and SaV are significant pathogens responsible for diarrheal outbreaks. In China, diarrheal outbreaks primarily occur in schools, with the GII.2 [P16] genotype being the main type associated with

Frontiers in Microbiology

only 1.3% of NV-induced outbreaks (Jin et al., 2020). The genotypes of SaV causing diarrheal outbreaks differ across various regions of China: GI.6 is predominantly associated with outbreaks in Zhejiang province (Su et al., 2023), while Beijing mainly sees GI.1 and GI.2 genotypes (Jiao et al., 2023), and Shenzhen is primarily affected by GI.2 and GII.3 genotypes (Cheng et al., 2023).

This study conducted an epidemiological investigation, pathogen identification, and whole-genome sequencing analysis of a diarrheal outbreak that occurred after heavy rainfall. The results indicated that 26 individuals exhibited symptoms of diarrhea, with 21 stool specimens collected, among which 18 tested positive for either NV or SaV. Additionally, NV GI and SaV were detected in two samples of self-sourced well water. The outbreak may have resulted from pollution of the well water due to sewage backflow or leakage following the rainfall, leading to NV GI and SaV transmission between individuals. The outbreak suggests that there may be a certain level of distribution of rare genotypes, such as NV GI.6 [P11] and SaV GI.6, in the environment. Under specific conditions, such as heavy rainfall, enhanced environmental monitoring should be implemented to provide early warnings for NV and SaV outbreaks.

In Jining, China, GII genotypes of NV are primarily prevalent, with GI genotypes being relatively rare; however, the 11 complete genomic sequences of NV obtained in this outbreak were all identified as NV GI.6 [P11]. Phylogenetic trees constructed from the whole genome sequence, along with ORF1, ORF2, and ORF3 sequences, revealed that these 11 NV GI strains formed a distinct



Sapovirus sp. isolate S6

TABLE 4 Analysis of variant sites in 2 SaV protein strains associated with the current diarrheal outbreak.

Protein	Compared to Sapovirus GI Strain Hu/GI.6/CHN/2021/RC-21020, the variant sites of SaV associated with the current diarrheal outbreak show differences in comparison to Sapovirus GI Strain Hu/GI.6/CHN/2021/RC-21020
NS1	L6F (2), T45A (2)
NS6-7	V42A (2), T554S (2), E613Q (2)
VP2	A153T

The numbers in parentheses indicate the number of mutations observed in the two SaV strains associated with the current diarrheal outbreak. VP2 represents the mutation sites in the diarrheal outbreak strain Sapovirus GI Strain Hu/GI.6/CHN/2024/JN01, while the VP2 sequence for the diarrheal outbreak strain Sapovirus GI Strain Hu/GI.6/CHN/2024/JN02 has not been obtained.

evolutionary cluster. Compared to the isolated strain MW243609.1 from Shanxi, the 11 NV GI strains from this outbreak exhibited no amino acid variations in Pro and VP2, while differences were observed in the amino acid sequences of P48, NTPase, p22, VPg, RdRp, and VP1. Notably, the polymerase active site of RdRp, the receptor binding sites of VP1 for human monoclonal antibodies, and the critical amino acids Y30 and C terminal of VPg involved in replication and translation remained unchanged. Conversely, amino acid mutations were detected in the H-box/NC domain of p48, which is involved in regulating lipid synthesis and metabolism, and in LTM-2 of NTPase related to ER localization and the induction of cellular apoptosis. Further research is needed to ascertain whether these mutations affect the functional characteristics of the associated proteins.

The two SaV strains obtained from this diarrheal outbreak were both identified as GI.6 genotype, displaying high homology. In phylogenetic trees of the whole genome and ORF1 sequences, these two SaV strains also formed a distinct evolutionary cluster. When compared to the Sapovirus GI strain Hu/GI.6/CHN/2021/RC-21020 from Jiangsu, none of the amino acid sequences of NS2, NS3, NS4, NS5, and VP1 underwent any changes, although some degree of amino acid mutations were observed in NS1, NS6-7, and VP2. Importantly, the protease catalytic sites in NS6-7 and the active structure domains of the polymerase remained unchanged.

