

# 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

Y Matsushima<sup>1,2</sup>, M Ishikawa<sup>1</sup>, T Shimizu<sup>1</sup>, A Komane<sup>1</sup>, S Kasuo<sup>3</sup>, M Shinohara<sup>4</sup>, K Nagasawa<sup>5</sup>, H Kimura<sup>5</sup>, A Ryo<sup>2</sup>, N Okabe<sup>1</sup>, K Haga<sup>6</sup>, Y H Doan<sup>6</sup>, K Katayama<sup>6</sup>, H Shimizu (shimizu-h@city.kawasaki.jp)<sup>2</sup>

1. Division of Virology, Kawasaki City Institute for Public Health, Kanagawa, Japan
2. Department of Microbiology, Yokohama City University School of Medicine, Kanagawa, Japan
3. Division of Infectious Diseases, Nagano Environmental Conservation Research Institute, Nagano, Japan
4. Division of Virology, Saitama Institute of Public Health, Saitama, Japan
5. Infectious Disease Surveillance Center, National Institute of Infectious Diseases, Tokyo, Japan
6. Department of Virology II, National Institute of Infectious Diseases, Tokyo, Japan

## Citation style for this article:

Matsushima Y, Ishikawa M, Shimizu T, Komane A, Kasuo S, Shinohara M, Nagasawa K, Kimura H, Ryo A, Okabe N, Haga K, Doan YH, Katayama K, Shimizu H. 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 Surveill.* 2015;20(26):pii=21173. Available online: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=21173>

Article submitted on 03 June 2015 / published on 02 July 2015

**A novel GII.P17-GII.17 variant norovirus emerged as a major cause of norovirus outbreaks from December 2014 to March 2015 in Japan. Named Hu/GII/JP/2014/GII.P17-GII.17, this variant has a newly identified GII.P17 type RNA-dependent RNA polymerase, while the capsid sequence displays amino acid substitutions around histo-blood group antigen (HBGA) binding sites. Several variants caused by mutations in the capsid region have previously been observed in the GII.4 genotype. Monitoring the GII.17 variant's geographical spread and evolution is important.**

The present study uses complete genome sequences and phylogenetic and *in silico* analyses to characterise GII.P17 norovirus strains contributing to gastroenteritis outbreaks in Japan from December 2014 to March 2015.

## Norovirus outbreaks from October 2014 to March 2015 in Japan

In Japan, numbers of norovirus cases are reported to the Infectious Agents Surveillance Report (IASR) system, which is the national surveillance system overseen by the National Institute of Infectious Diseases (<http://www.nih.go.jp/niid/ja/iasr-noro.html>). In the six months from October 2014 to March 2015, a total of 2,133 norovirus cases in the country were reported to IASR, including 373 cases caused by genotype GII.4, 146 cases caused by GII.3 and 100 cases caused by GII.17. Other genotypes (GI.2, GI.3, GI.6, GI.7, GII.2, GII.7, GII.12, GII.13 and GII.14) were also detected in this season. Although for most of the six months of the 2014/15 winter season GII.4 norovirus predominated, the

number of GII.17 cases presented a dramatic increase compared to the previous winter, whereby only three GII.17 cases had been detected from October 2013 to March 2014. In the previous five years, the average and standard deviation (SD) numbers of norovirus cases during the same months were moreover 2,727±340 for all norovirus cases, 589±256 for GII.4 cases, 130±216 for GII.3 cases and 1±1 for GII.17 cases. The first GII.17 cases in the 2014/15 winter season were observed in December 2014. In the subsequent months, more cases with this genotype continued to occur across the whole country. A sharp rise in GII.17 cases was moreover noted between January (n=11 cases) and February (n=55 cases) 2015, making GII.17 the most prevalent genotype in March (n=31 cases) 2015.

