Epidemiol. Infect. (2017), **145**, 2231–2240. © Cambridge University Press 2017
doi:10.1017/S0950268817001273

Full-genome sequence analysis of an uncommon norovirus genotype, GII.21, from South Korea

J. W. LEE[1], S. G. LEE[2], H. G. CHO[3], Y. J. PARK[4], J. W. YUN[5], C. J. PARK[5] AND S. Y. PAIK[1]*

[1] *Department of Microbiology, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea*
[2] *Korea Zoonosis Research Institute, Chonbuk National University, Iksan, Republic of Korea*
[3] *Division of Public Health Research, Gyeonggi Province Institute of Health and Environment, Suwon, Republic of Korea*
[4] *Department of Laboratory Medicine, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea*
[5] *Department of Dermatology, Bucheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea*

Received 16 November 2016; Final revision 21 May 2017; Accepted 31 May 2017; first published online 27 June 2017

SUMMARY

Noroviruses (NoVs) are major causal agents of acute gastroenteritis in humans. NoV GII.4 is the predominant genotype globally. However, uncommon and minor types of NoVs are consistently detected and some have been shown to dominate over GII.4. Therefore, the prevalence of dominant and uncommon NoVs makes the identification of these viruses important for the prediction and prevention of pandemics. In this study, the full-genome sequence of a NoV (strain JW) detected in Korea was extensively characterized. The full-length genome was 7510 nucleotides long, and phylogenetic analysis based on the whole-genome sequences, including open reading frame (ORF)1, ORF2, and ORF3, indicated that it belonged to the GII.21 genotype. Strain JW showed maximum identity with strain YO284; however, comparison of the amino acid sequence of ORF2, which functions as an antigen, showed substitutions in several amino acids. GII.21 is not a prevalent epidemiological agent of acute gastroenteritis in humans, but it is consistently found in gastroenteritis patients from several countries. The present study provides the first full-genome sequence analysis of NoV GII.21 isolated from a patient in Korea. Our findings provide not only valuable genome information but also data for epidemiology studies, epidemic prevention, and vaccine development strategies.

Key words: Norovirus, NoV, sequence analysis, uncommon NoV, GII.21.

INTRODUCTION

Since noroviruses (NoVs) were first identified as causal agents of viral gastroenteritis, they have been found to be the most common cause of epidemic gastroenteritis in all age groups across the globe and to be the leading cause of foodborne illness [1–3]. NoVs are highly infectious and constitute a major cause of gastroenteritis-related hospitalization [4]. In the USA alone, the annual disease incidence due to NoVs leads to about 570–800 deaths, 56 000–71 000 hospitalizations, 400 000 emergency department visits, 1·7–1·9 million outpatient visits, and 19–21 million total illnesses [5]. Furthermore,

* Author for correspondence: S. Y. Paik, Department of Microbiology, College of Medicine, The Catholic University of Korea, 222, Banpo-daero, Seocho-gu, Seoul 06591, Republic of Korea. (Email: paik@catholic.ac.kr)

outbreaks of acute gastroenteritis due to NoVs pose a considerable economic burden as a result of healthcare costs and lost productivity [6, 7]. In the USA, the cost of hospitalization for NoV illnesses has been estimated at approximately $500 million, and the cost of healthcare and lost productivity has been estimated at $2 billion [8]. Studies to reduce the burden of NoV illness have been hindered by the lack of complete understanding of NoVs. Part of this inadequate understanding of NoVs is attributable to the lack of *in vitro* cell-culture systems and *in vivo* small-animal models for human NoVs [9].

The typical transmission of NoV is directly via person-to-person transmission (fecal–oral and vomit–oral routes) or indirectly through waterborne, foodborne, and environmental transmission [10]. The average incubation period is 24–48 h, and symptoms include vomiting ($\geqslant$50% of cases), diarrhea, nausea, abdominal cramps, malaise, and low-grade fever. The illness usually resolves in 12–72 h but can last longer in young children, elderly people, hospitalized patients, and immunocompromised people [9].

NoVs belong to the genus *Norovirus* in the *Caliciviridae* family. The viral genome consists of a positive-sense single-stranded RNA about ~7·7 kb long, comprising three open reading frames (ORFs). The genome is protein-linked at the 5′-end and polyadenylated at the 3′-end. ORF1 encodes the non-structural proteins p48, NTPase, p22, VPg, 3C-like protease, and RNA-dependent RNA polymerase. ORF2 encodes the major capsid protein VP1, and ORF3 encodes the minor structural protein VP2. VP1 is a structural protein containing antigenic determinants that define strain specificity. VP1 is comprised of three domains, the *N*-terminal; the shell domain, *S*; and the protruding domain, *P*. The *P* domain, which forms the protrusions on the virus, is divided into two subdomains, *P*1 and *P*2. The *P*2 domain, which is located at the distal surface, is a hypervariable region that plays an important role in receptor binding and immune reactivity [11]. NoVs are divided into seven genogroups (GI, GII, GIII, GIV, GV, GVI, and GVII) and are classified into over 40 genotypes. Among them, three genogroups, GI, GII, and GIV, are known to infect humans [12, 13].

At present, GII.4-type viruses are believed to be the major cause of gastroenteritis in humans, being responsible for >70% of the outbreaks [14, 15]. Although GII.21 is not a dominant type of NoV, it is consistently detected in hospitalized patients with gastroenteritis and environments in many countries [14, 16–24]. In some countries, NoV GII.21 is a dominant cause of gastroenteritis [25]. In Korea, NoV GII.21 was found during the winter season of 2012–2013 [26].

The purpose of this study was to characterize an uncommon type of NoV GII.21 isolated from a stool sample obtained from a Korean patient with acute gastroenteritis.