Post-translational modifications (PTMs) can significantly impact viral replication and virulence. Phosphorylation and glycosylation are common PTMs in NV that affect processes such as replication and infection. However, research on the glycosylation and phosphorylation of SaV is relatively limited (Eden et al., 2011; Cheng et al., 2023; Mallagaray et al., 2019; Olspert et al., 2016). The NV strains from this diarrheal outbreak exhibited changes in glycosylation and phosphorylation patterns, including an increase in the phosphorylation site S162 in p22, loss of the S26 phosphorylation site in VPg, and loss of the N131 glycosylation site in RdRp. The NS6-7 of the two SaV strains lost the T554 phosphorylation site. The implications of these modifications on the functional characteristics of the related proteins remain unclear.

It is worth noting that although NV GI and SaV were detected in the self-sourced well water. The concentration of norovirus and SaV in the water samples from this outbreak may have been low. Only 500 mL of water was collected per sample, and largevolume samples were not taken, resulting in insufficient viral concentration after concentration, which hindered sequencing. The water samples only yielded a 75-bp gene sequence for NV, and this 75-bp fragment was identical to the sequences from 11 NV GI strains involved in the diarrhea outbreak. However, it showed some differences compared to the norovirus sequences in the NCBI database, providing supporting evidence that the norovirus in this outbreak may have been contaminated from the self-sourced well water. However, genomic sequences of SaV from the water samples were not obtained, preventing a comparative analysis between the genomic sequences of the outbreak cases and those from the well water. Rare genotypes such as NV GI.6 [P11] and SaV GI.6 are not the main types responsible for viral diarrhea. The immune antibody levels in the population against these genotypes may be low, and under specific conditions, these rare genotypes may cause an outbreak.

In conclusion, this study investigated a diarrheal outbreak that occurred in a school post-heavy rainfall, conducting epidemiological assessments, pathogen detection, and wholegenome sequencing analyses. The results confirmed that contamination of well water by NV GI and SaV, following the heavy rain, led to the outbreak. This further elucidates the correlation between heavy rainfall and the incidence of diarrhea. Therefore, enhancing awareness and preventive measures for diarrheal diseases during the rainy season should be prioritized to effectively prevent such outbreaks.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Ethics statement

The studies involving humans were approved by the Jining Center for disease control and prevention. The studies were conducted in accordance with the local legislation and institutional requirements. The human samples used in this study were acquired from primarily isolated as part of your previous study for which ethical approval was obtained. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

ZF: Resources, Software, Writing – original draft. CW: Formal analysis, Resources, Software, Writing – original draft. YY: Data curation, Formal analysis, Resources, Writing – original draft. JZ: Formal analysis, Methodology, Software, Writing – review & editing. YD: Data curation, Formal analysis, Methodology, Writing – review & editing. WL: Data curation, Methodology, Software, Writing – review & editing. HD: Data curation, Formal analysis, Software, Writing – review & editing. XS: Formal analysis, Writing – review & editing. CC: Formal analysis, Software, Writing – review & editing. BaJ: Formal analysis, Investigation, Software, Writing – original draft, Writing – review & editing. TL: Conceptualization, Formal analysis, Resources, Software, Validation, Writing – original draft. BoJ: Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the National Key Research and Development Program of China (2021YFC2302003), the Science and Technology Development Funds for Shandong Medical and Health (202112060725, 202412060875), the Key Research and Development Funds for Jining Medical and Health (2023YXNS130), and the Shandong Province Medical Staff Science and Technology Innovation Program (SDYWZGKCJHLH202467).

References

Azage, M., Kumie, A., Worku, A., C Bagtzoglou, A., and Anagnostou, E. (2017). Effect of climatic variability on childhood diarrhea and its high risk periods in northwestern parts of Ethiopia. *PLoS ONE* 12:e0186933. doi: 10.1371/journal.pone.0186933

Cheng, C. C., Ke, G. M., Chu, P. Y., and Ke, L. Y. (2023). Elucidating the implications of norovirus N- and O-Glycosylation, O-GlcNAcylation, and phosphorylation. *Viruses* 15:798. doi: 10.3390/v15030798

Chhabra, P., de Graaf, M., Parra, G. I., Chan, M. C., Green, K., Martella, V., et al. (2019). Updated classification of norovirus genogroups and genotypes. *Journal of general virology* 100, 1393–1406. doi: 10.1099/jgv.0.001318

