Infection, Genetics and Evolution 53 (2017) 199–205

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research paper

Variation analysis of norovirus among children with diarrhea in rural Hebei Province, north of China

Niu Qiao [a,f,1], Song-Mei Wang [b,1], Jin-Xia Wang [c,d,1], Bin Kang [a,f], Shan-Shan Zhen [a,f], Xin-Jiang Zhang [e], Zhi-Yong Hao [e], Jing-Chen Ma [d], Chao Qiu [a,f], Yu-Liang Zhao [d], Lei Liu [a,f,*], Xuan-Yi Wang [a,f,**]

[a] Key Laboratory of Medical Molecular Virology of MoE & MoH, Fudan University, Shanghai 200032, People's Republic of China
[b] Laboratory of Molecular Biology, Training Center of Medical Experiments, School of Basic Medical Sciences, Fudan University, Shanghai 200032, People's Republic of China
[c] College of Public Health, North China University of Science and Technology, Tangshan 063009, People's Republic of China
[d] Hebei Province Center for Disease Control and Prevention, Shijiazhuang 050021, People's Republic of China
[e] Zhengding County Center for Disease Control and Prevention, Zhengding 050800, People's Republic of China
[f] Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, People's Republic of China

ARTICLE INFO

Article history:
Received 12 March 2017
Received in revised form 3 June 2017
Accepted 6 June 2017
Available online 7 June 2017

Keywords:
Norovirus
Genotyping
Phylogenetic tree
Haplotype network
Evolutionary rates
China

ABSTRACT

To understand the distribution of genotyping, as well as evolution of norovirus circulating among children < 5 yrs., a population-based diarrhea surveillance targeted children < 5 yrs. was conducted in rural Zhengding County, Hebei Province, China between October 2011 and March 2012. RT-PCR was used to amplify the capsid-encoding region of GI and GII norovirus to identify norovirus infection. All PCR products were sequenced and analyzed for genotyping and constructing phylogenetic tree. Dynamic distribution network was constructed by TempNet to illustrate the genetic relationships at two different time points. Bayesian evolutionary inference techniques were applied by BEAST software to study the norovirus evolution rate. During the 6-month surveillance period, 1091 episodes of diarrhea were reported from 5633 children under 5 years of age lived in catchment area. 115 of 1091 stool specimens were detected as norovirus positive (10.54%). Five genotypes based on capsid gene sequences were identified, including GII.2 (11), GII.3 (52), GII.4 (47), GII.6 (4) and GII.7 (1). An identical haplotype of GII.4 circulated between 2006 and 2011 in Hebei Province. A mean rate of 6.29×10^{-2} nucleotide substitutions/site/year (s/s/y) was obtained for GII.3 viruses in Hebei, while the GII.4 viruses evolved at a mean rate of 3.67×10^{-2} s/s/y. In conclusions, GII.3 (45.22%) and GII. 4(40.87%) are the predominant strain in Hebei Province in the winter season of 2011 and 2012. Different from the current consensus, our study shows that GII.3 noroviruses in Hebei Province evolved at a faster rate than GII.4 viruses.

1. Introduction

Acute gastroenteritis is a major public health problem among children worldwide, and noroviruses (NoVs) is recognized as the major cause of viral gastroenteritis among all age groups (Ahmed et al., 2014). NoVs are genetically classified into seven genogroups (GI-GVII) that further subdivided into genotypes. GI and GII NoVs are the genogroups primarily responsible for human illness, although GIV has also been detected in humans (Ao et al., 2014; Eden et al., 2014; Kroneman et al., 2013; Martella et al., 2013; Baehner et al., 2016). Based on viral protein 1 (VP1) sequences, the GI and GII genogroups can be further divided into 9 and 22 genotypes, respectively. In addition, GIV.1 from GIV group is found also infected humans (Zheng et al., 2006; Mesquita et al., 2010). Despite the broad genetic diversity, the majority of acute gastroenteritis outbreaks due to NoVs infection is caused by GII.4 NoVs (Siebenga et al., 2009; Lindesmith et al., 2008), while GII.3 NoVs are one of the most common genotypes associated with sporadic NoV infection, particularly in children, where they often are identified as the dominant genotype (Siebenga et al., 2009; Boon et al., 2011; Barreira et al., 2010). It is generally believed that sequence mutations and homologous genome recombination are the two main mechanisms for the current NoV variations (Bull and White, 2011).

