

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Norovirus GII.17: The Emergence and Global Prevalence of a Novel Variant

Yongxin Yu, Yingjie Pan, Shuling Yan and
Yongjie Wang

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.76256>

Abstract

A rare norovirus (NoV) genotype GII.17 has recently emerged and rapidly became predominant in most East Asian countries in the winters of 2014–2015. In this study, we report the diversity of NoV GII.17 in detail; a total of 646 GII.17 sequences obtained during 1978–2015 were analyzed and subjected to meta-analysis. At least five major recombinant GII.17 clusters were identified. Each recombinant variant group appeared to have emerged following the time order: GII.P4-GII.17 (1978–1990), GII.P16-GII.17 (2001–2004), GII.P13-GII.17 (2004–2010), GII.Pe-GII.17 (2012–2015) and GII.P3-GII.17 (2011–2015). The newly emerged GII.P3-GII.17 variant, which exhibited significant sequence and structure variations, is evolving toward a unique lineage. Our results indicate that circulation of GII.17 appears to change every 3–5 years due to replacement by a newly emerged variant and that the evolution of GII.17 is sequentially promoted by inter-genotype recombination, which contributes to the exchange between non-GII.17 and GII.17 RdRp genes and drives the evolution of GII.17 capsid genes.

Keywords: emergence, evolution, genotyping, global prevalence, norovirus, recombination

1. Introduction

Norovirus (NoV) is the predominant etiological viral agent of acute gastroenteritis across all ages; usually old people and young kids are more susceptible to these viruses [1]. Currently, the genus of *norovirus* can be divided into at least seven genogroups (GI–GVII). Of these, GI, GII and

GIV can be detected in the samples from human gastroenteritis [2, 3]. Within each genogroup, NoV strains can be further subdivided into diverse (over 40) genotypes based on sequence similarity of the RNA-dependent RNA polymerase (RdRp) and major capsid genes [2, 3]. Globally, GII strains have already contributed to over 75% of human NoV cases [4, 5], whereas a special genotype GII.4 has been found to be responsible for the majority of outbreaks since 1990s [6, 7], and novel GII.4 variants emerged every 2–4 years [6–8].

In the winter of 2014–2015, a rare NoVs genotype, GII.17, emerged in most of the East Asian regions including China (Guangdong [9–13], Jiangsu [14], Zhejiang [15], Hebei [16], Hong Kong [17], Taiwan [18], Beijing [19], and Shanghai [20], Japan [21] and South Korea [22]). Soon after, the novel GII.17 strain became the most predominant NoV strain, replacing the pandemic strain GII.4 Sydney 2012, responsible for the majority of gastroenteritis outbreaks in this region [23]. A limited number of cases were also reported in other countries (Italy [24], Romania [25] and the USA [26]).

Recombination allows a substantial exchange of genetic materials and is a major driving force of viral evolution [2, 27]. However, although multiple studies have reported on the evolutionary dynamics of NoVs, the role of recombination in shaping NoV evolutionary history is still very significant. Given the importance of GII.17 NoV as a cause of epidemic gastroenteritis in recent years (2014–2015), it is crucial to better understand how this genotype has evolved over time.

This study aimed to determine the mechanisms of evolution in norovirus GII.17 strain from 1978 to 2015, with particular focus on the effects of recombination events on the acquisition of non-GII.17 RdRps.

2. Materials and methods

2.1. NoV GII.17 sequence datasets

Sequence datasets were constructed following the strategy described by Yu et al. [28]. Briefly, NoV GII.17 sequences were all collected from two divided pathways (**Figure 1**): firstly, search results from the GenBank nucleotide database using combination keywords such as “norovirus and GII.17” and “norovirus and 17” in December 2015 and secondly, publications from the PubMed and Google scholar literature databases (papers published between 2003 and 2015) that contain a combination of “norovirus” and “GII.17” in the titles, keywords and abstracts [28]. Non-English literatures were excluded [28].

Subsequently, all the NoV GII.17 sequences were downloaded in FASTA format. Then the corresponding information of each sequence was edited by using Geneious [29] as in a uniform format which included the sequence name (or accession number) and length, sample source, sampling time and sites [28]. Duplicated sequences were removed [28]. To delete the non-GII.17 sequences, all the candidate sequences were analyzed with the online-based genotyping tool [30], which is designed to identify norovirus genotypes based on phylogenetic analysis.

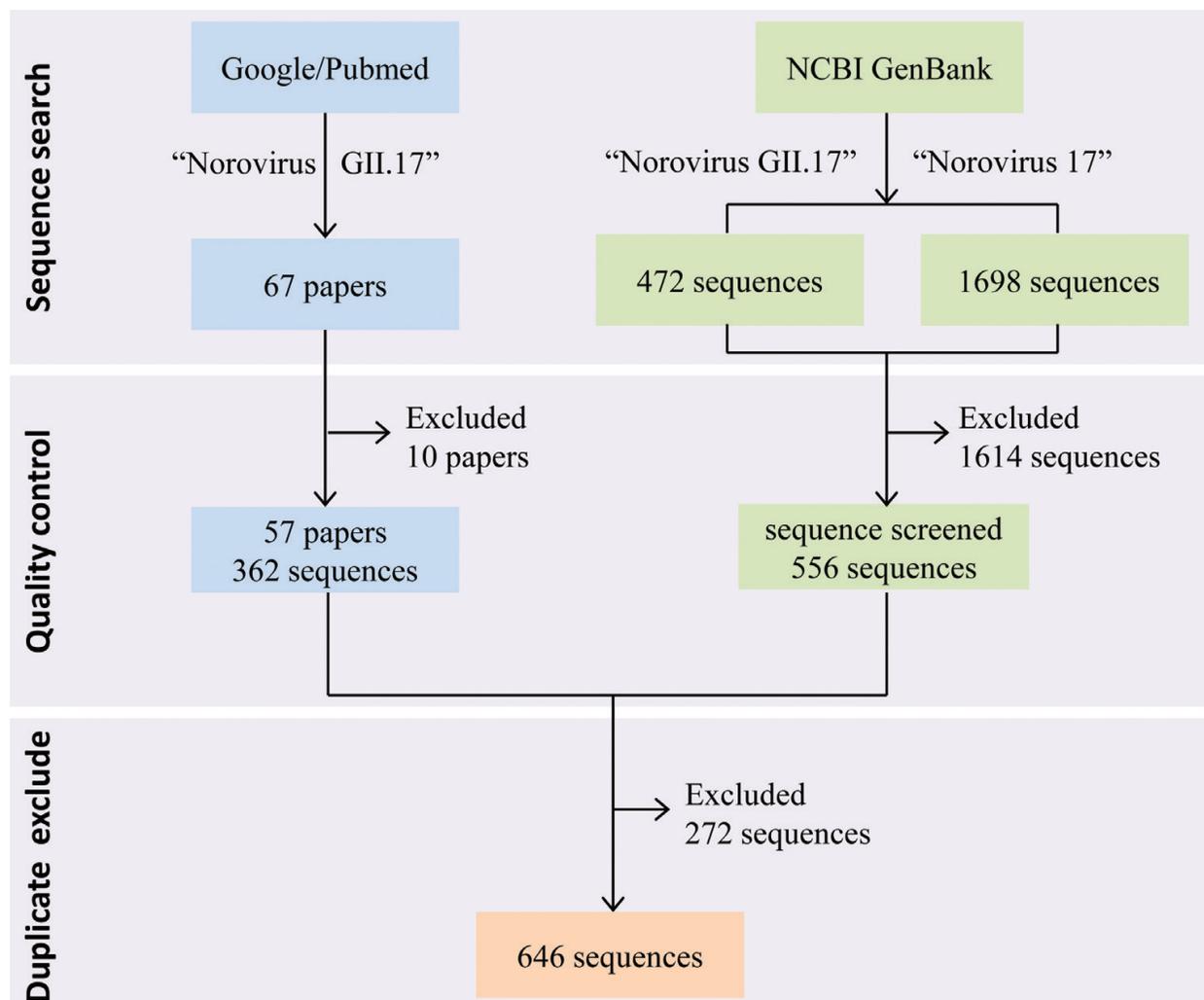


Figure 1. Flow chart of sequence collection strategy.

2.2. Phylogenetic analysis

Nucleotide sequences were aligned using the ClustalW program. Phylogenetic analysis was performed with MEGA 5.1 package [31] based on partial sequences of ORF1 (241 nt) and ORF2 (181 nt). The reference strains were retrieved from NCBI nucleotide database [32]. Phylogenetic trees were reconstructed using the Tamura-Nei model and maximum likelihood methods. Bootstrap was calculated with 1000 pseudo-replicate datasets. The distance scale represents the number of nucleotide substitution per position.

2.3. Recombination analysis

To detect recombination events, sequences were aligned with ClustalW program in the MEGA 5.1 package [31] and then checked manually. The reference strains were retrieved from NCBI nucleotide database. NoV strains were defined as recombinants if they were grouped into different genotypic clusters on the phylogenetic trees, which were reconstructed using full-length sequences of ORF1 (5107 nt) and ORF2 (1623 nt), respectively.

Simplot method [33] was employed to identify the recombination breakpoint site and to further verify the NoV recombinants. The bootstrap values were plotted for a window of 400 nt, moving in increments of 10 nt along the alignment.

2.4. Homology modeling

The tertiary structures of the capsid P domains of NoV GII.17 were modeled using SWISS-MODEL online server [34]. The recently published GII.17 domain dimer X-ray crystal structure (PDB accession: 5F4O) [35] was used as the template for generating homology models. The constructed models were examined and edited using PyMoL [36].

