Norovirus GII.2[P16] Strain in Shenzhen, China: A Retrospective Study

Jing Wang
Wuchang Hospital  https://orcid.org/0000-0002-7074-8351

Miao Jin
National Institute for Viral Disease Control and Prevention

Hailong Zhang
Shenzhen Center for Disease and Prevention

Yanan Zhu
Department of Epidemiology, School of Public health, Southern Medical University

Hong Yang
Shenzhen Center for Disease Control and Prevention

Xiangjie Yao
Shenzhen Center for Disease Control and Prevention

Long Chen
Shenzhen Center for Disease Control and Prevention

Jun Meng
Shenzhen Center for Disease Control and Prevention

Guifang Hu
Department of Epidemiology, School of Public health, Southern Medical University

Yaqing He (✉ heyaqing1019@126.com)
https://orcid.org/0000-0001-6608-8541

Zhaojun Duan
National Institution for Disease Control and Prevention, Chinese Center Control and prevention

Research article

Keywords: Norovirus, Epidemiology, GII.P16/GII.2 strain, Recombinant, Evolution

DOI: https://doi.org/10.21203/rs.3.rs-31347/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Background: Norovirus (NoV) is the main cause of non-bacterial acute gastroenteritis (AGE) outbreaks worldwide. From September 2015 through August 2018, 203 NoV outbreaks involving 2,500 cases were reported to the Shenzhen Center for Disease Control and Prevention.

Methods: Faecal specimens for 203 outbreaks were collected and epidemiological data were obtained through the AGE outbreak surveillance system in Shenzhen. Genotypes were determined by sequencing analysis. To gain a better understanding of the evolutionary characteristics of NoV in Shenzhen, recombination events were analysed and molecular evolution was evaluated based on time-scale evolutionary phylogeny and amino acid mutations.

Results: A total of nine counties reported NoV outbreaks and the reported NoV outbreaks peaked from November to March. Among the 203 NoV outbreaks, 150 were sequenced successfully. Most of these outbreaks were associated with the NoV GII.2[P16] strain (45.3%, 92/203) and occurred in school settings (91.6%, 186/203). The GII.2[P16] strain is a recombinant strain that is relatively stable. In these GII.2[P16] recombination strain outbreaks, the non-structural proteins of the recombination strain might have played a more significant role than VP1.

Conclusions: This study illustrates the molecular epidemiological patterns in Shenzhen, China, from September 2015 to August 2018 and provides evidence that the GII.2[P16] strain was relatively stable and that epidemic trend has weakened.

Background

Norovirus (NoV), which is the main cause of non-bacterial acute gastroenteritis (AGE) worldwide and could infect all age groups, especially children under 5 years of age. According to estimations, NoV is annually associated with 900,000 clinic visits amongst children in industrialized countries and up to 200,000 deaths of children in developing countries [1, 2]. In general, NoV circulates in colder weather and causes gastrointestinal symptoms such as vomiting, diarrhoea and abdominal pain. NoV outbreaks are frequently reported in semi-closed institutions, such as hospitals, nursing homes, schools, and childcare centres [3].

NoV belongs to the Caliciviridae family and can be divided into 10 genogroups (GI~GX), of which GI, GII and GIV infect humans. GI and GII are responsible for the majority of human diseases and can be further divided into nine (GI.1-GI.9) and 27 (GII.1-GII.27) genotypes based on the diversity of VP1[4]. The full-length of the single stranded RNA genome is approximately 7.5~7.7 kb, with three open reading forms (ORFs)[5]. The first 5 kb closest to the 5’ end of the genome is ORF1, which encodes non-structural proteins, including N terminal protein (P48), NTPase, 3A protein (P22), VPg (viral genomic junction protein), 3 C-like protein (Pro) and RNA-dependent RNA polymerase (RdRp)[6]. These proteins are important for replication of NoV. ORF2 is 1.6 kb in length and encodes the major structural protein VP1, which composes the main capsid structure and is responsible for the infectivity and antigenicity of
NoV[7]. The VP1 contains a well-conserved shell (S) domain and a protruding (P) domain, and the latter is divided into two sub-domains, P1 and P2[8]. Furthermore, the P2 region is considered to be a hypervariable part of the genome because the domain encodes the receptor binding domain, which is responsible for histoblood group antigen (HBGA) binding, and important epitopes targeted by antibodies that inhibit binding[9, 10]. ORF3 is 0.6 kb and encodes the minor structural protein (VP2) [11].

The dominant epidemic variant strain is generally GII.4, since 2002, new GII.4 variants have emerged every 2-3 years and replaced the previously predominant GII.4 strains, resulting in epidemics and sometimes global pandemics of AGE including Hunter2004, Yerseke2006a, Den Haag2006b, New Orleans2009 and Sydney 2012 [12]. However, during the winter of 2014-2015, a novel GII.17 strain initially emerged in Guangdong Province, surpassing GII.4-caused NoV infections[13]. Moreover, in late 2016, the re-emergence GII.2[P16] led to rapidly increasing AGE outbreaks in China[14] and during a short time, the GII.2[P16] recombinant strain swept through Japan, Italy, Germany[15-17]. The first GII.2[P16]-positive sample was also detected in Guangdong Province[14].

Shenzhen is one of the most important cities in Guangdong Province. However, the information about NoV outbreaks in this region is limited. The retrospective study aimed to determine the genotypic diversity of NoV strains in outbreaks and the genetic characteristics of the GII.2[P16] strain in Shenzhen, China, from September 2015 to August 2018.

