Epidemiology and Infection

cambridge.org/hyg

Complete sequence analysis of human norovirus GII.17 detected in South Korea

H. Kim[1], Y. J. Won[1], L. H. Kang[1], A. R. Lee[1], J. I. Han[2], C. I. Suh[3] and S. Y. Paik[1]

[1]Department of Microbiology, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea; [2]Laboratory of Wildlife Diseases, College of Veterinary Medicine, Chonbuk National University, Iksan, Korea and [3]Department of Research and Development, LumiMac, Inc., Songpa-gu, Seoul, Republic of Korea

Original Paper

Cite this article: Kim H, Won YJ, Kang LH, Lee AR, Han JI, Suh CI, Paik SY (2019). Complete sequence analysis of human norovirus GII.17 detected in South Korea. *Epidemiology and Infection* **147**, e203, 1–9. https://doi.org/10.1017/S0950268819000943

Received: 29 January 2019
Revised: 3 April 2019
Accepted: 1 May 2019

Key words:
Gastroenteritis; GII.17.; norovirus; phylogenetic tree; sequence analysis

Author for correspondence:
S. Y. Paik, E-mail: paik@catholic.ac.kr

Abstract

Norovirus, a major cause of gastroenteritis in people of all ages worldwide, was first reported in South Korea in 1999. The most common causal agents of pediatric acute gastroenteritis are norovirus and rotavirus. While vaccination has reduced the pediatric rotavirus infection rate, norovirus vaccines have not been developed. Therefore, prediction and prevention of norovirus are very important. Norovirus is divided into genogroups GI–GVII, with GII.4 being the most prevalent. However, in 2012–2013, GII.17 showed a higher incidence than GII.4 and a novel variant, GII.P17-GII.17, appeared. In this study, 204 stool samples collected in 2013–2014 were screened by reverse transcriptase-polymerase chain reaction; 11 GI (5.39%) and 45 GII (22.06%) noroviruses were identified. GI.4, GI.5, GII.4, GII.6 and GII.17 were detected. The whole genomes of the three norovirus GII.17 were sequenced. The whole genome of GII.17 consists of three open reading frames of 5109, 1623 and 780 bp. Compared with 20 GII.17 strains isolated in other countries, we observed numerous changes in the protruding P2 domain of VP1 in the Korean GII.17 viruses. Our study provided genome information that might aid in epidemic prevention, epidemiology studies and vaccine development.

Introduction

Norovirus and rotavirus are the most common causative agents of pediatric acute gastroenteritis worldwide, including in South Korea. The rotavirus infection rate in children has decreased since the introduction of a vaccine in 2007–2008 [1]. Norovirus-related gastroenteritis has an important impact on public health worldwide. However, official public health reports likely largely underestimate the actual incidence. Therefore, public health surveillance and thorough statistical data analysis are needed [2].

Norovirus infection is accompanied by vomiting and diarrhoea after a latency period of 12–48 h and the virus is transmitted person to person via the fecal-oral route or via a contaminated environment [3]. Norovirus outbreaks have been reported frequently in closed environments, such as facilities, schools, hospitals and cruise ships and often originate from contaminated groundwater [4]. The incidence of norovirus among children is very high; norovirus causes 200 000 deaths in children younger than 5 years of age annually in developing countries [5]. The elderly and immunocompromised patients are also vulnerable to norovirus and are at risk of dying from infection [5].

According to the Centers for Disease Control and Prevention, there are 21 million norovirus cases in the USA annually, accounting for 60% of patients with acute gastroenteritis [6]. In the USA alone, each year, 570–800 people die, 56 000–71 000 people are hospitalised and the total hospitalisation cost is about 500 million. Medical costs and lost productivity are estimated at 2 billion [7]. According to the Foodborne Viruses in Europe report, 7637 norovirus outbreaks occurred in 13 European countries from July 2001 to June 2006, mostly in Germany and the UK, with 3808 (49.9%) and 1938 (25.4%) outbreaks, respectively [8].

Norovirus was reported in South Korea for the first time in 1999 and the National Institute of Health has been monitoring acute diarrhea diseases since 2000 [9]. In 2006, norovirus caused 2400 infection cases at 31 school cafeterias throughout the country [10]. In 2008, approximately 130 infection cases in an elementary school in Incheon were reported [11]. According to the Korea Food and Drug Administration, the number of norovirus outbreaks and patients in Korea has increased significantly since 2000 [12].

The development of a vaccine or therapy for norovirus is hampered by the lack of a suitable cell culture system and animal models [13]. Previous studies on norovirus have mainly relied on electron microscopy, antibody investigation and reverse transcription-polymerase chain reaction (RT-PCR) [10].

CAMBRIDGE UNIVERSITY PRESS

Table 1. Sequence information for primers used for RT-PCR assays

Genotype	Primer	Sequence (5′→3′)	Region	Position	Reference
I	GI-F1M	CTG CCC GAA TTY GTA AAT GAT GAT	ORF1/2	5342–5671	[7]
	GI-R1M	CCA ACC CAR CCA TTR TAC ATY TG			
II	GII-F1M	GGG AGG GCG ATC GCA ATC T	ORF1/2	5049–5389	
	GII-R1M	CCR CCT GCA TRI CCR TTR TAC AT			