3. Results

3.1. NoV GII.17 sequence dataset

A total of 472 and 1698 sequences were obtained from the GenBank nucleotide database using “norovirus and GII.17” and “norovirus and 17” as search keywords, respectively. After manually checking sequence genotyping results, the non-GII.17 sequences ($n = 1614$) of the above two sequence datasets were excluded from subsequent analysis (**Figure 1**). A literature search yielded 67 citations, among which 57 of the research articles (85%) reported norovirus GII.17 sequences ($n = 362$). All screened sequences from these two independent sources were combined, and duplicated sequences were removed. Finally, a total of 646 sequences belonging to NoV GII.17 were obtained (**Figure 1**).

These 646 sequences ranged from 205 to 7570 nt in length. Over 60% of them ($n = 427$) were shorter than 400 nt, while 57 sequences covered the nearly complete or the complete viral genome of more than 7000 nt. All the NoV GII.17 sequences were located within ORF1, ORF2 or the region overlapping ORF1 and ORF2. Specifically, 3.41% of sequences ($n = 22$) belonged to ORF1, and 79.26% ($n = 512$) belonged to ORF2. The remaining 112 sequences contained regions from both ORF1 and ORF2.

3.2. Genetic diversity of NoV GII.17

Genotyping analysis revealed that many of the GII.17 sequences were genetic recombinants since two distinct genotypic regions, ORF1 (RdRp) and ORF2 (VP 1), were identified in the same sequence. All the recombinants contained the same ORF2 genotype of GII.17 while exhibiting varying ORF1 genotypes of GII.P3, GII.P13, GII.P16, GII.Pe and GII.P4. Phylogenetically, all the GII.17 ORF2 sequences were categorized into at least five major clusters. Interestingly, each cluster composed of only one type of GII.17 recombinant, for example, GII.P4-GII.17, GII.P16-GII.17, GII.P13-GII.17, GII.Pe-GII.17 and GII.P3-GII.17 (**Figure 2A**). The collection dates of the corresponding strains within each cluster indicated a temporally sequential clustering and were distributed as follows: GII.P4-GII.17 cluster during 1978–1990, GII.P16-GII.17 cluster during 2001–2004, GII.P13-GII.17 cluster during 2004–2010, GII.Pe-GII.17 cluster during 2012–2015 and GII.P3-GII.17 cluster during 2011–2015 (**Figure 2A**).

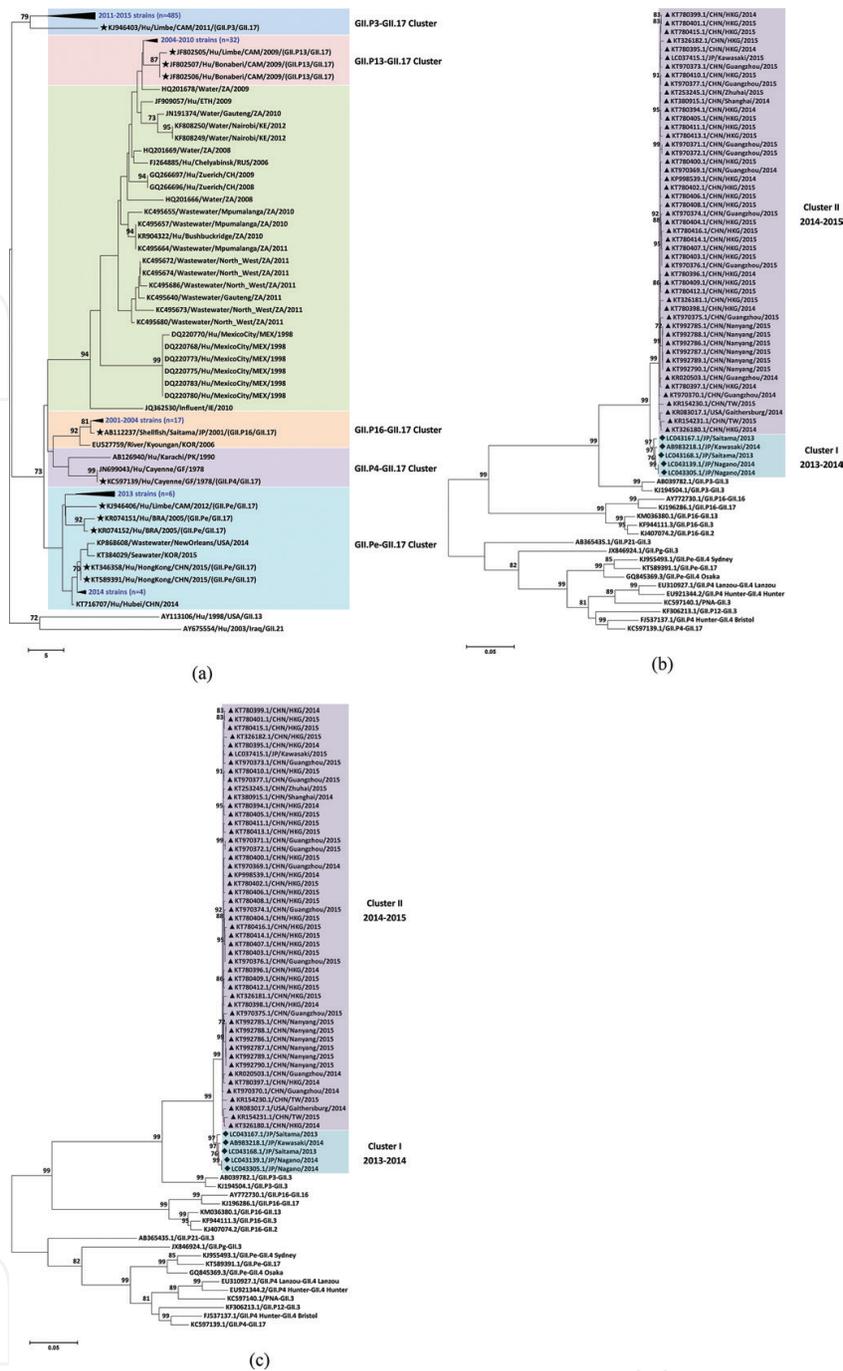


Figure 2. Phylogenetic trees of NoV GII.17. The trees were generated using the maximum likelihood method. Gaps in alignment were ignored in the analysis. Bootstrap (1000 replicates) analysis was used for statistical support; values >70% are shown. Tamura-Nei was used as the model of substitution. The distance scale represents the number of nucleotide substitutions per position. The sequences were obtained from GenBank and named according to the GenBank accession ID, strain name, followed by the year and country of isolation (AUS, Australia; BD, Bangladesh; BRA, Brazil; CAM, Cameroon; CHN, China; ETH, Ethiopia; FRA, France; GF, French Guiana; IE, Ireland; ITA, Italy; JP, Japan; KE, Kenya; KOR, Korea; MEX, Mexico; MOR, Morocco; NIC, Nicaragua; PK, Pakistan; RUS, Russia; SIG, Singapore; ZA, South Africa; CH, Switzerland; TH, Thailand; UY, Uruguay; the USA, United States of America). (A) The rooted tree was reconstructed using the 5' ends of VP1 nucleotide sequences (189 nt) of all GII.17 NoVs, corresponding to the location of LC037415: 5141–5329 nt. The recombinant strains are indicated by black pentagrams. The recombinant variant clusters are highlighted in different colors. GII.13 (AY113106) and GII.21 (AY675554) were used as the out-groups. Two un-rooted trees were reconstructed using near-complete ORF1 (B) and ORF2 (C) nucleotide sequences (n = 72). The recently emerged NoV GII.17 sequences (2013–2015) are indicated with diamonds (cluster I, 2013–2014) and triangles (Cluster II, 2014–2015).

Notably, 27 sequences did not belong to any of the above-described lineages (**Figure 2A**); however, they were more closely related to the GII.P13-GII.17 cluster. Coincidentally, all of these sequences were isolated in a similar time period (1998–2012) as the GII.P13-GII.17 cluster (2004–2010) (**Figure 2A**).

The GII.P3-GII.17 cluster, containing 486 sequences, was the largest recombinant lineage of the five recombinant clusters, with most of the sequences discovered in the past 2 years (2014 and 2015) (**Figure 2A**). In addition, the VP1 sequences (viral protein 1, the major capsid protein) of the recently emerged epidemic strains during 2014 and 2015 were also classified into this GII.P3-GII.17 cluster (**Figure 2A**).