**Methods**

**The surveillance of NoV outbreaks**

Faecal specimens in AGE outbreaks inspected to the Shenzhen Center for Disease Control and Prevention (Shenzhen CDC) by County Centers for Disease Control and Prevention (County-level CDCs) from September 2015 to August 2018 were obtained. County-level CDCs are responsible for conducting outbreak investigations, including providing epidemiological and clinical information. The Shenzhen CDC performs NoV detection and genotyping on the specimens. The NoV outbreaks were identified as >5 acute gastroenteritis cases within 3 days after exposure in a common setting where >2 samples (whole fecal, rectal swab, or vomitus) had been laboratory confirmed as NoV.

**Detection of NoV by real-time RT-PCR**

For faecal specimen analysis, a 10% suspension was prepared by mixing 0.1 g stool with 1 mL phosphate-buffered saline (pH 7.2). Viral RNA was extracted from the clarified stool suspension using the Viral Nucleic Acid Extraction Kit II (Geneaid, China), after which the viral RNA was examined by real-time reverse transcription polymerase chain reaction (real-time RT-PCR) using Ag-Path Kit (Applied Biosystems, USA) with primers (Cog1F, Cog1R, Cog2F, and Cog2R) and TaqMan probe (Ring 1E and Ring 2) (S1
The cycling conditions were described previously [18]. A negative control containing DEPC water and 2 positive controls containing RNA of NoV GI and GII were included in each experiment. Samples were scored as positive if the cycle threshold values were $\leq 40$ and the positive and negative controls showed the expected values.

**Genotyping of NoV by conventional RT-PCR**

NoV-positive samples were then amplified by conventional Reverse transcription and PCR (RT-PCR) using one-step RT-PCR Kit (QIAGEN, Germany). Before October 2016, the primer sets G1SKF/G1SKR and COG2F/G2SKR were used for VP1 genotyping to detect GI(330bp) and GII(387bp), respectively [19] (S1 Table). After October 2016, the primer sets MON432/G1SKR and MON431/G2SKR were used to amplify both of partial RdRp region and VP1 sequence of GI (543 bp) or GII (557 bp), respectively [20] (S1 Table).

**Genotyping analysis**

Genotypes were confirmed by the BLAST and an automated online NoV genotyping tool offered by the Netherlands National Institute for Public Health and the Environment (RIVM, http://www.rivm.nl/mpf/norovirus/typingtool)[21].

**Genome amplification of strains GII.2[P16] of NoV**

The genomes of the strains genotyped as GII.2[P16] were further amplified. cDNA was gained by reverse transcription of viral RNA by SuperScript III kit (invitrogen, USA). Six primer sets (S1 Table) were designed based on the whole genome KY421121.

**Intergenic recombination**

To confirm that the GII.2[P16] strain from Shenzhen is a recombinant strain as other areas in China, the nearly full-length genome sequences of the viral strains in this study were analysed with the reference strains retrieved from GenBank by using a Simplot software v.3.5.1 to plots the percent identity[22].

**Phylogenetic analysis of the RdRp region and VP1**

To evaluate the evolution of the NoV GII.2[P16] strain in Shenzhen, the full-length RdRp region or VP1 sequence from this study and all the sequences of the full-length RdRp region or VP1 sequence we found in GenBank as of September 2016 were collected. Phylogenetic trees were constructed using the Markov chain Monte Carlo (MCMC) method with the strict molecular clock in BEAST software v 1.8.2. The best substitution model was TN93 (Tamura-Nei)+G (Gamma) and TN93 (Tamura-Nei)+G (Gamma)+I (Invariable) of RdRp region and VP1 sequence, selected by MEGA 6.0 using the BIC method[23] Analysis of the phylogenetic data was confirmed until the effective sample sizes were greater than 200 based on Tracer. The final result was visualized using the FigTree software v1.4.3.

**Recombination variant**
To evaluate the impact of the intergenic recombination of non-structural region and VP1, the amino acid mutations of non-structural region and VP1 among different genotypes were analysed by MEGA6.0.

**Statistical analysis**

The difference between GII.2 NoV detection rates in the dominant setting distribution were compared using Fisher's exact test in SPSS Statistics software v.22.0, and a \( p \)-value less than 0.05 was considered statistically significant.

**Nucleotide sequence accession numbers**

The GenBank accession numbers for the GII.2[P16] strain sequences obtained in this study are MK729081, MK681452, MK614124-MK614161, MK720506-MK720583, and MK692738-MK692654.

**Results**

**NoV outbreak settings and geographical locations**

According to ten county-level CDCs, there were 203 NoV outbreaks in Shenzhen between the period September 2015 and August 2018. The most outbreaks were from Nanshan and no outbreak was from Yantian (Figure 1). Information on the outbreak size was reported for 197 (97.0%), ranging from 5 to 115 cases per outbreak (Table 1). Of the 203 outbreaks, 91.6% occurred in school settings, with 8.4% occurring in non-school settings (Table 2). The reported outbreaks peaked in the cold season, especially from November to March (Figure 2).

**Genotypic distribution of identified NoV**

Of the 203 outbreaks detected as caused by NoV according to real-time RT-PCR from September 2015 to August 2018, 150 were successfully genotyped. Of these 150 outbreaks with genotype information, 137 (91.3%,137/150) and 12 (8.0%,12/150) were classified into GII and GI genogroups. 1 (0.6%,1/150) outbreak involved both GI- and GII-positive samples. A total of 15 capsid genotypes and 15 polymerase genotypes were identified. The dominant genotype was GII.2[P16]. In addition, we identified a novel recombinant genotype GII.Pe/GII.17 that had not been previously found in Shenzhen before (Table 3).

**Genotype distribution and outbreak characteristics**

For outbreaks caused by the GII.2 strain, most occurred in school settings: 73 (79.3%, 73/92) outbreaks occurred in child care centre and the dominant setting distribution (child care centre, primary school, middle school) of GII.2 infection showed no significant differences (Fisher's exact test=3.595, \( p=0.177 \)). Of the thirteen outbreaks caused by the GII.3 strains, most (86%, 11/13) also occurred in child care centre.

