



Research article

Molecular evolution of the capsid (VP1) region in human norovirus genogroup II genotype 3



Mariko Saito^a, Hiroyuki Tsukagoshi^a, Hirotaka Ishigaki^b, Jumpei Aso^{b,c}, Haruyuki Ishii^c, Kaori Okayama^b, Akihide Ryo^d, Taisei Ishioka^b, Makoto Kuroda^e, Nobuhiro Saruki^a, Kazuhiko Katayama^f, Hirokazu Kimura^{b,d,*}

^a Gunma Prefectural Institute of Public Health and Environmental Sciences, 378 Kamioki-machi, Maebashi-shi, Gunma 371-0052, Japan

^b Department of Health Science, Gunma Paz University Graduate School of Health Sciences, 1-7-1 Tonyamachi, Takasaki-shi, Gunma 370-0006, Japan

^c Kyorin University Hospital, 6-20-2 Shinkawa, Mitaka-shi, Tokyo 181-8611, Japan

^d Department of Microbiology, Yokohama City University School of Medicine, 3-9 Fukuura, Yokohama-shi, Kanagawa 236-0004, Japan

^e Pathogen Genomics Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

^f Laboratory of Viral Infection I, Kitasato Institute for Life Sciences Graduate School of Infection Control Sciences, Kitasato University, 5-9-1 Shirokane, Minato-ku, Tokyo 108-8641, Japan

ARTICLE INFO

Keywords:

Bioinformatics

Microbiology

Virology

Viral disease

Viral genetics

Norovirus

GII.3

Molecular evolution

VP1

ABSTRACT

Norovirus GII.3 has been suggested to be a prevalent genotype in patients with acute gastroenteritis. However, the genetic properties of the VP1 region encoding the major GII.3 antigen remain unclear. Here, we performed molecular evolutionary analyses of the GII.3 VP1 region detected in various countries. We performed time-scaled phylogenetic analyses, selective pressure analyses, phylogenetic distance analyses, and conformational epitope analyses. The time-scaled phylogenetic tree showed that an ancestor of the GII.3 VP1 region diverged from the common ancestors of the GII.6, GII.11, GII.18, and GII.19 approximately 70 years ago with relatively low divergence. The evolutionary rate of the GII.3 VP1 region was rapid (4.82×10^{-3} substitutions/site/year). Furthermore, one positive site and many negative selection sites were observed in the capsid protein. These results suggest that the GII.3 VP1 region rapidly evolved with antigenic variations.

1. Introduction

Noroviruses (NoVs) are nonenveloped RNA viruses belonging to the family *Caliciviridae* and genus *Norovirus*. Numerous previous reports have indicated that NoV is a major causative agent of acute gastroenteritis in humans of all ages (Patel et al., 2008; Robilotti et al., 2015; Mans, 2019). Indeed, NoV was responsible for >90% of patients with viral gastroenteritis, rendering this infection a major disease burden worldwide (Patel et al., 2008).

The NoV genome length is approximately 7.6 kb, comprising a single plus-strand RNA; the genome has three open reading frames (ORFs) and encodes eight proteins (Robilotti et al., 2015). ORF1 encodes six nonstructural proteins, including a protease and RNA-dependent RNA polymerase, whereas ORF2 encodes a capsid protein, VP1, as the major antigen (Robilotti et al., 2015). Moreover, ORF3 encodes another capsid protein, VP2, although its function remains unclear (Green, 2013). A previous report showed that NoV is classified into two genogroups

[genogroup I (GI) and GII], which are detailed assigned to over 30 genotypes (GI.1–GI.9 and GII.1–GII.22) (Kroneman et al., 2013; Vinjé, 2015). Among them, compared with NoV GI viruses, NoV GII viruses are more dominantly detected in patients with acute gastroenteritis (Patel et al., 2008; Green, 2013). Moreover, some genotypes of GII viruses, such as GII.3, GII.4, GII.6, and GII.7, are frequently detected in these patients (Green, 2013). In particular, the genotype GII.4 suddenly emerged during the 2006/07 season and caused a pandemic of gastroenteritis worldwide (Tu et al., 2007; Motomura et al., 2008; Siebenga et al., 2009; Lam et al., 2012). Some previous reports have suggested that GII.3 is the predominant genotype in infants and children (Boon et al., 2011; Wangchuk et al., 2017; Boonchan et al., 2018; Liu et al., 2018). Indeed, GII.3 was the most frequently detected virus in children with gastroenteritis in Japan during 2003–2004 (Phan et al., 2006). To date, several molecular epidemiology and molecular evolutionary studies regarding certain prevalent genotypes, such as GII.2, GII.4, and GII.17, have been reported (Motoya et al., 2017; Parra et al., 2017; Siqueira et al., 2017; Mizukoshi et al., 2017;

* Corresponding author.

E-mail address: h-kimura@paz.ac.jp (H. Kimura).

<https://doi.org/10.1016/j.heliyon.2020.e03835>

Received 7 September 2019; Received in revised form 7 November 2019; Accepted 21 April 2020

2405-8440/© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Nagasawa et al., 2018); however, studies investigating GII.3 evolution in detail are scarce. Here, to understand the characteristics of GII.3 evolution, we performed a molecular evolutionary study of the VP1 region of GII.3 detected in various countries.

2. Materials and methods

2.1. Strains

The full-length NoV VP1 nucleotide coding region were collected from GenBank in November 11, 2019. The collected sequences were assigned using a NoV genotyping tool (Kroneman et al., 2011), and the GII.3 genotype strains were selectively collected. Among them, those with ambiguous sequences and an unknown year of collection were omitted from the data. Moreover, the Recombination Detection Program v. 4.95 as previously described (7 methods, BOOTSCAN/RECSCAN, CHIMAERA, GENECONV, MAXCHI, RDP, SISCAN, and 3SEQ) was used for recombination analysis (Martin et al., 2015). Significance of the p -value was less than 0.001. Recombinant was determined when they were confirmed by > four of these methods. However, no recombinant regions were found. In this study, we used 239 full-length of the GII.3 VP1 sequences in the dataset (Table S1).