3.3. Recombination confirms the emergence of GII.17

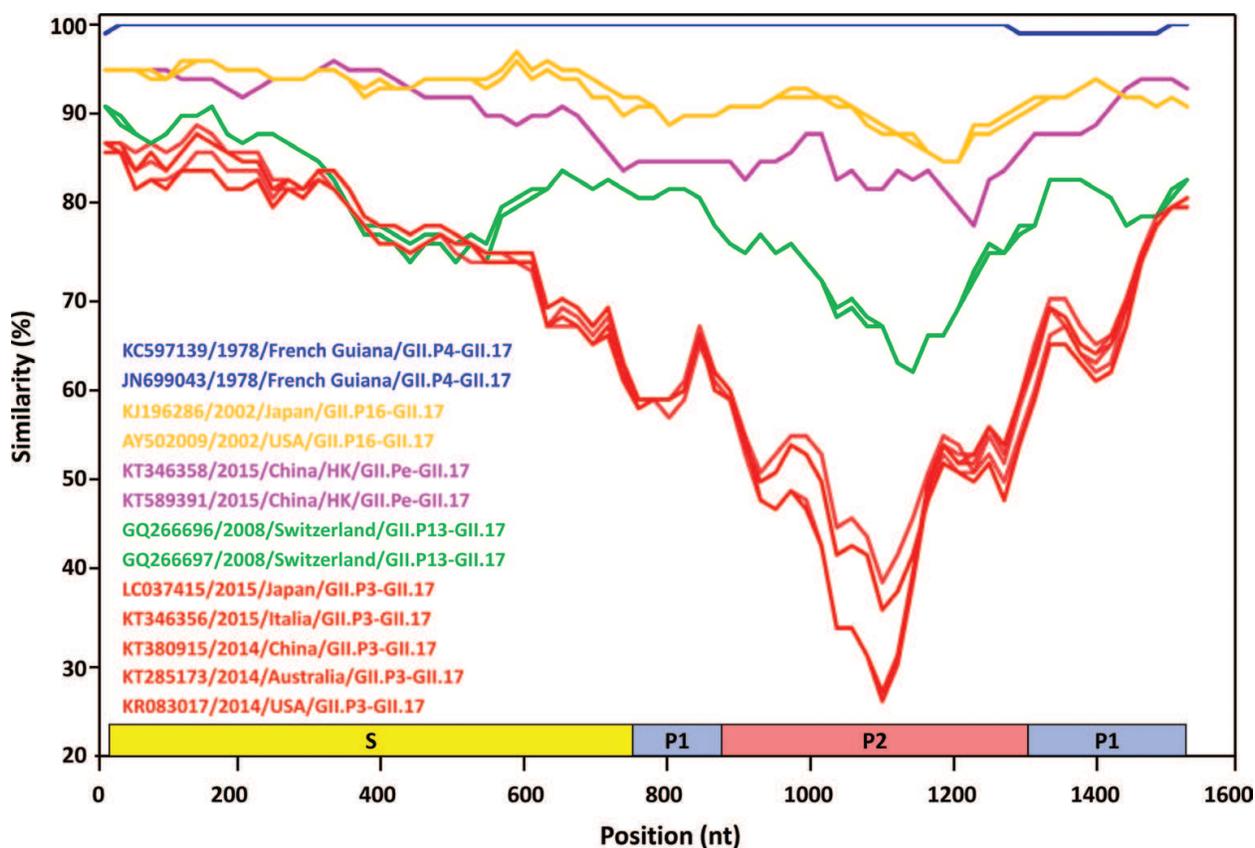
Given that only a small portion of the ORF2 sequences (181 nt) was included in the phylogenetic tree (**Figure 2A**) and that the bootstrap value of the GII.P3-GII.17 cluster was less than 80%, two more phylogenetic trees were constructed using the full-length sequences of ORF1 and ORF2 ($n = 54$) in order to further verify the recombination in sequences from the GII.P3-GII.17 cluster. On the ORF1 tree (**Figure 2B**), all the sequences were grouped into GII.3 genotype clusters with a high bootstrap value (99%). On the ORF2 tree (**Figure 2C**), sequences were grouped into GII.17 genotype clusters with a high bootstrap value (99%). These results demonstrated that the new epidemic strains were in fact GII.P3-GII.17 intergenetic recombinants.

Simplot analysis was also performed to confirm the recombination events. The candidate recombinant sequences shared high similarities with GII.3 in ORF1 and with GII.17 in ORF2, respectively. In addition, recombination breakpoints were identified within the overlapping regions of ORF1 and ORF2.

3.4. Sequence variations of the GII.17 P domain

To explore the evolution of NoV GII.17 capsid genes and the importance of intergenic recombination, five representatives of the P domain sequences (GII.P3, GII.P13, GII.P16, GII.Pe and GII.P4) from distinct clusters discovered between 1978 and 2015 were subjected to analysis. Sequence similarity among all complete genomic sequences was analyzed using Simplot. GII.P16-GII.17 and GII.Pe-GII.17 variants shared high similarities with the GII.17 variant that emerged the earliest (GII.P4-GII.17 in 1978) at the identification of 92.73 and 90.45%, respectively (**Figure 3B**). However, the variant GII.P13-GII.17 that emerged after revealed a decline in similarity with GII.P4-GII.17 (82.96% of identity). The recently emerged GII.P3-GII.17 variant showed the largest divergence (76.12% of identity) from the GII.P4-GII.17 variant (**Figure 3B**). Interestingly, a typical “V” shape was observed in our Simplot analysis (**Figure 3A**), suggesting that the P2 subdomain is the most hypervariable region in the GII.17 capsid and continues to evolve (**Figure 3A**).

A structure-based sequence alignment of the P domain (amino acid) was performed to examine any potential differences among these five GII.17 variants. Overall, P1 subdomain sequences were relatively conserved (**Figure 4**); however, significant sequence differences were observed in the outer loop regions in the P2 subdomain that includes the B-loop (aa 291–299), P-loop (aa



(a)

| | KC597139 | KJ196286 | KT589391 | GQ266696 | LC037415 |
|----------|----------|----------|----------|----------|----------|
| KC597139 | 100.00 | 92.73 | 90.45 | 82.96 | 76.12 |
| KJ196286 | 92.73 | 100.00 | 88.72 | 83.77 | 75.81 |
| KT589391 | 90.45 | 88.72 | 100.00 | 82.15 | 76.24 |
| GQ266696 | 82.96 | 83.77 | 82.15 | 100.00 | 77.18 |
| LC037415 | 76.12 | 75.81 | 76.24 | 77.18 | 100.00 |

(b)

Figure 3. Sequence similarity of P domain (nucleotides). (A) Simplot analysis: The Simplot method was performed to show the similarity between complete NoV GII.17 VP1 nucleotide sequences. Thirteen representative sequences were picked from five recombinant clusters and highlighted in different colors. The horizontal axis represents the nucleotide positions of VP1, and the vertical axis represents the sequence similarity compared to the GII.P4-GII.17 (KC597139) strain. (B) Similarity analysis: The pair-wise identity of five representative GII.17 sequences from each recombinant cluster, GII.P4-GII.17 (KC597139), GII.P16-GII.17 (KJ96286), GII.Pe-GII.17 (KT589391), GII.P13-GII.17 (GQ266696) and GII.P3-GII.17 (LC037415), was analyzed.

339–352), A-loop (aa 375–383) and T-loop (aa 395–400) (**Figure 4**). In addition, slight changes were also found in the S-loop (aa 438–445) located in the P1 subdomain and in the U-loop (aa 408–417) located within the junctional region of the P1 and P2 subdomains (**Figure 4**).

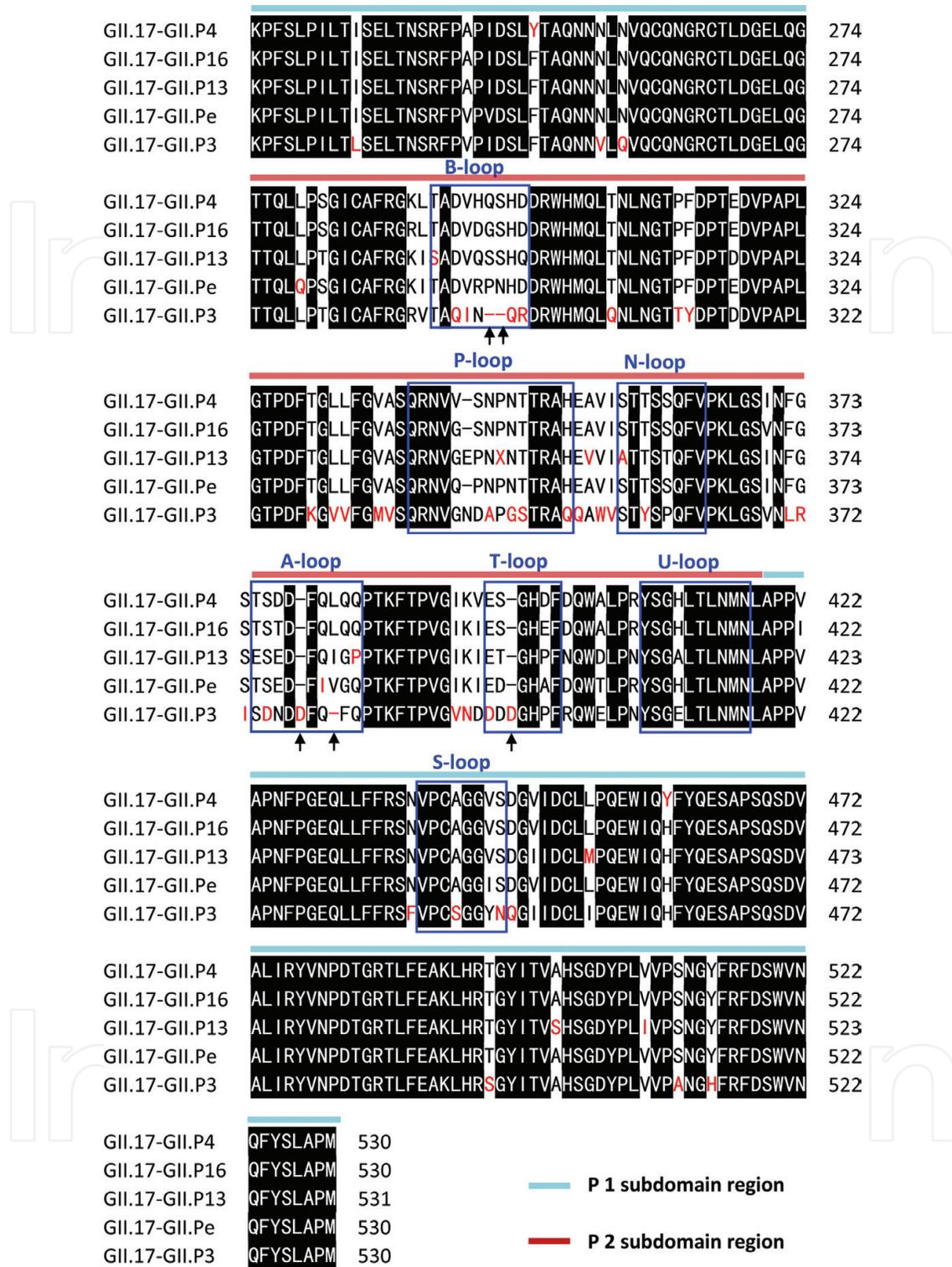


Figure 4. Sequence alignment of P domain (amino acids). The structure-based sequence alignment of the P domains of GII.P4-GII.17 (AGI17592), GII.P16-GII.17 (AII73747), GII.Pe-GII.17 (ALD83748), GII.P13-GII.17 (ACT68315) and GII.P3-GII.17 (BAR42289) from five recombinant clusters was performed. Regions spanning the P1 and P2 subdomains are indicated by stripes. Identical residues are highlighted with black background, while distinct residues are shown in red characters. The conventional GII genogroup HBGA surface binding loops (A-, B-, P-, S-, T-, U- and N-) are indicated by blue frames.