Chigor, V. N., Digwo, D. C., Adediji, A., Chidebelu, P. E., Chigor, C. B., Ugwu, K. O., et al. (2024). Epidemiology of norovirus infection in Nigeria: a systematic review and meta-analysis. *Arch. Virol.* 169:138. doi: 10.1007/s00705-024-06056-w

Cohen, A. L., Platts-Mills, J. A., Nakamura, T., Operario, D. J., Antoni, S., Mwenda, J. M., et al. (2022). Aetiology and incidence of diarrhoea requiring hospitalisation in children under 5 years of age in 28 low-income and middle-income countries: findings from the Global Pediatric Diarrhea Surveillance network. *BMJ Global Health* 7:e009548. doi: 10.1136/bmjgh-2022-009548

Daughenbaugh, K. F., Fraser, C. S., Hershey, J. W., and Hardy, M. E. (2003). The genome-linked protein VPg of the Norwalk virus binds eIF3, suggesting its role in translation initiation complex recruitment. *EMBO J.* 22, 2852–2859. doi: 10.1093/emboj/cdg251

Daughenbaugh, K. F., Wobus, C. E., and Hardy, M. E. (2006). VPg of murine norovirus binds translation initiation factors in infected cells. *Virol. J.* 3:33. doi: 10.1186/1743-422X-3-33

de Man, H., van den Berg, H. H., Leenen, E. J., Schijven, J. F., Schets, F. M., van der Vliet, J. C., et al. (2014). Quantitative assessment of infection risk from exposure to waterborne pathogens in urban floodwater. *Water Res.* 48, 90–99. doi: 10.1016/j.watres.2013.09.022

Deval, J., Jin, Z., Chuang, Y. C., and Kao, C. C. (2017). Structure(s), function(s), and inhibition of the RNA-dependent RNA polymerase of noroviruses. *Virus Res.* 234, 21–33. doi: 10.1016/j.virusres.2016.12.018

Dhimal, M., Bhandari, D., Karki, K. B., Shrestha, S. L., Khanal, M., Shrestha, R. R. P., et al. (2022). Effects of climatic factors on diarrheal diseases among children below 5 years of age at national and subnational levels in Nepal: an ecological study. *Int. J. Environ. Res. Public Health* 19:6138. doi: 10.3390/ijerph19106138

Eden, J. S., Sharpe, L. J., White, P. A., and Brown, A. J. (2011). Norovirus RNAdependent RNA polymerase is phosphorylated by an important survival kinase, Akt. J. Virol. 85, 10894–10898. doi: 10.1128/JVI.05562-11

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Fullerton, S. W., Blaschke, M., Coutard, B., Gebhardt, J., Gorbalenya, A., Canard, B., et al. (2007). Structural and functional characterization of sapovirus RNA-dependent RNA polymerase. *J. Virol.* 81, 1858–1871. doi: 10.1128/JVI.01462-06

GBD 2021 Diarrhoeal Diseases Collaborators (2024). Global, regional, and national age-sex-specific burden of diarrhoeal diseases, their risk factors, and aetiologies, 1990-2021, for 204 countries and territories: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet Infect. Dis.* 18:S1473. doi: 10.1016/S1473-3099(24)00691-1

Hung, C. H., Yen, J. B., Chang, P. J., Chen, L. W., Huang, T. Y., Tsai, W. J., et al. (2023). Characterization of human norovirus nonstructural protein NS1.2 involved in the induction of the filamentous endoplasmic reticulum, enlarged lipid droplets, LC3 Recruitment, and Interaction with NTPase and NS4. *Viruses* 15:812. doi: 10.3390/v15030812

Jiao, Y., Han, T., Qi, X., Gao, Y., Zhao, J., Zhang, Y., et al. (2023). Genotypes diversity of acute gastroenteritis outbreaks caused by human sapovirus - Beijing Municipality, China, 2015-2021. *China CDC Weekly* 5, 625–631. doi: 10.46234/ccdcw2023.119

Jin, M., Wu, S., Kong, X., Xie, H., Fu, J., He, Y., et al. (2020). Norovirus outbreak surveillance, China, 2016-2018. *Emer. Infect. Dis.* 26, 437–445. doi: 10.3201/eid2603.191183