2.2. Time-scaled phylogeny and estimation of the rate of evolution using the Bayesian Markov chain Monte Carlo (MCMC) method

To perform the molecular evolutionary analysis of the present strains, we constructed phylogenetic trees of the NoV VP1 region using the MCMC method. Evolutionary dynamics of the molecular clock was examined by the MCMC method in the BEAST package v2.4.8 (Suchard et al., 2001; Drummond and Rambaut, 2007; Bouckaert et al., 2014). To make the accurate phylogenetic tree, we further used the all GII genotypes sequences, including porcine NoV GII (GII.11, GII.18) and other human NoV genogroup II genotypes (19 strains), and an outgroup strain human NoV genogroup I genotype (GI.1) in the present dataset (total 261 strains in Table S1). First, we used the suitable substitution model using the jModelTest 2.1.10 program (Guindon and Gascuel, 2003; Darriba et al., 2012). Thereafter, the best of 4 clock models (strict clock, relaxed clock exponential, relaxed clock log normal, and random local clock) and 2 tree prior models (coalescent constant population and coalescent exponential population) were estimated using the path-sampling/stepping stone-sampling marginal likelihood estimation method. Finally, the optimal dataset was estimated as the relaxed clock exponential and exponential tree prior models. The MCMC chains were 250,000,000 steps with sampling every 5,000 steps. The convergence of all parameters (Effective Sample Size > 200), through inspection with Tracer v1.6. After discarding the 10% burn-in, phylogenetic trees were generated with TreeAnnotator v2.4.8 and visualized with FigTree v1.4.0. The credibility of each branch was supported by the 95% highest posterior density (HPD). In addition, the evolution rates of GII.3 were assumed by suitable models selected for each dataset as described above.

2.3. Calculation of the phylogenetic distance

To calculate the phylogenetic distance, we constructed phylogenetic tree of all GII.3 strains based on the maximum likelihood method using MEGA 7.0 software (Kumar et al., 2016). jModelTest 2.1.10 was used to select the most suitable evolutionary model. We calculated the phylogenetic distances in the present phylogenetic trees using Patristic program (Fourment and Gibbs, 2006).

2.4. Creation of the capsid protein structure and estimation of conformational epitopes for the B-cell in the VP1 protein

The NoV GII.3 VP1 dimer structural model of Hu/NoV/GII.3/Toronto 24/1991/CA (GenBank accession no. U02030) was created by

MODELLER v9.20 (Webb and Sali, 2014). The templates for homology modeling were made by the crystal structures of 7 strains (PDB ID: 1IHM, 5F4M, 4RPD, 3PUM, 3PA1, 4RPB, and 3SEJ). The capsid structure of GI (PDB ID: 1IHM) was used as a template to construct a shell domain in whole capsid protein (VP1) structures. The sequences of amino acid of each strain were aligned using MAFFTash (Katoh et al., 2002; Standley et al., 2007). The created structures were minimized using the GROMOS96 (van Gunsteren et al., 1996), included by Swiss PDB Viewer v4.1 (Guex and Peitsch, 1997), and predicted by Ramachandran plots via an available server RAMPAGE (Lovell et al., 2003). The final model was made/colored using Chimera v1.13.1 (Pettersen et al., 2004). BEPro (Sweredoski and Baldi, 2008), DiscoTope 2.0 (Kringelum et al., 2012), EPSVR (Liang et al., 2010), and EPCES (Liang et al., 2007) were employed for the prediction of conformational B-cell epitopes for the capsid VP1 protein of the GII.3 strain (Hu/NoV/GII.3/Toronto 24/1991/CA) with cut-off values of the epitopes were set at 1.3 (BEPro), -3.7 (DiscoTope 2.0), and 70 (EPSVR, EPCES). Consensus sites in the 4 tools and regions with close residues more than 2 of the sites on the dimer structures of capsid protein were estimated as conformational epitopes.

2.5. Positive and negative selection analyses

The rates of synonymous (dS) and non-synonymous (dN) substitution at each codon were estimated by the Datamonkey server to assume the selective sites in the capsid protein coding (VP1) region of GII.3 (Pond and Frost, 2005; Delpont et al., 2010). We used FEL, IFEL, and SLAC methods and estimated negative ($dN < dS$) and positive ($dN > dS$) selection sites. A significance level was $p < 0.05$. The two-tail extended binomial distribution (SLAC method) was applied to estimate the p -value. The single-degree-of-freedom likelihood ratio test (FEL and IFEL methods; the chi-squared asymptotic distribution was used) to estimate a site as positive or negative selection.

2.6. Estimation of the similarities

Similarities between the nucleotide sequences and Hu/NoV/GII.3/Toronto 24/1991/CA were estimated by an available software (SimPlot program 3.5.1) as previously described (Lole et al., 1999). A reference strain Hu/NoV/GII.3/Toronto 24/1991/CA was used as the query sequence. The similarity was examined using the Kimura 2-parameter method with window size of 200 nucleotides in length and a step size of 20 nucleotides in the full-length VP1 coding region.

2.7. Phylodynamics based on the Bayesian skyline plots (BSP)

To examine the phylodynamics of the GII.3 strains based on BSP analyses were made by BEAST v2.4.8. as previously described (Suchard et al., 2001; Drummond and Rambaut, 2007; Bouckaert et al., 2014). The best clock and substitution models for these analyses were estimated as described above. Visualization of the analyzed plots were made by the 95% HPD with Tracer.