Table 2. Newly designed primers used

Primer name	Sequence (5′→3′)	Location	Size (bp)	Polarity	Reference
GSP1	GAG TGT AGA CCA CTC TCC A	525–543		–	This study
GSP2	GGT GGT TTG TGT ACT CCG A	458–476		–	
Nested GSP	GTG GTA GAT CTC ACC ATC C-	408–468		–	
ORF1_1F	GTG AAT GAA GAT GGC GTC T	1–19	671	+	This study
ORF1_1R	CAG GAG TCG TTG AGG TCT A	651–670		–	
ORF1_2F	CAT ACA CGG CCT TTG ACA AC	603–622	645	+	
ORF1_2R	CCA GGA CTG CAT CCT CAA T	1229–1247		–	
ORF1_3F	CAA GAA GGA GGA GGC TAA C	1189–1207	631	+	
ORF1_3R	GAC ATA GTC TAG TGG TGC T	1801–1819		–	
ORF1_4F	GAA TTG GCT GAC ACC TGT C	1694–1712	722	+	
ORF1_4R	CTT GCA GAT CTC AGG TGG T	2395–2415		–	
ORF1_5F	GCA CCT ATA CCA TGG AAT C	2298–2316	713	+	
ORF1_5R	CAC GCT CCT GTT GTC ATC A	2992–3010		–	
ORF1_6F	CCA GAG AAT CTT CCG ACC	2866–2883	681	+	
ORF1_6R	GCA CAT ATG ACC GTG TTG C	3528–3546		–	
ORF1_7F	GCA TGG ATC TAG GCA CAA C	3423–3441	629	+	
ORF1_7R	CAT GTG CTT GCC TTC CTC A	4033–4051		–	
ORF1_8F	CAC GTC CGG AAG AAT GAC T	3950–3968	679	+	
ORF1_8R	CTG TGC TCA CAA TCT CAT C	4610–4628		–	
ORF1_9F	CCT TGC ACC TCA CAA TGG A	4487–4505	630	+	
ORF1_9R	CAT TCA TTC GAC GCC ATC T	5098–5116		–	
ORF2_1F	GTG CCC AGA CAA GAG TCA A	5003–5022	695	+	
ORF2_1R	GAC ATC ATC ACC TGA GCC A	5680–5698		–	
ORF2_2F	CCA ATT ATG ATC CCA CTC C	5570–5580	584	+	
ORF2_2R	GAG CCA GGC GCA TCA TTA	6136–6153		–	
ORF2_3F	GTG GCA CAT GCA ACT GCA A	6010–6028	730	+	
ORF2_3R	CAT TTA CTG AGC CCT CCT T	6720–6740		–	
ORF3_1F	GGC TCA CTC TGG AGA CTA T	6592–6628	682	+	
ORF3_1R	GAG AGG AGC CAG AGC TTA A	7274–7292		–	
ORF3_2F	CCT AGG TCT CTA GCT CCT T	7154–7172	363	+	
ORF3_2R	CTA AAC ACG TGA CTC CCC	7500–7516		–	
3′-Oligo This study (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			–	

The primers were based on the kawasaki323 strain (AB983218), CAU-55 strain (KU561250).
GSP1, GSP2 and Nested GSP were based on the CMC-01 strain.

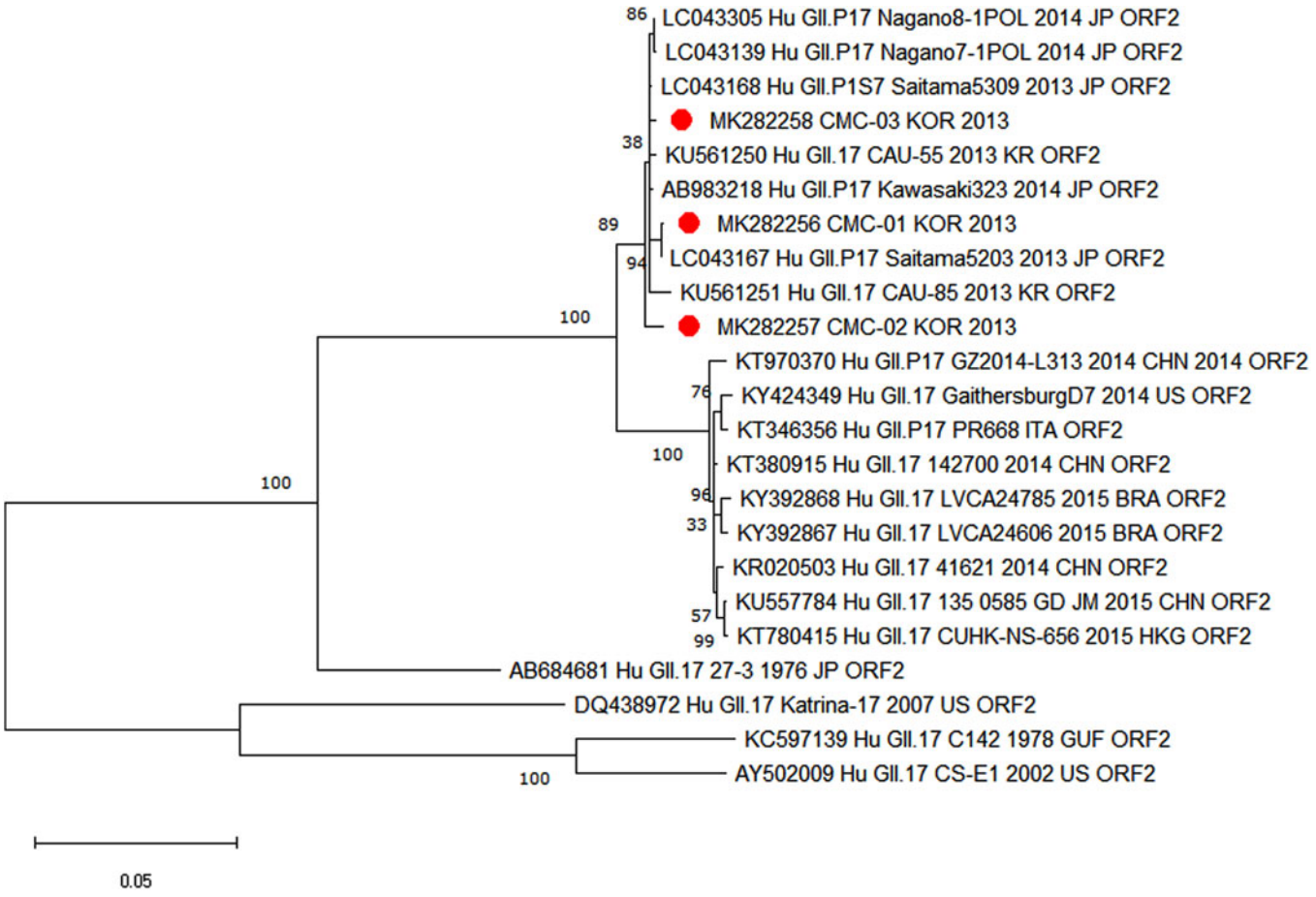

Fig. 1. Phylogenetic analysis of the full-length sequences of the GII type detected determined by genotyping and reference strains isolated worldwide.