NoVs are highly genetically and antigenically diverse and their epidemiology and transmission patterns are similar to that of influenza viruses. Epidemiologic studies have shown that GII.4 NoVs has a rapid local transmission and that novel epidemic strains emerged every 2 to 3 years and spread globally in months (Siebenga et al., 2009; Lindesmith et al., 2008). The surface-exposed host ligand binding site on the NoV capsid is under heavy immune selection and likely evolves

* Corresponding author at: Institutes of Biomedical Sciences, Shanghai Medical College, Fudan University, 138 Yi Xue Yuan Rd., Shanghai, People's Republic of China.
** Correspondence to: X.-Y. Wang Key Laboratory of Medical Molecular Virology of MoE & MoH, Shanghai Medical College, Fudan University, 138 Yi Xue Yuan Rd., Shanghai, People's Republic of China.
E-mail addresses: liulei@fudan.edu.cn (L. Liu), xywang@shmu.edu.cn (X.-Y. Wang).
[1] Niu Qiao, Song-Mei Wang, and Jin-Xia Wang contributed equally.

http://dx.doi.org/10.1016/j.meegid.2017.06.007
1567-1348/

by antigenic drift in the face of human herd immunity (Lindesmith et al., 2008; Bull and White, 2011). Fast mutations, like genetic drift, in the surface proteins allow virus to elude host immunity, resulting in an ineffective immune protection produced by previous infection. As influenza viruses, vaccines could be targeted to protect against NoV infections (Lindesmith et al., 2008). A thorough understanding of the evolutionary pattern, evolutionary rate, genetic diversity, and epidemic cycle of NoVs may help interpret how these viruses change, evade the host immune response, and adapt to the host environment.

NoV genome consists of three open reading frames (ORFs) (Glass et al., 2009). Current classification of NoV genotypes is based on the genetic diversity within the polymerase (regions A and B) and the major capsid (VP1; regions C, D, and E) gene (Eden et al., 2014; Martella et al., 2013; Ando et al., 2000; Kirsten et al., 2009) Based on overall performance, the region C is recommended for routine genotyping of NoVs, while the region D may be useful for identifying new GII.4 variants (Kirsten et al., 2009).

In this article, the genetic characterization and phylogenetic analyses of the NoVs detected in Hebei Province were conducted to determine their genogroup and genotypes and constructed the corresponding phylogenetic tree. The genetic relationships of these NoVs in Hebei Province at two different time points were also illustrated by constructing a haplotype network, which is a useful tool for understanding the number, the relative frequency and dissimilarity between haplotypes within a single population. Furthermore, Bayesian genealogical inference of time-measured trees was performed using Markov chain Monte Carlo (MCMC (Green, 1995)) sampling to calculate the evolutionary rates of the dominant NoV strains.

2. Material and methods

2.1. Study population and surveillance

In total, 34 villages located in five townships in Zhengding County, Hebei Province, China, were selected as the catchment area for the population- and health care facility-based viral diarrhea surveillance targeted children < 5 years of age. All health-care providers that offered health care for the children in the catchment area constitute the surveillance system, including 101 village clinics, 5 township hospitals, and 1 county hospital. The surveillance was conducted from October 1, 2011 through March 31, 2012, the peak season for viral diarrheal illness in children. Bulk stool samples were collected from children with diarrhea by health care providers during their visits to the hospitals or clinics. The total number of children who contributed to the cohort were 6441. Of these, 1211 diarrhea episodes were reported, and 1091 (90.1%) children provided stool samples for NoV test through surveillance system, resulting in a diarrhea incidence rate of 188.0/1000 person/year (Zhen et al., 2015). This study was reviewed and approved by the Institutional Review Board of Hebei Center for Disease Control and Prevention. Written inform consent was obtained from the parent/guardian of each child. The study was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

2.2. Laboratory diagnosis of NoVs

Stool samples were prepared as 20% fecal suspensions, and total viral RNA was extracted using the immunomagnetic virus DNA/RNA extraction kit (Xi'an Tianlong Science & Technology Co. Ltd), and reverse transcribed into cDNA using GoScript Reverse Transcription System A5001 (Promega Corporation) according to the manufacturer's instructions.

GI and GII NoVs were detected in separate reactions by conventional Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) using consensus primers, G1SKF/G1SKR and COG2F/G2SKR (Kojima et al., 2002) for both GI and GII NoVs, respectively. These primers amplify a region spanning from position 5342 to 5671 (330 bp) of the GI NoV genome (Norwalk/68, GenBank accession No. M87661) containing 17 bp of 3′ end ORF1 and 313 bp of 5′ end ORF2, or spanning from position 5003 to 5389 (387 bp) of GII NoV genome (Lordsdale/93, GenBank accession No. X86557) containing 83 bp of 3′ end ORF1 and 304 bp of 5′ end ORF2. The GoTaqDNA polymerase (including 5× buffer) was purchased from Promega Company. Sanger dideoxy termination sequencing was applied based on the amplified PCR products representing corresponding norovirus genomic fragments. The sequencing was carried out by the Biosune Co., Ltd. in Shanghai, China. All obtained sequences were in high quality. Manual editing was performed to delete indels that occurred in >50% of the sequences, to allow better and reliable outcomes of the sequence comparison.