Interestingly, various mutations were also identified in the P2 subdomain, especially in the loops on the outer surface of the protein. Comparing to the other four earlier variants, three deletions (aa 349, 350 and 381) and two insertions (aa 379 and 397) were observed in the recently emerged GII.P3-GII.17 variant (**Figure 4**). These mutations, which are located in the B-loop (aa 349 and 350), A-loop (aa 379 and 381) and T-loop (aa 397), might have altered the binding capacity of NoV to host human histo-blood group antigen (HBGA) (**Figure 4**). These results also suggested that the GII.17 P domain might have evolved from a non-prevalent strain into an epidemic GII.17 variant strain (GII.P3-GII.17).

The position of amino acids in VP1 corresponds to AGI17592. Arrows indicate deletions (aa 349, 350 and 381) and insertions (aa 379 and 397) in the recently emerged GII.P3-GII.17 variant, as compared to the other GII.17 variants.

3.5. Structure shift of the GII.17 P domain

To better understand how this novel NoV GII.17 spread so rapidly and widely, the P domain structures of all five GII.17 variants were constructed based on homology modeling. Comparison of the overall structures of the five GII.17 variants revealed that most amino acid substitutions occurred in the P2 subdomain, while the P1 subdomain remained relatively conserved.

Figure 5 shows the surface loops that formed the conventional GII NoV binding interface (B-, T-, N-, P-, U-, S- and A-loops). The front-side view indicates that the outer structures of the N- and U-loops of all five GII.17 variants remained relatively unchanged (**Figure 5**). However, compared to the other four GII.17 variants, distinct structural changes were observed in the A-, B- and T-loops of the GII.P3-GII.17 variant (**Figure 5**). The top-side view revealed that the P-loops of the GII.P13-GII.17 and the GII.P3-GII.17 variants are protruding from the surface of the proteins (**Figure 5**), a feature clearly different from those observed in the other three GII.17 variants.

To clarify the specific changes observed in the loops, superpositions of the P domain structures of the five representative GII.17 variants were constructed (**Figure 6**). The most prominent difference was observed in the P-loop (**Figure 6**). Three distinct structures of the P-loop were observed among the five GII.17 variants, with the GII.P4-GII.17, GII.P16-GII.17 and GII.Pe-GII.17 variants sharing the same structure. In contrast, the GII.P13-GII.17 and GII.P3-GII.17 variants each presented a unique structure (**Figure 6**). In addition, the T-, B-, U- and A-loops also exhibited considerable differences. Notably, the GII.P4-GII.17, GII.P16-GII.17, GII.Pe-GII.17 and GII.P13-GII.17 variants shared similar T-, B-, U- and A-loops; while the GII.P3-GII.17 variant displayed a unique shift in structure in these four loops (**Figure 6**).

3.6. Global distribution of NoV GII.17 variants

Sequences of the five distinct variants of GII.17 (n = 646) were obtained from at least 24 countries and regions from five continents: Asia (China, Hong Kong, Taiwan, Japan, Bangladesh, Singapore, South Korea, Pakistan and Thailand), Europe (Italy, France and Russia), America (Brazil, Mexico, Nicaragua, Uruguay, French Guiana and the USA), Africa (Cameroon,

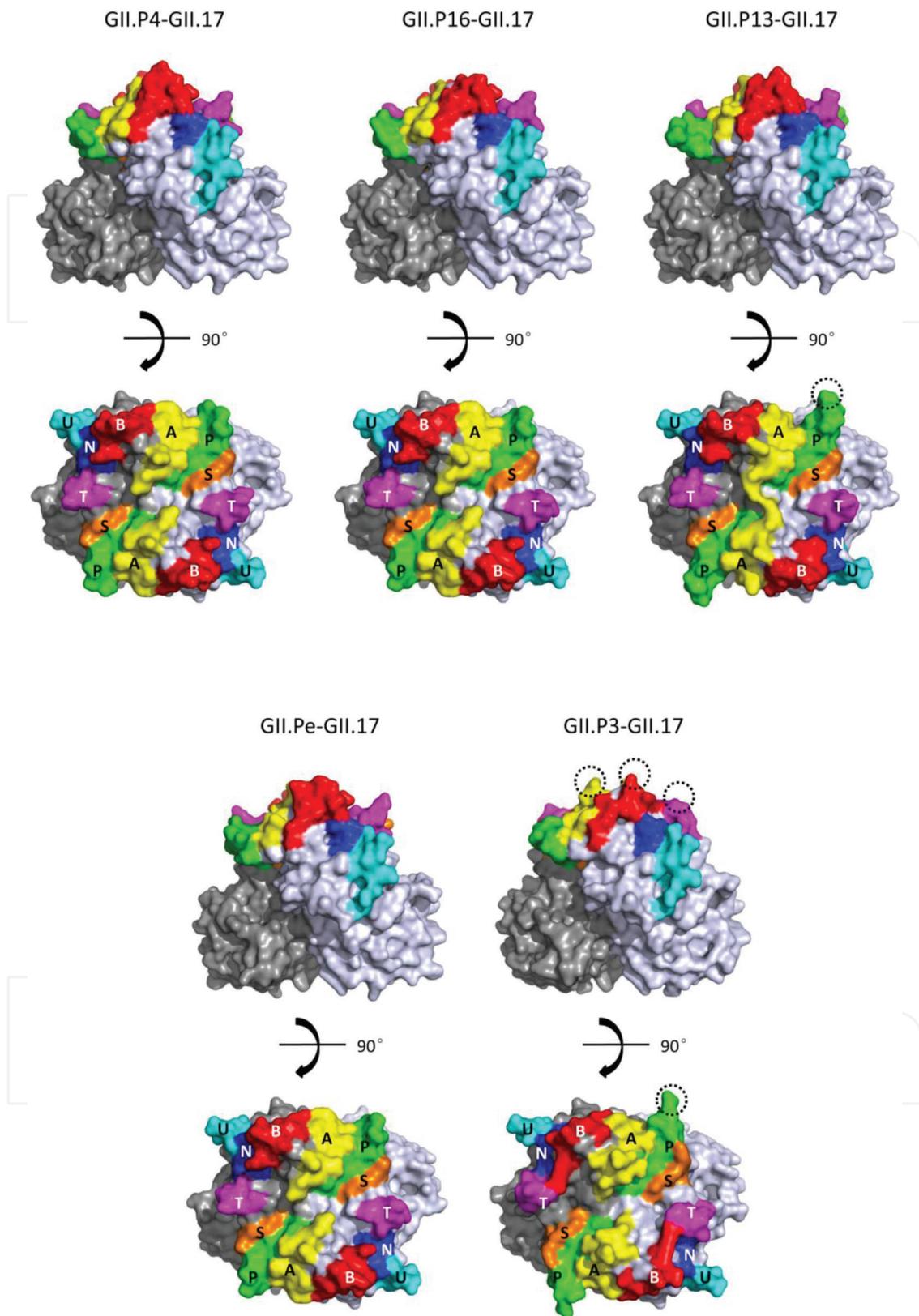


Figure 5. Three-dimensional structure of P domain. Three-dimensional P domain structures of the five GII.17 variants were predicted by homologous modeling. The seven surface-binding loops, that is, A-, B-, P-, S-, T-, U- and N-loops, are indicated in yellow, red, green, orange, pink, sky blue and blue, respectively. The black circles highlight the changes present in the binding loops among different GII.17 variants.