**Intergenic recombination**
To determine whether all the GII.2[P16] strains in Shenzhen are recombinant strains, 21 nearly full-length genome sequences from Shenzhen were analysed using Simplot software v.3.5.1, with GII.P16/GII.4 Sydney2012 (KY887605.1), GII.2[P16] (LC20944.8) and GII.2[P16] (MG746027.1) as the reference strain sequences. The results showed that the ORF1 region of GII.2[P16] in Shenzhen is highly homologous to that of GII.P16/GII.4 Sydney2012. ORF2 and ORF3 are highly homologous with LC20944.8 and MG746027.1. The breakpoint was found to be 5003 nt, close to the ORF1/ORF2 overlap region (Figure 3).

**Phylogenetic analysis of RdRp region and VP1 sequence of GII.2[P16] Strain**

To examine strain evolution, 52 full-length RdRp region of GII.2[P16] strain sequences from Shenzhen and 95 reference sequences from GenBank were collected for analysis. MCMC chains were run for $1.0 \times 10^8$ steps for the RdRp region sequences. Effective sample sizes greater than 200 were confirmed by the Tracer. According to the MCC (Maximum clade credibility) tree, the evolutionary rate of the RdRp region of the GII.P16/GII.2 strain was estimated as $2.1 \times 10^{-3}$ substitutions/site/year (95% HPD interval, $1.7 \times 10^{-3}$ - $2.5 \times 10^{-3}$ substitutions/site/year). The common ancestors of the GII.2[P16] strain from Shenzhen and GII.P16/GII.4 Sydney 2012 diverged from 2011 to 2012, and the prototype of the RdRp region of the GII.2[P16] strain from Shenzhen formed during 2012 to 2013. The phylogenetic analyses suggested that the RdRp region evolved from GII.P16/GII.4 Sydney2012 (Figure 4A).

Simultaneously, 72 full-length VP1 sequences of GII.2[P16] retrieved from Shenzhen and 146 GenBank reference sequences were used to explore the evolutionary rate. MCMC chains were run for $2.0 \times 10^8$ steps for the VP1 sequence, Effective sample sizes were greater than 200, as confirmed by Tracer. The evolutionary rate of the VP1 sequence of the GII.2[P16] strain was estimated at $2.7 \times 10^{-3}$ substitutions/site/year (95% HPD interval, $2.4 \times 10^{-3}$-3.1$ \times 10^{-3}$ substitutions/site/year) based on the MCC tree. The common ancestors of the GII.2[P16] strain from Shenzhen and previous GII.2[P16] diverged from 2003 to 2004, and the prototype of VP1 sequence of GII.2[P16] strain in Shenzhen was formed during 2013 to 2014. The phylogenetic analyses suggested that VP1 evolved from GII.2[P16] (2010-2012) (Figure 4B).

**Amino acid mutations of non-structural region of GII.2[P16]**

To explore the amino acid mutations within the non-structural region of the recombinant strains, 14 nearly full-length non-structural protein sequences and 22 full-length reference sequences, including GII.17[P16] (2002), GII.2[P16] (2009-2014), GII.2[P16] (2010-2012), GII.13[P16] (2015), GII.3[P16] (2012-2013), GII.4[P16] (2015-2016) and GII.17[P16] (2016-2018), from GenBank were aligned. Sequence data revealed 102 (6%) parsimony-informative sites, but no amino acid mutation in non-structural region of the GII.2[P16] recombinant strain. Furthermore, 6 amino acid substitutions (*77E, R750K, P845Q, H1310Y, K1546Q, T1549A) were found only in recent strains (GII.4 Sydney 2012[P16] and the GII.2[P16] recombinant strain), 2 sites (A644P, A1521V) were substituted in the GII.2[P16] recombinant strains and 1 site (S/T753T) reverted. The result showed that amino acid 1310 (Motifs G) was substituted (Table 4).
Amino acid mutations of HBGA-binding and epitopes sites of the GII.2[P16]

To explore HBGA-Binding profile, predicted epitopes and epitope A to E sites of the GII.2[P16] recombinant strain[24-26], 72 full-length VP1 sequences from this study and 65 reference sequences, including GII.2[Pc] (1976-1978), GII.2[Ph] (1997), GII.2[P2] (1987-2015), GII.2[P12] (2004-2006), GII.2[P21] (2010), GII.2[Pe] (2014), GII.2[P16] (2010-2012), GII.2[P16] (2008-2014) and GII.2[P16] (2016-2018), from 1975 to 2018 were collected and aligned. Sequencing data revealed 29 parsimony-informative sites but there were no mutations in the HBGA-binding profile, predicted epitopes and epitope A to E of the GII.2[P16] strain (S2 Table).