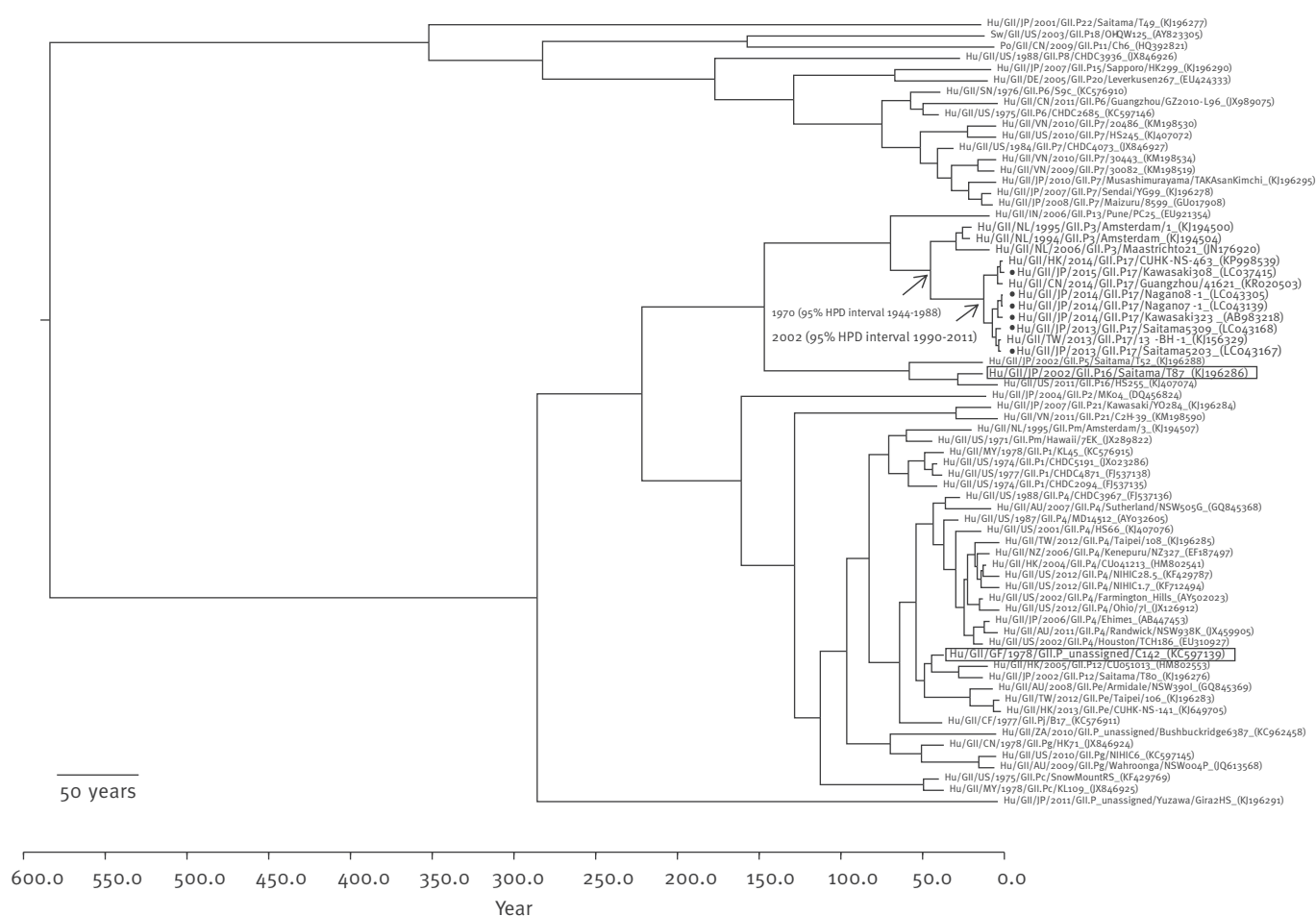
To investigate the GII.P17 norovirus strains responsible for outbreaks in Japan, and to also look for any changes in the viral genome of strains circulating between 2013 and 2015, six specimens from the Kanagawa, Nagano and Saitama prefectures that were available for further characterisation were used in this study.

