Infection, Genetics and Evolution 70 (2019) 15–23

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Genome characterization and temporal evolution analysis of a non-epidemic norovirus variant GII.8

Liang Xue[a], Weicheng Cai[a], Junshan Gao[a], Le Zhang[a], Haoming Wu[a], Rui Pang[a], Ying Li[a], Juan Wang[b], Jumei Zhang[a], Shi Wu[a], Qinghua Ye[a], Tao Lei[a], Qingping Wu[a,*]

[a] Guangdong Institute of Microbiology, State Key Laboratory of Applied Microbiology Southern China, Guangdong Provincial Key Laboratory of Microbial Culture Collection and Application, Guangdong Open Laboratory of Applied Microbiology, Guangzhou 510070, PR China
[b] College of Food Science, South China Agricultural University, Guangzhou 510642, PR China

ARTICLE INFO

Keywords:
Norovirus
Genome
GII.8
Genetic diversity
Phylogenetic analysis
Homology modeling

ABSTRACT

Noroviruses are the primary cause of non-bacterial acute gastroenteritis worldwide, and GII.8 belongs to a non-epidemic genotype with a limited understanding currently. In this study, we assembled the first GII.8 norovirus genome from China and clarified the temporal evolutionary process of this non-epidemic variant. Using the "4 + 1 + 1" application strategy with newly designed primer sets, the genome of one GII.8 strain GZ2017-L601 from China was firstly sequenced that comprised 7476 nucleotides. The homology of the new genome and the previous only GII.8 genome reached 93.8% identity at the nucleotide level, but only 10, 6, 7 amino acid mutations occurred in three ORFs. When compared the new strain with other GII reference strains, p22 and P2 were calculated as the variable encoding regions, and NTPase, VPg, 3CL, RdRp and S were shown as the conserved ones. We then reconstructed the evolutionary process of the GII.8 genotype using other available sequences in GenBank. Based on the partial N/C region, all GII.8 strains could be subdivided chronologically into four clusters, which spans 1967–1994, 1997–2005, 2003–2009, and 2007–2017, respectively. Moreover, differences of capsid P proteins between GII.8 strains and the epidemic GII.4 strain VA387 were also compared. There existed 147/310 distinct amino acid sites in the alignment, including two insertion and three deletion mutations. Distribution of antigen epitopes of two GII.8 variants was comparable, but the numbers of antigenic sites of GII.8 strains were less than that of VA387. In summary, the first GII.8 genome from China was assembled and extensively characterized, and a time-order evolutionary process of this genotype was identified with a static pattern of antigenic variations.

1. Introduction

Norovirus (NoV) is an important pathogen of acute gastroenteritis in the world. It almost caused 18% (95% CI: 17%–20%) diarrheal disease and 212,000 deaths annually (Lopman et al., 2016). It has also been estimated that NoV infections resulted in huge medical economic ($4.2 billion, 95% UI: $3.2–5.7 billion) and social losses ($60.3 billion, 95% UI: $44.4–83.4 billion) globally per year (Bartsch et al., 2016). However, there are still no licensed virus vaccines or effective medicine treatments against NoV infections as of this writing (Bernstein et al., 2015). NoVs could be classified into seven genogroups, and GI, GII, and GIV could infection humans. Each genogroup could be further divided into different genotypes. > 30 human NoV genotypes have been reported, but only GII.4 has been the predominant genotype around the world since the mid-1990s. It has been reported that over 80% of sporadic infections and 60% of outbreaks were caused by this genotype (Siebenga et al., 2009).

Non-GII.4 epidemic strains have caused widespread concerns in recent years. Especially in the winter of 2014/2015, a novel GII.17 NoV variant (Kawasaki 2014) emerged and replaced the previous dominant GII.4 Sydney 2012 variant in Asian countries (de Graaf et al., 2015; Xue et al., 2016c). Other genotypes, including GII.2, GII.3, GII.6, have also been reported as the local epidemic genotypes exceeding GII.4 during gastroenteritis investigations (Ao et al., 2017; Iritani et al., 2012; Lu et al., 2015). Potential pandemics of the non-epidemic NoV variants based on their continuous evolution has become an interesting issue (Xue et al., 2017; Xue et al., 2016c). However, due to their low detection rates, genomic information of non-epidemic strains is usually limited. For example, there were not any GII.17 genome sequences before the emergence of GII.17 Kawasaki 2014. The lack of genetic

* Corresponding author at: No. 100, Xianlie Zhong Road, Guangzhou 510070, PR China.
E-mail address: wuqp203@163.com (Q. Wu).

https://doi.org/10.1016/j.meegid.2019.02.010
Received 29 October 2018; Received in revised form 25 January 2019; Accepted 14 February 2019
Available online 15 February 2019
1567-1348/

Table 1
Primers designed for amplifying and sequencing GII·P8/GII.8 NoV genomes in this study.

Primer	Sequence 5′→3′	Position[a]	Tm (°C)	GC %	Reference
II-1F	GTGAATGAAGATGGCGTCTAAC	1–22	51.8	45.5	This study
II.8-1R	CTGTAGGCATCCCAGTGATC	1629–1610	51.7	55.0	This study
II.8-2F	TCATCTTCGCCTGACATCGT	1373–1392	53.4	50.0	This study
II.8-2R	TCCTCTTCACAGAAGTTCTC	2841–2822	48.3	45.0	This study
II.8-3F	CACACTGCATTCTCCAGCAA	2669–2688	53.0	50.0	This study
II.8-3R	CAATCCTCCAAATGCCCGA	4177–4159	52.0	53.0	This study
II.8-4F	GAGGAAGCCAAGAAGACAGT	3821–3840	51.3	50.0	This study
II-4R	CCRCCNGCATRHCCRTTRTACAT	5389–5367	50.0–64.3	39.1–65.2	(Kojima et al., 2002)
II-5F	GGAGGGCGATCGCAATC	5050-5066	52.4	64.7	(Tu et al., 2008)
II-5R	CCRGCRAAGAAAGCTCCAGCCAT	6720-6698	57.8–61.6	52.2–60.9	(Tu et al., 2008)
II.8-6F	TCYAACAGTGGGGACCACC	6573–6591	52.3–55.5	558.0–63.0	This study
II.8-6R	TCACTAAGCCCGTGACTCC	7474–7456	53.0	58.0	This study
II.8-seq1R	AACTCACCATCCCACATCTC	429–410	51.5	50.0	This study
II.8-seq6F	GACAAGGAGATGTTGAATGC	6842–6861	49.0	45.0	This study

[a] Primer mapping position in the norovirus GII.8 genome (GenBank Accession No. AB039780).