Discussion

In this study, NoV-associated AGE outbreaks in Shenzhen, China, from September 2015 to August 2018 were analysed. A total of 203 NoV outbreaks were reported to the Shenzhen CDC. The NoV infection was initially described as "winter vomiting disease" due to its seasonal characteristic[27]. Analysis of monthly distribution also indicated that the peak of the outbreak in Shenzhen occurred from November to March. Previous studies have found a link between climate or weather and increased NoV abundance, and low absolute humidity provides an ideal condition for NoV persistence and transmission during cold months[28]. Indeed, NoV rapidly loses viability and infectivity with the increase in increasing temperature, therefore NoV appears to be more stable in a cold climate and thus is transmitted more easily among people at cold times of the year[29, 30]. The peak in this study was in December, when Shenzhen began to become cold, and March, when the temperature began to turn warm, suggesting that that climate changes has an impact on NoV transmission. The NoV outbreaks usually occur in hospital, nursing home, school, child care centre, hotels and other semi-enclosed places [3]. A study in United States reported 3,960 NoV outbreaks between 2009 and 2013 and found that long-term care homes were the most frequent sites of NoV outbreaks[31]. Another study from Qin et al[32] showed that middle school was the most important setting of NoV outbreaks in China, followed by primary school between 2006 and 2016. In this study, we classified the outbreak settings into 12 categories, and the results showed that most were occurred in child care centre, followed by primary school. This suggests that school remains the most common setting for NoV outbreaks in Shenzhen, but that the current high incidence is occurring among younger children who are under six years of age. Combining the results of the monthly distribution of NoV outbreaks in Shenzhen, we suspect that the decrease in the number of NoV outbreaks in January and February is related to school holidays. When the scale of the outbreaks was analysed, the average number of people involved per outbreak in Shenzhen was nine, smaller than the 18 persons reported in the United States[31]. Shenzhen is one of the cities where the economy is most developed, which may be a benefit of the local public health system and highly effective handling of public health emergencies in Shenzhen (http://www.szemo.gov.cn). Regarding genotype detection, both GI and GII genogroups were found, as were 15 capsid types and 15 polymerase types. Among the genotypes, the most common was GII.2, followed by GII.3. GII.4 Sydney 2012 only accounted for 3.4%. In this study, we
identified the GII.2 strain was GII.2[P16] recombinant strain as other regions in China [33, 34]. Moreover, the first outbreak identified as caused by the GII.2[P16] recombinant strain in Shenzhen was in September 30, 2016 after which the GII.2[P16] strain caused a steep rise in AGE in Shenzhen in the ensuing months. In general, recombination is thought to be important and common in virus evolution[35]. Most recombination occurs within ORF1/ORF2 overlapping regions or near the RdRp region, resulting in different capsid and RdRp genotypes[36]. In this study, we calculated the evolutionary rates of RdRp region and VP1 sequence, which were 2.1×10^{-3} substitutions/site/year and 2.7×10^{-3} substitutions/site/year, respectively, indicating that the polymerase and capsid regions of NoV GII.2[P16] strains had evolved independently, which was consistent with the results of previous studies [37]. The evolution rate of NoV GII.2 was much lower than that of GII.4 NoV (4.4–7.4 × 10^{-3} substitutions/site/year), it suggested GII.2 was was relatively stable in Shenzhen. Based on the evolutionary divergence time, the GII.2[P16] strains in Shenzhen might have recombined in 2013-2014, providing a better understanding of the formation of GII.2[P16] recombinant strains in Shenzhen.

The results of sequence alignment showed that important sites of VP1, including the HBGA-binding profile and epitopes predicted sites, have not mutated. This suggested that the reason for the prevalence of NoV GII.2[P16] strains in the population is different from that of the previous pandemic NoV GII.4, which was mainly due to changes in the capsid region leading to changes in blocking antibody epitopes to cause population among people [38, 39]. Parra et al [40] analyzed the GII.2 capsid sequences over a 40-year period and found only small differences, which agree with our results, indicating that the GII.2 strain is more genetically stable than is the GII.4 strain. At the same time, the lack of variation in the antigen regions of strains may also explain their short duration. These results indicate that the presence of a structure other than the VP1 contributes significantly to the prevalence of the GII.2[P16] strain[41], which may help to reveal the reasons of GII.P17/GII.17 epidemic that caused the outbreak of acute gastroenteritis in many countries in the winter of 2014-2015. Tohma et al [42] summarized the reasons for the epidemic caused by GIIP17/GII.17 and believed it to be related to the non-structural region. In this study, amino acid substitutions were found within the nonstructural regions including P48, NTPase, P22 and RdRp. These non-structural proteins played important roles in NoV replication, damaging host cells and promoting virus synthesis by interfering with intracellular protein transport, vesicle misorientation and Golgi disintegration [43-45]. The RdRp region can be divided into three highly conserved segments according to function and structure, including the fingers, thumb, and palm subdomains, which could be organized into Motifs A to G [46]. The results of amino acid mutation of non-structural protein sites of the GII.2[P16] recombinant strain suggest that the non-structural region may provide materials for virus replication, accelerate apoptosis in host cells and enhance fitness by changing the interaction mode.

Another study also reported that the GII.2[P16] strain leads to higher viral load than GII.Pe/GII.4 and GII.P17/GII.17 in patients [47]. However, not all changes in the non-structural region would cause epidemics. The study of Tohma et al calculated the amino acid substitution sites in the RdRp region of GII.P2/GII.2 and found that the replacement rate of GII.P2 was higher than that of GII.P16[40]. Regardless, NoV GII.P2/GII.2 outbreaks have not resulted in pandemics indicating that the RdRp region plays a crucial role in the GII.2[P16] epidemic.
This study showed that the GII.2[P16] outbreaks have decreased in Shenzhen, though the continuous surveillance to monitor genotypes is still necessary to identify new variants in a timely manner. The limitations of this study were as follows: First, genotyping was only successful for 150 (73.9%) of the positive NoV cases in our study. Second, our study lacked of clinical information and epidemiological data for outbreaks. In future studies, epidemiological surveillance should be more comprehensive and molecular analysis for different NoV genotypes should be developed.

**Conclusions**

In conclusion, this study reported the epidemiological patterns and genetic characteristics of NoV in Shenzhen from September 2015 through August 2018 and the main cause was GII.P16/GII.2 strain. This study also provided the evidence that the NoV GII.P16/GII.2 strain was static in Shenzhen.

**Abbreviations**

AGE  Acute gastroenteritis

NoV  Norovirus

RdRp  RNA-dependent RNA polymerase

HBGA  Histoblood group antigen

**Declarations**

**Competing interests**

The authors declare that they have no competing interests.