## Genotyping GII.P17 strains

From the norovirus GII.17 specimens collected from January 2013 to March 2015, six available specimens (i.e. Kawasaki308 (collected in February, 2015), Kawasaki323 (collected in March, 2014), Nagano7-1 (collected in August, 2014), Nagano8-1 (collected in August, 2014), Saitama5203 (collected in April, 2013) and Saitama5309 (collected in July, 2013)) were selected for full genome analyses by next-generation

**FIGURE 1**

Time-scaled phylogenetic tree obtained with the Bayesian Markov chain Monte Carlo method of RNA-dependent RNA polymerase (RdRp) sequences (1,283 nt) of norovirus GII strains



The phylogenetic analysis includes the nucleotide sequences of six Japanese GII.P17-GII.17 strains (indicated by black dots). The oldest GII.17 strain (C142) detected in 1978 and a previous Japanese GII.17 (Saitama T87) detected in 2002 are indicated by rectangular boxes. The arrows point to some nodes, for which the node ages are indicated with 95% highest posterior density (HPD) intervals. The scale bar represents time in years.

sequencing as described [1,2]. The data analysis was performed with CLC Genomics Workbench v8.0.1 (CLC Bio). Contigs were assembled from the obtained sequence reads by de novo assembly. The nucleotide sequences for the GII.P17 strains in this study were deposited in GenBank and assigned accession numbers AB983218, LC037415, LC043139, LC043167, LC043168, and LC043305. When the Norovirus Genotyping Tool was used (<http://www.rivm.nl/mpf/norovirus/typing-tool>) [3], the capsid genotypes of all six strains were assigned to GII.17, but the RNA-dependent RNA polymerase (RdRp) genotypes of some strains could not be assigned to any known genotype in the database. This observation suggested the genetic novelty of the virus in this region. Upon first noticing this with the sequence of Kawasaki323 strain in June 2014, we sought the advice of NoroNet, who coordinate norovirus nomenclature through a global network of research scientists, clinicians and public health officials [4]. After discussions with them, this variant was assigned to the RdRp genotype, GII.P17, in August, 2014. Finally,

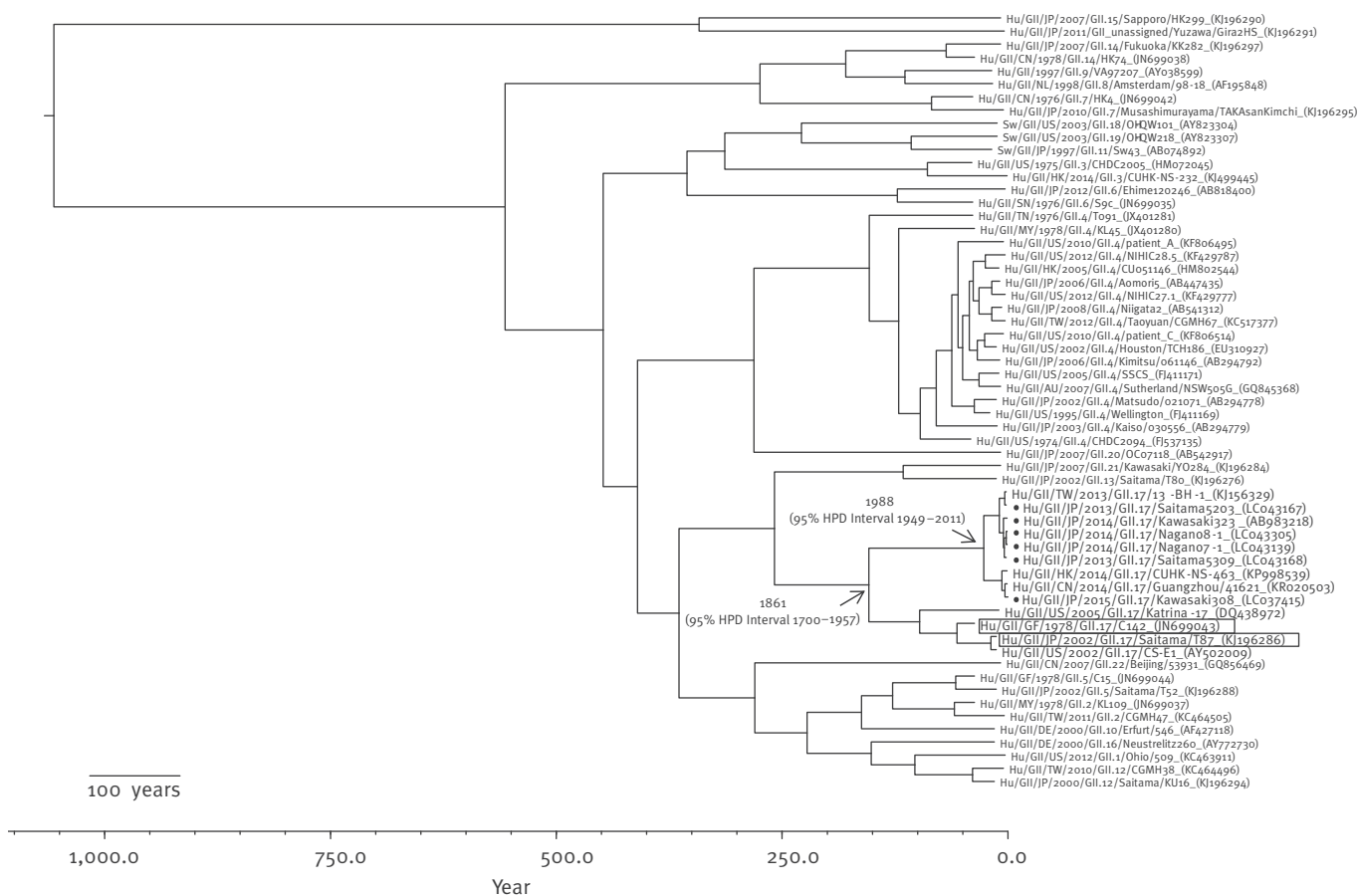
we named the emergent variants of norovirus as Hu/GII/JP/2014/GII.P17-GII.17.