Norovirus belongs to the family *Caliciviridae* [14]. It has a single-stranded, positive-sense RNA genome of approximately 7.5 kb [15]. The genome comprises three open reading frames (ORF). ORF1 encodes non-structural proteins, such as NTpase, protease and RNA-dependent RNA polymerase (RdRp). ORF2 encodes major capsid protein (VP1). ORF3 encodes minor capsid protein (VP2) [16]. Noroviruses are divided into seven genogroups (GI–GVII) and are classified into more than 40 genotypes [17]. GI, GII, GIV can infect humans. GII is the most common genogroup [18] and GII.4 is the most predominant genotype, with new variants occurring in 2–3-year cycles. GII.4 viruses have a larger susceptible population than viruses from other genotypes as they can bind a wider range of histo-blood group antigens, which have been suggested to be attachment factors [19]. The high rate of evolution of GII.4 viruses leads to the emergence of new antigenic variants [20].

GII.17 occurred at a low frequency for many years [21]. Recently, a GII.17 norovirus variant emerged [22] and globally became more dominant than GII.4, which had been dominant in China and Hong Kong for more than 20 years. The novel variant has been sporadically reported in Italy and the USA [23]. In Japan in 2014, strains with an RdRp sequence different from that of existing GII.17 were found. This variant with newly identified RdRp was assigned GII.P17-GII.17 by the web-based norovirus typing tool NoroNet [24]. In this study, we sequenced the whole genome of GII.P17-GII.17 strains discovered in 2013 and we confirmed amino acid mutations through various comparative analyses with other global isolates.

Materials and methods

Ethics statement

Stool samples of patients with acute gastroenteritis symptoms were provided by Waterborne Virus Bank. Due to difficulties in tracing the exact records of the young patients at the donation hospital, informed consent could not be obtained from the patients' parents. The institutional review board of Songeui Medical Campus, The Catholic University of Korea (MC18SESI0063) reviewed and approved the use of the samples for research purposes as this study had no impact on the patients. All experimental work and sample collection were supervised by the institutional review board.

Stool sample preparation

In total, 204 stool samples were collected from children of 0–13 years of age with fever and diarrhea between January 2013 and December 2013. Frozen stool samples were diluted with 10% phosphate-buffered saline and centrifuged at 13 000 × ***g*** for 10 min at 4 °C. The supernatant was separated and stored at −80 °C until analysis.

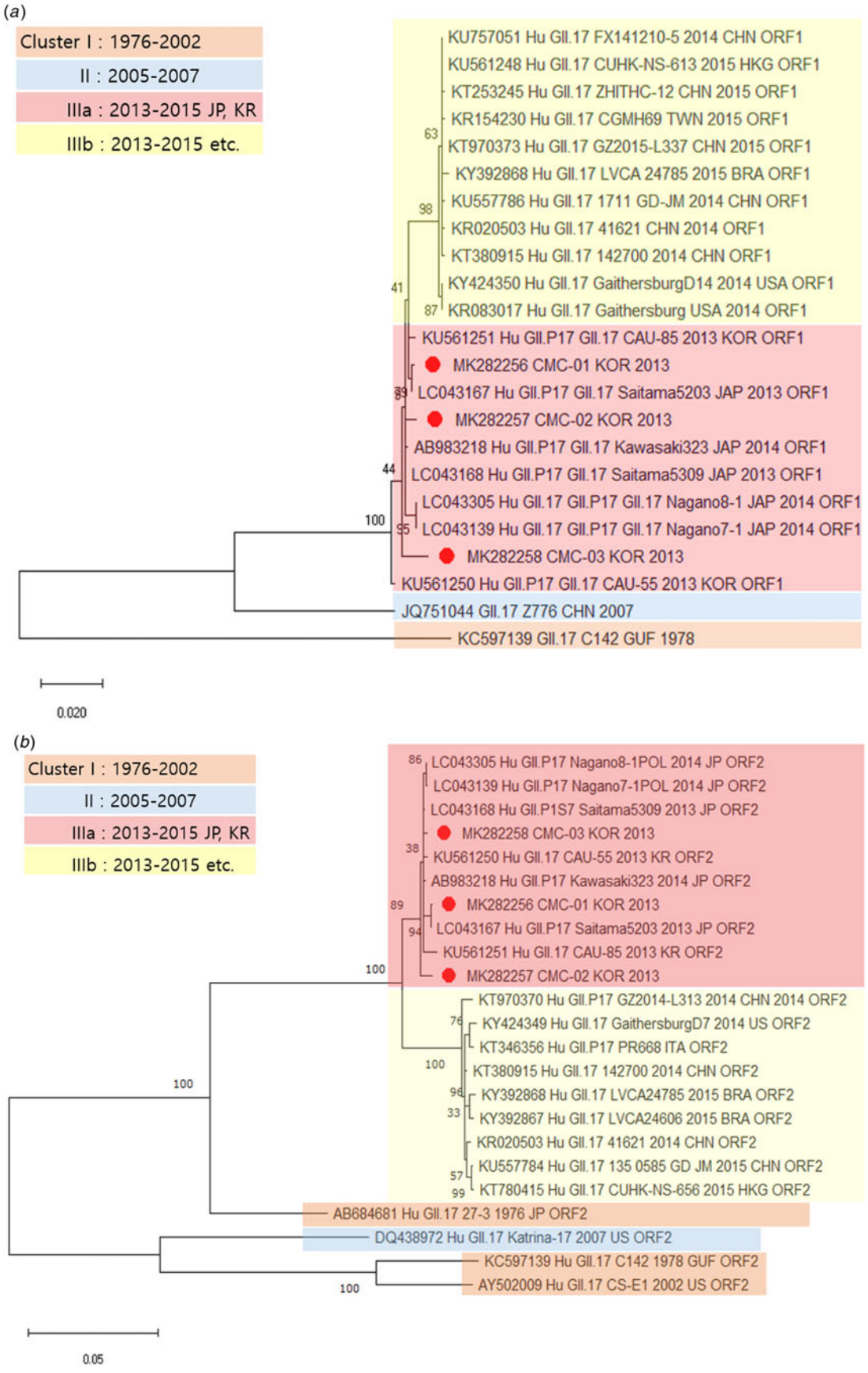

Fig. 2. Phylogenetic analysis of norovirus based on nucleotide sequences. The trees were constructed with the neighbour-joining method. Phylogenetic trees based on (a) amino acid sequence of ORF1, (b) amino acid sequence of ORF2 and (c) amino acid sequence of ORF3. CMC-01, CMC-02 and CMC-03 are indicated by red circles.