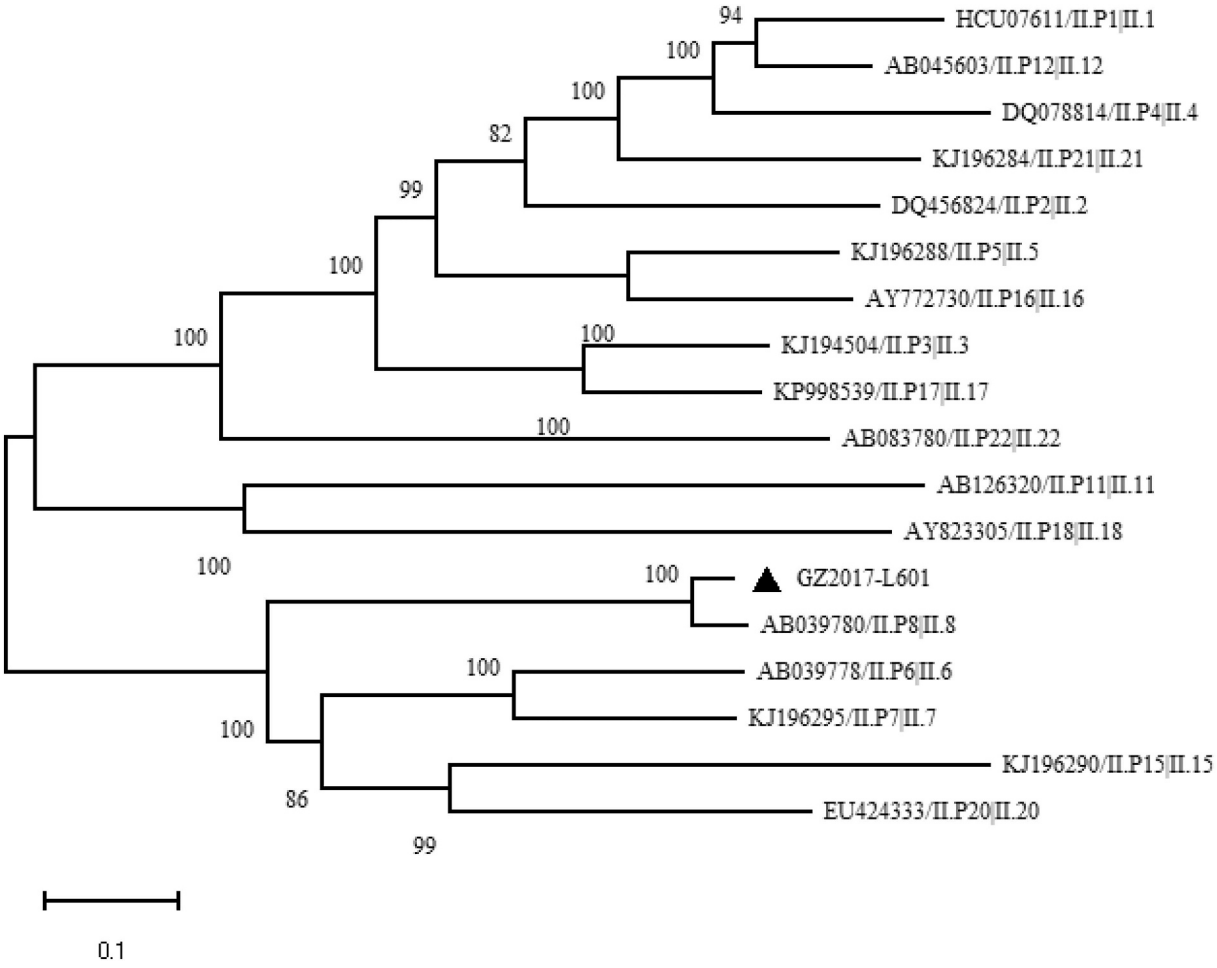


Fig. 1. Phylogenetic analyses based on full genome nucleotide sequences of representative strains for each human NoV genotype. Maximum likelihood trees were constructed using MEGA v7.0 with the best model GTR + I + G. Bootstrap values higher than 70% are shown in the corresponding branches. The scale bar represents the unit for the expected number of substitutions per site. GZ2017-L601 is highlighted with a black triangle.

information in non-epidemic strains will have a serious impact on predicting their prevalence threats and analyzing their evolutionary processes (van Beek et al., 2018).

GII.8 belongs to a non-epidemic genotype which was often detected with low positive rates. It has been reported that GII.8 caused NoV outbreaks in restaurants and hotels (Lysen et al., 2009). Besides, GII.8 strains were also detected in the monitoring on water and shellfish (Nishida et al., 2003). However, there is little information about the GII.8 gene sequence accumulated so far, including only one genome and five VP1 sequences (Katayama et al., 2002). During our surveillance of NoV-associated acute gastroenteritis recently, the GII.8 strain was detected from a 29-year-old female patient in South China. For further understanding the genetic characteristics of the genotype, we assembled the first Chinese GII.8 NoV genome by a newly established method and clarified the evolutionary process for this genotype.

Table 2
Genomic structure of GZ2017-L601 and its similarity with the reference sequence AB039780

GZ2017-L601	Position	Length (nt/aa)	Pairwise Identity (nt) %	Mutation nt	Pairwise Identity (aa) %	Mutation aa[a]
Genome	–	7476/–	93.80	–	–	–
ORF1	6–5105	5100/1700	93.60	327	99.40	*n* = 10, A60K, T72A, I86V, S242P, I527V, D623E, K753R, I1080V, N1347D, L1382 V
ORF2	5086–6699	1614/538	94.4	91	98.9	*n* = 6, E193D, I337V, N374H, Q394E, Q491R, I536V
ORF3	6699–7472	774/258	94.1	46	97.3	*n* = 7, L6F, T121I, K133R, R155K, V209I, I249M, K251R

[a] Location of amino acid mutations in the amino residues of three ORFs in the norovirus GII.8 genome (GenBank Accession No. AB039780).