### Phylogenetic and molecular dating analyses

The phylogenetic analyses, the time of most recent common ancestor (tMRCA), and the divergence times were estimated for emergent GII.P17-GII.17 variants, along with other representative norovirus GII strains, by the Bayesian Markov chain Monte Carlo (MCMC) method implemented in Bayesian Evolutionary Analysis Sampling Trees (BEAST) v1.8.1 [5]. As a result of the marginal likelihood calculation in the four clock (strict clock, uncorrelated lognormal relaxed clock, uncorrelated exponential relaxed clock and random local clock) and demographic models (constant size, exponential growth, logistic growth and expansion growth), the datasets were analysed using the Tamura and Nei 1993 (TN93)+Gamma+Proportion Invariant (for RdRp) and the generalised time-reversible (GTR)+Gamma+Proportion Invariant (for capsid) nucleotide substitution models, with an uncorrelated exponential relaxed clock model under a constant size

**FIGURE 2**

Time-scaled phylogenetic tree obtained with the Bayesian Markov chain Monte Carlo method of full length capsid gene (virus protein 1 gene) sequences of norovirus GII strains



The phylogenetic analysis includes the nucleotide sequences of six Japanese GII.P17-GII.17 strains (indicated by black dots). The oldest GII.17 strain (C142) detected in 1978 and previous Japanese GII.17 (Saitama T87) detected in 2002 are indicated by rectangular boxes. The arrows point to some nodes, for which the node ages are indicated with 95% highest posterior density (HPD) intervals. The scale bar represents time in years.

tree prior. Convergence was assessed by the effective sample size (ESS) after a 2% burn-in. Only parameters with an ESS above 150 were accepted. Phylogenetic analysis using the maximum likelihood (ML) method showed no differences of topology to the Bayesian MCMC trees (data not shown).

In terms of the RdRp gene, the emergent Japanese GII.P17 belonged to a single cluster including other strains detected in Asia in 2013 (Taiwan: KJ156329) and in 2014 (Guangzhou: KR020503; Hong Kong: KP998539). The tMRCA for the emergent GII.P17 cluster was estimated to be 2002 (95% highest posterior density (HPD) interval: 1990–2011

This suggests that the emergent GII.P17 has been circulating around Asia for ca 13 years (Figure 1). In terms of the RdRp gene, the emergent GII.P17 was most closely related to GII.P3, and evolved from a common ancestor in 1970 (95% HPD interval: 1944–1988). Nearly 32 years elapsed between the tMRCA of the emergent GII.P17 cluster, and the divergence of emergent GII.P17 strains

from the GII.P3 cluster (Figure 1). Interestingly, the oldest GII.17 (C142 detected in French Guiana in 1978) and the oldest Japanese GII.17 (Saitama T87 detected in 2002) had genotype combinations that were different from the emergent GII.17 forms in this study, the GII.P\_unassigned-GII.17 for C142 and GII.P16-GII.17 for Saitama T87 (Figure 1). These results suggested that capsid GII.17 genotype evolved by exchanging the RdRp gene through at least two recombination events.