Joosten, R., Sonder, G., Parkkali, S., Brandwagt, D., Fanoy, E., Mughini-Gras, L., et al. (2017). Risk factors for gastroenteritis associated with canal swimming in two cities in the Netherlands during the summer of 2015: a prospective study. *PLoS ONE* 12:e0174732. doi: 10.1371/journal.pone.0174732

Kimura-Someya, T., Katsura, K., Kato-Murayama, M., Hosaka, T., Uchikubo-Kamo, T., Ihara, K., et al. (2024). Structural analyses of the GI.4 norovirus by cryoelectron microscopy and X-ray crystallography revealing binding sites for human monoclonal antibodies. J. Virol. 98:e0019724. doi: 10.1128/jvi.00197-24

Leen, E. N., Kwok, K. Y., Birtley, J. R., Simpson, P. J., Subba-Reddy, C. V., Chaudhry, Y., et al. (2013). Structures of the compact helical core domains of feline calicivirus and murine norovirus VPg proteins. *J. Virol.* 87, 5318–5330. doi: 10.1128/JVI.03151-12

Leen, E. N., Sorgeloos, F., Correia, S., Chaudhry, Y., Cannac, F., Pastore, C., et al. (2016). A Conserved Interaction between a C-Terminal Motif in Norovirus VPg and the HEAT-1 Domain of eIF4G Is Essential for Translation Initiation. *PLoS Pathog.* 12:e1005379. doi: 10.1371/journal.ppat.1005379

Levy, K., Woster, A. P., Goldstein, R. S., and Carlton, E. J. (2016). Untangling the impacts of climate change on waterborne diseases: a systematic review of relationships between diarrheal diseases and temperature, rainfall, flooding, and drought. *Environ. Sci. Technol.* 50, 4905–4922. doi: 10.1021/acs.est.5b06186

Li, T. C., Kataoka, M., Doan, Y. H., Saito, H., Takagi, H., Muramatsu, M., et al. (2022). Characterization of a human sapovirus genotype GII.3 strain generated by a

reverse genetics system: VP2 is a minor structural protein of the virion. *Viruses* 14:1649. doi: 10.3390/v14081649

Li, T. F., Hosmillo, M., Schwanke, H., Shu, T., Wang, Z., Yin, L., et al. (2018). Human norovirus NS3 Has RNA helicase and chaperoning activities. *J. Virol.* 92:17. doi: 10.1128/JVI.01606-17

Lin, S. C., Bai, G. H., Lin, P. C., Chen, C. Y., Hsu, Y. H., Lee, Y. C., et al. (2023). Molecular and genetics-based systems for tracing the evolution and exploring the mechanisms of human norovirus infections. *Int. J. Mol. Sci.* 24:9093. doi: 10.3390/ijms24109093

Mallagaray, A., Creutznacher, R., Dulfer, J., Mayer, P. H. O., Grimm, L. L., Orduna, J. M., et al. (2019). A post-translational modification of human Norovirus capsid protein attenuates glycan binding. *Nature Commun.* 10:1320. doi: 10.1038/s41467-019-09251-5

Oka, T., Mori, K., Iritani, N., Harada, S., Ueki, Y., Iizuka, S., et al. (2012). Human sapovirus classification based on complete capsid nucleotide sequences. *Arch. Virol.* 157, 349–352. doi: 10.1007/s00705-011-1161-2

Oka, T., Wang, Q., Katayama, K., and Saif, L. J. (2015). Comprehensive review of human sapoviruses. *Clin. Microbiol. Rev.* 28, 32–53. doi: 10.1128/CMR.00011-14

Oka, T., Yamamoto, M., Yokoyama, M., Ogawa, S., Hansman, G. S., Katayama, K., et al. (2007). Highly conserved configuration of catalytic amino acid residues among calicivirus-encoded proteases. *J. Virol.* 81, 6798–6806. doi: 10.1128/JVI.02840-06

Olspert, A., Hosmillo, M., Chaudhry, Y., Peil, L., Truve, E., Goodfellow, I., et al. (2016). Protein-RNA linkage and posttranslational modifications of feline calicivirus and murine norovirus VPg proteins. *PeerJ* 4:e2134. doi: 10.7717/peerj.2134