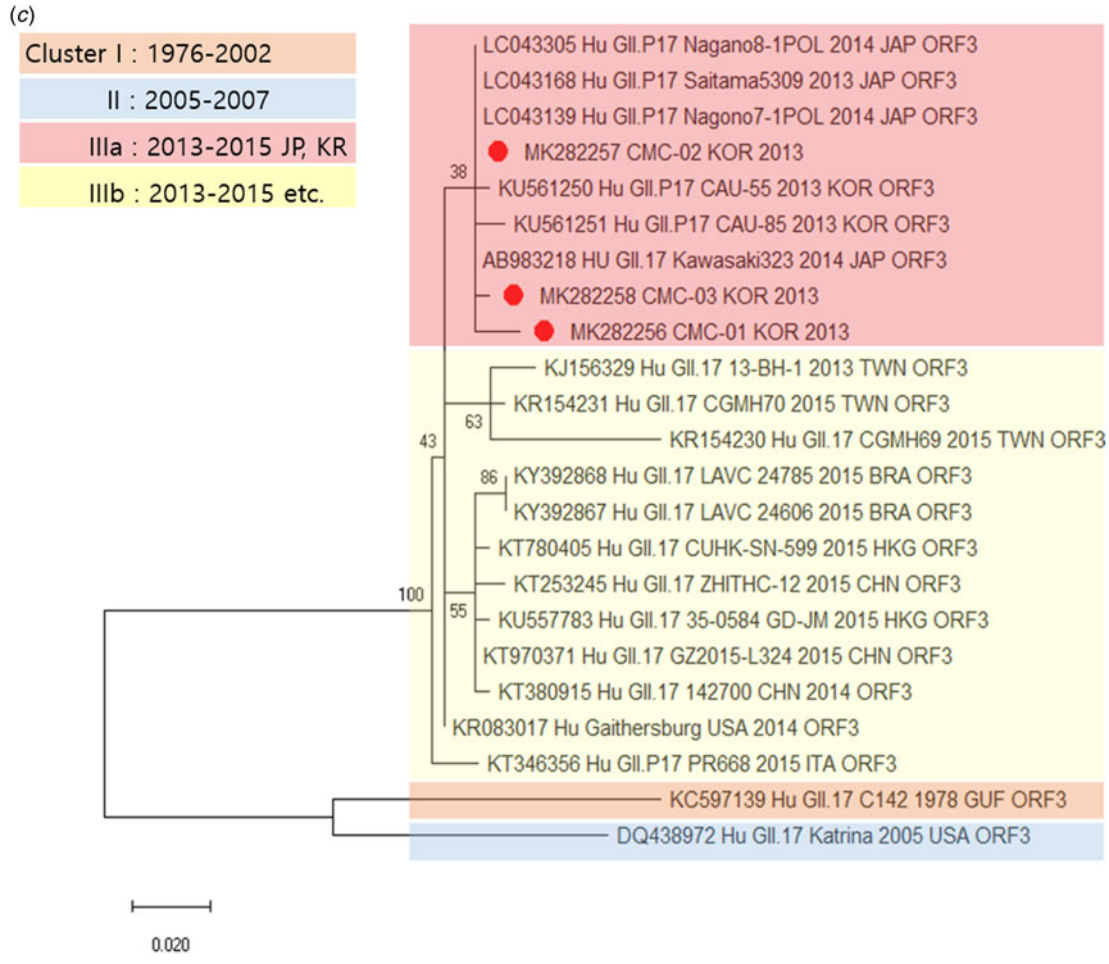

Fig. 2c. (Continued).

Viral RNA extraction

Viral RNA was extracted from 140 µl of stool sample using a QIAamp Viral RNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The RNA was stored at −80 °C until use in RT-PCR.

RT-PCR analysis

Norovirus-positive specimens were detected by RT-PCR using the GI-F1M–GI-R1M and GII-F1M–GII-R1M primer sets [7] (Table 1) and a One-Step RT-PCR kit (Qiagen, Hilden, Germany). Thermal cycles were as follows: 50 °C for 30 min, 95 °C for 15 min, followed by 40 cycles of 94 °C for 60 s, 55 °C for 60 s, and 72 °C for 60 s and a final step at 72 °C for 10 min. To analyse the whole-genome sequence of GII.17 norovirus, 14 pairs of primers were newly designed (Table 2). Thermal cycles were the same as above. PCR products were analysed by electrophoresis in ethidium bromide-stained 2% agarose gels.

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

To determine the 5′-ends of norovirus genomic RNA, rapid amplification of cDNA ends (RACE) was performed using the 5′ RACE System for Rapid Amplification of cDNA Ends Version 2.0 Kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's recommendations. Three primers (GSP1, GSP2 and nested GSP) were designed based on ORF1 of CMC-1 for RACE (Table 2). To obtain the exact sequence of the 3′-end of the genomic RNA, cDNA was synthesised by reverse transcription using 3′-oligo (dT)-anchor-R (Table 2). The second PCR was conducted using the ORF3_2F and 3′-anchor-R primers (Table 2) under the following conditions: 35 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 1 min, followed by 72 °C for 7 min.