In terms of the capsid gene, the emergent GII.P17-GII.17 strains also belonged to a single cluster that was different from the cluster formed by the older GII.17 strains (Figure 2). Notably, the emerging GII.17 cluster diverged from the old GII.17 cluster around the year 1861 (95% HPD interval: 1700–1957), yet the tMRCA was in the year 1988 (95% HPD interval: 1949–2011) (Figure 2). The two independent clusters formed by the old and emerging GII.17 strains partially arose from changes in the amino acids of the major epitopes (Figure 2 and Table). Thus, we thought that the emerging GII.17

**TABLE**

**Amino acid substitutions in the major epitopes of virus protein 1 between novel GII.17 variants and previously identified GII.17 strains of norovirus**

Name of strains <sup>a</sup>	Amino acid number (major epitopes) <sup>b</sup>																																																		
	217–225 (I)					291–298 (II)					359–363 (III)					371–379 (IV)				390–396 (V)																															
	P	P	S	V	E	S	K	T	K	T	A	D	V	H	Q	S	H	D	D	S	S	Q	F	V	F	G	S	T	S	D	-	D	F	Q	I	K	V	E	S	G	-	H									
Saitama T87	.	.	.	.	.	.	.	.	.	.	.	.	.	D	G	.	.	D	D	.	.	.	.	.	.	.	.	.	T	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.				
CS-E1	.	.	.	.	.	.	.	.	.	.	.	.	.	D	G	.	.	D	D	.	.	.	.	.	.	.	.	.	T	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		
Katrina-17	.	.	.	.	.	.	.	S	.	.	.	.	.	Q	N	.	.	Q	D	.	T	.	.	.	.	.	.	E	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Saitama5203 <sup>c</sup>	.	.	.	.	.	.	.	.	.	.	E	T	-	E	T	-	D	.	R	D	.	P	.	.	.	.	R	.	N	D	N	-	.	.	.	.	.	N	D	D	D	G	D	-	H	.	.	.			
Saitama5309 <sup>c</sup>	.	.	.	.	.	.	.	.	.	.	E	T	-	E	T	-	D	.	R	D	.	P	.	.	.	.	R	.	N	D	N	-	.	.	.	.	.	N	D	D	D	G	D	-	H	.	.	.			
13-BH-1	.	.	.	.	.	.	.	.	.	.	E	T	-	E	T	-	D	.	R	D	.	P	.	.	.	.	R	.	N	D	N	-	.	.	.	.	.	N	D	D	D	G	D	-	H	.	.	.			
Kawasaki323 <sup>c</sup>	.	.	.	.	.	.	.	.	.	.	E	T	-	E	T	-	D	.	R	D	.	P	.	.	.	.	R	.	N	D	N	-	.	.	.	.	.	N	D	D	D	G	D	-	H	.	.	.			
Nagan07-1 <sup>c</sup>	.	.	.	.	.	.	.	.	.	.	E	T	-	E	T	-	D	.	R	D	.	P	.	.	.	.	R	.	N	D	N	-	.	.	.	.	.	N	D	D	D	G	D	-	H	.	.	.			
Nagan08-1 <sup>c</sup>	.	.	.	.	.	.	.	.	.	.	E	T	-	E	T	-	D	.	R	D	.	P	.	.	.	.	R	.	N	D	N	-	.	.	.	.	.	N	D	D	D	G	D	-	H	.	.	.			
CUHK-NS-463	.	.	.	.	.	.	.	.	.	.	Q	I	-	Q	I	-	N	Q	R	D	.	P	.	.	.	.	L	R	I	S	D	N	D	.	.	.	V	N	D	D	D	D	D	G	H	.	.	.			
Guangzhou41621	.	.	.	.	.	.	.	.	.	.	Q	I	-	Q	I	-	N	Q	R	D	.	P	.	.	.	.	L	R	I	S	D	N	D	.	.	.	V	N	D	D	D	D	D	G	H	.	.	.			
Kawasaki308 <sup>c</sup>	.	.	.	.	.	.	.	.	.	.	Q	I	-	Q	I	-	N	Q	R	D	.	P	.	.	.	L	R	I	S	D	N	D	.	.	.	V	N	D	D	D	D	D	D	G	H	.	.	.			