2.8. Statistical analysis

The Kruskal-Wallis test was applied to determine significant differences between the clusters for the evolutionary rates and phylogenetic distance.

3. Results

3.1. Phylogeny and evolution rates of the capsid protein (VP1) coding region in NoV GII.3

To make time-scale evolutionary analysis, we created a time-scaled phylogenetic tree based on the MCMC method. The tree showed that

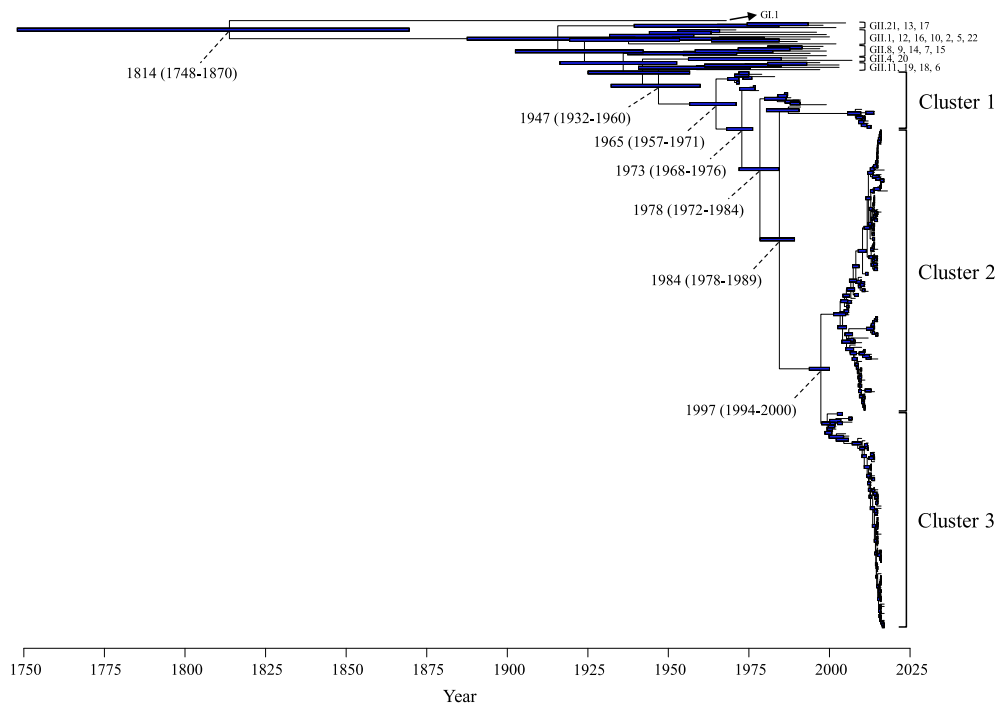


Figure 1. Time-scaled phylogenetic tree for the GII.3 VP1 region constructed using the Bayesian Markov chain Monte Carlo method. The scale bar represents time (years). Blue bars indicate the 95% highest posterior density for the branched year.

the VP1 region is divided into three clusters (Figure 1). The most recent common ancestor of the GII.3 strains was assumed at 1947 (95% HPD, 1932–1960) and formed three clusters after 1965 (95% HPD, 1957–1971). Cluster 1 emerged in 1965 (95% HPDs, 1957–1971). Subsequently, clusters 2 and 3 emerged in 1997 (95% HPD, 1994–2000). The evolutionary rate of the VP1 region in GII.3 strains was estimated as 4.82×10^{-3} substitutions/site/year (95% HPD, $4.14\text{--}5.48 \times 10^{-3}$ substitutions/site/year). The nucleotide substitution rates showed a significant difference among each cluster in GII.3 (Kruskal–Wallis test; $p < 0.001$). Comparing the evolutionary rates among the clusters, that of cluster 2 was the most rapid estimated at 5.29×10^{-3} substitutions/site/year (95% HPD, $4.29\text{--}6.34 \times 10^{-3}$ substitutions/site/year) (Table S2).

3.2. Phylogenetic distances (p-distance) of the VP1 region in the GII.3 strains

We estimated the phylogenetic distance (p-distance) of the VP1 region in the GII.3 strains. The p-distance value of all GII.3 strains was 0.112 ± 0.069 (mean \pm standard deviation, SD) (Figure 2). The mean values of each cluster were within the range 0.029–0.116, and the p-distances were significantly different among GII.3 clusters. Cluster 1 had the largest phylogenetic distance [0.116 ± 0.061 (mean \pm SD)].

3.3. Mapping of conformational epitopes and positive selection sites on the capsid (VP1) protein structures of GII.3

To comprehensively estimate between the conformational epitopes on the GII.3 VP1 protein and positive/selection sites, we mapped them in the protein using *in silico* methods (Figure 3). Five sites were estimated as conformational epitopes. Of them, 2 epitopes were located in the exterior surface of the P2 domain. Moreover, several amino acid substitutions were estimated on and/or around these epitopes.

One positively selected site (aa385) was estimated as follows: Gly385Asp, Ser, His, Asp385Gly, Ser, and Ser385Gly, Asp. This was located in the P2 domain, while many negatively selected sites were estimated (257 sites).

3.4. Similarity plot analysis of the VP1 region in the present GII.3 strains

Using SimPlot, we performed similarity plot analysis based on the full-length VP1 region (Figure 4). The results revealed that the P1 and P2 domains yielded more divergence of the VP1 region in the present GII.3. In cluster 2 of the GII.3 strains, the similarities between the P1 and P2 domains (80.1%–91.5%) were lower than those that occurred in the shell domain (84.7%–94.2%).