Cloning and sequencing of the complete genome

GII.17 norovirus PCR product was extracted from a 2% agarose gel using the HiYield Gel/PCR DNA Fragments Extraction Kit (Real Biotech, Taipei, Taiwan) and was cloned into the pGEM-T Easy Vector (Promega, Madison, WI, USA). The cloned vectors were transformed into *Escherichia coli* DH5α competent cells (RBC Bioscience) according to the manufacturer's instructions and selected at 37 °C for 16–18 h on Luria–Bertani agar plates (Duchefa, Haarlem, the Netherlands) containing 40 mg/ml X-gal, 0.1 mM isopropyl-β-D-thiogalactoside and 50 mg/ml ampicillin. Selected clones were inoculated into LB broth (Duchefa) and cultured overnight at 37 °C under shaking at 200 rpm. Plasmid DNA was purified using a HiYield Plasmid Mini Kit (RBC Bioscience) and sequenced (Macrogen, Seoul, South Korea). The sequencing results were analysed using the

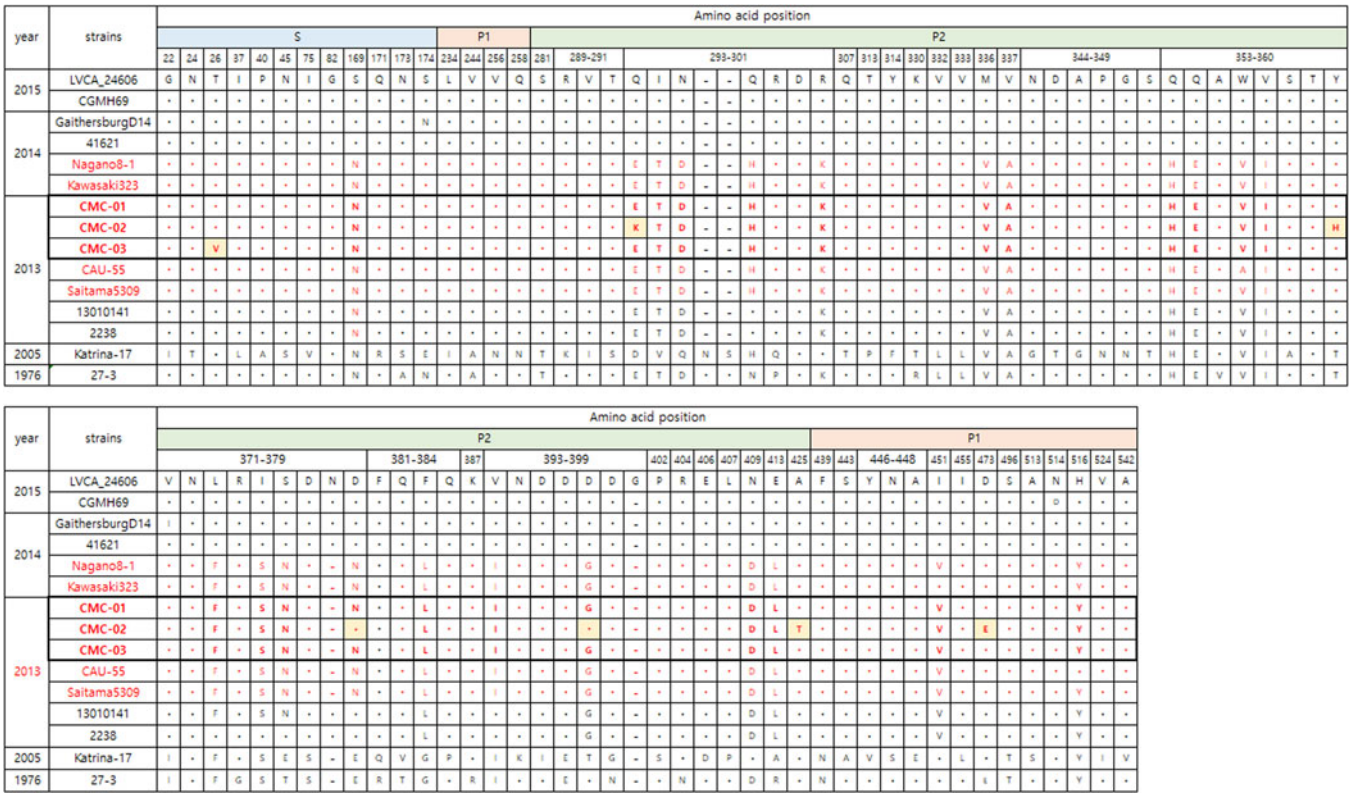

Fig. 3. Amino acid substitutions in the viruses in this study (GII.17; black colour, GII.P17-GII.17; red colour). Alignment of VP1 amino acid sequences of CMC-01, CMC-02 and CMC-03 strains. Dots indicate sequence identity among sequences presented. Dashes indicate deletions/insertions of amino acid residues. LVCA24606 (KY392867), CGMH69 (KR154230), GaithersburgD14 (KY424350), 41621 (KR020503), Nagano8-1 (LC043305), Kawasaki323 (AB983218), CAU-55 (KU561250), Saitama5309 (LC043168), 13010141 (KU757046), 2238 (KU557788), Katrina-17 (DQ438972), 27-3 (AB684681).

Basic Local Alignment Search Tool (BLAST; National Center for Biotechnology Information (NCBI); National Institutes of Health, Bethesda, MD, USA).

Phylogenetic analysis

Nucleotide sequences of GII.17 were aligned and classified by ORF1, ORF2 and ORF3. Phylogenetic analysis of the GII.17 genome sequences was performed using sequences of norovirus strains published in the NCBI database, using Molecular Evolutionary Genetic Analysis software (MEGA version 7.0).