New GII.17-GII.17 variants are highlighted in red. Previous GII.17 strains are highlighted in black. Dots indicate sequence identity among sequences presented. A dash represents the relative deletion of an amino acid at a certain position, compared to other sequences in the alignment.

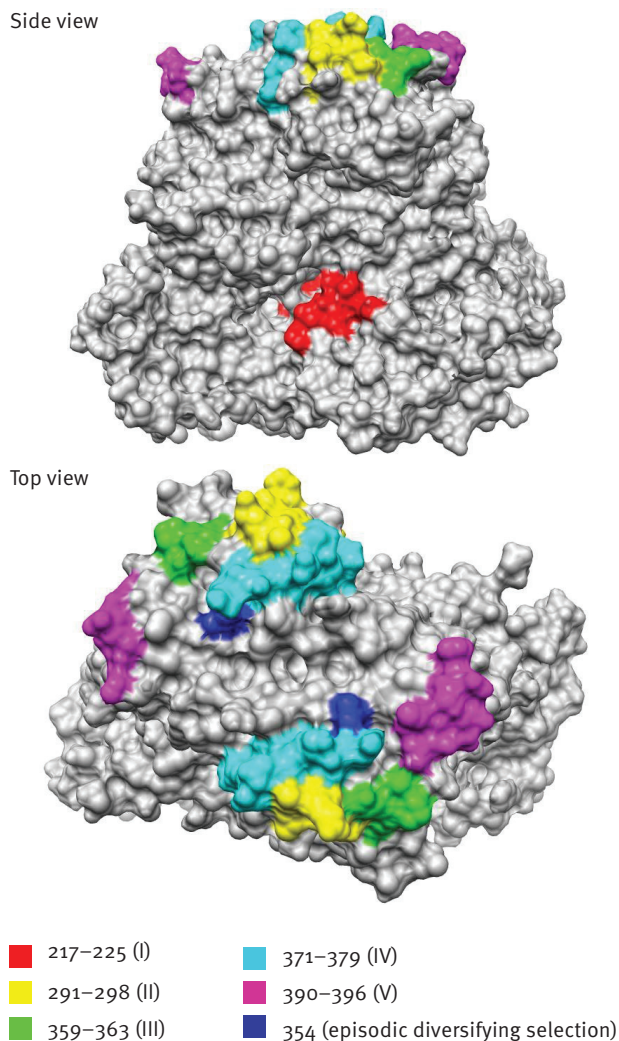
<sup>a</sup> The GenBank accession numbers of the sequences are as follows: C142 (JN699043); CS-E1 (AY502009); Katrina-17 (DQ438972); Saitama T87 (KJ196286); Kawasaki308 (LC037415); Kawasaki323 (AB983218); Nagan07-1 (LC043139); Nagan08-1 (LC043305); Saitama5203 (LC043167); Saitama5309 (LC043168).

<sup>b</sup> Amino acid numbering is based on the sequence of the Kawasaki323 strain.

<sup>c</sup> Japanese GII.17-GII.17 strains in this study.

### FIGURE 3

Structural model of the dimer formed by virus protein 1 (VP1) of Kawasaki323 GII.17 strain



The GII.17 VP1 model was constructed by homology modelling with the crystal structure of norovirus capsid (1IHM). Five predicted epitopes (I–V) are shown by each colour with the amino acid positions based on the sequence of Kawasaki323 strain. Amino acids with episodic diversifying selection are highlighted in dark blue.

strains may have variants of the capsid gene, as has been observed with some GII.4 strains [6].