MATERIALS AND METHODS

Ethics statement

The stool sample for the sequencing study was provided by the Waterborne Virus Bank (WAVA, Seoul, South Korea). Because of difficulties in tracking the exact records of the pediatric patient from the donor hospital, informed consent from the parent of the said patient could not be acquired. The Institutional Review Board reviewed and approved the use of this sample for the purpose of research, as this study does not affect the patient. All the experimental work and sample collections were supervised by the Catholic Medical Center Office of the Human Research Protection Program (CMC OHRP) of South Korea (approval number MC15EASE0117).

Specimen preparation and viral RNA extraction

One hundred and ninety-five stool samples obtained from patients who presented with fever and diarrhea from January 2013 to December 2013 were screened from the WAVA. Most stool samples were from hospitalized infants ($\leqslant$3 years old), and the GII.21 sample used in this study was also from an infant patient (1 year old). The NoV-positive GII.21 sample obtained from the WAVA was stored at −70 °C until RNA extraction. The frozen stool sample was thawed and diluted with 10% phosphate-buffered saline, after which it was centrifuged. NoV RNA was extracted from the supernatant using a QIAamp Viral RNA mini kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Isolated RNA was used as a template for reverse transcription–polymerase chain reaction (RT-PCR) and stored at −70 °C until further use.

RT–PCR

For the detection of NoV-positive samples, RT–PCR was performed with the OneStep RT–PCR kit (Qiagen) using GII-F1M and GII-R1M primers

https://doi.org/10.1017/S0950268817001273 Published online by Cambridge University Press

(Table 1). To analyze the whole-genome sequence of NoV, nine more primer pairs were newly designed on the basis of the YO284 strain (GenBank accession number KJ196284) (Table 1). RT–PCR was performed with an S1000 thermal cycler (Bio-Rad, Hercules, California, USA), and the steps comprised reverse transcription (50 °C for 30 min), initial PCR activation (95 °C for 15 min), 30 cycles of three-step cycling (94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min), and final extension (72 °C for 10 min). All RT-PCR products were examined by electrophoresis in ethidium bromide-stained 1·5% agarose gels.

Determination of the 5′- and 3′-ends of NoV genomic RNA

To determine the 5′-end of NoV genomic RNA, rapid amplification of cDNA ends (RACE) was performed with the 5′ RACE System for Rapid Amplification of cDNA Ends Version 2.0 Kit according to the manufacturer's recommendations (Invitrogen, Carlsbad, California, USA). Three primers (GSP1, GSP2, and nested GSP) were designed based on the ORF1 sequence for 5′-end RACE PCR (Table 1). To obtain the exact sequence of the 3′-end of the NoV genomic RNA, cDNA was synthesized using reverse transcription with 3′-oligo (dT)-anchor-R (Table 1). The second PCR was conducted using the ORF3-F and 3′-anchor-R primers (Table 1) under the following conditions: 30 cycles of three-step cycling (94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min) and 72 °C for 10 min.

Cloning and sequencing of the complete genome

All PCR products obtained were extracted using the HiYield Gel/PCR DNA Fragments Extraction Kit (RBC, Taipei, Taiwan) and were cloned into pGEM-T Easy Vectors (Promega, Madison, Wisconsin, USA). The cloned vector was transformed into *Escherichia coli* DH5α competent cells (RBC) according to the manufacturer's instructions and was selected from Luria–Bertani (LB) agar plates (Duchefa, Haarlem, The Netherlands) containing 40 mg/ml X-gal, 0·1 mM isopropyl-β-D-thiogalactoside, and 50 mg/ml ampicillin at 37 °C for 16–18 h. Selected clones were inoculated in LB broth (Duchefa) and incubated overnight in a shaking incubator (IS-971R, Jeiotech, Daejeon, South Korea) at 37 °C and 200 rpm. Plasmid DNA was purified using the HiYield Plasmid Mini Kit (RBC) and sequenced (Cosmo Genetech, Seoul, South Korea). The sequencing results were analyzed using BLAST (National Center for Biotechnology Information, NCBI).

Phylogenetic analysis

Comparative sequence analysis, including sequence alignments and estimation of genetic distances, was performed with Clustal W using the Molecular Evolutionary Genetic Analysis software (MEGA soft version 6.0) [27]. Phylogenetic trees were constructed using the neighbor-joining method with a Kimura two-parameter model in MEGA [28], and branch support was calculated based on 1000 bootstrap replicates. The complete genome sequences and partial genome sequences were collected from NCBI.

RESULTS

NoV detection and full-length genome of strain JW

Of the 195 stool samples screened, 56 (28·72%) were positive for NoV genogroup II. Only one sample (1/56; 1·79%) was a GII.21-type NoV. The full-length genome of strain JW (GenBank accession number KX079488) was 7510 nucleotides (nt) long and contained three ORFs: ORF1 (5–5104; 5100 nt), ORF2 (5085–6707; 1623 nt), and ORF3 (6707–7465; 759 nt). ORF1 and ORF2 had an overlap of 20 nt, whereas ORF2 and ORF3 had a single-nucleotide overlap. Genotyping by the online Norovirus Genotyping Tool (http://www.rivm.nl/mpf/norovirus/typingtool) [29] showed that strain JW belonged to genogroup II and genotype GII.P21 (by ORF1) or GII.21 (by ORF2). BLAST results with the full-length genome sequence of strain JW confirmed the genotyping result and showed the highest identity with strain YO284 (GenBank accession number KJ196284; query cover =99% and identity =96%). ORF1 and ORF3 BLAST results also showed the highest identity with strain YO284 (GenBank accession number KJ196284; query covers, 100% each; identities, 97% and 96%, respectively). ORF2 BLAST results showed the highest identity with strain CUHK-NS-626 (GenBank accession number KR921942; query cover =100% and identity =99%).