3.5. Phylodynamics of the GII.3 VP1 coding regions

To estimate the changes of the phylodynamics of the GII.3, we calculated the effective population sizes (EPS) using the BSP method. The EPS values of the VP1 region in the GII.3 strains constant up to 2013. Thereafter, the EPS slightly dropped (Figure 5). In cluster 3 of the GII.3 strains, the mean EPS gradually reduced since 2007, although the EPS

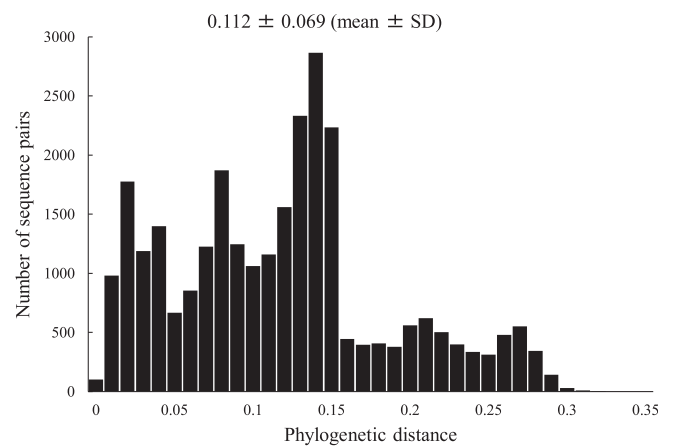


Figure 2. Phylogenetic distance of intra-genotypes in GII.3 strains. The y-axis represents the number of sequence pairs corresponding to each distance. The x-axis shows phylogenetic distances.

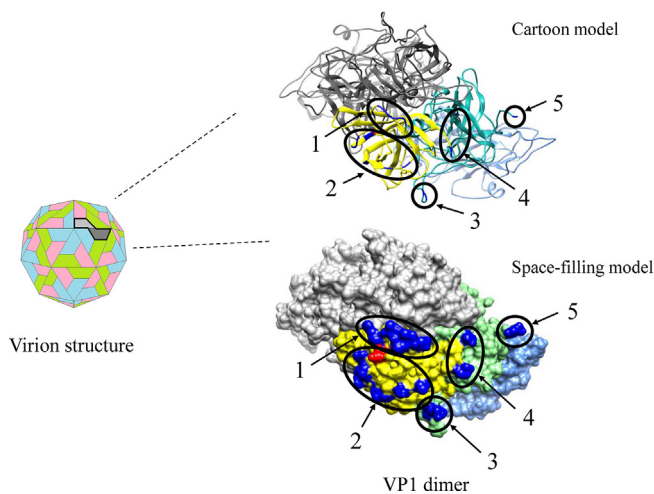


Figure 3. Structural models for the capsid VP1 protein of GII.3. The virion structure and three-dimensional VP1 dimer structures (cartoon and space-filling models) for the Hu/NoV/GII.3/Toronto 24/1991/CA strain are shown. Chains comprising the dimer structures are colored as follows: yellow: P2 domain (chain A); green: P1 domain (chain A); light blue: shell domain (chain A); gray: chain B; blue: predicted epitopes (chain A); red: positive selection site (aa385).

slightly increased by 2012. No changes in the mean EPS of the other sample clusters were observed.

4. Discussion

In this study, we performed detailed molecular evolutionary analyses of the GII.3 VP1 region detected in various countries. Thus, an ancestor of the GII.3 VP1 region diverged from the common ancestors of the GII.6, GII.11, GII.18, and GII.19 strains approximately 70 years ago and formed several clusters with relatively low genetic divergence in the time-scaled phylogenetic tree (Figure 1). The evolution rates of the GII.3 VP1 coding region was estimated at 4.82×10^{-3} substitutions/site/year. Furthermore, one positive site and many negative selection sites were found in the GII.3 capsid protein. These results suggest that the GII.3 VP1 region rapidly evolved and had antigenic variations.

The present phylogenetic tree showed that a common ancestor of the GII.3 VP1 region diverged at around 70 years ago and further diverged at 1984 and 1997, resulting in the formation of three clusters (Figure 1).

The common ancestor of this genotype was emerged after another major genotype, GII.4, while GII.3 emerged prior to GII.2 (Figure 1) (Motoya et al., 2017; Mizukoshi et al., 2017; Nagasawa et al., 2018). Previous studies have also suggested that the GII.3 VP1 region is phylogenetically classified into three major clusters (clusters I to III) (Boon et al., 2011; Mahar et al., 2013), a finding that is compatible with the present data. However, the present strains belonging to cluster 1 correspond to clusters I and II, while the present strains assigning to clusters 2 and 3 correspond to cluster III (Boon et al., 2011; Mahar et al., 2013). These results suggest that the GII.3 strain continuously evolved and adapted to humans.

In the present study, the evolutionary rate of the GII.3 VP1 region collected from various countries was assumed at 4.82×10^{-3} substitutions/site/year. Previous studies have reported the evolutionary rates in some genotypes of the NoV VP1 region (Boon et al., 2011; Mahar et al., 2013; Kobayashi et al., 2016; Motoya et al., 2017; Parra et al., 2017; Mizukoshi et al., 2017; Nagasawa et al., 2018). For example, Boon et al. showed that the evolutionary rate of 7 GII.3 strains detected in the USA and 56 GII.3 strains available in GenBank was 4.16×10^{-3} substitutions/site/year (Boon et al., 2011). Moreover, another report estimated that the rate of evolution in 6 GII.3 strains collected from Australian children and 66 GII.3 strains from GenBank was 4.16×10^{-3} to 6.97×10^{-3} substitutions/site/year (Mahar et al., 2013). Current data was similar to previous reports, while there were some differences in the sequence numbers, methodology, and collection date (Boon et al., 2011; Mahar et al., 2013). These may reflect the differences in evolutionary rates between the previous (3.76×10^{-3} substitutions/site/year; Kobayashi et al., 2016) and present data. In addition, the evolutionary rate of the VP1 region in another major prevalent type GII.4, which was detected in various countries, was estimated at approximately 7.7×10^{-3} substitutions/site/year (Motoya et al., 2017), and this value was higher than that in the GII.3 strain. Moreover, genetic divergence (phylogenetic distance) of the GII.3 VP1 region was shorter than that of GII.4 (0.112 ± 0.069 vs. 0.210 ± 0.105) (Motoya et al., 2017). These results suggested that the divergence of the GII.3 VP1 coding region may be relatively low.