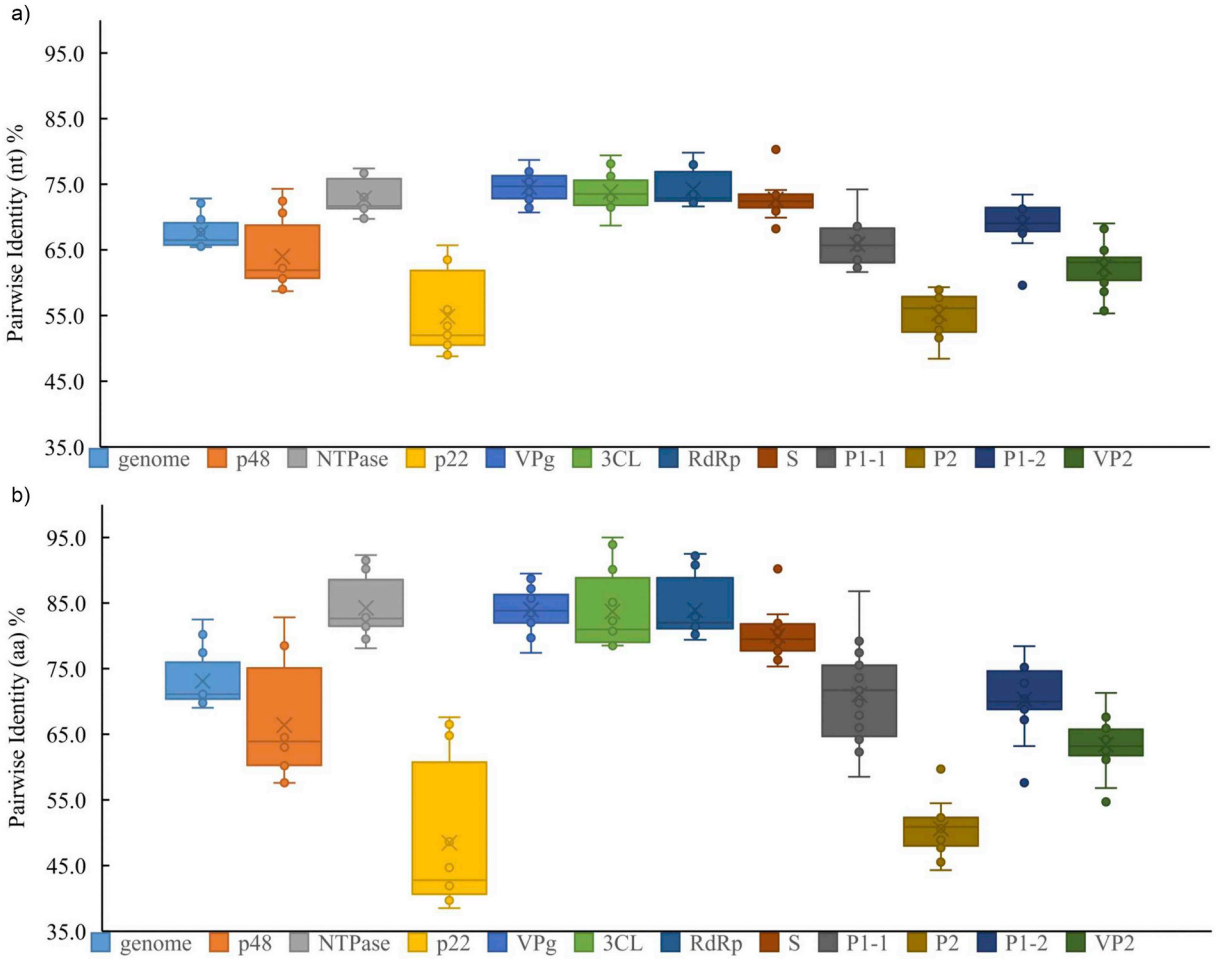


Fig. 2. Box plots of pairwise identities of protein encoding regions of the GII.8 strain GZ2017-L601 and representative strains of different GII NoVs at (A) nucleotide and (B) amino acid levels. Viral genome and each protein are represented by the colors of the legend. The genome at the amino acid level is composed of the amino acid sequences of each protein.

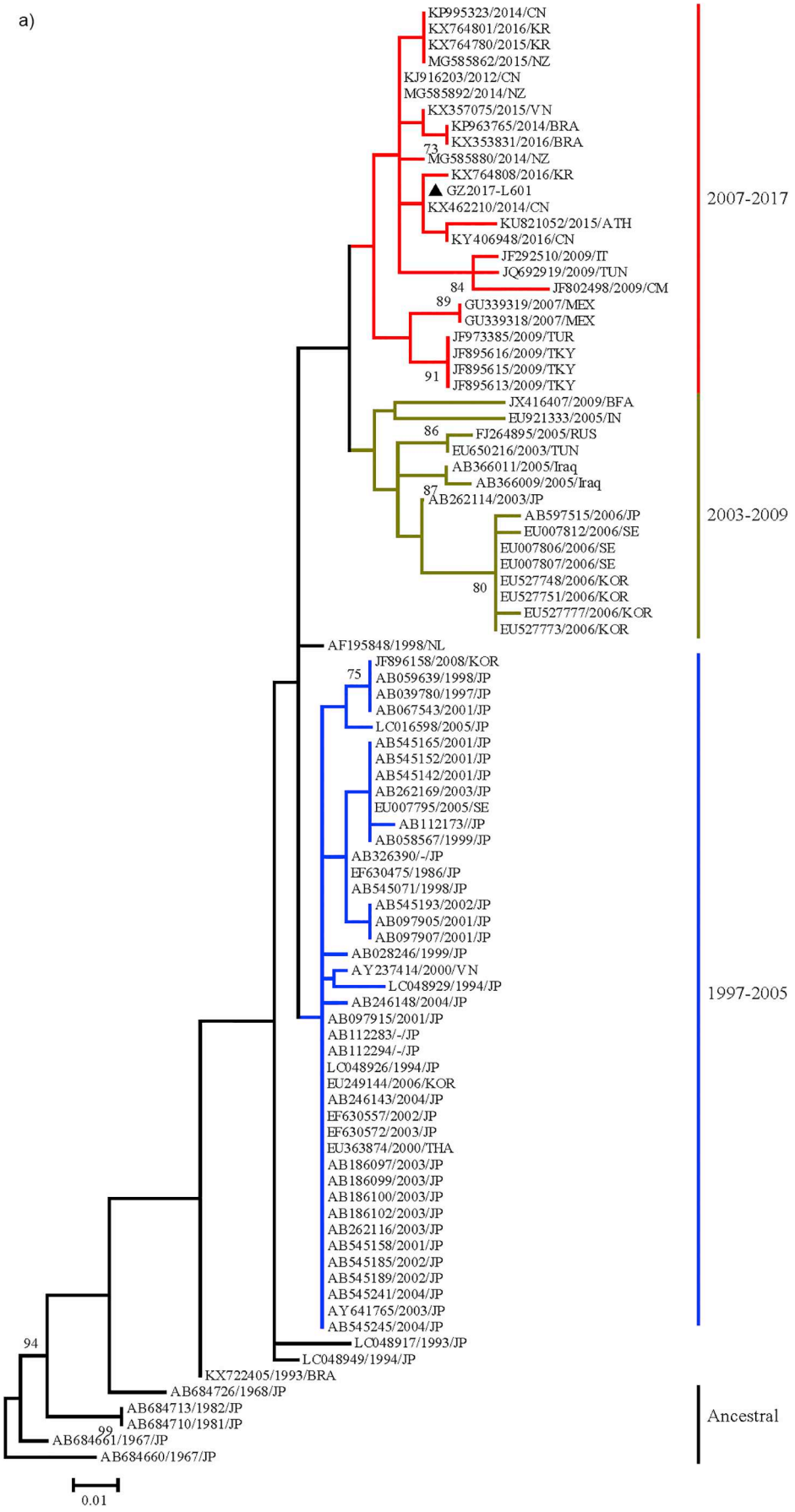
2. Materials and methods

2.1. Primer designing

GII·P8 or GII.8 NoV sequences with length > 500 bp in GenBank were chosen as reference sequences. The previous "4 + 1 + 1" strategy was employed for primer designing (Xue et al., 2016a,b). Six amplifying primer pairs and two sequencing primers for GII·P8/GII.8 NoV genomes were designed by PRIMER PREMIER v5.0 (Premier Biosoft International, Palo Alto, CA, USA) according to the following criteria: primer length 18–22 bp, PCR product size 1400–1600 bp, DNA melting temperature 50.0–65.0 °C, GC content 40–60% (Table 1).