Phylogenetic analysis

We next attempted to understand the genetic relationship between strain JW and other published full-length genomes of NoV reference strains, which are available in GenBank. As shown in the tree, the whole-genome

https://doi.org/10.1017/S0950268817001273 Published online by Cambridge University Press

Table 1. *Primers used in this study*

Primer	Sequence (5′→3′)	Location	Size (bp)	Polarity	Reference
GSP1	CAG CCA TGA ATC GTT CAA GTC AAG A	640–664		–	
GSP2	GTC AAA GGC TGT GTA AGG GAA C	589–610		–	This study
Nested GSP	CAG GGG AAA TGA GGT ACT GTG G	539–560		–	
ORF1-1F	ATG GCG TCT AAC GAC GCT TC	11–30	940	+	
ORF1-1R	CAT CTG GGG GTG TCC AGA AAA C	929–950		–	
ORF1-2F	CCT AAA TAT CCT CAA CAT CCT AGC CA	841–866	937	+	
ORF1-2R	GGT GGT TAT TAT TAT GGC GTC ACT GTC	1751–1777		–	
ORF1-3F	GCC CCC TAA CAC TTA ATT GTG A	1701–1722	922	+	
ORF1-3R	TTG ATG TCA TCA GAG GAG ATG A	2601–2622		–	
ORF1-4F	CAA GTG GAA GAT TTG GAG GAG A	2531–2552	951	+	
ORF1-4R	GTA GTC GTT CCC CCT CTT GTA	3461–3481		–	
ORF1-5F	GAG TAT GGA TTT GGG CAC AAC TC	3412–3434	951	+	This study*
ORF1-5R	TTG ACC ATT ATT TCC AGG GCT G	4341–4362		–	
ORF1-6F	GAC TCG ACA CAG CAA AGA GCA G	4310–4331	856	+	
ORF1-6R	CAA GGT CTC AGT GTT GAT CTC TG	5143–5165		–	
ORF2-1F	ATG AAG ATG GCG TCG AAT GA	5085–5104	727	+	
ORF2-1R	GAA ATC TAG AGT TGG TCA ACT C	5790–5811		–	
ORF2-2F	CTG TGG AGT CTA AAA CCA AAC	5740–5760	965	+	
ORF2-2R	TTA TTG AAT TCT CCT GCG CC	6685–6704		–	
ORF3-F	GGC AGG GAT TTA TCA CCA TAG C	6559–6580		+	
3′-Oligo (dT)-anchor-R	CAA TGA GGT TAT GGC TTT GGA ACT TTT TTT TTT TTT TT	3′-end poly A tail		–	This study
3′-Anchor-R	CAA TGA GGT TAT GGC TTT GGA AC			–	
GII-F1M	GGG AGG GCG ATC GCA ATC T	5049–5067	341	+	[47]†
GII-R1M	CCR CCT GCA TRI CCR TTR TAC AT	5367–5389		–	

* The primers were based on the YO284 strain (GenBank accession number KJ196284).
† The primers were for NoV detection (GenBank accession number X86557).

sequences of strain JW were classified as NoV GII.21. The full-length nucleotide sequences of strain JW showed maximum identity with those of strain YO284 (GenBank accession number KJ196284), with 96·3% similarity (Fig. 1*a*). The tree based on the amino acid sequences of ORF1, ORF2, and ORF3 showed 98·6%, 96·3%, and 94·9% identities, respectively, with strain YO284 (Fig. 1*b–d*).

Phylogenetic analysis was performed using the partial VP1 genes of strain JW and other GII.21 strains available in GenBank. Strain JW was confirmed to cluster in GII.21.b1 (Fig. 2*a*) [30], which comprises strains detected in water, clinical samples, and seafood. Strain JW showed a very high identity (95–97%) with other strains clustering in GII.21.b1.

Alignment analysis of amino acid sequences of VP1

Figure 2*b* shows the comparison of partial VP1 sequences that indicate amino acid substitution patterns in *S* domain of VP1, which represents the most conserved region of the sequence that is used for genotyping [31, 32]. In the GII.21.a cluster, seven strains had the Ala36Val substitution, six strains had the Asn45Thr substitution, and five strains had the Ile47Met substitution. In the GII.21.b2 cluster, five strains had the Asn54Thr or Asn54Ser substitution (Fig. 2*b*). The consensus sequence was based on the corresponding amino acid sequence of the most highly conserved sequence.

The functional VP1 protein of NoV, which functions as an antigen and determines the specific genotype, was analyzed [32, 33]. Particularly, the *P*2 subdomain of VP1 (residues 279–405), is the highly variable and most exposed region of the structure as an antigen [31]. Five strains were used for alignment analysis of the amino acid sequences. A total of 19 amino acid alterations (red arrow) were seen in the recently collected strains (cluster B, Fig. 3). Among these, three amino acid substitutions and one amino acid insertion (in red box) were located within the *P*1 subdomain, one amino acid substitution was located in the *S* domain, and 14 amino acid alterations were located within the *P*2 subdomain. A comparison of VP1 amino acid sequences of previously collected strains (cluster A) and the recently collected strains (cluster B)

https://doi.org/10.1017/S0950268817001273 Published online by Cambridge University Press