Next, we analyzed conformational epitopes and mapped them on the GII.3 capsid protein (Figure 3). Previous reports have suggested that most epitopes were estimated in the P2 domain (Motoya et al., 2017; Nagasawa et al., 2018). Motoya et al. showed that four epitopes were found in the GII.4 P2 domain (Motoya et al., 2017). Moreover, Nagasawa et al. showed that epitopes were mainly found in the GII.2 P2 domain (Nagasawa et al., 2018). These reports also suggested that the differences in the amino acid substitutions may reflect differences in the antigenicity of these genotypes (Motoya et al., 2017; Nagasawa et al., 2018). In this

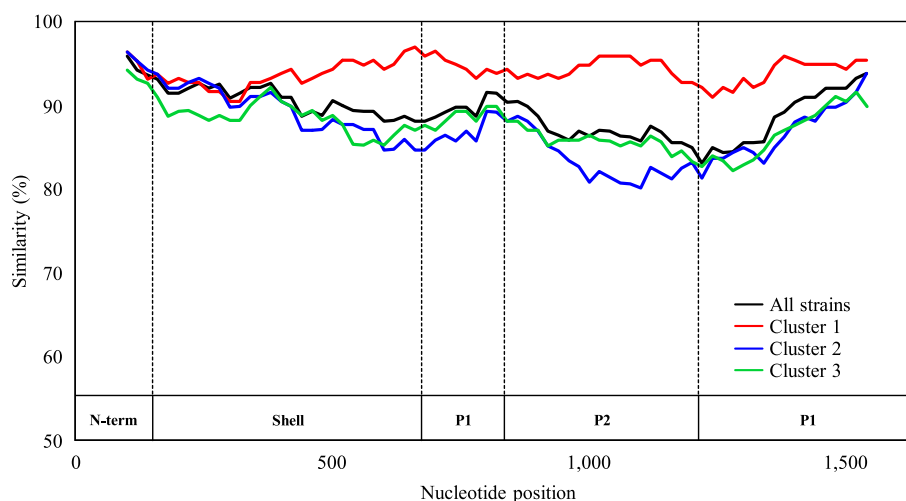


Figure 4. Similarity plot analysis of the VP1 region across GII.3 strains. Nucleotide similarity to the Hu/NoV/GII.3/Toronto 24/1991/CA strain was calculated using SimPlot analysis. Each cluster similarity to the Hu/NoV/GII.3/Toronto 24/1991/CA is represented. The positions of the shell, P1, and P2 domains in VP1 region are shown below the graph.

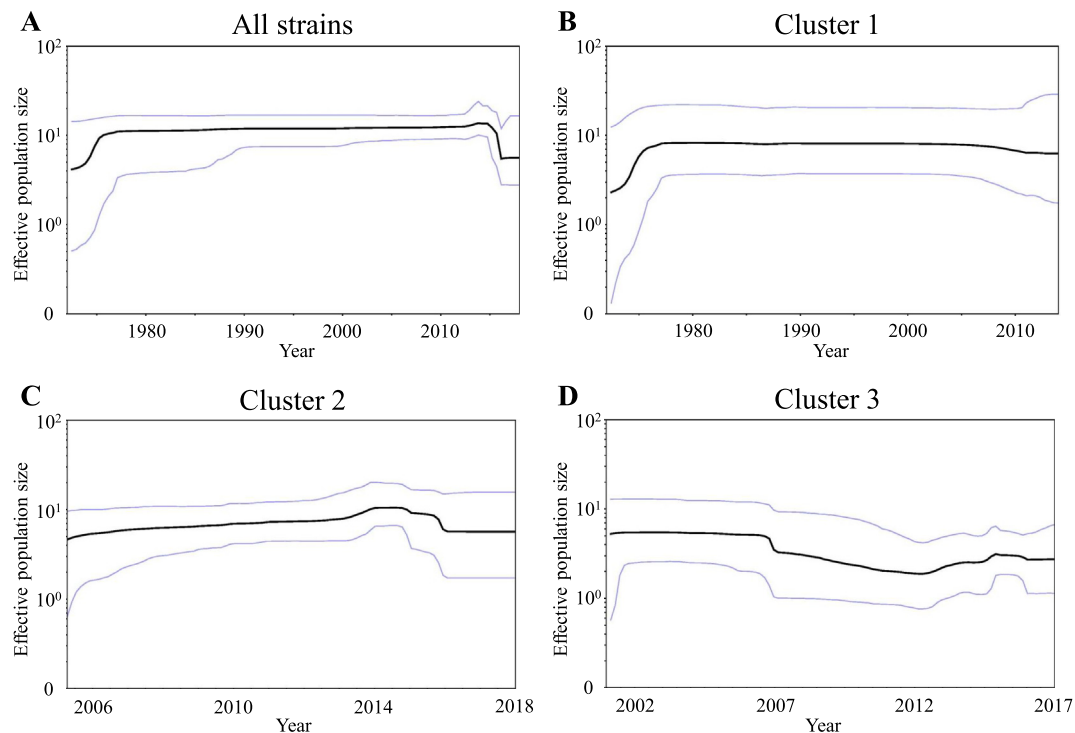


Figure 5. Bayesian skyline plot for the VP1 sequences of GII.3. Plots for all GII.3 strains (A), cluster 1 (B), cluster 2 (C), and cluster 3 (D) are shown. The y-axis represents the effective population size on the logarithmic scale, whereas the x-axis denotes the time in years. The solid black line indicates the median posterior value. The intervals with the highest posterior densities (95%) are indicated by blue lines.