Pang, X. L., Honma, S., Nakata, S., and Vesikari, T. (2000). Human caliciviruses in acute gastroenteritis of young children in the community. *J. Infect. Dis.* 181, S288–294. doi: 10.1086/315590

Robel, I., Gebhardt, J., Mesters, J. R., Gorbalenya, A., Coutard, B., Canard, B., et al. (2008). Functional characterization of the cleavage specificity of the sapovirus chymotrypsin-like protease. *J. Virol.* 82, 8085–8093. doi: 10.1128/JVI.00693-08

Sharp, T. M., Crawford, S. E., Ajami, N. J., Neill, F. H., Atmar, R. L., Katayama, K., et al. (2012). Secretory pathway antagonism by calicivirus homologues of Norwalk virus nonstructural protein p22 is restricted to noroviruses. *Virol. J.* 9:181. doi: 10.1186/1743-422X-9-181

Sharp, T. M., Guix, S., Katayama, K., Crawford, S. E., and Estes, M. K. (2010). Inhibition of cellular protein secretion by norwalk virus nonstructural protein p22 requires a mimic of an endoplasmic reticulum export signal. *PLoS ONE* 5:e13130. doi: 10.1371/journal.pone.0013130

Su, L., Mao, H., Sun, Y., Yan, H., Ge, Q., Gong, L., et al. (2023). The analysis of the genotype of Sapovirus outbreaks in Zhejiang Province. *Virol. J.* 20:268. doi: 10.1186/s12985-023-02202-z

Subba-Reddy, C. V., Goodfellow, I., and Kao, C. C. (2011). VPg-primed RNA synthesis of norovirus RNA-dependent RNA polymerases by using a novel cell-based assay. *J. Virol.* 85, 13027–13037. doi: 10.1128/JVI.06191-11

Tan, M., and Jiang, X. (2010). Norovirus gastroenteritis, carbohydrate receptors, and animal models. *PLoS Pathog.* 6:e1000983. doi: 10.1371/journal.ppat.1000983

Winder, N., Gohar, S., and Muthana, M. (2022). Norovirus: an overview of virology and preventative measures. *Viruses* 14:2811. doi: 10.3390/v14122811

Wu, C., Fu, Z., Xie, C., Zhao, J., He, F., Jiao, B., et al. (2024). Epidemiological Characteristics and Genotypic Features of Rotavirus and Norovirus in Jining City, 2021-2022. *Viruses* 16:925. doi: 10.3390/v16060925

Yen, J. B., Chen, L. W., Wei, L. H., Hung, C. H., Wang, S. S., Lin, C. L., et al. (2021). Identification and characterization of human norovirus NTPase regions required for lipid droplet localization, cellular apoptosis, and interaction with the viral P22 protein. *Microbiol. Spectr.* 9:e0042221. doi: 10.1128/Spectrum. 00422-21

Yinda, C. K., Conceicao-Neto, N., Zeller, M., Heylen, E., Maes, P., Ghogomu, S. M., et al. (2017). Novel highly divergent sapoviruses detected by metagenomics analysis in straw-colored fruit bats in Cameroon. *Emer. Microbes Infect.* 6:e38. doi: 10.1038/emi.2017.20

Zamyatkin, D. F., Parra, F., Alonso, J. M., Harki, D. A., Peterson, B. R., Grochulski, P., et al. (2008). Structural insights into mechanisms of catalysis and inhibition in Norwalk virus polymerase. *J. Biol. Chem.* 283, 7705–7712. doi: 10.1074/jbc.M709563200

Zamyatkin, D. F., Parra, F., Machin, A., Grochulski, P., and Ng, K. K. (2009). Binding of 2/-amino-2/-deoxycytidine-5/-triphosphate to norovirus polymerase induces rearrangement of the active site. *J. Mol. Biol.* 390, 10–16. doi: 10.1016/j.jmb.2009.04.069

Zhao, W., Gao, Z., Guo, C., Zhang, Y., Zhang, Y., Wang, Q., et al. (2024). dual typing system establishment and global diversity analysis for sapoviruses. *BMC Genomics* 25:1131. doi: 10.1186/s12864-024-11048-w