2.2. NoV samples, genome amplifying and sequencing

The GII.8 NoV sample GZ2017-L601 was detected during the previous monitoring of NoV-associated acute gastroenteritis, which was stored at − 80 °C. The NoV genome was amplified and sequenced according to the previous protocol. In brief, viral RNA was extracted from the 10% (w/v) fecal supernatant using the Viral RNA Mini Kit (QIAgen, Hilden, Germany). Six overlapping fragments encompassing the whole genome were generated with the newly designed primer sets via one-step reverse transcription PCR (TAKARA, Dalian, China). Amplicons were sequenced directly using an automated sequencer (ABI 3730XL, Applied Biosystems, Foster City, CA, USA), and viral genome was then constructed by connecting overlapped sections as a single sequence.



(caption on next page)

Fig. 3. Phylogenetic analyses of (A) GII.8 NoV strains based on parial capsid N/C region and (B) GII·P8 NoV strains based on partial polymerase region. Maximum likelihood trees were constructed using MEGA v7.0 with the best models (N/C, K2 + G; polymerase, K2). Bootstrap values higher than 70% are shown in the corresponding branches. The scale bar represents the unit for the expected number of substitutions per site. GZ2017-L601 is highlighted with a black triangle.

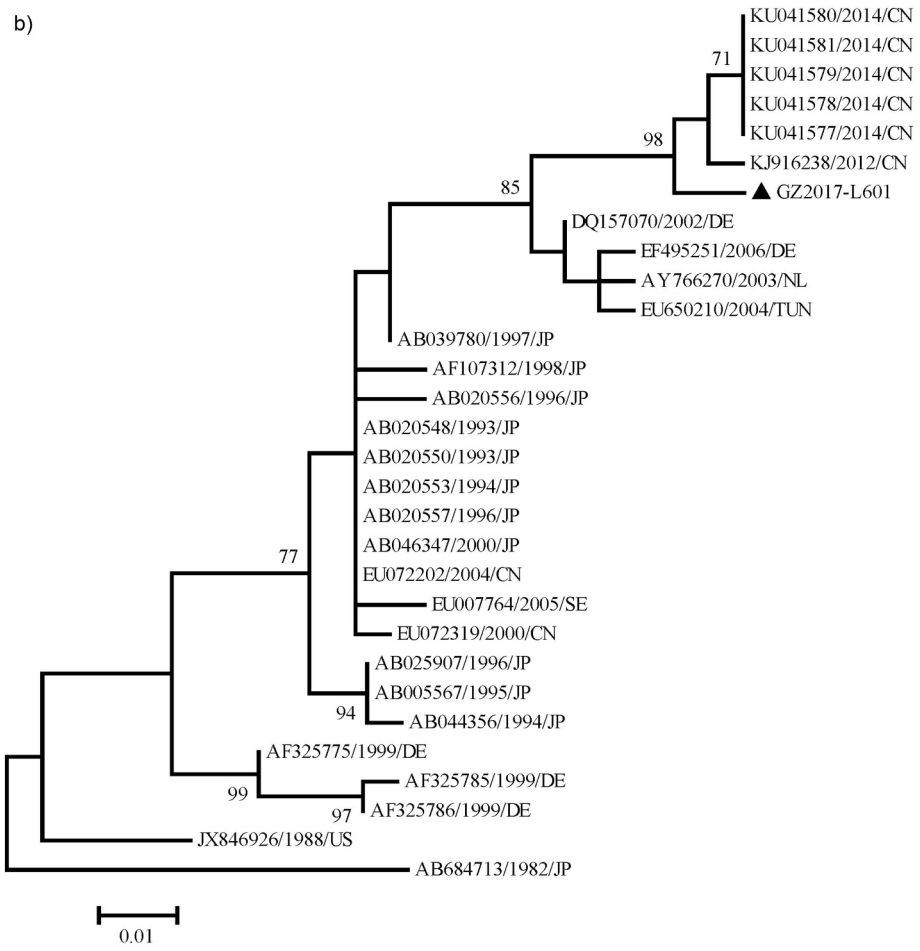


Fig. 3. (*continued*)

2.3. Sequence analyses

All nucleotide sequences were first edited with the BioEdit® Sequence Alignment Editor software (v.7.0.1). By compared with the representative sequence AB039778, positions of ORFs and protein-encoding regions in the new genome were verified. Multiple alignments of different protein sequences were performed using ClustalX v1.83 with the default parameters. Pairwise similarities of the new genome and reference sequences were calculated using DNASTAR Lasergene MegAlign program (DNASTAR Inc., WI, USA). Phylogenetic analysis of aligned sequences was performed in Molecular Evolutionary Genetics Analysis (MEGA) v7.0 (Kumar et al., 2016). Best substitution models for the dataset were chosen based with the lowest Bayesian Information Criterion scores. The reliability of the phylogenetic tree was assessed by bootstrap sampling of 1000 replicates. Nucleotide sequences were also submitted to the online NoV Genotyping Tool v2.0 (http://www.rivm.nl/mpf/norovirus/typingtool) to verify phylogenetic results (Kroneman et al., 2011).

2.4. Homology modeling and structure-based epitope prediction

Homology modeling of the viral capsid was performed by generating a Protein Data Bank (PDB) file from the amino acid sequence by using the Swiss Model server (Arnold et al., 2006). Modeling was carried out using the published crystal structure of GII.4 NoV strain VA387 (PDB: 2OBT) as a template. Protein structures were visualized and manipulated using PyMOL v1.4.1 (DeLano Scientific, LLC, San Francisco, CA). By submitting modeled 3D structures of capsid P proteins, potential B cell epitopes were predicted using the DiscoTope v2.0 (http://tools.iedb.org/discotope/) with the default threshold −3.7 (Kringelum et al., 2012).

3. Results

3.1. The genome structure of the GII.8 NoV strain from China

The entire genome nucleotide sequence of the GII.8 NoV strain GZ2017-L601 was 7476 bp in size (GenBank Accession Number JX989075). Phylogenetic analysis with NoV representative genome sequences showed that GZ2017-L601 belongs to GII.8 genotype (Fig. 1). By compared with the reference GII.8 genome (AB039780), three ORFs were identified in the new genome, spanning 6–5105, 5086–6699, and 6699–7472 (Table 2). There is only one GII.8 genome sequence reported as of this writing (AB039780), and the new genome shared 93.80% nucleotide similarity with it. In detail, a total of 464 different nucleotide sites were verified between above two GII.8 genomes (327 in ORF1, 91 in ORF2, 46 in ORF3), but only 23 different amino acid sites were generated (10 in ORF1, 6 in ORF2, 7 in ORF3) (Table 2).