study, we showed that the GII.3 strain includes two epitopes in the P2 domain and three epitopes in the P1 domain (Figure 3). Among them, aa385 in the P2 domain was estimated as a positive selection site (Gly385Asp, Ser, His, Asp385Gly, Ser, and Ser385Gly, Asp). This site may act as a surface-exposed residue, resulting in a pivotal binding site of a neutralizing antibody (Mahar et al., 2013). Thus, it is possible that the positive selection site plays important roles in the changes in the GII.3 capsid protein antigenicity (Boon et al., 2011; Mahar et al., 2013, 2014).

Moreover, we performed phylodynamic analysis of the GII.3 VP1 region (Figure 5). Overall, phylodynamic fluctuations of the gene was minimal. However, in cluster 3 of the GII.3 strains, the size was slightly increased by 2012. Very recent molecular epidemiological studies have reported outbreaks of gastroenteritis due to the GII.3 strain (Boonchan et al., 2018; Kim et al., 2018; Liu et al., 2018). Thus, the fluctuation of cluster 3 may reflect the transient prevalence of the GII.3 strain (Kim et al., 2018).

Recent studies have suggested that NoV pandemics may be associated with recombination between ORF1 and ORF2, as well as with the properties of the RNA-dependent RNA polymerase (RdRp) (Bull et al., 2010; Eden et al., 2013; Ozaki et al., 2018). In this study, we only performed the molecular evolutionary analysis in the GII.3 VP1 region. Previous reports showed that the recombination between NoV ORF1 (including the RdRp region) and ORF2 (including the VP1 region) resulted in the emergence of the distinct RdRp genotype of NoV (Bidalot et al., 2017; Lu et al., 2017; Nagasawa et al., 2018; Niendorf et al., 2017; Tohma et al., 2017). Furthermore, the VP1 region co-evolved with the RdRp region (Mizukoshi et al., 2017; Ozaki et al., 2018). Therefore, the RdRp protein may affect the evolutionary process of the capsid proteins (VP1 and VP2 proteins) (Mizukoshi et al., 2017). To deeply study the molecular evolutionary analyses of the NoV, the analyses of both the RdRp and VP1 regions may be considered (Mizukoshi et al., 2017).

In conclusion, NoV GII.3 strains are frequently detected following other genotypes GII.4 and GII.2. Although there were some sporadic cases where GII.3 did not cause pandemics of acute gastroenteritis like

GII.4 (Bull et al., 2010; White, 2014). However, whether the prevalence of the GII.3 is in the near future may not be known yet. Therefore, continuous molecular epidemiological and molecular evolutionary analyses of GII.3 may be required.

Declarations

Author contribution statement

Mariko Saito, Hirokazu Kimura: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Hiroyuki Tsukagoshi: Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Hirohisa Ishigaki, Jumpei Aso, Kaori Okayama, Nobuhiro Saruki: Performed the experiments.

Haruyuki Ishii, Kazuhiko Katayama: Conceived and designed the experiments.

Akihide Ryo, Taisei Ishioka: Analyzed and interpreted the data.

Makoto Kuroda: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data.

Funding statement

This work was supported by a commissioned project for Research on Emerging and Re-emerging Infectious Diseases from the Japan Agency for Medical Research and Development, AMED (grant number JP20fk0108120h0401).

Competing interest statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at <https://doi.org/10.1016/j.heliyon.2020.e03835>.

Acknowledgements

We thank Dr. Kiyotaka Fujita for valuable discussion. We also thank Keita Ozaki for technical support.