**Authors’ contributions**

JW, MJ, YQH and GFH designed the study. ZHL and ZYN participated in the Norovirus detection. HY, XJ Y, LC, JM participated in the Sequence analysis and phylogenetic analysis. ZJD coordinated in the design of the study.

All authors read and approved the final version of the article.

**Acknowledgments**

We are particularly grateful to the local districts CDC for epidemiological data and sample collection in this study.
Author details

1 Wuhan Wuchang Hospital, Wuchang Hospital Affiliated to Wuhan University of Science and Technology, Wuhan, 430063, China

2 NHC Key Laboratory of Medical Virology and Viral Diseases, National Institute for Viral Disease Control and Prevention, Chinese Center Control and Prevention, Beijing 102206, China

3 Shenzhen Center for Disease Control and Prevention, Shenzhen 518055, China

4 Department of Epidemiology, School of Public Health, Southern Medical University, Guangzhou 510515, China

Funding

Study on the interaction between molecular evolution and histoblood group antigen (HBGA) of GII.2 norovirus in Shenzhen. Scientific Research Project of Shenzhen Municipal Commission of Health and Family Planning in 2017 (SZGW2017017)

Study on the interaction between molecular evolution and histoblood group antigen (HBGA) of GII.2 norovirus in Shenzhen. Basic Research Project (Free-Exploration) of Shenzhen Science and Technology Innovation Committee(JCYJ20170306160421143)

Key Project of Science and Technology (2017ZX10104001-003)

China–US Collaborative Program on Emerging and Re-emerging Infectious Disease (Number 6)

Availability of data and materials

The datasets used in the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

This study was approved by the Ethics committee of Chinese Center Control and Prevention. Informed consent for the fecal specimens was obtained from the patients or their guardians.

Consent for publication

Not applicable.

Contributor Information

Jing Wang, Email:wangjing5561@126.com

Miao Jin, Email:jinmiao37@126.com
Hailong Zhang, Email: 22899360@qq.com
Yanan Zhu, Email: 1963805943@qq.com
Hong Yang, Email: 78641474@qq.com
Xiangjie Yao, Email: Yaoxiangjie26@qq.com
Long Chen, Email: chen_l_2011@163.com
Jun Meng, Email: 269380997@qq.com
Guifang Hu, Email: hgf@smu.edu.cn
Yaqing He, Email: heyaqing1019@126.com
Zhaojun Duan, Email: zhaojund@126.com

References

1. Glass RI, Parashar UD, Estes MK: Norovirus gastroenteritis. N Engl J Med 2009, 361(18):1776-1785.
2. Patel MM, Widdowson MA, Glass RI, Akazawa K, Vinje J, Parashar UD: Systematic literature review of role of noroviruses in sporadic gastroenteritis. EMERG INFECT DIS 2008, 14(8):1224-1231.
3. Hall AJ, Wikswo ME, Manikonda K, Roberts VA, Yoder JS, Gould LH: Acute gastroenteritis surveillance through the National Outbreak Reporting System, United States. EMERG INFECT DIS 2013, 19(8):1305-1309.
4. Chhabra P, de Graaf M, Parra GI, Chan MC, Green K, Martella V, Wang Q, White PA, Katayama K, Vennema H et al: Updated classification of norovirus genogroups and genotypes. J GEN VIROL 2019, 100(10):1393-1406.
5. Hardy ME, Estes MK: Completion of the Norwalk virus genome sequence. VIRUS GENES 1996, 12(3):287-290.
6. Belliot G, Sosnovtsev SV, Mitra T, Hammer C, Garfield M, Green KY: In vitro proteolytic processing of the MD145 norovirus ORF1 nonstructural polyprotein yields stable precursors and products similar to those detected in calicivirus-infected cells. J VIROL 2003, 77(20):10957-10974.
7. Debbink K, Lindesmith LC, Donaldson EF, Baric RS: Norovirus immunity and the great escape. PLOS PATHOG 2012, 8(10):e1002921.
8. Hardy ME: Norovirus protein structure and function. FEMS MICROBIOL LETT 2005, 253(1):1-8.
9. Bu W, Mamedova A, Tan M, Xia M, Jiang X, Hegde RS: Structural basis for the receptor binding specificity of Norwalk virus. J VIROL 2008, 82(11):5340-5347.
10. Cao S, Lou Z, Tan M, Chen Y, Liu Y, Zhang Z, Zhang XC, Jiang X, Li X, Rao Z: Structural basis for the recognition of blood group trisaccharides by norovirus. J VIROL 2007, 81(11):5949-5957.

11. Glass PJ, White LJ, Ball JM, Leparc-Goffart I, Hardy ME, Estes MK: Norwalk virus open reading frame 3 encodes a minor structural protein. J VIROL 2000, 74(14):6581-6591.

12. Qiao N, Wang XY, Liu L: Temporal Evolutionary Dynamics of Norovirus GII.4 Variants in China between 2004 and 2015. PLOS ONE 2016, 11(9):e163166.

13. Lu J, Sun L, Fang L, Yang F, Mo Y, Lao J, Zheng H, Tan X, Lin H, Rutherford S et al: Gastroenteritis Outbreaks Caused by Norovirus GII.17, Guangdong Province, China, 2014-2015. EMERG INFECT DIS 2015, 21(7):1240-1242.

14. Ao Y, Cong X, Jin M, Sun X, Wei X, Wang J, Zhang Q, Song J, Yu J, Cui J et al: Genetic Analysis of Reemerging GII.P16-GII.2 Noroviruses in 2016-2017 in China. J INFECT DIS 2018, 218(1):133-143.

15. Nagasawa K, Matsushima Y, Motoya T, Mizukoshi F, Ueki Y, Sakon N, Murakami K, Shimizu T, Okabe N, Nagata N et al: Genetic Analysis of Human Norovirus Strains in Japan in 2016-2017. FRONT MICROBIOL 2018, 9:1.