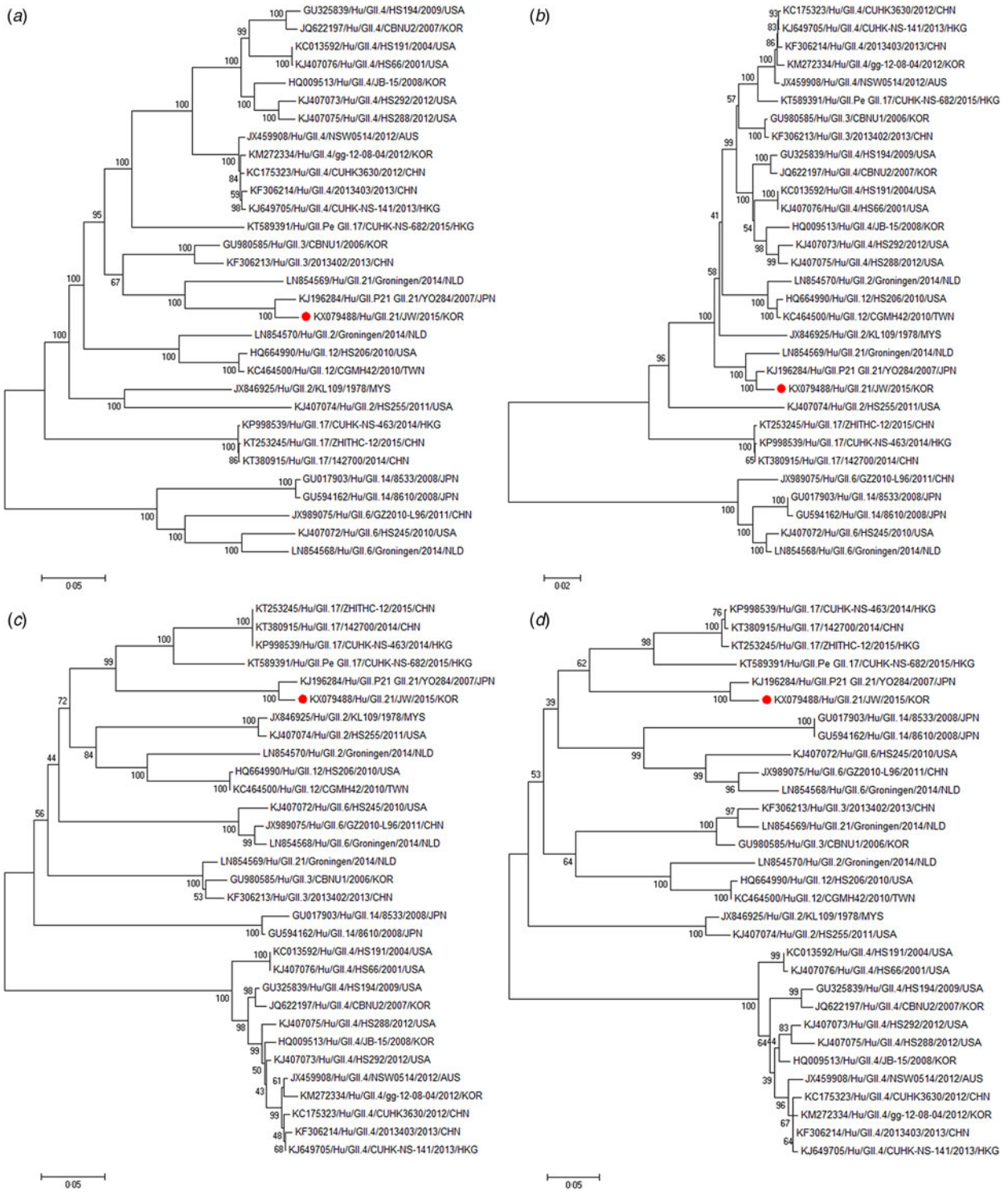

Fig. 1. Phylogenetic analysis of NoV based on nucleotide and amino acid sequences. The neighbor-joining method in MEGA was used to construct phylogenetic trees. The numbers associated with each branch indicate the bootstrap values for the genotype. Phylogenetic trees based on (*a*) full-length nucleotide sequence, (*b*) amino acid sequence of ORF1, (*c*) amino acid sequence of ORF2, and (*d*) amino acid sequence of ORF3. Strain JW is highlighted with a solid red circle. The representative strains are named by 'accession number/host/genogroup' and 'genotype/strain/collection date/country'. Country codes: CHN, China; HKG, Hong Kong; JPN, Japan; USA, United States of America; AUS, Australia; TWN, Taiwan; NLD, Netherlands; MYS, Malaysia; KOR, Korea.

https://doi.org/10.1017/S0950268817001273 Published online by Cambridge University Press