Results

Norovirus detection and whole genome sequencing

As a result of screening 204 stool samples, 11 samples (5.39%) were determined to contain norovirus belonging to genogroup GI and 60 samples (29.41%) were found to contain norovirus belonging to GII. In GI samples, we detected genotypes GI.4 ($n = 7$; 3.43%) and GI.5 ($n = 4$; 1.96%). In GII samples, we identified GII.4 ($n = 45$; 22.06%), GII.6 ($n = 12$; 5.88%) and GII.17 ($n = 3$; 1.47%). The three GII.17 viral specimens were named CMC-01, CMC-02 and CMC-03. All three viruses were subjected to whole genome sequencing. The genome of GII.17 has a total length of 7512 nt and consists of ORF 1 (5109 nt), ORF 2 (1623 nt) and ORF 3 (780 nt). The genome sequences obtained in this study have been deposited in NCBI (MK282256, MK282257 and MK282258).

Phylogenetic analysis

Genetic relationships between the viral specimens from this study and norovirus strains reported in GenBank were determined on the basis of phylogenetic analysis. The whole genome sequences of CMC-01, CMC-02 and CMC-03 were classified as norovirus GII.17 and showed maximum identity (99.5%) with the strain Saitama5203 (LC043167) (Fig. 1). The sequences of CMC-01, CMC-02, CMC-03 were compared and analysed with the sequences of isolates reported in various countries, including Brazil, Japan and China, from 1976 to 2016. Prior to the analysis, we divided cluster I (1976–2002), cluster II (2005–2007), cluster IIIa (2013–2015) and cluster IIIb (2013–2015) by year and country. IIIa strains are found in Korea and Japan, IIIb strains are found in countries other than Korea and Japan, and CMC-01, CMC-02 and CMC-03 were classified in cluster IIIa. Next, we constructed phylogenetic trees based on amino acid sequences. The ORF1 sequence of CMC-01 showed 99.5% identity with that of Saitama5203 (LC043167). The ORF1 sequences of CMC-02 and CMC-03 showed 99.5% and 99.7% identity with that of Saitama5309 (LC043168), respectively (Fig. 2a). For ORF2, CMC-01 showed the highest identity of 99.9% with Saitama5203 and CMC-02 and CMC-03 showed 99.3% and 99.8% identity with Saitama5309 (Fig. 2b). ORF3 of CMC-01 (99.3%), CMC-02 (99.7%) and CMC-03 (99.7%) showed the highest similarity with that of Saitama5309 (LC043168), Kawasaki323 (AB983218) (Fig. 2c). A phylogenetic tree was constructed with strains collected from 1976 to 2015. ORF1, ORF2 and ORF 3 of the three detected in this study all fell in cluster

year	strains	Amino acid number																										
		S	P2																								P1	
		169	293-297					334-335		351-352		354-355		358	371	373-374		376-377		381	391	395	397	407	411	423	449	514
2016	Ljubljana535	S	Q	I	N	Q	R	M	V	Q	Q	W	V	Y	L	I	S	N	D	F	V	D	G	N	E	A	I	H
	Osaka15-377	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	X27_16NV149	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	1613179	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	1613305	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2015	X15_15NV796	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	23438	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	Y
	LVCA_24606	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	23233	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	E12972	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	15-EN-10	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	CGMH69	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
2014	GD-GZ	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	CUHK-NS-502	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	14200	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	41621	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	Gaithersersburgd14	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	GZ2014-L313	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	14-BQ-2	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
	Nagano8-1	N	E	T	D	H	K	V	A	H	E	V	I	·	F	S	N	-	N	L	I	G	-	D	L	·	V	Y
	Kawasaki323	N	E	T	D	H	K	V	A	H	E	V	I	·	F	S	N	-	N	L	I	G	-	D	L	·	V	Y
2013	Saitama5309	N	E	T	D	H	K	V	A	H	E	V	I	·	F	S	N	-	N	L	I	G	-	D	L	·	V	Y
	CAU-55	N	E	T	D	H	K	V	A	H	E	V	I	·	F	S	N	-	N	L	I	G	-	D	L	·	V	Y
	CMC-01	**N**	**E**	**T**	**D**	**H**	**K**	**V**	**A**	**H**	**E**	**V**	**I**	·	**F**	**S**	**N**	-	**N**	**L**	**I**	**G**	-	**D**	**L**	·	**V**	**Y**
	CMC-02	**N**	**K**	**T**	**D**	**H**	**K**	**V**	**A**	**H**	**E**	**V**	**I**	**H**	**F**	**S**	**N**	-	·	**L**	**I**	·	-	**D**	**L**	**T**	**V**	**Y**
	CMC-03	**N**	**E**	**T**	**D**	**H**	**K**	**V**	**A**	**H**	**E**	**V**	**I**	·	**F**	**S**	**N**	-	**N**	**L**	**I**	**G**	-	**D**	**L**	·	**V**	**Y**
	13-BH-1	N	E	T	D	H	K	V	A	H	E	V	I	·	F	S	N	-	N	L	I	G	-	D	L	·	V	Y
	2238-GD-JM	N	E	T	D	H	K	V	A	H	E	V	I	·	F	S	N	-	N	L	I	G	-	D	L	·	V	Y
	13010141	N	E	T	D	H	K	V	A	H	E	V	I	·	F	S	N	-	N	L	I	G	-	D	L	·	V	Y

Fig. 4. Comparison of major amino acid substitutions in viral specimens of 2013–2016 (GII.17; black colour, GII.P17-GII.17; red colour). Dots indicate sequence identity among sequences presented. Dashes indicate deletions/insertions of the amino acid residues. Amino acid numbering is based on the sequence of the Saitama5309 strain. Ljubljana535(KX134671), Osaka15-377(LC148852), X27-16NV149 (KX371112), 1613179(KU953395), 1613305(KU953397), X15_15NV796 (KX371109), 23438 (KX216804), LVCA_24606 (KY392867), 23233 (KX216793), E12972 (KU587628), 15-EN-10 (KT732275), CGMH69 (KR154230), GD-GZ (KU557801), CUHK-NS-502 (KT780399), 14200 (KT380915), 41621 (KR020503), GaithersburgD14 (KY424350), GZ2014-L313 (KT970370), 14-BQ-2 (KT906670), Nagano8-1 (LC043305), Kawasaki323 (AB983218), Saitama5309 (LC043168), CAU-55 (KU561250), 13-BH-1 (KJ156329), 2238-GD-JM(KU557788), 13010141 (KU757046).