Moreover within the emergent GII.P17-GII.17 cluster, diversification of strains further led to two sub-clusters in the capsid and RdRp genes, with changes to the amino acids of the major epitopes of the virus protein 1 (VP1) protein (Figures 1 and 2 and Table). One sub-cluster comprised the Kawasaki308 strain, while the other sub-cluster included the Kawasaki323 strain as well as all the other four GII.P17-GII.17 strains from this study. These results clearly suggest that the new GII.P17-GII.17 variants have different evolutionary histories than previously identified GII.17 strains, and that rapid evolution may occur within the emergent GII.P17-GII.17 variants. The GII.17 genotype may produce other variants whereby mutations lead to changes

in the antigenicity of the P2 domain while still being constrained by host immunity, in the same way as has been observed for GII.4 [6].

### Estimation of positive selection sites and B-cell epitopes in virus protein 1 sequence of the GII.P17-GII.17 variant

We analysed the evolutionary constraints on the GII.P17-GII.17 variants from host-immune pressure, based on the single likelihood ancestor counting (SLAC), fixed effect likelihood (FEL), internal FEL (IFEL), fast unconstrained bayesian approximation for inferring selection (FUBAR), random effects likelihood (REL) and mixed effects model of evolution (MEME) with only the GII.17 sequences in the dataset used in phylogenetic analyses for VP1 gene, but using another alignment [7]. One positive selection (V354W) (V, sub-cluster including Kawasaki323 strain; W, sub-cluster including Kawasaki308) was identified with the MEME analysis. This non-synonymous mutation was only observed in the sub-cluster of Kawasaki308 strain, suggesting that the selection was episodic but not pervasive. Moreover, the BepiPred and DiscoTope servers [8,9] predicted B-cell epitopes associated with humoral immunity at the amino acid positions 217–225 (I), 291–298 (II), 359–363 (III), 371–379 (IV), and 390–396 (V) in the VP1 protein of the Kawasaki323 strain (Table). Epitope I was conserved in all sequences within the GII.17 cluster, whereas the others (II–V) were variable with amino acid substitutions, deletions, and insertions not only between the novel GII.P17-GII.17 variant and old GII.17 strains, but also within the GII.P17-GII.17 variant cluster, between sub-clusters represented by the Kawasaki323 and Kawasaki308 strains (Table).

To identify these amino acid positions on the capsid structure, we calculated and constructed a capsid 3D model of the Kawasaki323 strain with the MODELLER 9.13 programme [10]. The epitopes were located on the exterior surface of the shell (epitope I) and the protruding 2 (P2) domains (epitope II–V), including the binding pocket for the histo-blood group antigens (Figure 3). Additionally, the mutation (V354W) (V, sub-cluster including Kawasaki323 strain; W, sub-cluster including Kawasaki308) associated with episodic positive selection was close to the 372R within the epitope IV (Figure 3). These results indicate that our GII.P17-GII.17 variants might have the potential to escape from host neutralising antibody by amino acid alterations of four putative B-cell epitopes in the P2 domain top, and to improve the binding capacity of the histo-blood group antigens. During the winter 2014/15, noroviruses similar to the Kawasaki308 strain became extremely prevalent in Japan and China [11], and this observation may imply a high infectivity of the strain because a number of mutations, including positive selection, were observed between strains detected in the 2013/14 and 2014/15 seasons.

## Conclusions

During the season from October 2014 to March 2015, a novel norovirus variant GII.P17-GII.17 was prevalent in Japan from December 2014 onwards. While, in general, the total number of norovirus cases during this season was lower than previous years beyond a range of average $\pm$ SD, the number of cases affected by the GII.17 genotype appeared to be higher and increased dramatically in February 2015, making this the predominant genotype in the country in March 2015. Further characterisation GII.17 available strains from January 2013 to March 2015 assigned these to the novel GII.P17-GII.17 variant. Because this variant was detected from a few cases in Japan and Taiwan, prior to becoming prevalent and causing outbreaks during the 2014/15 winter season in Japan and China [11], early surveillance of sporadic cases caused by this or any other potential variants may assist in anticipating outbreaks. Molecular and phylogenetic analyses conducted here show that the novel GII.17 variant has a different evolutionary history to previously identified GII.17 strains. As it might have the potential to spread globally in the near future, presumably by escaping host immunity as GII.4 variants do [6], monitoring trends in the geographical spread and evolution of the variant is important.