References

- Bidalot, M., Théry, L., Kaplon, J., De Rougemont, A., Ambert-Balay, K., 2017. Emergence of new recombinant noroviruses GII.p16-GII.4 and GII.p16-GII.2, France, winter 2016 to 2017. *Euro Surveill.* 22, 30508.
- Boon, D., Mahar, J.E., Abente, E.J., Kirkwood, C.D., Purcell, R.H., Kapikian, A.Z., Green, K.Y., Bok, K., 2011. Comparative evolution of GII.3 and GII.4 norovirus over a 31-year period. *J. Virol.* 85, 8656–8666.
- Boonchan, M., Guntapong, R., Sripirom, N., Ruchusatsawat, K., Singchai, P., Rungnobbhakhun, P., Tacharoenmuang, R., Mizushima, H., Tatsumi, M., Takeda, N., Sangkitporn, S., Mekmullica, J., Motomura, K., 2018. The dynamics of norovirus genotypes and genetic analysis of a novel recombinant GII.P12-GII.3 among infants and children in Bangkok, Thailand between 2014 and 2016. *Infect. Genet. Evol.* 60, 133–139.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10, e1003537.
- Bull, R.A., Eden, J.S., Rawlinson, W.D., White, P.A., 2010. Rapid evolution of pandemic noroviruses of the GII.4 lineage. *PLoS Pathog.* 6, e1000831.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772.
- Delpoit, W., Poon, A.F., Frost, S.D., Kosakovsky Pond, S.L., 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* 26, 2455–2457.
- Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7, 214.
- Eden, J.S., Tanaka, M.M., Boni, M.F., Rawlinson, W.D., White, P.A., 2013. Recombination within the pandemic norovirus GII.4 lineage. *J. Virol.* 87, 6270–6282.
- Fourment, M., Gibbs, M.J., 2006. PATRISTIC: a program for calculating patristic distances and graphically comparing the components of genetic change. *BMC Evol. Biol.* 6, 1.
- Green, K.Y., 2013. *Caliciviridae: the noroviruses*. In: Knipe, D.M., Howley, P.M., Cohen, J.L., Griffin, D.E., Lamb, R.A., Martin, M.A., Racaniello, V.R., Roizman, B. (Eds.), *Fields Virology*, sixth ed. Lippincott Williams & Wilkins, Philadelphia, pp. 582–608.
- Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704.
- Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18, 2714–2723.
- Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066.
- Kim, Y.E., Song, M., Lee, J., Seung, H.J., Kwon, E.Y., Yu, J., Hwang, Y., Yoon, T., Park, T.J., Lim, I.K., 2018. Phylogenetic characterization of norovirus strains detected from sporadic gastroenteritis in Seoul during 2014–2016. *Gut Pathog.* 10, 36.
- Kobayashi, M., Matsushima, Y., Motoya, T., Sakon, N., Shigemoto, N., Okamoto-Nakagawa, R., Nishimura, K., Yamashita, Y., Kuroda, M., Saruki, N., Ryo, A., Saraya, T., Morita, Y., Shirabe, K., Ishikawa, M., Takahashi, T., Shinomiya, H., Okabe, N., Nagasawa, K., Suzuki, Y., Katayama, K., Kimura, H., 2016. Molecular evolution of the capsid gene in human norovirus genogroup II. *Sci. Rep.* 6, 29400.
- Kringelum, J.V., Lundegaard, C., Lund, O., Nielsen, M., 2012. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput. Biol.* 8, e1002829.
- Kroneman, A., Vega, E., Vennema, H., Vinjé, J., White, P.A., Hansman, G., Green, K., Martella, V., Katayama, K., Koopmans, M., 2013. Proposal for a unified norovirus nomenclature and genotyping. *Arch. Virol.* 158, 2059–2068.
- Kroneman, A., Vennema, H., Deforche, K., v d Avoort, H., Peñaranda, S., Oberste, M.S., Vinjé, J., Koopmans, M., 2011. An automated genotyping tool for enteroviruses and noroviruses. *J. Clin. Virol.* 51, 121–125.
- Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874.
- Lam, T.T., Zhu, H., Smith, D.K., Guan, Y., Holmes, E.C., Pybus, O.G., 2012. The recombinant origin of emerging human norovirus GII.4/2008: intra-genotypic exchange of the capsid P2 domain. *J. Gen. Virol.* 93, 817–822.
- Liang, S., Liu, S., Zhang, C., Zhou, Y., 2007. A simple reference state makes a significant improvement in near-native selections from structurally refined docking decoys. *Proteins* 69, 244–253.
- Liang, S., Zheng, D., Standley, D.M., Yao, B., Zacharias, M., Zhang, C., 2010. EPSVR and EPMeta: prediction of antigenic epitopes using support vector regression and multiple server results. *BMC Bioinf.* 11, 381.
- Liu, J., Li, S., Wang, C., Zheng, L., Ma, J., Li, C., Huo, Y., Wang, Y., 2018. Genomic characterization of GII.3 noroviruses isolated from children in Zhengzhou city, China, 2015/16. *Arch. Virol.* 163, 2737–2742.
- Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G., Ingersoll, R., Sheppard, H.W., Ray, S.C., 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* 73, 152–160.
- Lovell, S.C., Davis, I.W., Arendall III, W.B., de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., Richardson, D.C., 2003. Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins* 50, 437–450.
- Lu, J., Fang, L., Sun, L., Zeng, H., Li, Y., Zheng, H., Wu, S., Yang, F., Song, T., Lin, J., Ke, C., Zhang, Y., Vinjé, J., Li, H., 2017. Association of GII.P16-GII.2 recombinant norovirus strain with increased norovirus outbreaks, Guangdong, China, 2016. *Emerg. Infect. Dis.* 23, 1188–1190.
- Mahar, J.E., Bok, K., Green, K.Y., Kirkwood, C.D., 2013. The importance of intergenic recombination in norovirus GII.3 evolution. *J. Virol.* 87, 3687–3698.
- Mahar, J.E., Donker, N.C., Bok, K., Talbo, G.H., Green, K.Y., Kirkwood, C.D., 2014. Identification and characterization of antibody-binding epitopes on the norovirus GII.3 capsid. *J. Virol.* 88, 1942–1952.
- Mans, J., 2019. Norovirus infections and disease in lower-middle- and low-income countries, 1997–2018. *Viruses* 11, 341.
- Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol.* 1 (vev003).
- Mizukoshi, F., Nagasawa, K., Doan, Y.H., Haga, K., Yoshizumi, S., Ueki, Y., Shinohara, M., Ishikawa, M., Sakon, N., Shigemoto, N., Okamoto-Nakagawa, R., Ochi, A., Murakami, K., Ryo, A., Suzuki, Y., Katayama, K., Kimura, H., 2017. Molecular evolution of the RNA-dependent RNA polymerase and capsid genes of human norovirus genotype GII.2 in Japan during 2004–2015. *Front. Microbiol.* 8, 705.
- Motomura, K., Oka, T., Yokoyama, M., Nakamura, H., Mori, H., Ode, H., Hansman, G.S., Katayama, K., Kanda, T., Tanaka, T., Takeda, N., Sato, H., 2008. Identification of monomorphic and divergent haplotypes in the 2006–2007 norovirus GII/4 epidemic population by genomewide tracing of evolutionary history. *J. Virol.* 82, 11247–11262.
- Motoya, T., Nagasawa, K., Matsushima, Y., Nagata, N., Ryo, A., Sekizuka, T., Yamashita, A., Kuroda, M., Morita, Y., Suzuki, Y., Sasaki, N., Katayama, K., Kimura, H., 2017. Molecular evolution of the VP1 gene in human norovirus GII.4 variants in 1974–2015. *Front. Microbiol.* 8, 2399.
- Nagasawa, K., Matsushima, Y., Motoya, T., Mizukoshi, F., Ueki, Y., Sakon, N., Murakami, K., Shimizu, T., Okabe, N., Nagata, N., Shirabe, K., Shinomiya, H., Suzuki, W., Kuroda, M., Sekizuka, T., Suzuki, Y., Ryo, A., Fujita, K., Oishi, K., Katayama, K., Kimura, H., 2018. Genetic analysis of human norovirus strains in Japan in 2016–2017. *Front. Microbiol.* 9, 1.
- Niendorf, S., Jacobsen, F., Faber, M., Eis-Hübinger, A.M., Hofmann, J., Zimmermann, O., Höhne, M., Bock, C.T., 2017. Steep rise in norovirus cases and emergence of a new recombinant strain GII.P16-GII.2, Germany, winter 2016. *Euro Surveill.* 22, 30447.
- Ozaki, K., Matsushima, Y., Nagasawa, K., Motoya, T., Ryo, A., Kuroda, M., Katayama, K., Kimura, H., 2018. Molecular evolutionary analyses of the RNA-dependent RNA polymerase region in norovirus genogroup II. *Front. Microbiol.* 9, 3070.
- Parra, G.L., Squires, R.B., Karangwa, C.K., Johnson, J.A., Lepore, C.J., Sosnovtsev, S.V., Green, K.Y., 2017. Static and evolving norovirus genotypes: implications for epidemiology and immunity. *PLoS Pathog.* 13, e1006136.
- Patel, M.M., Widdowson, M.A., Glass, R.I., Akazawa, K., Vinjé, J., Parashar, U.D., 2008. Systematic literature review of role of noroviruses in sporadic gastroenteritis. *Emerg. Infect. Dis.* 14, 1224–1231.
- Petersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612.
- Phan, T.G., Kuroiwa, T., Kaneshi, K., Ueda, Y., Nakaya, S., Nishimura, S., Yamamoto, A., Sugita, K., Nishimura, T., Yagyu, F., Okitsu, S., Müller, W.E., Maneekam, N., Ushijima, H., 2006. Changing distribution of norovirus genotypes and genetic analysis of recombinant GIIb among infants and children with diarrhea in Japan. *J. Med. Virol.* 78, 971–978.
- Pond, S.L., Frost, S.D., 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21, 2531–2533.
- Robilotti, E., Deresinski, S., Pinsky, B.A., 2015. Norovirus. *Clin. Microbiol. Rev.* 28, 134–164.
- Siebenga, J.J., Vennema, H., Zheng, D.P., Vinjé, J., Lee, B.E., Pang, X.L., Ho, E.C., Lim, W., Choudhary, A., Broor, S., Halperin, T., Rasool, N.B., Hewitt, J., Greening, G.E., Jin, M., Duan, Z.J., Lucero, Y., O’Ryan, M., Hoehne, M., Schreiber, E., Ratcliff, R.M., White, P.A., Iritani, N., Reuter, G., Koopmans, M., 2009. Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001–2007. *J. Infect. Dis.* 200, 802–812.
- Siqueira, J.A.M., Bandeira, R.D.S., Oliveira, D.S., Dos Santos, L.F.P., Gabbay, Y.B., 2017. Genotype diversity and molecular evolution of noroviruses: a 30-year (1982–2011) comprehensive study with children from northern Brazil. *PLoS One* 12, e0178909.
- Standley, D.M., Toh, H., Nakamura, H., 2007. ASH structure alignment package: sensitivity and selectivity in domain classification. *BMC Bioinf.* 8, 116.
- Suchard, M.A., Weiss, R.E., Sinsheimer, J.S., 2001. Bayesian selection of continuous-time Markov chain evolutionary models. *Mol. Biol. Evol.* 18, 1001–1013.
- Sweredski, M.J., Baldi, P., 2008. PEPITO: improved discontinuous B-cell epitope prediction using multiple distance thresholds and half sphere exposure. *Bioinformatics* 24, 1459–1460.