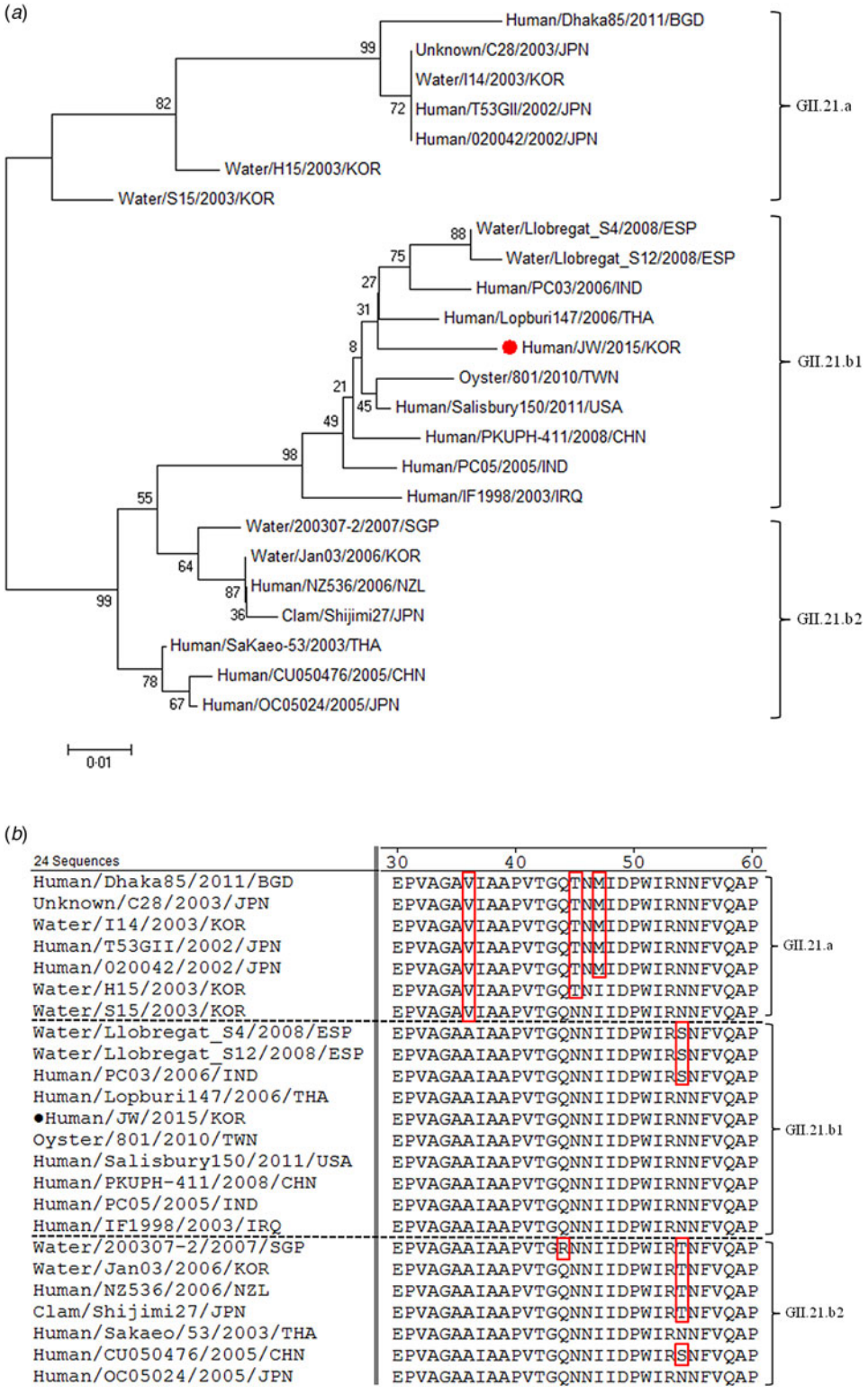

Fig. 2. Phylogenetic analysis and amino acid sequence comparison of partial ORF2 gene. (*a*) Neighbor-joining phylogenetic analysis of the partial nucleotide sequences of the VP1 genes for strain JW (solid red circle). Strain JW clustered in GII.21.b1. (*b*) The comparison of the partial amino acid sequence of VP1 genes is shown as amino acid substitution patterns. The consensus sequence was based on the corresponding amino acid sequence that was most highly conserved. The substituted amino acids are in red boxes. The representative strains are named by 'isolation source/strain/collection date/country'. Country codes: CHN, China; JPN, Japan; USA, United States of America; TWN, Taiwan; BGD, Bangladesh; THA, Thailand; IND, India; IRQ, Iraq; SGP, Singapore; NZL, New Zealand; ESP, Spain; KOR, Korea.

https://doi.org/10.1017/S0950268817001273 Published online by Cambridge University Press