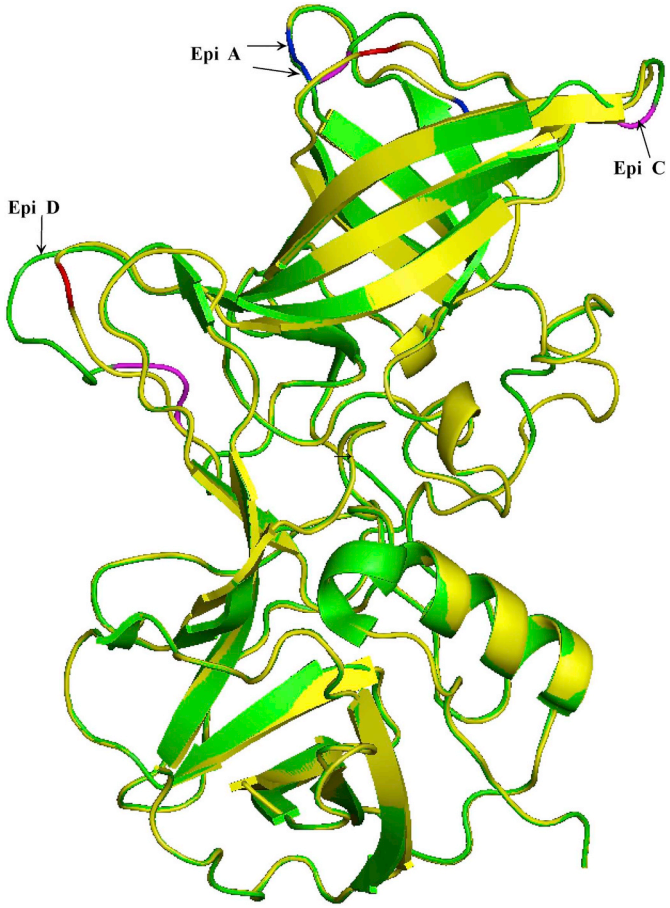


Fig. 4. Homology models of the capsid P protein of the GII.8 strain GZ2017-L601. The structural superposition of the predicted structure for GZ2017-L601 P domain (yellow) and the crystal structure of GII.4 NoV strain VA387 (green) was performed and displayed in cartoon mode. Insertion regions of GZ2017-L601 and VA387 are highlighted in blue and magenta, and two variable sites (H374 and E394)of GII.8 NoVs in the P2 region were colored in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. Genome homology of the GII.8 NoV strain from China and other GII NoV genotypes

Sequence homology of the new GII.8 strain GZ2017-L601 and representative strains of other GII genotypes were investigated, which were between 65.4% (GII.18) -72.8% (GII.7) for complete nucleotide genomes **(**Fig. 2**A)**, and 69.0% (GII.18) -82.5% (GII.7) for combined amino acid sequences of all viral proteins **(**Fig. 2**B)**. When calculating pairwise identities for different protein-encoding regions, GZ2017-L601 and the GII.8 representative strain showed high homology, which ranged from 91.5% (NTPase) to 96.2% (P1-1) at the nucleotide level and from 97.6% (VP2) to 100.0% (VPg and P1-1) at the amino acid level. For other representative strains, p22 and P2 were found as the most variable proteins/regions. Average nucleotide similarities of two proteins reached 54.86% (48.8%–65.7%) and 55.28% (48.4%–59.3%), and the average amino acid similarities reached 48.46% (38.5%–67.6%) and 50.58% (44.3%–59.7%). On the other hand, NTPase, VPg, 3CL, RdRp and capsid S protein were found to be the most conserved proteins, and their average pairwise identities reached 72.86% (69.4%–77.4%), 74.51% (70.7%–78.7%), 73.84% (68.7%–79.4%), 74.18% (71.6%–79.8%) and 72.68% (68.2%–80.3%) at the nucleotide level, and 84.23% (78.1%–92.3%), 83.98% (77.4%–89.5%), 83.68% (78.5%–95.0%), 83.95% (79.4%–92.5%) and 79.98% (75.3%–90.2%) at the amino acid level.

3.3. The evolutionary process of GII.8 NoVs based on partial capsid protein VP1 and RdRp regions

For understanding the evolutionary process of GII.8 NoVs, all available GII.8 sequences were collected and phylogenetically analyzed together with GZ2017-L601. Based on the N/C regions of 282 bp, a total of 90 GII.8 strains could be divided into four clusters with an obvious time-ordered evolutionary process **(**Fig. 3**A)**. The strains detected from the 1960s to the 1980s were defined as an ancestral cluster (GII.8-ancestral). Epidemics periods of the other three clusters spanned 1997–2005 (GII.8–1997), 2003–2009 (GII.8–2003), and 2007–2017 (GII.8–2007), respectively. Meantime, the phylogenetic tree based on partial RdRp regions (258 bp) from 30 GII·P8 strains was also conducted, and the result revealed the similar evolutionary process as above **(**Fig. 3**B)**. The Guangzhou NoV strain GZ2017-L601 was identified as the GII·P8–2007/GII.8–2007 variant.

3.4. Differences of sequence, structure and antigen distribution between GII.8 and GII.4 capsid proteins

The differences of sequence, structure and epitope distribution between capsid P proteins of GII.8 strains and GII.4 strain VA387 were compared. In addition to GZ2017-L601, there are only 5 GII.8 capsid protein sequences reported, of which three belong to the GII.8–1997 cluster, one belongs to GII.8–2007, and one is unassigned **(**Fig. 3**A)**. No insertion or deletion mutations were found within GII.8 strains, and there existed 15 amino acid replacements, only 4 of which were in the P2 region (Fig. 5). When compared with GII.4 strainVA387, two insertion and three deletions were found in the P2 region of the GII.8 NoV strain. In detail, two insertion mutations (aa 291, aa 297–298) and two deletion mutations (aa 341/342, aa 372/373) mainly located near the epitope A and epitope C, and the remaining deletion occurred between aa 398 and 399. Homology modeling of GII.8 capsid proteins was performed to determine the potential effect of these mutations. Two insertions resulted in two expanding loops near the epitope A, whereas the deletion between 398 and 399 truncated the original loop of VA387 (Fig. 4). Epitope distribution of the above structures was also compared (Fig. 5). It was found that the number of predicted antigen sites on the capsid of VA387 (n = 92) was more than those of six GII.8 strains (69, 66, 77, 74, 77, 69), but their distribution was comparable. Moreover, no discernible differences of epitope distribution were also found between GII.8 strains.

4. Discussion

NoVs are an important pathogen of acute gastroenteritis in the world, and GII.8 belongs to a non-epidemic genotype which was often detected with low positive rates. However, the understanding of this genotype is still limited, and little genetic information has been accumulated so far. In this study, we assembled the first GII.8 NoV genome from China and clarified the genetic diversity and evolutionary process for this genotype.

GII.8 is a commonly detected NoV genotype. During the surveillance of sporadic NoV infections, GII.8 was reported to be detected in children with diarrhea in Korea and Thailand and in asymptomatic children in Brazil (Barreira et al., 2010; Malasao et al., 2008; Yoon et al., 2008). Besides, GII.8 could also cause NoV outbreaks. In the United States and Japan, GII.8 was detected in samples from diarrhea outbreaks from the 1960s to the 1980s (Bok et al., 2009; Mori et al., 2017). Recently, it was reported that GII.8 outbreaks occurred in nursing care center, hotels, families, schools, and catering (Lysen et al., 2009; Ozawa et al., 2007). In addition, GII.8 were also found during food and environmental monitoring, such as Han River in Korea (GenBank Accession Number EU527748, EU527751, EU527773, EU527777) and oysters in Japan (Nishida et al., 2003).