16. Medici MC, Tummolo F, Martella V, De Conto F, Arcangeletti MC, Pinardi F, Ferraglia F, Chezzi C, Calderaro A: Emergence of novel recombinant GII.P16_GII.2 and GII. P16_GII.4 Sydney 2012 norovirus strains in Italy, winter 2016/2017. NEW MICROBIOL 2018, 41(1):71-72.

17. Niendorf S, Jacobsen S, Faber M, Eis-Hubinger AM, Hofmann J, Zimmermann O, Hohne M, Bock CT: Steep rise in norovirus cases and emergence of a new recombinant strain GII.P16-GII.2, Germany, winter 2016. Euro Surveill 2017, 22(4).

18. Trujillo AA, McCaustland KA, Zheng DP, Hadley LA, Vaughn G, Adams SM, Ando T, Glass RI, Monroe SS: Use of TaqMan real-time reverse transcription-PCR for rapid detection, quantification, and typing of norovirus. J CLIN MICROBIOL 2006, 44(4):1405-1412.

19. Kojima S, Kageyama T, Fukushi S, Hoshino FB, Shinhara M, Uchida K, Natori K, Takeda N, Katayama K: Genogroup-specific PCR primers for detection of Norwalk-like viruses. J VIROL METHODS 2002, 100(1-2):107-114.

20. Richards GP, Watson MA, Fankhauser RL, Monroe SS: Genogroup I and II noroviruses detected in stool samples by real-time reverse transcription-PCR using highly degenerate universal primers. Appl Environ Microbiol 2004, 70(12):7179-7184.

21. Kroneman A, Vennema H, Deforce K, V DAH, Penaranda S, Oberste MS, Vinje J, Koopmans M: An automated genotyping tool for enteroviruses and noroviruses. J CLIN VIROL 2011, 51(2):121-125.

22. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkami 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. J VIROL 1999, 73(1):152-160.

23. Schwarz G. Estimating the dimension of a model. Annals of Statistics 1978,6:461–464.

24. Kobayashi M, Matsushima Y, Motoya T, Sakon N, Shigemoto N, Okamoto-Nakagawa R, Nishimura K, Yamashita Y, Kuroda M, Saruki N et al: Molecular evolution of the capsid gene in human norovirus genogroup II. Sci Rep 2016, 6:29400.
25. Lindesmith LC, Beltramello M, Donaldson EF, Corti D, Swanstrom J, Debbink K, Lanzavecchia A, Baric RS: Immunogenetic mechanisms driving norovirus GII.4 antigenic variation. *PLOS PATHOG* 2012, 8(5):e1002705.

26. Bu W, Mamedova A, Tan M, Xia M, Jiang X, Hegde RS: Structural basis for the receptor binding specificity of Norwalk virus. *J VIROL* 2008, 82(11):5340-5347.

27. Cubitt WD, McSwiggan DA, Moore W: Winter vomiting disease caused by calicivirus. *J CLIN PATHOL* 1979, 32(8):786-793.

28. Colas DLNA, Estienney M, Aho S, Perrier-Cornet JM, de Rougemont A, Pothier P, Gervais P, Belliot G: Absolute Humidity Influences the Seasonal Persistence and Infectivity of Human Norovirus. *Appl Environ Microbiol* 2014, 80(23):7196-7205.

29. Lopman B, Armstrong B, Atchison C, Gray JJ: Host, weather and virological factors drive norovirus epidemiology: time-series analysis of laboratory surveillance data in England and Wales. *PLOS ONE* 2009, 4(8):e6671.

30. Kauppinen A, Miettinen IT: Persistence of Norovirus GII Genome in Drinking Water and Wastewater at Different Temperatures. *Pathogens* 2017, 6(4).

31. Vega E, Barclay L, Gregoricus N, Shirley SH, Lee D, Vinje J: Genotypic and epidemiologic trends of norovirus outbreaks in the United States, 2009 to 2013. *J CLIN MICROBIOL* 2014, 52(1):147-155.

32. Qin SW, Chan TC, Cai J, Zhao N, Miao ZP, Chen YJ, Liu SL: Genotypic and Epidemiological Trends of Acute Gastroenteritis Associated with Noroviruses in China from 2006 to 2016. *Int J Environ Res Public Health* 2017, 14(11).

33. Han J, Wu X, Chen L, Fu Y, Xu D, Zhang P, Ji L: Emergence of norovirus GII.P16-GII.2 strains in patients with acute gastroenteritis in Huzhou, China, 2016-2017. *BMC INFECT DIS* 2018, 18(1):342.

34. Fu JG, Shi C, Xu C, Lin Q, Zhang J, Yi QH, Zhang J, Bao CJ, Huo X, Zhu YF et al: Outbreaks of acute gastroenteritis associated with a re-emerging GII.P16-GII.2 norovirus in the spring of 2017 in Jiangsu, China. *PLOS ONE* 2017, 12(12):e186090.

35. Bull RA, Tanaka MM, White PA: Norovirus recombination. *J GEN VIROL* 2007, 88(Pt 12):3347-3359.

36. Bull RA, Hansman GS, Clancy LE, Tanaka MM, Rawlinson WD, White PA: Norovirus recombination in ORF1/ORF2 overlap. *EMERG INFECT DIS* 2005, 11(7):1079-1085.

37. Nagasawa K, Matsushima Y, Motoya T, Mizukoshi F, Ueki Y, Sakon N, Murakami K, Shimizu T, Okabe N, Nagata N et al: Genetic Analysis of Human Norovirus Strains in Japan in 2016-2017. *FRONT MICROBIOL* 2018, 9:1.

38. Bruggink LD, Moselen JM, Roberts JA, Marshall JA: Evolutionary changes in the capsid P2 region of Australian strains of the norovirus GII.Pe_GII.4. *J MED MICROBIOL* 2017, 66(7):1014-1022.