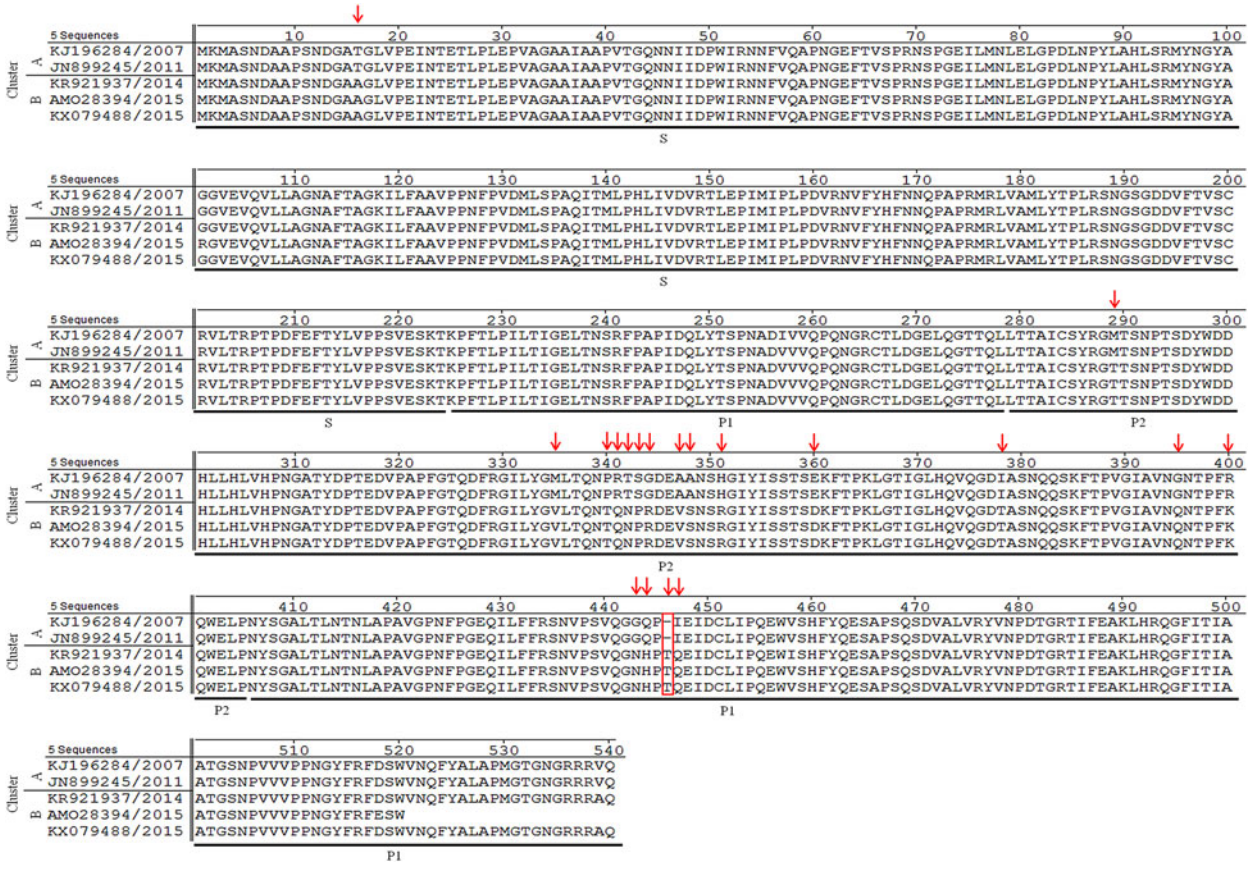

Fig. 3. Comparison of amino acid substitutions of five strains. Alignment of VP1 amino acid sequences of NoV strains YO284 (GenBank accession number KJ196284), Salisbury150 (GenBank accession number JN899245), CUHK-NS-293 (GenBank accession number KR921937), GL02BLPV2 (GenBank accession number AMO28394), and JW (GenBank accession number KX079488). A total of 19 amino acid alterations (red arrows) were noted in the previously (cluster A) and recently collected strains (cluster B). Most substitutions were present within the hypervariable *P*2 subdomain (residues 279–405).

indicated that most substitutions were within the hypervariable region of the *P*2 subdomain (Fig. 3).

The HBGA-binding pocket is conserved in NoV GII.21

The GII.21 lineage has major residues that form the novel histo-blood group antigen (HBGA)-binding pocket (containing the *B*, *T*, and *N* loops) and the conventional GII-binding interface (containing *P*, *S*, and *A* loops) [34]. The alignment of HBGA-binding pocket amino acid sequences in strain JW and three other strains showed very high identity. However, two amino acids were substituted in the *N* and *T* loops (in red box). The conserved residues forming the binding interface are shown (Fig. 4).

DISCUSSION

NoVs are a crucial cause of viral gastroenteritis-related hospitalization [4]. Clarity on the virus and its pathogenesis and epidemiology is affected by the lack of *in vitro* cell-culture systems and *in vivo* small-animal models for human NoVs. In-depth information on this virus is essential for reducing the burden of NoV illness and hospitalization costs [9]. The molecular epidemiology of emerging NoV strains is of particular interest to researchers because NoV recombination, variants, and mutations occur frequently [35, 36]. Furthermore, uncommon and minor types of NoVs, which has the potential to dominate, have been consistently detected. In this study, NoV was detected in 64 out of 195 clinical samples from Korea, and positive samples were classified by genotyping using BLAST. The NoV genotypes GI.4 ($n = 6$), GI.6 ($n = 2$), GII.2 ($n = 1$), GII.3 ($n = 2$), GII.4 ($n = 43$), GII.6 ($n = 3$), GII.16 ($n = 1$), GII.17 ($n = 5$), and GII.21 ($n = 1$) were identified. Although GII.21 may have been consistently detected in many countries, it has not been detected as a dominant cause of gastroenteritis in certain countries [14, 16–26].

https://doi.org/10.1017/S0950268817001273 Published online by Cambridge University Press

4 Sequences		B-loop * * **		N-loop * *		T-loop * * **
AFC89665/Hu/GII.21/Salisbury150/2011/USA	291	SNPTSDYWD	356	SSTSEKFT	394	NGNTPF
AY675554/Hu/GII.21/IF1998/2003/IRQ	289	SNPTRDYWD	354	SSTSEKFT	392	NGNTPF
KJ196284/Hu/GII.P21_GI.21/YO284/2007/JPN	291	SNPTSDYWD	356	SSTSEKFT	394	NGNTPF
KX079488/Hu/GII.21/JW/2015/KOR	291	SNPTSDYWD	356	SSTSDKFT	394	NQNTPF

Fig. 4. Amino acid sequence alignment of the surface loops of the *P* domains forming HBGA-binding interfaces. Sequences of the three surface loops on the HBGA-binding interfaces (*B*, *N*, and *T* loops) representing the genotype of the GII.21 [34]. The conserved residues forming the binding interface are indicated with asterisks. Two substituted amino acids in the *N* and *T* loops are shown in red boxes. The representative strains were named by 'accession number/host/genogroup' and 'genotype/strain/collection date/country'. Country codes: JPN, Japan; USA, United States of America; IRQ, Iraq; KOR, Korea.

The data generated from the current full-genome sequencing are expected to be helpful in the disease surveillance of sporadic gastroenteritis caused by non-dominant strains, such as NoV GII.21. BLAST results of the full-genome sequence of strain JW confirmed it to be a GII.21-type NoV, and ORF1, ORF2, and ORF3 genes showed very high similarity with other NoV GII.21 strains. Phylogenetic analysis based on the full-length genomic sequence and ORF1, ORF2, and ORF3 sequences showed the highest identity with NoV GII.21 (strain YO284). The amino acid sequences of ORF1, ORF2, and ORF3 also indicated the highest identity with strain YO284 (Fig. 1*a*–*d*). However, alignment analysis of amino acid sequences of the VP1 region showed 21 sequence substitutions; of these, 14 amino acids were included in the *P*2 subdomain. In addition, alignment analysis of amino acid sequences of the five former strains and the recent strains indicated that most substitutions were present within the *P*2 subdomain (Fig. 3). The *P*2 subdomain is the highly variable part and the most exposed region of the structure [31]. This suggests the possible formation of different epitopes or different protein structures because of a high rate of substitution in the *P*2 subdomain even among strains with high identity.