NoVs have a rich genetic diversity, and the homology of its different

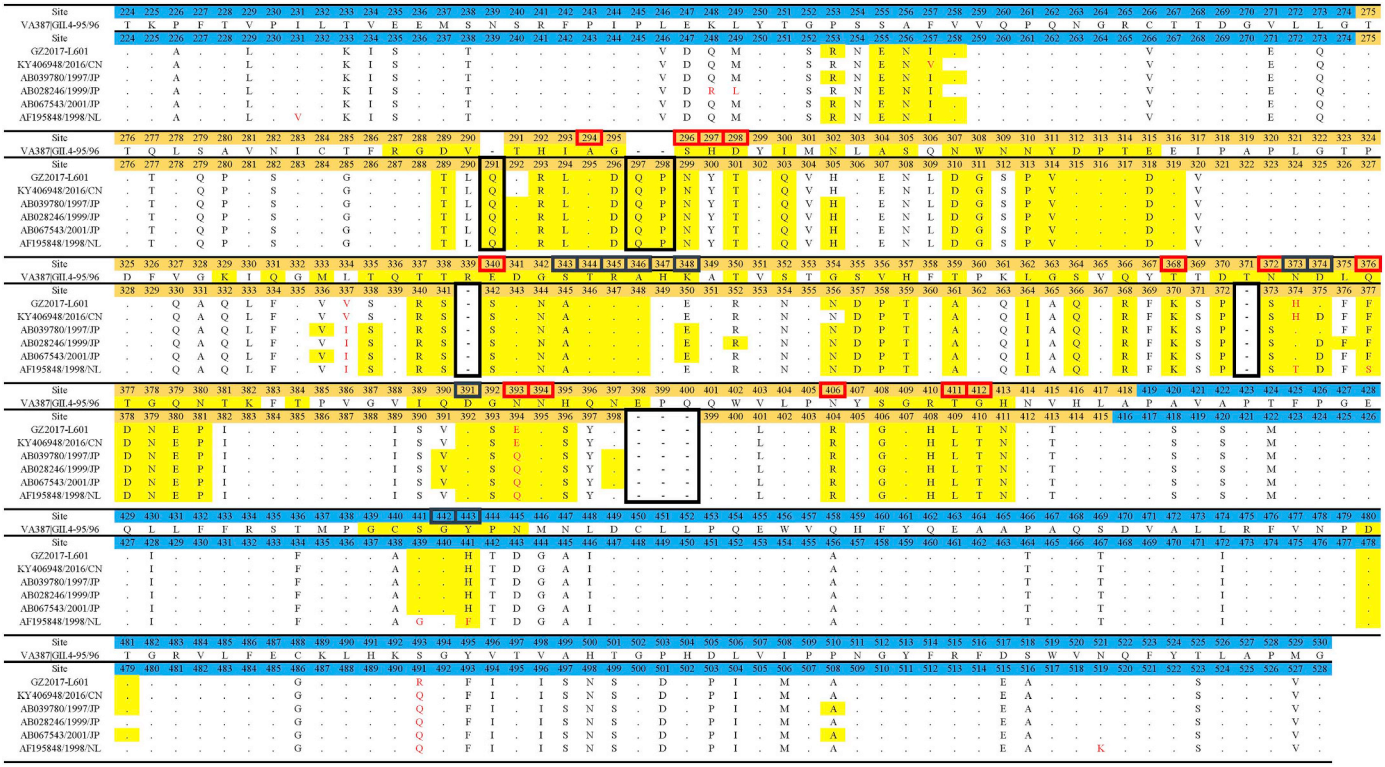


Fig. 5. Differences of sequences and epitope distribution between capsid P proteins of GII.8 strains and the GII.4 VA387 strain. A black dot indicates the same usage of amino acid as that of the GII.4 strain, and a dash indicates a deletion-mutant. The potential spatial epitopes are highlighted in yellow, and the positions with different amino acids among GII.8 strains are colored in red. The predicted antigen epitope and receptor-binding sites of VA387 are in red and dark blue boxes, respectively. Insertion and deletion mutations of GII.8 strains are highlighted in black boxes. The numbers of predicted antigen sites on capsid P proteins of above strains were 92 (VA387|GII.4–95/96), 69 (GZ2017-L601), 66 (KY406948/2016/CN), 77 (AB039780/1997/JP), 74 (AB028246/1999/JP), 77 (AB067543/2001/JP), 69 (AF195848/1998/NL), respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

protein-encoding regions in the viral genome is inconsistent. It has reported that p48, p22, P2, and VP2 were verified to be variable proteins compared to others (Cotten et al., 2014; Xue et al., 2016c). Homology between GZ2017-L601 and other genotypes at the genomic level differed between 65.4%–72.8% at the nucleotide level and 69.0%–82.5% at the amino acid level. When analyzing for different protein-coding regions, the lowest homology could reach 54.86% (nt level) and 48.46% (aa level) for p22 protein, and the highest homology could reach 74.51% (nt level) for VPg and 84.23% (aa level) for NTPase. It is necessary to verify different viral proteins for fully understanding NoVs. In addition, some GII.8 strains also undergo recombination, and recombinant types of GII·P4/GII.8 and GII·P8/GII.7 (GenBank Accession Number JX846926) have been reported (Nayak et al., 2009). Currently, the dual NoV nomenclature based on partial RdRp and VP1 genes has been proposed and widely used (Kroneman et al., 2013), but genome-based NoV researches are an inevitable direction in the future.

The limitation of GII.8 genetic information hindered the understanding of GII.8 evolutionary process. The divergence time of GII.8 was predicted as in 1948 (1921–1970, 95% HPDs) based on the capsid protein VP1 (Kobayashi et al., 2016). The earliest detected GII.8 strains were collected in 1967 (GenBank Accession Number AB684660 and AB684661) (Mori et al., 2017). However, only a few GII.8 NoV sequences have been accumulated as of this writing (van Beek et al., 2018), and only one available GII.8 genome sequence was available in GenBank, which belonged to a Japanese strain collected in 1997. Therefore, in this study, we established a new genome cloning method and obtained the first GII.8 genome from China, which would contribute to the accumulation of GII.8 genetic information in the future.

Based on the available sequences of partial capsid protein or RdRp encoding regions, GII.8 showed a similar time-ordered evolutionary process as GII.4 and GII.17 genotypes (Parra et al., 2017). Since the 1960s, all reported GII.8 strain could be further divided into four clusters chronologically. However, when we compared the only two GII.8 genomes that were separated by two decades, there existed 464 nucleotide mutations between them, but only 23 of them caused amino acid mutations (Table 2). Even for alignments of all available GII.8 capsid proteins from three clusters, there were 206 nucleotide variable sites, but only 15 amino acid variable sites. In previous studies, it was found that GII.4 strains not only had a higher mutation rate (5.54×10^{-4} to 9.06×10^{-4} nucleotide substitutions/site) and rate of evolution (3.9×10^{-3} nucleotide substitutions/site/year) compared to the less frequently detected strains, but also had a greater number of non-synonymous changes (Ka/Ks ratio = 0.0912 ± 0.0322) compared to other NoVs (Bull et al., 2010). Similarly, we also compared the differences between new identified GII.17 sequences and other reported genomes of the same genotype previously. Although only 84 nucleotide changes were identified, it resulted in 15 amino acid changes (Xue et al., 2016c). Low nonsynonymous mutation rate suggested a decreased incidence/epidemiological fitness, which may be through less antigenic drift, a consequence of the lower mutation rate of the viral RdRp (Bull et al., 2010). Therefore, more GII.8 sequences should to be accumulated in the future, which helps to elucidate the viral evolution rate and identify the mutation rate of its RNA dependent RNA polymerase.