- Tohma, K., Lepore, C.J., Ford-Siltz, L.A., Parra, G.I., 2017. Phylogenetic analyses suggest that factors other than the capsid protein play a role in the epidemic potential of GII.2 norovirus. *mSphere* 2 e00187-17.
- Tu, E.T., Nguyen, T., Lee, P., Bull, R.A., Musto, J., Hansman, G., White, P.A., Rawlinson, W.D., McIver, C.J., 2007. Norovirus GII.4 strains and outbreaks, Australia. *Emerg. Infect. Dis.* 13, 1128–1130.
- van Gunsteren, W.F., Billeter, S.R., Eising, A.A., Hünenberger, P.H., Krüger, P., Mark, A.E., Scott, W.R.P., Tironi, I.G., 1996. *Biomolecular Simulation: the GROMOS96 Manual and User Guide*. Vdf Hochschulverlag AG an der ETH Zürich, Zürich, pp. 1–1042.
- Vinje, J., 2015. Advances in laboratory methods for detection and typing of norovirus. *J. Clin. Microbiol.* 53, 373–381.
- Wangchuk, S., Matsumoto, T., Iha, H., Ahmed, K., 2017. Surveillance of norovirus among children with diarrhea in four major hospitals in Bhutan: replacement of GII.21 by GII.3 as a dominant genotype. *PloS One* 12, e0184826.
- Webb, B., Sali, A., 2014. Protein structure modeling with MODELLER. *Methods Mol. Biol.* 1137, 1–15.
- White, P.A., 2014. Evolution of norovirus. *Clin. Microbiol. Infect.* 20, 741–745.