NoV GII.21 strains identified in Bangladesh were divided into three clusters based on the partial sequence of the *S* domain of the VP1 gene, which determines NoV genotype [30]. Strain JW was included in GII.21.b1, and each cluster showed different amino acid substitution patterns (Fig. 2*a*, *b*).

The *P* domain of VP1 interacts with HBGA in a strain-specific manner. GII.21 is divided into unique lineages on the basis of the binding interface with HBGAs, which is distinct from the conventional GII-binding interface [34]. Strain JW and three other GII.21 lineage strains had very high amino acid sequence similarities in *B*, *N*, and *T* loops, and the residues appeared to be well conserved. However, two amino acid sites were found to be substituted (Fig. 4). Such amino acid substitutions resulting from NoV point mutations occur frequently [35].

Eight strains were isolated from the Hong Kong group in 2014–2015 (GenBank accession numbers KR921935, KR921936, KR921937, KR921938, KR921939, KR921940, KR921941, KR921942), and their alignment was analyzed with the full sequence of the ORF2 region. The strains were isolated around the same time as strain JW. The alignment of the amino acid sequence of the ORF2 region shows 100% identity with that of strain JW, and only one strain (GenBank accession number KR921940) showed one amino acid difference (data not shown). However, at the nucleotide level, only 2 nt differed between strain JW and strain CUHK-NS-626 (GenBank accession number KR921942). The third nucleotide of the codon was altered; however, this did not affect the encoded amino acid. NoV GII.21 isolated around the same time in Korea and Hong Kong exhibited high capsid protein sequence identity. In Korea, 10 strains were isolated from stream water in 2015 (GenBank accession numbers AMO28393, AMO28394, AMO28395, AMO28396, AMO28397, AMO28398, AMO28388, AMO28390, AMO28399, and AMO28400); these strains were isolated around the same time as strain JW. Amino acid sequences of the ORF2 region of the strains were aligned; the results showed high sequence identity ($\geqslant$98%) (data not shown). This indicates the epidemic potential of GII.21-type NoV.

Recently, a novel NoV recombinant strain, GII.4/GII.21, was isolated in Bangladesh [30]. Furthermore, a variant of genotype GII.17 has been reported to be more predominant than GII.4 for acute gastroenteritis outbreaks in several countries [37–46]. GII.17 is genetically the closest to GII.21 [34], indicating that GII.21 also has the potential to dominate over GII.4. These studies highlight the importance of molecular

https://doi.org/10.1017/S0950268817001273 Published online by Cambridge University Press

epidemiology studies bolstered by whole-genome analysis to characterize non-dominant strains in addition to prevalent strains, and identify and track sporadic cases of NoV gastroenteritis globally.

To our knowledge, this is the first report of the full-genome sequence analysis of NoV GII.21 from Korea. This is a valuable contribution to the databases that enable viral evolutionary studies and molecular epidemiology studies. Furthermore, the information generated might facilitate the development of diagnostic tools and effective vaccines.

ACKNOWLEDGEMENTS

This research was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI15C1781)), as well as by the Korea Ministry of Environment (MOE) as a Public Technology Program based on Environmental Policy (grant number: 2016000210002).

DECLARATION OF INTEREST

None.

REFERENCES

1. **Adler JL, Zickl R.** Winter vomiting disease. *Journal of Infectious Diseases* 1969; **119**: 668–673.
2. **Glass RI, Parashar UD, Estes MK.** Norovirus gastroenteritis. *New England Journal of Medicine* 2009; **361**: 1776–1785.
3. **Scallan E, *et al.*** Foodborne illness acquired in the United States – major pathogens. *Emerging Infectious Diseases* 2011; **17**: 7–15.
4. **Koopmans M.** Progress in understanding norovirus epidemiology. *Current Opinion in Infectious Diseases* 2008; **21**: 544–552.
5. **Hall AJ, *et al.*** Norovirus disease in the United States. *Emerging Infectious Diseases* 2013; **19**: 1198–1205.
6. **Navas E, *et al.*** Economic costs of outbreaks of acute viral gastroenteritis due to norovirus in Catalonia (Spain), 2010–2011. *BMC Public Health* 2015; **15**: 999.
7. **Lopman BA, *et al.*** Increasing rates of gastroenteritis hospital discharges in US adults and the contribution of norovirus, 1996–2007. *Clinical Infectious Diseases* 2011; **52**: 466–474.
8. **Bartsch SM, *et al.*** The potential economic value of a human norovirus vaccine for the United States. *Vaccine* 2012; **30**: 7097–7104.
9. **Lee BE, Pang XL.** New strains of norovirus and the mystery of viral gastroenteritis epidemics. *CMAJ: Canadian Medical Association Journal* 2013; **185**: 1381–1382.
10. **Barclay L, *et al.*** Infection control for norovirus. *Clinical Microbiology and Infection* 2014; **20**: 731–740.
11. **Hardy ME.** Norovirus protein structure and function. *FEMS Microbiology Letters* 2005; **253**: 1–8.
12. **Zheng DP, *et al.*** Norovirus classification and proposed strain nomenclature. *Virology* 2006; **346**: 312–323.
13. **Vinje J.** Advances in laboratory methods for detection and typing of norovirus. *Journal of Clinical Microbiology* 2015; **53**: 373–381.
14. **Vega E, *et al.*** Genotypic and epidemiologic trends of norovirus outbreaks in the United States, 2009 to 2013. *Journal of Clinical Microbiology* 2014; **52**: 147–155.
15. **Kim JS, *et al.*** Molecular epidemiology of human norovirus in Korea in 2013. *BioMed Research International* 2015; **2015**: 468304.
16. **Mans J, *et al.*** Norovirus diversity in children with gastroenteritis in South Africa from 2009 to 2013: GII.4 variants and recombinant strains predominate. *Epidemiology and Infection* 2016; **144**: 907–916.
17. **Alam A, *et al.*** Genetic characterization of norovirus strains in hospitalized children from Pakistan. *Journal of Medical Virology* 2016; **88**: 216–223.
18. **Montazeri N, *et al.*** Pathogenic enteric viruses and microbial indicators during secondary treatment of municipal wastewater. *Applied and Environmental Microbiology* 2015; **81**: 6436–6445.
19. **Zhirakovskaia EV, *et al.*** Molecular epidemiology of noroviruses associated with sporadic gastroenteritis in children in Novosibirsk, Russia, 2003–2012. *Journal of Medical Virology* 2015; **87**: 740–753.
20. **de Andrade Jda S, *et al.*** Noroviruses associated with outbreaks of acute gastroenteritis in the State of Rio Grande do Sul, Brazil, 2004–2011. *Journal of Clinical Virology* 2014; **61**: 345–352.
21. **Moyo S, *et al.*** Genetic diversity of norovirus in hospitalised diarrhoeic children and asymptomatic controls in Dar es Salaam, Tanzania. *Infection, Genetics and Evolution* 2014; **26**: 340–347.
22. **Bruggink LD, Marshall JA.** Altered patterns of norovirus GII.b recombinant forms in gastroenteritis outbreaks in Victoria, Australia, 2002–2005 compared to 2006–2011. *Journal of Medical Virology* 2013; **85**: 1433–1443.
23. **Ferreira MS, *et al.*** Genotyping of gastroenteric viruses in hospitalised children: first report of norovirus GII.21 in Brazil. *Memorias do Instituto Oswaldo Cruz* 2012; **107**: 1064–1067.
24. **Pérez-Sautu U, *et al.*** Human norovirus occurrence and diversity in the Llobregat river catchment, Spain. *Environmental Microbiology* 2012; **14**: 494–502.
25. **Yahiro T, *et al.*** Norovirus GII.21 in children with Diarrhea, Bhutan. *Emerging Infectious Diseases* 2015; **21**: 732–734.
26. **Kim HS, *et al.*** Emergence of GII.4 Sydney norovirus in South Korea during the winter of 2012–2013. *Journal of Microbiology and Biotechnology* 2013; **23**: 1641–1643.
27. **Tamura K, *et al.*** MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* 2013; **30**: 2725–2729.