Receptor binding function and antigenic epitopes of capsid proteins are regarded as important drivers of NoV evolution. It was reported that different GII.4 epidemic variants emerged through constant variation of epitopes in the P2 region (Lindesmith et al., 2008). And GII.17 has also undergone variations in antigenic epitopes, which caused its sudden emergence and epidemics in 2014/2015 in East Asia (Chan et al., 2015). However, few non-synonymous mutations on the GII.8 capsid proteins occurred. Only 4 variable sites were found in the P2 region, but

none of them occurred in the predicted receptor binding sites or epitopes (Tan and Jiang, 2010). Conservation of capsid proteins at the amino acid level may be one of the reasons for the non-epidemics of GII.8 NoVs. In addition, it has been reported that there was weak cross-reactivity of GII.8 antiserum for other genotypes, and the existing immunoassay kits could not be used to detect GII.8 NoVs, indicating that the prevalence of other genotypes might not affect GII.8 evolution (Gray et al., 2007; Hansman et al., 2006). Therefore, it is necessary to monitor the distribution of anti-GII.8 in serum to investigate the limiting effects of herd immunity on the new GII.8 variants in the future.

Although GII.4 NoV has been recognized as the dominant genotype for over two decades, the sudden emergence of the novel GII.17 variant shows the necessity to strengthen the understanding of other non-epidemic genotypes. In this study, based on accessing the first GII.8 genome from China, we reconstructed the time-order evolution process of GII.8 variants over fifty years, and described the variation of its capsid protein structure and immunogenicity via the predominant genotype GII.4. The results of this study could not only provide reference data for NoV researches in the future, but also deepen the understanding of evolution mechanisms of the non-predominant NoV variants.

Disclosure statement

No competing financial interests exist.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (31701717), the National Key Research and Development Program of China (2017YFC1601200), Pearl River S&T Nova Program of Guangzhou (201610010020), Implementation of Innovation Driven Development Capacity Building Special Funds of Guangdong Academy of Sciences (2017GDASCX-0814), and Science and Technology Planning Project of Guangdong Province (2017A070702018). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Ao, Y., Wang, J., Ling, H., He, Y., Dong, X., Wang, X., Peng, J., Zhang, H., Jin, M., Duan, Z., 2017. Norovirus GII.P16/GII.2-associated gastroenteritis, China, 2016. Emerg. Infect. Dis. 23, 1172–1175.

Arnold, K., Bordoli, L., Kopp, J., Schwede, T., 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22, 195–201.

Barreira, D.M.P.G., Ferreira, M.S.R., Fumian, T.M., Checon, R., de Sadovsky, A.D.I., Leite, J.P.G., Miagostovich, M.P., Spano, L.C., 2010. Viral load and genotypes of noroviruses in symptomatic and asymptomatic children in Southeastern Brazil. J. Clin. Virol. 47, 60–64.

Bartsch, S.M., Lopman, B.A., Ozawa, S., Hall, A.J., Lee, B.Y., 2016. Global economic burden of norovirus gastroenteritis. PLoS One 11, e0151219.

van Beek, J., de Graaf, M., Al-Hello, H., Allen, D.J., Ambert-Balay, K., Botteldoorn, N., Brytting, M., Buesa, J., Cabrerizo, M., Chan, M., Cloak, F., Di Bartolo, I., Guix, S., Hewitt, J., Iritani, N., Jin, M., Johne, R., Lederer, I., Mans, J., Martella, V., Maunula, L., McAllister, G., Niendorf, S., Niesters, H.G., Podkolzin, A.T., Poljsak-Prijatelj, M., Rasmussen, L.D., Reuter, G., Tuite, G., Kroneman, A., Vennema, H., Koopmans, M.P.G., NoroNet, 2018. Molecular surveillance of norovirus, 2005-16: an epidemiological analysis of data collected from the NoroNet network. Lancet Infect. Dis. 18, 545–553.

Bernstein, D.I., Atmar, R.L., Lyon, G.M., Treanor, J.J., Chen, W.H., Jiang, X., Vinje, J., Gregoricus, N., Frenck, R.W., Moe, C.L., Al-Ibrahim, M.S., Barrett, J., Ferreira, J., Estes, M.K., Graham, D.Y., Goodwin, R., Borkowski, A., Clemens, R., Mendelman, P.M., 2015. Norovirus vaccine against experimental human GII.4 virus illness: a challenge study in healthy adults. J. Infect. Dis. 211, 870–878.

Bok, K., Abente, E.J., Realpe-Quintero, M., Mitra, T., Sosnovtsev, S.V., Kapikian, A.Z., Green, K.Y., 2009. Evolutionary dynamics of GII.4 noroviruses over a 34-year period. J. Virol. 83, 11890–11901.

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.

Chan, M., Lee, N., Hung, T., Kwok, K., Cheung, K., Tin, E., Lai, R., Nelson, E., Leung, T., Chan, P., 2015. Rapid emergence and predominance of a broadly recognizing and fast-evolving norovirus GII.17 variant in late 2014. Nat. Commun. 6, 10061.

Cotten, M., Petrova, V., Phan, M.V.T., Rabaa, M.A., Watson, S.J., Ong, S.H., Kellam, P., Baker, S., 2014. Deep sequencing of norovirus genomes defines evolutionary patterns in an urban tropical setting. J. Virol. 88, 11056–11069.

de Graaf, M., van Beek, J., Vennema, H., Podkolzin, A.T., Hewitt, J., Bucardo, F., Templeton, K., Mans, J., Nordgren, J., Reuter, G., Lynch, M., Rasmussen, L.D., Iritani, N., Chan, M.C., Martella, V., Ambert-Balay, K., Vinje, J., White, P.A., Koopmans, M.P., 2015. Emergence of a novel GII.17 norovirus - End of the GII.4 era? Eur. Secur. 20, 8–15.

Gray, J.J., Kohli, E., Ruggeri, F.M., Vennema, H., Sanchez-Fauquier, A., Schreier, E., Gallimore, C.I., Iturriza-Gomara, M., Giraudon, H., Pothier, P., Di Bartolo, I., Inglese, N., de Bruin, E., van der Veer, B., Moreno, S., Montero, V., de Llano, M.C., Hoehne, M., Diedrich, S.M., 2007. European multicenter evaluation of commercial enzyme immunoassays for detecting norovirus antigen in fecal samples. Clin. Vaccine Immunol. 14, 1349–1355.