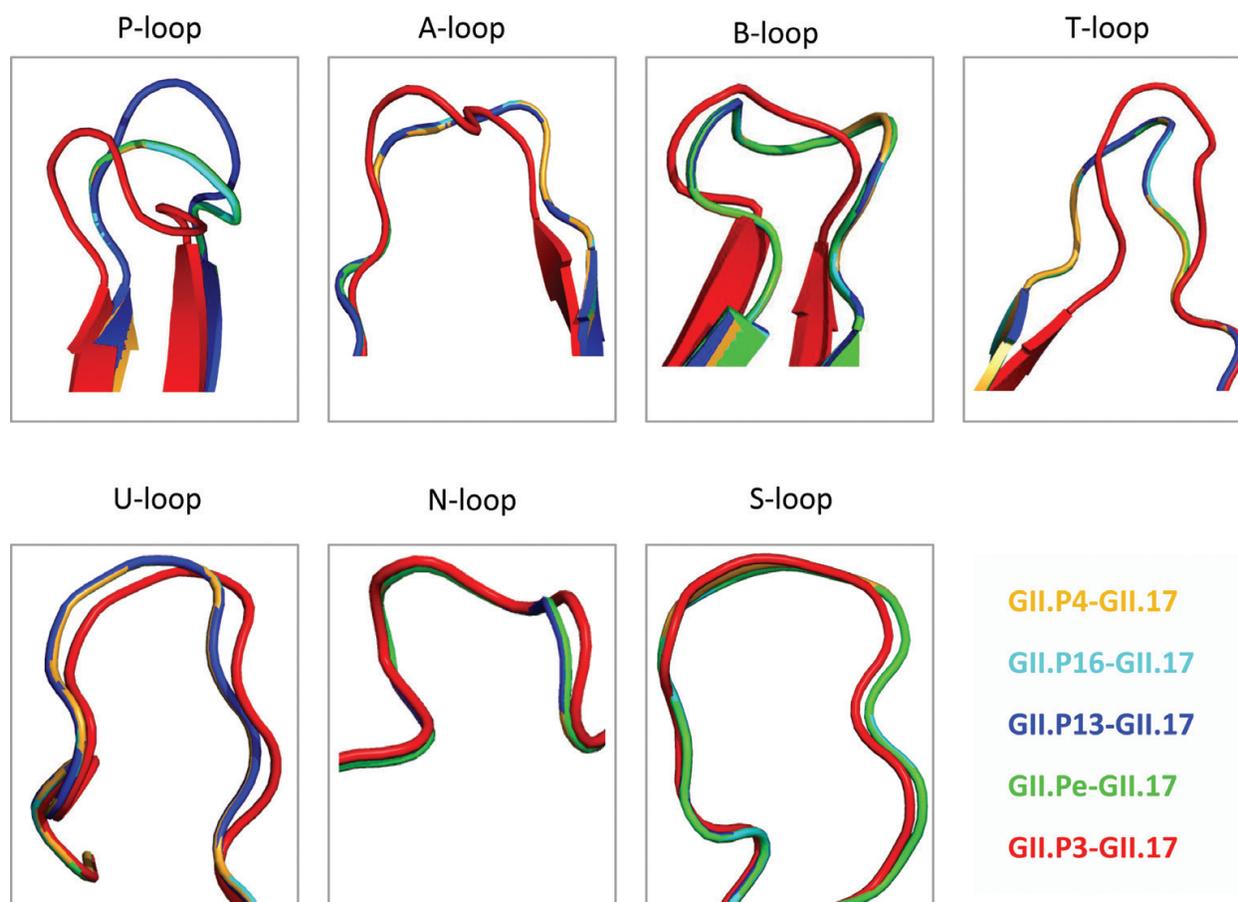


Figure 6. Comparison of the seven major surface loops of the GII.17 variants. Structures of the seven surface-binding loops located on the P domain were compared by superposition. Color schemes: wheat, GII.P4-GII.17 (AGI17592); sky blue, GII.P16-GII.17 (AII73747); green, GII.Pe-GII.17 (ALD83748); blue, GII.P13-GII.17 (ACT68315) and red, GII.P3-GII.17 (BAR42289).

Ethiopia, Kenya, Morocco and South Africa) and Oceania (Australia). The number of sequences obtained in these regions was unevenly distributed, with certain countries being over-represented compared to others. Most of the sequences were obtained in Asia (89.30%) and Europe (7.31%). The number of sequences collected in China was the highest (71.12%), followed by South Korea (9.80%), Kenya (4.99%) and Japan (4.81%).

Since most of these NoV GII.17 sequences were categorized into five recombination variant clusters; it is important to understand how the different GII.17 variants are distributed globally. Only three sequences belonging to the GII.P4-GII.17 variant were collected from French Guiana and Pakistan (**Figure 7A**). Most of the sequences of the GII.P16-GII.17 variant were isolated in Japan (n = 17) (**Figure 7A**). In addition, sequences of the GII.Pe-GII.17 variant were well dispersed among different countries including Cameroon (n = 1), Ethiopia (n = 6), China (n = 7), South Korea (n = 1), Brazil (n = 2) and the USA (n = 1) (**Figure 7A**). Sequences from the GII.P13-GII.17 variant were even more widely distributed: China (n = 3), Japan (n = 1), Bangladesh (n = 10), Singapore (n = 3), South Korea (n = 3), Thailand (n = 2), France (n = 2), Russia (n = 2), Brazil (n = 2), Mexico (n = 1), Nicaragua (n = 1), Uruguay (n = 1), Cameroon (n = 3) and South Africa (n = 1) (**Figure 7A**).

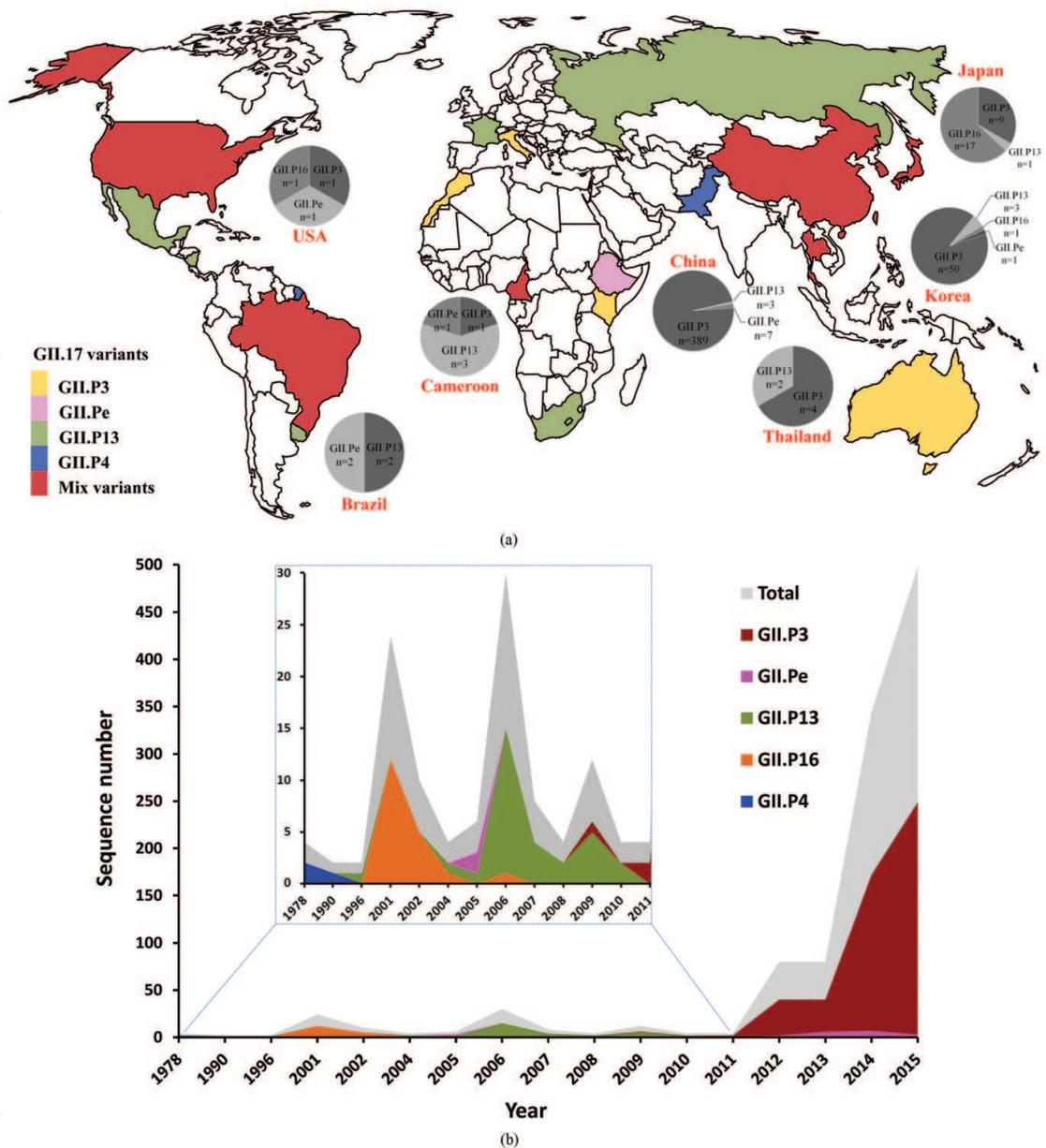


Figure 7. Global distribution and prevalence of GII.17 variants. (A) Geographical distribution of the GII.17 variants from 1978 to 2015: Different GII.17 variants are shown in different colors. The distributions of multiple variants in the same regions are shown in pie-charts (map template from digital vector maps) and (B) Yearly prevalence of GII.17 variants during 1978–2015: Different GII.17 variants are shown in different colors. Part of the graph representing data from the period between 1978 and 2011 was partially enlarged to allow the examination of detailed results.

As for the GII.P3-GII.17 variant, its sequence distribution was also quite scattered, with most of the sequences found in China ($n = 389$), South Korea ($n = 28$), Japan ($n = 9$) and Thailand ($n = 4$) (**Figure 7A**).

Interestingly, multiple GII.17 variants were often observed in one country or region (**Figure 7A**). For example, over 70% of the GII.17 sequences (399 in 561) were detected in China, including sequences from the GII.P3-GII.17, GII.P13-GII.17 and GII.Pe-GII.17 variants (**Figure 7A**). Except for GII.P4-GII.17, the other four GII.17 variants were found in South Korea, with GII.P3-GII.17 being the most dominant. Different GII.17 variants were also observed in Japan

(GII.P3-GII.17, GII.P13-GII.17 and GII.P16-GII.17), Cameroon (GII.P3-GII.17, GII.P13-GII.17 and GII.Pe-GII.17), the USA (GII.P3-GII.17, GII.P16-GII.17 and GII.Pe-GII.17), Thailand (GII.P3-GII.17 and GII.P13-GII.17) and Brazil (GII.P13-GII.17 and GII.Pe-GII.17) (**Figure 7A**).

3.7. Yearly prevalence of GII.17 variants

Figure 7B showed the yearly distribution of the GII.17 variants detected worldwide. The GII.17 sequences isolated from 1978 to 2015 were unevenly distributed. The number of isolated sequences peaked in 2001, 2006 and 2009, then continued to increase significantly from 2011 to 2015 and reached the highest peak in 2015 (**Figure 7B**). This result suggested that circulation of GII.17 might change every 3–5 years due to replacement by a newly emerged variant. The GII.P4-GII.17 variant, which emerged the earliest, only appeared in 1978 and 1990 (**Figure 7B**). About a dozen years later, the GII.P16-GII.17 variant was discovered in 2002 and persisted till 2004, when the GII.P13-GII.17 variant emerged and replaced it. A few of the GII.P3-GII.17 variants emerged between 2009 and 2013, and the number of sequences that belonged to this variant sharply increased during 2014 and 2015. Over 70% of the GII.17 sequences belonged to this GII.P3-GII.17 variant and were isolated during this period (**Figure 7B**). It is worth noting that the GII.Pe-GII.17 variant was also observed between 2012 and 2015, although the number of sequences identified during this period was far less than that of the GII.P3-GII.17 variant (**Figure 7B**).