39. van Beek J, Ambert-Balay K, Botteldoorn N, Eden JS, Fonager J, Hewitt J, Iritani N, Kroneman A, Vennema H, Vinje J et al: Indications for worldwide increased norovirus activity associated with emergence of a new variant of genotype II.4, late 2012. *Euro Surveill* 2013, 18(1):8-9.
40. Parra GI, Squires RB, Karangwa CK, Johnson JA, Lepore CJ, Sosnovtsev SV, Green KY: Static and Evolving Norovirus Genotypes: Implications for Epidemiology and Immunity. *PLOS PATHOG* 2017, 13(1):e1006136.

41. Tohma K, Lepore CJ, Ford-Siltz LA, Parra GI: Phylogenetic Analyses Suggest that Factors Other Than the Capsid Protein Play a Role in the Epidemic Potential of GII.2 Norovirus. *MSPHERE* 2017, 2(3).

42. Tohma K, Lepore CJ, Ford-Siltz LA, Parra GI: Evolutionary dynamics of non-GII genotype 4 (GII.4) noroviruses reveal limited and independent diversification of variants. *J GEN VIROL* 2018, 99(8):1027-1035.

43. Li TF, Hosmillo M, Schwanke H, Shu T, Wang Z, Yin L, Curry S, Goodfellow IG, Zhou X: Human Norovirus NS3 Has RNA Helicase and Chaperoning Activities. *J VIROL* 2018, 92(5).

44. Sharp TM, Crawford SE, Ajami NJ,Neill FH, Atmar RL, Katayama K, Utama B, Estes MK: Secretory pathway antagonism by calicivirus homologues of Norwalk virus nonstructural protein p22 is restricted to noroviruses. *VIROL J* 2012, 9:181.

45. Hardy ME, Crone TJ, Brower JE, Ettayebi K: Substrate specificity of the Norwalk virus 3C-like proteinase. *VIRUS RES* 2002, 89(1):29-39.

46. Deval J, Jin Z, Chuang YC, Kao CC: Structure(s), function(s), and inhibition of the RNA-dependent RNA polymerase of noroviruses. *VIRUS RES* 2017, 234:21-33.

47. Cheung S, Kwok K, Zhang LY, Mohammad KN, Lui G, Lee N, Nelson E, Lai R, Leung TF, Chan P et al.: Higher Viral Load of Emerging Norovirus GII.P16-GII.2 than Pandemic GII.4 and Epidemic GII.17, Hong Kong, China. *EMERG INFECT DIS* 2019, 25(1):119-122.

**Tables**

**Table 1** Number of people with NoV gastroenteritis per outbreaks according to genotype
| Genotype          | No. of ill people\(\text{range}\) | No. of outbreaks involved ill people | Median no. of ill people\(\text{range}\) |
|-------------------|------------------------------------|-------------------------------------|-----------------------------------------|
| GII.2             | 3-73                               | 92                                  | 10                                      |
| GII.3             | 3-45                               | 13                                  | 7.5                                     |
| GII.4 Sydney 2012 | 9-14                               | 7                                   | 11                                      |
| GII.4             | 3-13                               | 3                                   | 11                                      |
| GII.6             | 4-8                                | 8                                   | 6                                       |
| GII.8             | 11-13                              | 2                                   | 12                                      |
| GII.17            | 3-12                               | 8                                   | 5.5                                     |
| GII.13            | 14                                 | 1                                   | -                                       |
| GII.21            | 11-23                              | 2                                   | 17                                      |
| Multiple genotype | 3-10                               | 2                                   | 6.5                                     |
| GI.1              | 7                                  | 1                                   | -                                       |
| GI.2              | 5-9                                | 3                                   | 7                                       |
| GI.3              | 5-84                               | 3                                   | 6                                       |
| GI.5              | 3-5                                | 2                                   | 4                                       |
| GI.6              | 6-64                               | 3                                   | 12                                      |
| GII               | 3-115                              | 53                                  | 9                                       |
| Total             | 3-115                              | 203                                 | 9                                       |

**Table 2** Genotype distribution of norovirus outbreaks from 2015 to 2018 according to setting
| Institute | Total No.(%) of outbreaks | No.(%) of outbreaks |
|-----------|-------------------------|---------------------|
|           | child care centers | primary school | middle school | university hospital | hotel | restaurant company | Nine-year school | Fifteen-year school | Institution Community |
| GII.2     | 92(45.3%) | 73(79.3%) | 7(7.6%) | 4(4.3%) | 10(1.1%) | 10(1.1%) | - | - | - | - |
| GII.3     | 13(6.4%) | 11(84.6%) | 1(7.7%) | - | - | - | - | - | 1(7.7%) | - |
| GII.4Sydney 2012 | 7(3.4%) | 6(85.7%) | 1(14.3%) | - | - | - | - | - | - |
| GII.4     | 3(1.5%) | 3(100.0%) | - | - | - | - | - | - | - |
| GII.6     | 8(3.9%) | 5(62.5%) | 2(25.0%) | - | - | - | - | - | 1(12.5%) | - |
| GII.8     | 2(1.0%) | 1(50.0%) | 1(50.0%) | - | - | - | - | - | - |
| GII.17    | 8(3.9%) | 2(25.0%) | 1(12.5%) | - | - | - | 1(12.5%) | 1(12.5%) | 2(25.0%) | - |
| GII.13    | 1(1.0%) | 1(100.0%) | - | - | - | - | - | - |
| GII.21    | 2(1.0%) | - | - | - | - | 1(50.0%) | - | - | - | 1(50.0%) |
| Multiple genotype | 2(1.0%) | - | - | 1(50.0%) | - | 1(50.0%) | - | - | - |
| GI.1      | 1(1.0%) | 1(100.0%) | - | - | - | - | - | - |
| GI.2      | 3(1.5%) | 1(33.3%) | 1(33.3%) | - | - | - | - | 1(33.3%) | - |
| GI.3      | 3(1.5%) | 1(50.0%) | 1(43.3%) | 1(14.3%) | - | - | - | - | - |
| GI.5      | 2(1.0%) | 2(100.0%) | - | - | - | - | - | - |
| GI.6      | 3(1.5%) | 2(66.7%) | - | - | - | 1(33.3%) | - | - |
| GII       | 53(26.1%) | 34(64.2%) | 8(15.1%) | 2(3.8%) | 2(3.8%) | 1(1.9%) | 1(1.9%) | 1(1.9%) | 1(1.9%) | 1(1.9%) |
| Total     | 203(100) | 143(70.4%) | 23(11.3%) | 7(3.4%) | 4(2.0%) | 2(1.0%) | 4(2.0%) | 2(1.0%) | 5(2.5%) | 8(3.9%) | 1(0.5%) | 3(1.5%) | 1(0.5%) |