Hansman, G.S., Natori, K., Shirato-Horikoshi, H., Ogawa, S., Oka, T., Katayama, K., Tanaka, T., Miyoshi, T., Sakae, K., Kobayashi, S., Shinohara, M., Uchida, K., Sakurai, N., Shinozaki, K., Okada, M., Seto, Y., Kamata, K., Nagata, N., Tanaka, K., Miyamura, T., Takeda, N., 2006. Genetic and antigenic diversity among noroviruses. J. Gen. Virol. 87, 909–919.

Iritani, N., Kaida, A., Abe, N., Sekiguchi, J., Kubo, H., Takakura, K.I., Goto, K., Ogura, H., Seto, Y., 2012. Increase of GII.2 norovirus infections during the 2009–2010 season in Osaka City, Japan. J. Med. Virol. 84, 517–525.

Katayama, K., Shirato-Horikoshi, H., Kojima, S., Kageyama, T., Oka, T., Hoshino, F., Fukushi, S., Shinohara, M., Uchida, K., Suzuki, Y., Gojobori, T., Takeda, N., 2002. Phylogenetic analysis of the complete genome of 18 Norwalk-like viruses. Virology 299, 225–239.

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.

Kojima, S., Kageyama, T., Fukushi, S., Hoshino, F.B., Shinohara, M., Uchida, K., Natori, K., Takeda, N., Katayama, K., 2002. Genogroup-specific PCR primers for detection of Norwalk-like viruses. J. Virol. Methods 100, 107–114.

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., Vennema, H., Deforche, K., von der Avoort, H., Penaranda, S., Oberste, S., Vinje, J., Koopmans, M., 2011. An automated genotyping tool for enteroviruses and noroviruses. J. Clin. Virol. 51, 121–125.

Kroneman, A., Vega, E., Vennema, H., Vinje, 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. 159, 2059–2068.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

Lindesmith, L.C., Donaldson, E.F., Lobue, A.D., Cannon, J.L., Zheng, D.P., Vinje, J., Baric, R.S., 2008. Mechanisms of GII.4 norovirus persistence in human populations. PLoS Med. 5, 269–290.

Lopman, B.A., Steele, D., Kirkwood, C.D., Parashar, U.D., 2016. The vast and varied global burden of norovirus: prospects for prevention and control. PLoS Med. 13, e1001999.

Lu, Q.B., Huang, D.D., Zhao, J., Wang, H.Y., Zhang, X.A., Xu, H.M., Qu, F., Liu, W., Cao, W.C., 2015. An increasing prevalence of recombinant GII norovirus in pediatric patients with diarrhea during 2010-2013 in China. Infect. Genet. Evol. 31, 48–52.

Lysen, M., Thorhagen, M., Brytting, M., Hjertqvist, M., Andersson, Y., Hedlund, K.O., 2009. Genetic diversity among food-borne and waterborne norovirus strains causing outbreaks in Sweden. J. Clin. Microbiol. 47, 2411–2418.

Malasao, R., Maneekarn, N., Khamrin, P., Pantip, C., Tonusin, S., Ushijima, H., Peerakome, S., 2008. Genetic diversity of norovirus, sapovirus, and astrovirus isolated from children hospitalized with acute gastroenteritis in Chiang Mai, Thailand. J. Med. Virol. 80, 1749–1755.

Mori, K., Nagano, M., Kimoto, K., Somura, Y., Akiba, T., Hayashi, Y., Sadamasu, K., Kai, A., 2017. Detection of enteric viruses in fecal specimens from nonbacterial foodborne gastroenteritis outbreaks in Tokyo, Japan between 1966 and 1983. Jpn. J. Infect. Dis. 70, 143–151.

Nayak, M.K., Chatterjee, D., Nataraju, S.M., Pativada, M., Mitra, U., Chatterjee, M.K., Saha, T.K., Sarkar, U., Krishnan, T., 2009. A new variant of Norovirus GII.4/2007 and inter-genotype recombinant strains of NVGII causing acute watery diarrhoea among children in Kolkata, India. J. Clin. Virol. 45, 223–229.

Nishida, T., Kimura, H., Saitoh, M., Shinohara, M., Kato, M., Fukuda, S., Munemura, T., Mikami, T., Kawamoto, A., Akiyama, M., Kato, Y., Nishi, K., Kozawa, K., Nishio, O., 2003. Detection, quantitation, and phylogenetic analysis of noroviruses in Japanese oysters. Appl. Environ. Microbiol. 69, 5782–5786.

Ozawa, K., Oka, T., Takeda, N., Hansman, G.S., 2007. Norovirus infections in symptomatic and asymptomatic food handlers in Japan. J. Clin. Microbiol. 45, 3996–4005.

Parra, G.I., 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.

Siebenga, J.J., Vennema, H., Zheng, D.P., Vinje, J., Lee, B.E., Pang, X.L., Ho, E.C.M., Lim, W., Choudekar, A., Broor, S., Halperin, T., Rasool, N.B.G., Hewitt, J., Greening, G.E., Jin, M., Duan, Z.J., Lucero, Y., O'Ryan, M., Hoehne, M., Schreier, 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.

Tan, M., Jiang, X., 2010. Norovirus gastroenteritis, carbohydrate receptors, and animal

models. PLoS Pathog. 6, e1000983.
Tu, E.T.V., Bull, R.A., Greening, G.E., Hewitt, J., Lyon, M.J., Marshall, J.A., McIver, C.J., Rawlinson, W.D., White, P.A., 2008. Epidemics of gastroenteritis during 2006 were associated with the spread of norovirus GII.4 variants 2006a and 2006b. Clin. Infect. Dis. 46, 413–420.
Xue, L., Cai, W., Wu, Q., Zhang, J., Guo, W., 2016a. Direct sequencing and analysis of the genomes of newly emerging GII.17 norovirus strains in South China. J. Appl. Microbiol. 120, 1130–1135.
Xue, L., Cai, W.C., Wu, Q.P., Kou, X.X., Zhang, J.M., Guo, W.P., 2016b. Development of a sensitive method for directly sequencing GII.4 norovirus genome. Diagn. Microbiol. Infect. Dis. 84, 200–202.
Xue, L., Wu, Q.P., Cai, W.C., Zhang, J.M., Guo, W.P., 2016c. Molecular characterization of new emerging GII.17 norovirus strains from South China. Infect. Genet. Evol. 40, 1–7.
Xue, L., Wu, Q., Dong, R., Cai, W., Wu, H., Chen, M., Chen, G., Wang, J., Zhang, J., 2017. Comparative phylogenetic analyses of recombinant noroviruses based on different protein-encoding regions show the recombination-associated evolution pattern. Sci. Rep. 7, 4976.
Yoon, J.S., Lee, S.G., Hong, S.K., Lee, S.A., Jheong, W.H., Oh, S.S., Oh, M.H., Ko, G.P., Lee, C.H., Paik, S.Y., 2008. Molecular epidemiology of norovirus infections in children with acute gastroenteritis in South Korea in November 2005 through November 2006. J. Clin. Microbiol. 46, 1474–1477.