4. Discussion

NoV, one of the leading causes of human acute gastroenteritis, is widely distributed around the world. In the past decades, NoV GII.4 was identified as the most predominant genotype involved in numerous epidemic outbreaks, for example, in 2002, 2004, 2006, 2009 and 2012 [37]. Recently, a novel NoV GII.17 variant has emerged and is responsible for multiple disease outbreaks mainly in China and Japan [9, 11–22]. GII.17 appeared as a dominant strain replacing the GII.4 strain in these regions during 2014–2015 [23].

NoV GII.17 strains have been circulating among various human populations for over 37 years, with previously emerged GII.17 sequences being sporadically detected in multiple regions in various continents including most of Africa, Asia, Europe, North America and South America (**Figure 7A**). However, it is still unclear as to why this recently emerged GII.17 variants spread so rapidly and widely within such a short time period.

4.1. Genetic diversity and recombination

Previous studies have suggested that GII.17 sequences could be classified into distinct clusters based on the divergence of VP1. However, only a few GII.17 sequences were analyzed in these studies [14, 15, 18, 20, 24, 26]. In order to investigate the diversity of NoV GII.17 in detail, large sequence dataset was constructed and subjected to meta-analysis and genotyping in this study (**Figure 1**). Surprisingly, at least five major recombinant GII.17 clusters, including GII.P4-GII.17, GII.P16-GII.17, GII.P13-GII.17, GII.Pe-GII.17 and GII.P3-GII.17, were identified. In addition, each recombinant group appeared to have emerged following a par-

ticular time order. For example, the earliest group of GII.P4-GII.17 was isolated in 1978, followed by GII.P16-GII.17 (2001–2004), GII.P13-GII.17 (2004–2010), GII.Pe-GII.17 (2012–2015) and GII.P3-GII.17 (2011–2015). This indicates that the evolution of GII.17 is sequentially promoted by inter-genotype genetic recombination, which contributes to the exchange between non-GII.17 and GII.17 RdRp genes and promotes the evolution of GII.17 capsid genes. Consequently, these genetic recombinations could have potentially affected the antigenic properties of NoVs and accelerated the emergence of novel epidemic variants or strains [17, 23], for example, GII.P3-GII.17. The highly diverse genome of the rare genotype of GII.17 demonstrated in this study is far beyond what has been reported previously.

Moreover, based on the phylogenetic trees of RdRp and VP1 (**Figure 2B** and **C**), all GII.P3-GII.17 sequences were further subdivided into cluster I and II. Cluster I contains the variants identified from 2013 to 2014, while the variants isolated from 2014 to 2015 comprise Cluster II. Our results suggest that after RdRp recombination with the pandemic genotype of GII.P3, the subsequently emerged GII.17 strains underwent various modifications at the capsid [23, 35], which, consequently, promoted the emergence of new variants (**Figure 3**).

Notably, the co-circulation of the GII.P3-GII.17 variant with the non-epidemic strains of GII.Pe-GII.17 indicates that the evolution of GII.17 strains was driven by multiple mechanisms in different directions within the human population in recent years (**Figure 2A**).

4.2. Shift and variation in capsid structure

Chan et al. [17] and Lu et al. [38] have previously reported that the GII.17 VP1 evolved faster than GII.4 VP1. The rapid evolution of this GII.P3-GII.17 VP1 at a rate of 5.68×10^{-3} nucleotide substitutions per site per year, which is comparable to that of GII.4 VP1 ($5.3\text{--}6.3 \times 10^{-3}$ substitutions per site per year), is faster than that in other NoV genotypes such as the epidemic genotypes GII.3 and GII.7 (1.961×10^{-3} and 2.36×10^{-3} nucleotide substitutions per site per year, respectively) [38]. In this study, compared with previous circulating GII.17 variants, the recently emerged GII.17 variant was found to exhibit significant sequence and structure variations.

First of all, the new GII.P3-GII.17 variant showed the most noticeable sequence divergence (76.12% of nucleotide identity) compared to the other GII.17 variants, for example, GII.P16-GII.17, GII.P13-GII.17 and GII.Pe-GII.17 (**Figure 3**), which could be a driving force of antigenic drift of the new GII.17 variant.

Secondly, at the amino acid level, the P domain of the GII.P3-GII.17 variant appeared to have evolved more rapidly than the other variants (**Figure 4**). Specifically, three deletions and two insertions were observed in the new GII.P3-GII.17 variant (**Figure 4**). The homology structure of the P domain also showed that many mutations were located in the human histo-blood group antigen (HBGA) binding loops of the newly emerged GII.P3-GII.17 variant (**Figure 5**). Notably, most amino acid substitutions were found in the P2 subdomain. These results suggest that these mutations most likely affect the HBGA binding property of the new GII.17 variant, which in turn, might expand its host range and prevalence [17, 21, 35]. It is important

to note that changes at the antigenic epitopes of the viral capsid might also lead to adaptive advantages that contribute to the rapid spread of the virus.

The observed structural variations led us to hypothesize that, in the beginning of GII.17 evolution, the binding interface of these viruses remained relatively stable from 1978 to 2004, as reflected in the variants GII.P4-GII.17 (1978–1990) and GII.P16-GII.17 (2001–2004). Soon after that, a significant change occurred at the surface loops of the capsid (GII.P13-GII.17, 2004–2010), followed by more dramatic changes in 2011–2015 (GII.P3-GII.17, 2011–2015) (**Figure 2A**). Interestingly, the loop structures of the GII.Pe-GII.17 variant remained relatively constant, even though it was in circulation during the 2012–2015 period, indicating that this variant retained the conventional GII.17 HBGA binding interfaces [35]. Our results also confirmed that the GII.P3-GII.17 variant might have evolved as a unique lineage separated from the other GII.17 variants.

Thirdly, it is also possible that variations in the rate of evolution of the capsid were promoted by recombination, since this recombination resulted in the same capsid lineage that was associated with different RdRp [39]. Recombination breakpoints between RdRp and capsid genes might have allowed GII.17 to acquire RdRp with lower fidelity and/or increased replication efficiency, resulting in a higher rate of evolution for the associated capsid gene [17]. Moreover, within a short time period, the human host might have yet to adapt to the new GII.17 capsid, which might also support the rapid spread of the new GII.17 variant.

4.3. Waterborne GII.17 strains

NoV particles are usually detected in environmental water or in effluents from the wastewater treatment plant throughout the year [40–42]. In fact, the attack rate of waterborne NoV (over 11%) was found to be significantly higher than that associated person to person or environmental transmission [5].

Interestingly, the dataset in this study revealed that over 20% (129 in 646) of GII.17 sequences were collected from water samples or water-related outbreaks. Such waterborne GII.17 outbreaks have been reported in many countries, for example, the USA, in 2005 [43]; South Korea, in 2008–2012 [44]; Guatemala, in 2009 [45]; and China, in 2014–2015 [16, 46]. Moreover, many GII.17 sequences were also detected in environmental water samples from around the world, such as Shandong (China) [47], Singapore [48, 49], Japan [50, 51], South Africa [52, 53], Kenya [54] and New Orleans (USA) [40].

Since the capsid protects genomic RNA from the environment, it is assumed that capsid degradation could lead to viral inactivation, probably followed by the degradation of the unprotected viral RNA [55]. Our data suggest that the capsid of GII.17 virus is more stable and therefore, can persist longer in water samples. Interestingly, Arthur et al. demonstrated that Tulane virus (a novel human NoV surrogate) also remained stable in surface water (<1 log₁₀ reduction) after 28 days; however, viral load reduced from ≥ 3.5 to 4 log₁₀ in ground-water by day 21 [56]. Consequently, improving surveillance and systematic monitoring of environmental water samples could provide valuable information on viral circulation and enable further assessment of the emergence of novel NoVs at an earlier stage.

4.4. Nomenclature of the new GII.17 variant

To coordinate the assignment of new genotypes and variants, a dual-typing system based on the complete capsid (VP1) and the partial polymerase (1300 nt) was proposed [3]. In 2014, the RdRp genotype of Kawasaki 323 strain emerged, but it was not assigned to any of the known genotypes in the database. Instead, it was assigned as the GII.P17 RdRp genotype, and the variants of NoVs were named Hu/GII/JP/2014/GII.P17-GII.17 [21]. Meanwhile, the Hu/GII.17/Gaithersburg/2014/US strain was categorized into the same group as the GII.3 strains based on its RdRp sequence. However, based on the bootstrap values (<70%) and genetic distances (<0.143), insufficient confidence was encountered when an attempt was made to classify these variants with any of the known RdRp genotypes [26]. In addition, Fu and coauthors reported that most NoV GII.17 strains detected in the 2000s were affiliated with a GII.3-like RdRp genotype and that the new GII.17 variant might be a recombinant strain that expresses a GII.3-like RdRp gene and a GII.17 capsid gene [14].

Viruses that exhibit a GII.17 VP1 genotype have previously been reported to harbor a GII.P13 ORF1 genotype, although recombinants expressing an ORF1 GII.P16, GII.P3 and GII.P4 genotype have also been identified. Sequence comparison showed that the ORF1 region of the novel GII.17 viruses, which are clustered between the GII.P3 and GII.P13 viruses, has never been detected before. Since this is the first orphan ORF1 sequence associated with GII.17, we decided to designate GII.P17 according to the criteria listed in the proposal for a unified NoV nomenclature and genotyping. The novel GII.17 virus was named Kawasaki 2014 after the first near-complete genome sequence (AB983218) was submitted to GenBank [21]. This typing tool was updated to ensure correct classification of both ORF1 and ORF2 sequences of the new GII.P17-GII.17 viruses [23].