IIIa and were further classified as GII.P17–GII.17, which is a variant of GII.17. All three viruses were very similar to Saitama5309 (LC043168), Kawasaki323 (AB983218), Nagano8-1 (LC043305) and Nagano7-1 (LC043139).

VP1 sequence alignment

Figure 3 shows the VP1 amino acid substitution pattern in our viral specimens compared with 12 strains collected from 1976 to 2015 based on alignment. CMC-01, CMC-02 and CMC-03 were classified with GII.P17-GII.17 strains collected in 2013–2014. Specifically, the GII.17 strains 14-BQ-2 and GZ2014-L313 discovered in 2014, had 28 amino acid alterations in the P2 domain (Fig. 4). CMC-01 and CMC-03 were similar to GII.P17-GII.17 strains and CMC-02 had five amino acid substitutions (Glu293Lys, Tyr360Lys, Asn379Asp, Gly397Asp and Ala425Thr). Comparative analysis of the 2013 and 2014 GII.17 strains revealed 21 substitutions, especially in the P2 domain (Glu293Gln, Thr294Ile, Asp295Asn, Gln298His, Asp301Asn, Val336Met, Ala337Val, His353Gln, Glu354Gln, Val356Trp, Ile357Val, Phe373Leu, −378Asn, Ser375Ile, Asn376Ser, Asn379Asp, Leu383Phe, Val393Ile, Gly397Asp, Asp409Asn and

Leu413Glu). When comparing GII.P17-GII.17 and GII.17 strains, the substitutions Gln298His, −378Asn and Val393Ile were specific to GII.P17-GII.17. In addition, substitutions including Asn169Ser in the shell (S) domain and Asp410Asn, Leu414Glu and Tyr516His in the P1 domain were detected.

Discussion

Norovirus globally is one of the important causal agents of acute gastroenteritis. Although norovirus causes economical and public-health problems, a thorough study of norovirus is hampered by the lack of cell culture systems or animal models. In infants and the elderly, illness can last longer [25]. GII.4 accounts for 70–80% of norovirus gastroenteritis since the mid-1990s [26]. Since 1996, GII.4 variants have periodically appeared every 2–3 years [19] and GII.4 is the major genotype circulating in the community and medical environment [27]. A GII.P17-GII.17 norovirus strain was discovered in Japan during the winter season of 2014–2015.

In this study, we screened 204 stool samples collected from children with symptoms of acute gastroenteritis between January 2013 and December 2013 for the presence of norovirus. The collected strains were classified as GII.17 by BLAST and were identified as GII.P17-GII.17 according to the sequences of viral RdRp. Other GII.P17-GII.17 strains include Saitama 5309 (LC043168) and Kawasaki 323 (AB983218). Sequence analysis of ORF1–3 revealed a very high similarity between our viruses and previously collected GII.P17-GII.17 strains. CMC-01 is very similar to Saitama5203 (LC043167) and CMC-02 and CMC-03 showed high sequence identity with Saitama5309 (LC043168). The GII.17 strain detected in South Korea was phylogenetically related with strains detected in the USA, Italy, China, Japan and Hong Kong. Interestingly, the GII.3 strain discovered in the 2000s has an evolutionary relationship with the recently discovered GII.17 strain [28]. Norovirus VP1 consists of an S domain that forms a scaffold enveloping the viral RNA and a P domain composed of subdomains P1 and P2 [29]. The S domain has the most stable conserved sequence. The P1 domain is less variable than P2 domain and the P2 domain is the most exposed and the most variable in structure [30]. Therefore, among highly similar strains, variations in the P2 domain sequence gives rise to different P2 domain structures. CMC-01 and CMC-03 showed very close similarity to previously collected strains, whereas CMC-02 showed six amino acid substitutions (Figure 3; Glu293Lys, Tyr360His, Asn379Asp, Gly397Asp, Ala425Thr and Glu473Asp). Figure 4 shows amino acid substitutions in the VP1 region of viruses collected in 2013–2016. GII.17 viruses collected after 2014 showed amino acid sequence variants different from those occurring before 2014 (Glu293Gln, Thr294Ile, Asp295Asn, His296Gln, Lys297Arg, Val334Met, Ala335Val, His351Gln, Glu352Gln, Val354Trp, Ile355Val, Phe371Leu, Asn373Ile, Asn374Ser, − 376Asn, Asn377Asp, Leu381Phe, Ile391Val, Gly395Asp, − 397Gly, Asp407Asn, Leu411Glu). The hypervariable P2 domains of VP1 of the CMC strains were different from those of GII.17 viruses reported after 2014, but similar to those of GII.P17-GII.17 and GII.17 viruses collected before 2014. Even though strains 14-BQ-2 (KT906670) and GZ2014-L313 (KT970370) were classified as GII.P17-GII.17, their P2 domain is similar to that of GII.17 viruses collected after 2014. Thus, it can be assumed that the timing of appearance of 14-BQ-2 (KT906670) and GZ2014-L313 (KT970370) may be similar to that of GII.17, based on P2 domain similarity. Thus, GII.P17-GII.17 was prevalent in 2013–2014 and it is likely that the P2 domain changed at that time. A GII.P17-GII.17 norovirus strain was discovered in Japan during the winter season of 2014–2015. Since then, it has been detected in Asian countries at high frequency and in other continents at low frequency [26]. However, GII.P17-GII.17 strains (KU561250, KU561251) were collected in Korea in 2013; thus, this type may have been prevalent in Korea before it occurred in Japan. This sequence data will be useful for comparisons with full-length norovirus sequences of other strains identified worldwide. Moreover, the information acquired from the whole-genome sequences of strains CMC-01 and CMC-03 may prove useful for advancing basic research toward the elucidation of gene functions, the prediction of newly appearing pandemic variants via comparison with norovirus in neighbouring countries and in vaccine development. Overall, broadening the information and genetic resources of noroviruses circulating globally will have important benefits for public health and help to identify new emerging strains of norovirus.

Author ORCIDs. S. Y. Paik, 0000-0003-2748-6699.