28. **Saitou N, Nei M.** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 1987; **4**: 406–425.
29. **Kroneman A, *et al.*** An automated genotyping tool for enteroviruses and noroviruses. *Journal of Clinical Virology* 2011; **51**: 121–125.
30. **Nahar S, *et al.*** A novel norovirus recombinant strain GII.4/GII.21 in Bangladesh, 2011. *Virus Genes* 2013; **46**: 538–541.
31. **Prasad BV, *et al.*** X-ray crystallographic structure of the Norwalk virus capsid. *Science* 1999; **286**: 287–290.
32. **Kroneman A, *et al.*** Proposal for a unified norovirus nomenclature and genotyping. *Archives of Virology* 2013; **158**: 2059–2068.
33. **Tan M, *et al.*** Noroviral P particle: structure, function and applications in virus-host interaction. *Virology* 2008; **382**: 115–123.
34. **Liu W, *et al.*** A unique human norovirus lineage with a distinct HBGA binding interface. *PLoS Pathogens* 2015; **11**: e1005025.
35. **Dingle KE.** Mutation in a Lordsdale norovirus epidemic strain as a potential indicator of transmission routes. *Journal of Clinical Microbiology* 2004; **42**: 3950–3957.
36. **Bull RA, Tanaka MM, White PA.** Norovirus recombination. *Journal of General Virology* 2007; **88**: 3347–3359.
37. **de Graaf M, *et al.*** Emergence of a novel GII.17 norovirus – end of the GII.4 era? *Euro Surveillance* 2015; **20**: pii: 21178.
38. **Fu J, *et al.*** Emergence of a new GII.17 norovirus variant in patients with acute gastroenteritis in Jiangsu, China, September 2014 to March 2015. *Euro Surveillance* 2015; **20**: pii: 21157.
39. **Han J, *et al.*** Emergence and predominance of norovirus GII.17 in Huzhou, China, 2014–2015. *Virology Journal* 2015; **12**: 139.
40. **Lee CC, *et al.*** Emerging norovirus GII.17 in Taiwan. *Clinical Infectious Diseases* 2015; **61**: 1762–1764.
41. **Lu J, *et al.*** Gastroenteritis outbreaks caused by norovirus GII.17, Guangdong Province, China, 2014–2015. *Emerging Infectious Diseases* 2015; **21**: 1240–1242.
42. **Matsushima Y, *et al.*** Genetic analyses of GII.17 norovirus strains in diarrheal disease outbreaks from December 2014 to March 2015 in Japan reveal a novel polymerase sequence and amino acid substitutions in the capsid region. *Euro Surveillance* 2015; **20**: pii: 21173.
43. **Medici MC, *et al.*** Identification of the novel Kawasaki 2014 GII.17 human norovirus strain in Italy, 2015. *Euro Surveillance* 2015; **20**: 30010.
44. **Parra GI, Green KY.** Genome of emerging norovirus GII.17, United States, 2014. *Emerging Infectious Diseases* 2015; **21**: 1477–1479.
45. **Gao Z, *et al.*** Increased norovirus activity was associated with a novel norovirus GII.17 variant in Beijing, China during winter 2014–2015. *BMC Infectious Diseases* 2015; **15**: 574.
46. **Chan MC, *et al.*** Rapid emergence and predominance of a broadly recognizing and fast-evolving norovirus GII.17 variant in late 2014. *Nature Communications* 2015; **6**: 10061.
47. **Won YJ, *et al.*** Full-genomic analysis of a human norovirus recombinant GII.12/13 novel strain isolated from South Korea. *PLoS ONE* 2013; **8**: e85063.