## Acknowledgments

This work was partially supported by a commissioned project for the Research on Emerging and Re-emerging Infectious Diseases from the Japanese Ministry of Health, Labour and Welfare and for the Research Program on Emerging and Re-emerging Infectious Diseases from Japan Agency for Medical Research and development (AMED). The research was approved by research and ethical committee in Kawasaki City Institute for Public Health.

## Conflict of interest

None declared.

## Authors' contributions

Yuki Matsushima and Hideaki Shimizu designed the study. Yuki Matsushima and Kazuhiko Katayama drafted the manuscript. Yuki Matsushima, Mariko Ishikawa and Koo Nagasawa performed bioinformatic analyses. Yuki Matsushima, Mariko Ishikawa, Tomomi Shimizu, Ayako Komane and Hideaki Shimizu conducted experiments using samples derived from Kawasaki City, Japan. Shizuko Kasuo performed experiments with samples collected in Nagano Prefecture, Japan. Michiyo Shinohara conducted experiments using samples derived from Saitama Prefecture, Japan. Kei Haga and Yen Hai Doan performed sequencing analyses of these strains using a next generation sequence system. Yuki Matsushima, Mariko Ishikawa, Hirokazu Kimura, Akihide Ryo, Nobuhiko Okabe, Yen Hai Doan, Kazuhiko Katayama and Hideaki Shimizu reviewed and revised the manuscript.

## References

1. Dennis FE, Fujii Y, Haga K, Damanka S, Lartey B, Agbembiesie CA, et al. Identification of novel Ghanaian G8P[6] human-bovine reassortant rotavirus strain by next generation

- sequencing. *PLoS ONE*. 2014;9(6):e100699. <http://dx.doi.org/10.1371/journal.pone.0100699> PMID:24971993
2. Ide T, Komoto S, Higo-Moriguchi K, Htun KW, Myint YY, Myat TW, et al. Whole Genomic Analysis of Human G12P[6] and G12P[8] Rotavirus Strains that Have Emerged in Myanmar. *PLoS ONE*. 2015;10(5):e0124965. <http://dx.doi.org/10.1371/journal.pone.0124965> PMID:25938434
3. Kroneman A, Vennema H, Deforche K, v d Avoort H, Peñaranda S, Oberste MS, et al. An automated genotyping tool for enteroviruses and noroviruses. *J Clin Virol*. 2011;51(2):121-5. <http://dx.doi.org/10.1016/j.jcv.2011.03.006> PMID: 21514213
4. Kroneman A, Vega E, Vennema H, Vinjé J, White PA, Hansman G, et al. Proposal for a unified norovirus nomenclature and genotyping. *Arch Virol*. 2013;158(10):2059-68. <http://dx.doi.org/10.1007/s00705-013-1708-5> PMID:23615870
5. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*. 2012;29(8):1969-73. <http://dx.doi.org/10.1093/molbev/mss075> PMID:22367748
6. Debbink K, Lindesmith LC, Donaldson EF, Baric RS. Norovirus immunity and the great escape. *PLoS Pathog*. 2012;8(10):e1002921. <http://dx.doi.org/10.1371/journal.ppat.1002921> PMID:23093932
7. Delpont W, Poon AF, Frost SD, Kosakovsky Pond SL. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*. 2010;26(19):2455-7. <http://dx.doi.org/10.1093/bioinformatics/btq429> PMID:20671151
8. Kringelum JV, Lundegaard C, Lund O, Nielsen M. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput Biol*. 2012;8(12):e1002829. <http://dx.doi.org/10.1371/journal.pcbi.1002829> PMID:23300419
9. JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res*. 2006;2(1):2. <http://dx.doi.org/10.1186/1745-7580-2-2> PMID:16635264
10. Webb B, Sali A. Protein structure modeling with MODELLER. *Methods Mol Biol*. 2014;1137:1-15. [http://dx.doi.org/10.1007/978-1-4939-0366-5\\_1](http://dx.doi.org/10.1007/978-1-4939-0366-5_1) PMID:24573470
11. Lu J, Sun L, Fang L, Yang F, Mo Y, Lao J, et al. Gastroenteritis outbreaks caused by norovirus GII.17, Guangdong Province, China. *Emerg Infect Dis*. 2015;21(7):1240-2. <http://dx.doi.org/10.3201/eid2107.150226> PMID:26080037