Acknowledgement. This work 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); and the Korea Ministry of Environment (MOE) as a Public Technology Program based on Environmental Policy (grant number 2016000210002).

Conflict of interest. None.

References

1. **Jin HI *et al.*** (2016) Recent viral pathogen in acute gastroenteritis: a retrospective study at a tertiary hospital for 1 year. *Korean Journal of Pediatrics* **59**, 120–125.
2. **LeBlanc JJ *et al.*** (2016) Outbreak of norovirus GII.P17-GII.17 in the Canadian Province of Nova Scotia. *The Canadian Journal of Infectious Diseases & Medical Microbiology = Journal Canadien des Maladies Infectieuses et de la Microbiologie Medicale* **2016**, 1280247.
3. **Fretz R *et al.*** (2005) Outbreaks of gastroenteritis due to infections with norovirus in Switzerland, 2001-2003. *Epidemiology and infection* **133**, 429–437.
4. **Cho HG *et al.*** (2015) Emergence of norovirus GII.4 variants in acute gastroenteritis outbreaks in South Korea between 2006 and 2013. *Journal of Clinical Virology : the Official Publication of the Pan American Society for Clinical Virology* **72**, 11–15.
5. **Pothier P and Kaiser L** (2014) Norovirus disease today. *Clinical Microbiology and Infection : the Official Publication of the European Society of Clinical Microbiology and Infectious Diseases* **20**, 716.
6. **Robilotti E, Deresinski S and Pinsky BA** (2015) Norovirus. *Clinical Microbiology Reviews* **28**, 134–164.
7. **Lee JW *et al.*** (2017) Full-genome sequence analysis of an uncommon norovirus genotype, GII.21, from South Korea. *Epidemiology and Infection* **145**, 2231–2240.
8. **Kroneman A *et al.*** (2008) Analysis of integrated virological and epidemiological reports of norovirus outbreaks collected within the Foodborne Viruses in Europe network from 1 July 2001 to 30 June 2006. *Journal of Clinical Microbiology* **46**, 2959–2965.
9. **Gong Y-W *et al.*** (2008) Molecular Epidemiologic Investigation of Norovirus Infections in Incheon City, Korea, from 2005 to 2007. *Journal of Bacteriology and Virology* **38**, 249–257.
10. **Koo HS *et al.*** (2016) Molecular epidemiology of norovirus in asymptomatic food handlers in Busan, Korea, and emergence of genotype GII.17. *Journal of Microbiology (Seoul, Korea)* **54**, 686–694.
11. **Yu JH *et al.*** (2010) Epidemiology of foodborne norovirus outbreak in Incheon, Korea. *Journal of Korean Medical Science* **25**, 1128–1133.

12. **Cheon DS *et al.*** (2010) Seasonal prevalence of asymptomatic norovirus infection in Korean children. *Foodborne Pathogens and Disease* **7**, 1427–1430.
13. **Duizer E *et al.*** (2004) Laboratory efforts to cultivate noroviruses. *The Journal of General Virology* **85**, 79–87.
14. **Patel MM *et al.*** (2008) Systematic literature review of role of noroviruses in sporadic gastroenteritis. *Emerging Infectious Diseases* **14**, 1224–1231.
15. **Vega E *et al.*** (2014) Genotypic and epidemiologic trends of norovirus outbreaks in the United States, 2009 to 2013. *Journal of Clinical Microbiology* **52**, 147–155.
16. **Kim JS *et al.*** (2015) Molecular epidemiology of human norovirus in Korea in 2013. *BioMed Research International* **2015**, 468304.
17. **Vinje J** (2015) Advances in laboratory methods for detection and typing of norovirus. *Journal of Clinical Microbiology* **53**, 373–381.
18. **White PA** (2014) Evolution of norovirus. *Clinical Microbiology and Infection : the Official Publication of the European Society of Clinical Microbiology and Infectious Diseases* **20**, 741–745.
19. **Eden J-S *et al.*** (2013) Recombination within the pandemic norovirus GII.4 lineage. *Journal of Virology* **87**, 6270–6282.
20. **Lindesmith LC, Donaldson EF and Baric RS** (2011) Norovirus GII.4 strain antigenic variation. *Journal of Virology* **85**, 231–242.
21. **Han J *et al.*** (2015) Emergence and predominance of norovirus GII.17 in Huzhou, China, 2014-2015. *Virology Journal* **12**, 139.
22. **Pogan R, Dulfer J and Uetrecht C** (2018) Norovirus assembly and stability. *Current Opinion in Virology* **31**, 59–65.
23. **Chan MCW *et al.*** (2017) Global spread of norovirus GII.17 Kawasaki 308, 2014–2016. *Emerging Infectious Diseases* **23**, 1359–1354.
24. **Matsushima Y *et al.*** (2015) 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 : Bulletin Europeen sur les Maladies Transmissibles = European Communicable Disease Bulletin* **20**, pii: 21173.
25. **Lee BE and Pang XL** (2013) New strains of norovirus and the mystery of viral gastroenteritis epidemics. *CMAJ : Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne* **185**, 1381–1382.
26. **Medici MC *et al.*** (2015) Identification of the novel Kawasaki 2014 GII.17 human norovirus strain in Italy, 2015. *Euro Surveillance : Bulletin Europeen sur les Maladies Transmissibles = European Communicable Disease Bulletin* **20**, 30010.
27. **Chan MC *et al.*** (2015) Rapid emergence and predominance of a broadly recognizing and fast-evolving norovirus GII.17 variant in late 2014. *Nature Communications* **6**, 10061.
28. **Dang Thanh H *et al.*** (2016) Emergence of norovirus GII.17 variants among children with acute gastroenteritis in South Korea. *PloS ONE* **11**, e0154284.
29. **Beier R *et al.*** (2014) Selection of a DNA aptamer against norovirus capsid protein VP1. *FEMS Microbiology Letters* **351**, 162–169.
30. **Prasad BV *et al.*** (1999) X-ray crystallographic structure of the Norwalk virus capsid. *Science (New York, NY)* **286**, 287–290.