According to the phylogenetic tree of ORF1 (**Figure 2B**), all the recently emerged GII.17 sequences are closely related to GII.P3, with a bootstrap value above 95% (=99%) (**Figure 2B**).

Moreover, Simplot analysis also supports that the new GII.17 sequences had undergone recombination between a GII.P3 RdRp and a GII.17 capsid gene, and the recombination breakpoint is located within the ORF1/ORF2 overlapping region. Taken together, our results led us to propose that the new GII.17 sequences should be named as a recombinant genotype, GII.P3-GII.17, or at least be classified as a new GII.17 variant.

5. Conclusion

In conclusion, the genetic diversity and global prevalence of NoV GII.17 from 1978 to 2015 were analyzed in this study. A highly diverse genome was discovered within this rare genotype of GII.17. For example, at least five major recombinant GII.17 clusters were found to have emerged following a particular time order. The circulation of GII.17 changes every 3–5 years due to replacement by a newly emerged variant, and the evolution of GII.17 is sequentially promoted by inter-genotype genetic recombination. Most of the GII.17 sequences were detected in Asian countries including China, South Korea and Japan, and multiple GII.17 variants were found in one country or region. Moreover, the recently emerged GII.P3-GII.17 variant exhibited significant sequence and structure variations and had evolved as a unique lineage. The results presented in this study contribute to the understanding of the evolution and persistence of NoV GII.17 in human population through analyzing recombination in the

viral genome. Our findings also provide important insights into the future monitoring of the global circulation of this novel GII.17 variant.

Acknowledgements

We thank the anonymous contributors for collecting and sharing the NoV sequences. This work was supported partially by the National Key Research and Development Program of China (SQ2017YFC160114), the National Natural Science Foundation of China (41376135 and 31570112), Doctoral Fund of Ministry of Education of China (20133104110006) and Innovation Program of Shanghai Municipal Education Commission (14ZZ144), China.

Conflict of interest

The authors have declared that no competing interests exist.

Author details

Yongxin Yu^{1,2}, Yingjie Pan^{1,2}, Shuling Yan^{1,3} and Yongjie Wang^{1,2*}

*Address all correspondence to: yjwang@shou.edu.cn

1 College of Food Science and Technology, Shanghai Ocean University, Shanghai, China

2 Laboratory of Quality and Safety Risk Assessment for Aquatic Products on Storage and Preservation (Shanghai), Ministry of Agriculture, China

3 Institute of Biochemistry and Molecular Cell Biology, University of Göttingen, Göttingen, Germany

References

- [1] Glass RI, Parashar UD, Estes MK. Norovirus gastroenteritis. *New England Journal of Medicine*. 2009;**361**(18):1776-1785. DOI: 10.1056/NEJMra0804575
- [2] Zheng D-P, Ando T, Fankhauser RL, Beard RS, Glass RI, Monroe SS. Norovirus classification and proposed strain nomenclature. *Virology*. 2006;**346**(2):312-323. DOI: 10.1016/j.virol.2005.11.015
- [3] Kroneman A, Vega E, Vennema H, Vinjé J, White PA, Hansman G, Green K, Martella V, Katayama K, Koopmans M. Proposal for a unified norovirus nomenclature and genotyping. *Archives of Virology*. 2013;**158**(10):2059-2068. DOI: 10.1007/s00705-013-1708-5
- [4] Hall AJ, Vinjé J, Lopman B, Park GW, Yen C, Gregoricus N, Parashar U. Updated Norovirus Outbreak Management and Disease Prevention Guidelines. US Department of Health and Human Services, Centers for Disease Control and Prevention, U.S.

Government Printing Office (GPO), Washington, DC, 20402-9371; 2011. Available from: <https://www.cdc.gov/mmwr/preview/mmwrhtml/rr6003a1.htm>

- [5] Matthews J, Dickey B, Miller R, Felzer J, Dawson B, Lee A, Rocks J, Kiel J, Montes J, Moe C. The epidemiology of published norovirus outbreaks: A review of risk factors associated with attack rate and genogroup. *Epidemiology and Infection*. 2012;**140**(07):1161-1172. DOI: 10.1017/S0950268812000234
- [6] Siebenga JJ, Vennema H, Zheng D-P, Vinjé J, Lee BE, Pang X-L, Ho EC, Lim W, Choudekar A, Broor S. Norovirus illness is a global problem: Emergence and spread of norovirus GII. 4 variants, 2001-2007. *Journal of Infectious Diseases*. 2009;**200**(5):802-812. DOI: 10.1086/605127
- [7] Van Beek J, Ambert-Balay K, Botteldoorn N, Eden J, Fonager J, Hewitt J, Iritani N, Kroneman A, Vennema H, Vinje J. Indications for worldwide increased norovirus activity associated with emergence of a new variant of genotype II 4, late 2012. *Euro Surveillance*. 2013;**18**(1):8-9. DOI: 10.2807/ese.18.01.20345-en
- [8] Bull RA, White PA. Mechanisms of GII. 4 norovirus evolution. *Trends in Microbiology*. 2011;**19**(5):233-240. DOI: 10.1016/j.tim.2011.01.002
- [9] Lu J, Sun L, Fang L, Yang F, Mo Y, Lao J, Zheng H, Tan X, Lin H, Rutherford S. Gastroenteritis outbreaks caused by norovirus GII. 17, Guangdong Province, China, 2014-2015. *Emerging Infectious Diseases*. 2015;**21**(7):1240. DOI: 10.3201/eid2107.150226
- [10] Zhang X-F, Huang Q, Long Y, Jiang X, Zhang T, Tan M, Zhang Q-L, Huang Z-Y, Li Y-H, Ding Y-Q. An outbreak caused by GII. 17 norovirus with a wide spectrum of HBGA-associated susceptibility. *Scientific Reports*. 2015;**5**:17687. DOI: 10.1038/srep17687
- [11] Wang H-B, Wang Q, Zhao J-H, Tu C-N, Mo Q-H, Lin J-C, Yang Z. Complete nucleotide sequence analysis of the norovirus GII. 17: A newly emerging and dominant variant in China, 2015. *Infection, Genetics and Evolution*. 2016;**38**:47-53. DOI: 10.1016/j.meegid.2015.12.007, 10.1111/jam.13052
- [12] Xue L, Cai W, Wu Q, Zhang J, Guo W. Direct sequencing and analysis of the genomes of newly emerging GII. 17 norovirus strains in South China. *Journal of Applied Microbiology*. 2016;**120**(4):1130-1135. DOI: 10.1111/jam.13052
- [13] Xue L, Wu Q, Cai W, Zhang J, Guo W. Molecular characterization of new emerging GII. 17 norovirus strains from South China. *Infection, Genetics and Evolution*. 2016;**40**:1-7. DOI: 10.1016/j.meegid.2016.02.026
- [14] Fu J, Ai J, Jin M, Jiang C, Zhang J, Shi C, Lin Q, Yuan Z, Qi X, Bao C. Emergence of a new GII. 17 norovirus variant in patients with acute gastroenteritis in Jiangsu, China, September 2014 to March 2015. *Euro Surveillance*. 2015;**20**(24):21157. DOI: 10.2807/1560-7917.ES2015.20.24.21157
- [15] Han J, Ji L, Shen Y, Wu X, Xu D, Chen L. Emergence and predominance of norovirus GII. 17 in Huzhou, China, 2014-2015. *Virology Journal*. 2015;**12**(1):1. DOI: 10.1186/s12985-015-0370-9

- [16] Qin M, Dong X-G, Jing Y-Y, Wei X-X, Wang Z-E, Feng H-R, Yu H, Li J-S, Li J. A water-borne gastroenteritis outbreak caused by Norovirus GII. 17 in a hotel, Hebei, China, December 2014. *Food and Environmental Virology*. 2016;**8**(3):180-186. DOI: 10.1007/s12560-016-9237-5
- [17] Chan MC, Lee N, Hung T-N, Kwok K, Cheung K, Tin EK, Lai RW, Nelson EAS, Leung TF, Chan PK. Rapid emergence and predominance of a broadly recognizing and fast-evolving norovirus GII. 17 variant in late 2014. *Nature Communications*. 2015;**6**:10061. DOI: 10.1038/ncomms10061
- [18] Lee C-C, Feng Y, Chen S-Y, Tsai C-N, Lai M-W, Chiu C-H. Emerging norovirus GII. 17 in Taiwan. *Clinical Infectious Diseases*. 2015;**61**(11):1762-1764. DOI: 10.1093/cid/civ647
- [19] Gao Z, Liu B, Huo D, Yan H, Jia L, Du Y, Qian H, Yang Y, Wang X, Li J. Increased norovirus activity was associated with a novel norovirus GII. 17 variant in Beijing, China during winter 2014-2015. *BMC Infectious Diseases*. 2015;**15**(1):1. DOI: 10.1186/s12879-015-1315-z
- [20] Chen H, Qian F, Xu J, Chan M, Shen Z, Zai S, Shan M, Cai J, Zhang W, He J. A novel norovirus GII. 17 lineage contributed to adult gastroenteritis in Shanghai, China, during the winter of 2014-2015. *Emerging Microbes and Infection*. 2015;**4**(11):e67. DOI: 10.1038/emi.2015.67
- [21] Matsushima Y, Ishikawa M, Shimizu T, Komane A, Kasuo S, Shinohara M, Nagasawa K, Kimura H, Ryo A, Okabe N. Genetic analyses of GII. 17 norovirus strains in diarrheal disease outbreaks from December 2014 to March 2015 in Japan reveal a novel polymerase sequence and amino acid substitutions in the capsid region. *Euro Surveillance*. 2015;**20**(26):21173. DOI: 10.2807/1560-7917.ES2015.20.26.21173
- [22] Thanh HD, Nguyen TH, Lim I, Kim W. Emergence of norovirus GII. 17 variants among children with acute gastroenteritis in South Korea. *PLoS One*. 2016;**11**(5):e0154284. DOI: 10.1371/journal.pone.0154284. eCollection 2016
- [23] De Graaf M, Van Beek J, Vennema H, Podkolzin A, Hewitt J, Bucardo F, Templeton K, Mans J, Nordgren J, Reuter G. Emergence of a novel GII. 17 norovirus-end of the GII. 4 era. *Euro Surveillance*. 2015;**20**(26):21178. DOI: 10.2807/1560-7917.ES2015.20.26.21178
- [24] Medici M, Tummolo F, Calderaro A, Chironna M, Giammanco G, De Grazia S, Arcangeletti M, De Conto F, Chezzi C, Martella V. Identification of the novel Kawasaki 2014 GII. 17 human norovirus strain in Italy, 2015. *Euro Surveillance*. 2015;**20**(35):30010. DOI: 10.2807/1560-7917.ES.2015.20.35.30010
- [25] Dinu S, Nagy M, Negru D, Popovici E, Zota L, Opreșan G. Molecular identification of emergent GII. P17-GII. 17 norovirus genotype, Romania, 2015. *Euro Surveillance*. 2016;**21**(7):30141. DOI: 10.2807/1560-7917.ES.2016.21.7.30141
- [26] Parra GI, Green KY. Genome of emerging norovirus GII. 17, United States, 2014. *Emerging Infectious Diseases*. 2015;**21**(8):1477. DOI: 10.3201/eid2108.150652
- [27] Bull RA, Tanaka MM, White PA. Norovirus recombination. *Journal of General Virology*. 2007;**88**(12):3347-3359. DOI: 10.1099/vir.0.83321-0