Note: Nine-year school indicated schools which included the primary school and middle school (especially junior high school). And the Fifteen-year school indicated schools which included occurred in schools which include primary school and middle school (junior high school and senior high school). Institution indicated units that belong to government.

**Table 3** Genotype distribution of identified NoV strains in Shenzhen, September 2015-August 2018
| Genotype | 2015.09-2016.08 | 2016.09-2017.08 | 2017.09-2018.08 |
|----------|----------------|----------------|----------------|
| Number(percentage) | Number(percentage) | Number(percentage) |
| Capsid | | | |
| GI.3 | 1(4.8) | 1(0.9) | - |
| GII.3 | 4(19.0) | 1(0.9) | - |
| GII.4 | - | 1(0.9) | 1(1.4) |
| GII.4 Sydney2012 | 4(19.0) | 3(2.7) | - |
| GII.6 | - | 1(0.9) | 1(1.4) |
| GII.8 | - | 1(0.9) | - |
| GII.17 | 3(14.3) | - | - |
| GII.21 | 2(9.5) | - | - |
| RdRp/Capsid | | | |
| GL.P1/GL.1 | - | 1(0.9) | - |
| GL.Pb/GL.6 | - | - | 3(4.3) |
| GL.P2/GL.2 | - | - | 3(4.3) |
| GL.Pd/GL.3 | - | - | 1(1.4) |
| GL.P5/GL.5 | - | - | 2(2.8) |
| GII.P16/GII.2 | - | 73(65.2) | 19(27.1) |
| GII.P12/GII.3 | - | 2(1.7) | 6(8.6) |
| GII.Pe/GII.17 | - | 1(0.9) | - |
| GII.P17/GII.17 | - | 2(1.7) | 2(2.8) |
| GII.P7/GII.6 | - | 1(0.9) | 5(7.1) |
| GII.Pe/GII.4 | - | - | 1(1.4) |
| GII.P8/GII.8 | - | - | 1(1.4) |
| GII.P21/GII.13 | - | - | 1(1.4) |
| Multiple genotype | 1(4.8) | - | 1(1.4) |
| GII | 6(28.6) | 24(21.4) | 23(32.9) |
| Total | 21(100) | 112(100) | 70(100) |

**Table 4** Nonstructural polymerase of GII.16/GII.2 recombination strain

| Variants | P48 | NTpase | P22 | RdRp |
|----------|-----|--------|-----|------|
| GII.17(2002) | * | A | K | S | P | H | V | K | T |
| GII.2/2009-2014 | * | A | K | S | P | H | V | K | T |
| GII.2/2010-2012 | * | S | K | I | P | H | V | K | T |
| GII.3(2012-2013) | * | P/S | K | V/I | P | H | V | K | T |
| GII.13(2015) | * | S | K | T | P | H | V | K | T |
| GII.4/2015-2016 | E | P | S | R | T | Q | H/Y | V/I | Q | A |
| GII.2(2016-2018) | E | P | R | T | Q | H/Y | I | Q | A |

**Additional File**

Additional file 1: Table S1.

Primes used in this study
HBGA-Binding profile and epitopes predicted of Reemerging NoV GII.P16/GII.2 were compared with previous GII.2 strain.

Reference sequences used in this study

**Figures**

**Figure 1**

Geographical distribution of NoV outbreaks in Shenzhen from September 2015 to August 2018.
Figure 2

Monthly distribution of NoV outbreaks in Shenzhen by genotype. Note: The numbers on the Y axis represent the number of NoV outbreaks.

Figure 3

The SimPlot analysis of the GII.2[P16] strain sequences Note: The reference strain sequences were GII.2[P16] (LC209448) and GII.2[P16] (MG746027.1). LC209448 was collected in 2011 and MG746027.1 was confirmed as a recombinant GII.2[P16] strain collected in
March 2017. At each position of the window, the query sequence was compared to each of the reference strains. The X-axis indicates the nucleotide positions in the multiple alignments of the NoV sequences; and the Y-axis indicates nucleotide identities (%) between the query sequence and the NoV reference strains.

Figure 4

4A Phylogenetic tree of the RdRp region of NoV GII.2[P16]. 4B Phylogenetic tree for the VP1 sequence of NoV GII.2[P16]. Note: The scale bars denote the actual time (years). Estimated divergence times are shown on the ancestral nodes. Phylogenetic clusters, including the previous GII.P16/GII.2 2010-2012 cluster, the previous GII.P16/GII.2 2008-2015 cluster and the GII.P16/GII.2 2016-2018 cluster, were marked. The sequences of the GII.P16/GII.2 2016-2018 cluster were all collected from Shenzhen.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- materialusedinthisstduy.rar