- [28] Yu Y, Cai H, Hu L, Lei R, Pan Y, Yan S, Wang Y. Molecular epidemiology of oyster-related human noroviruses and their global genetic diversity and temporal-geographical distribution from 1983 to 2014. *Applied and Environmental Microbiology*. 2015;**81**(21):7615-7624. DOI: 10.1128/AEM.01729-15
- [29] Kearse M, Moir R, Wilson A, Stoneshavas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C. Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;**28**(12):1647. DOI: 10.1093/bioinformatics/bts199
- [30] Kroneman A, Vennema H, Deforche K, Avoort H, Penaranda S, Oberste M, Vinjé J, Koopmans M. An automated genotyping tool for enteroviruses and noroviruses. *Journal of Clinical Virology*. 2011;**51**(2):121-125. DOI: 10.1016/j.jcv.2011.03.006
- [31] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*. 2011;**28**(10):2731. DOI: 10.1093/molbev/msr121
- [32] Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research*. 2007;**33**(Database Issue):501-504. DOI: 10.1093/nar/gki025
- [33] Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *Journal of Virology*. 1999;**73**(1):152-160. Available from: <http://jvi.asm.org/content/73/1/152.long>
- [34] Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L. Swiss-model: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*. 2014;**42**(Web Server Issue):W252. DOI: 10.1093/nar/gku340
- [35] Singh BK, Koromyslova A, Hefele L, Gürth C, Hansman GS. Structural evolution of the emerging 2014/15 GII. 17 noroviruses. *Journal of Virology*. 2016;**40**:1-7. DOI: 10.1016/j.meegid.2016.02.026
- [36] DeLano WL. PyMOL. San Carlos: DeLano Scientific; 2002. p. 700. Available from: <https://pymol.org/2/>
- [37] Eden J-S, Tanaka MM, Boni MF, Rawlinson WD, White PA. Recombination within the pandemic norovirus GII. 4 lineage. *Journal of Virology*. 2013;**87**(11):6270-6282. DOI: 10.1128/JVI.03464-12
- [38] Lu J, Fang L, Zheng H, Lao J, Yang F, Sun L, Xiao J, Lin J, Song T, Ni T. The evolution and transmission of epidemic GII. 17 noroviruses. *Journal of Infectious Diseases*. 2016;**214**(4):556-564. DOI: 10.1093/infdis/jiw208
- [39] de Graaf M, van Beek J, Koopmans MP. Human norovirus transmission and evolution in a changing world. *Nature Reviews Microbiology*. 2016;**14**(7):421-433. DOI: 10.1038/nrmicro.2016.48

- [40] Montazeri N, Goettert D, Achberger EC, Johnson CN, Prinyawiwatkul W, Janes ME. Pathogenic enteric viruses and microbial indicators during secondary treatment of municipal wastewater. *Applied and Environmental Microbiology*. 2015;**81**(18):6436-6445. DOI: 10.1128/AEM.01218-15
- [41] Myrmel M, Lange H, Rimstad E. A 1-year quantitative survey of noro-, adeno-, human boca-, and hepatitis E viruses in raw and secondarily treated sewage from two plants in Norway. *Food and Environmental Virology*. 2015;**7**(3):213-223. DOI: 10.1007/s12560-015-9200-x
- [42] Prevost B, Lucas FS, Ambert-Balay K, Pothier P, Moulin L, Wurtzer S. Deciphering the diversities of astroviruses and noroviruses in wastewater treatment plant effluents by a high-throughput sequencing method. *Applied and Environmental Microbiology*. 2015;**81**(20):7215-7222. DOI: 10.1128/AEM.02076-15
- [43] Yee EL, Palacio H, Atmar RL, Shah U, Kilborn C, Faul M, Gavagan TE, Feigin RD, Versalovic J, Neill FH. Widespread outbreak of norovirus gastroenteritis among evacuees of Hurricane Katrina residing in a large “megashelter” in Houston, Texas: Lessons learned for prevention. *Clinical Infectious Diseases*. 2007;**44**(8):1032-1039. DOI: 10.1086/512195
- [44] Cho H, Lee S, Kim W, Lee J, Park P, Cheon D, Jheong W, Jho E, Lee J, Paik S. Acute gastroenteritis outbreaks associated with ground-waterborne norovirus in South Korea during 2008-2012. *Epidemiology and Infection*. 2014;**142**(12):2604-2609. DOI: 10.1017/S0950268814000247
- [45] Arvelo W, Sosa SM, Juliao P, López MR, Estevéz A, López B, Morales-Betoulle ME, González M, Gregoricus NA, Hall AJ. Norovirus outbreak of probable waterborne transmission with high attack rate in a Guatemalan resort. *Journal of Clinical Virology*. 2012;**55**(1):8-11. DOI: 10.1016/j.jcv.2012.02.018
- [46] Wang X, Yong W, Shi L, Qiao M, He M, Zhang H, Guo B, Xie G, Zhang M, Jin M. An outbreak of multiple norovirus strains on a cruise ship in China, 2014. *Journal of Applied Microbiology*. 2016;**120**(1):226-233. DOI: 10.1111/jam.12978
- [47] Tao Z, Xu M, Lin X, Wang H, Song L, Wang S, Zhou N, Zhang D, Xu A. Environmental surveillance of genogroup I and II noroviruses in Shandong Province, China in 2013. *Scientific Reports*. 2015;**5**:17444. DOI: 10.1038/srep17444
- [48] Aw T, Gin KH. Environmental surveillance and molecular characterization of human enteric viruses in tropical urban wastewaters. *Journal of Applied Microbiology*. 2010;**109**(2):716-730. DOI: 10.1111/j.1365-2672.2010.04701.x
- [49] Aw T, Gin KH. Prevalence and genetic diversity of waterborne pathogenic viruses in surface waters of tropical urban catchments. *Journal of Applied Microbiology*. 2011;**110**(4):903-914. DOI: 10.1111/j.1365-2672.2011.04947.x
- [50] Kitajima M, Oka T, Haramoto E, Takeda N, Katayama K, Katayama H. Seasonal distribution and genetic diversity of genogroups I, II, and IV noroviruses in the Tamagawa River, Japan. *Environmental Science and Technology*. 2010;**44**(18):7116-7122. DOI: 10.1021/es100346a

- [51] Kitajima M, Haramoto E, Phanuwat C, Katayama H, Furumai H. Molecular detection and genotyping of human noroviruses in influent and effluent water at a wastewater treatment plant in Japan. *Journal of Applied Microbiology*. 2012;**112**(3):605-613. DOI: 10.1111/j.1365-2672.2012.05231.x
- [52] Mans J, Netshikweta R, Magwalivha M, Van Zyl WB, Taylor MB. Diverse norovirus genotypes identified in sewage-polluted river water in South Africa. *Epidemiology and Infection*. 2013;**141**(02):303-313. DOI: 10.1017/S0950268812000490
- [53] Murray TY, Mans J, Taylor MB. Human calicivirus diversity in wastewater in South Africa. *Journal of Applied Microbiology*. 2013;**114**(6):1843-1853. DOI: 10.1111/jam.12167
- [54] Kiulia N, Mans J, Mwenda J, Taylor M. Norovirus GII. 17 predominates in selected surface water sources in Kenya. *Food and Environmental Virology*. 2014;**6**(4):221-231. DOI: 10.1007/s12560-014-9160-6
- [55] de la Noue AC, Estienne M, Aho S, Perrier-Cornet J-M, de Rougemont A, Pothier P, Gervais P, Belliot G. Absolute humidity influences the seasonal persistence and infectivity of human norovirus. *Applied and Environmental Microbiology*. 2014;**80**(23):7196-7205. DOI: 10.1128/AEM.01871-14
- [56] Arthur SE, Gibson KE. Environmental persistence of Tulane virus—A surrogate for human norovirus. *Canadian Journal of Microbiology*. 2016;**62**(5):449-454. DOI: 10.1139/cjm-2015-0756