2.3. Reference sequences collection

A total of 25 NoV sequences that were originally isolated in Hebei Province, China, are available from GenBank Database. 24 of them that were isolated from 2006 to 2007 were downloaded for the analysis of the temporal network. In addition, 11 GII NoV sequences were obtained from the Database of the Netherlands National Institute for Public Health and the Environment (RIVM) to construct phylogenetic tree. The detailed accession numbers of these reference sequences download from GeneBank was shown in the Supplemental materials (Table S1).

2.4. Sequence analysis

The nucleotide sequences obtained from our study were firstly genotyped by NCBI BLAST (Pevsner, 2013) tool. The resulting NoV genotypes were then confirmed by the online NoV genotyping tool (http://www.rivm.nl/mpf/norovirus/typingtool) (Kroneman et al., 2011) based on ORF2 sequences offered by RIVM.

The nucleotide sequences were aligned using the multiple sequence alignment program MUSCLE (Edgar, 2004) implemented by Molecular Evolutionary Genetics Analysis (MEGA) software version 6.06 (Tamura et al., 2013). Manual editing was carried out to delete indels occurring in >50% of the sequences. Some of the sequences were truncated at both the 5′ and 3′ ends, possibly confounding the phylogeny of local clades.

2.5. Recombination analysis

The Genetic Algorithms for Recombination Detection (GARD (Kosakovsky Pond et al., 2006)) and the Single Break-Point (SBP (Kosakovsky Pond, 2006)) methods from Datamonkey Web Site (Delport et al., 2010; Pond and Frost, 2005) were utilized to determine if recombination occurred within the capsid protein-encoding sequences, which can therefore be used to determine if recombination was a factor in the evolution of the capsid protein-encoding sequences.

2.6. Phylogenetic analyses

To infer the evolutionary relationships among the NoV strains, Maximum Likelihood (ML) phylogenetic trees were constructed by using MEGA software v6.06 (Tamura et al., 2013; Hall, 2013). A separate Hasegawa-Kishino-Yano (HKY (Hasegawa et al., 1985)) substitution model with gamma-distributed rate variation among sites (Beth et al., 2006) and the complete deletion for gaps/missing data treatment (Hall, 2013) were applied. In the search process for the optimal evolutionary tree, Neighbor-Joining (NJ) tree as initial tree with Nearest-Neighbor Interchange (NNI) for ML heuristic method was chosen. Besides, a ML bootstrap analysis (1000 replicates) was used to evaluate the robustness of the phylogenetic grouping.

2.7. Temporal network analyses

To compare the genetic distributions of the NoVs isolated in Hebei Province, a haplotype network was constructed at two different time

points by TempNet (Prost and Anderson, 2011), which is written for the open-source statistical environment R (http://www.r-project.org/). The input included 24 Hebei NoV sequences identified from 2006 to 2007 and 115 sequences obtained in this study. The DNA sequences in standard fasta alignment format after quality control were imported, and each sequence was assigned to the corresponding time layer. After customizing the relative size of circles corresponding to the haplotypes both present and absent in the layer and assigning a scale length corresponding to one mutation on the links, a two-dimensional temporal network was produced.

2.8. Bayesian evolutionary analyses

In an effort to estimate the nucleotide substitution rate of the dominant NoVs strains, the Bayesian genealogical inference of time-measured trees was performed using Markov Chain Monte Carlo (MCMC (Green, 1995)) sampling implemented in the Bayesian Evolutionary Analysis Sampling Trees (BEAST) software v1.8.2 (Drummond et al., 2012) combined with BEAGLE (Ayres et al., 2011) package, which can improve the calculation performance.

Firstly, the evolutionary model and options for the MCMC (Green, 1995) analysis was set using the Bayesian Evolutionary Analysis Utility (BEAUTi) (Drummond et al., 2012) application, which is a part of the BEAST package. A separate Hasegawa-Kishino-Yano (HKY) (Hasegawa et al., 1985) substitution model with gamma-distributed rate variations among sites was applied. We used a strict molecular clock model to assume a constant rate of evolution across the tree and specified a flexible Bayesian Skyride coalescent (Minin et al., 2008) tree prior. Then, the prior for clock rate parameter to a gamma distribution with shape = 0.001 and scale = 1000 was set. The graphical representation of this prior distribution indicates that most prior mass was put on small values, but the density remains sufficiently diffuse. Lastly, MCMC analyses for 100 million generations, sampling every 1000th generation and removed 10% as chain burn-in were ran. A graphical tool, named Tracer version 1.6 (Rambaut et al., 2014), was used for visualization and diagnostics of MCMC output, making sure that effective sample sizes (ESS) for the continuous parameters were >200.

3. Results

3.1. The temporal distribution of NoVs genotypes

Over the study period, 115 NoV positive stool samples were identified from 1091 individuals (10.54%). Genotyping of the NoV isolates demonstrated that the predominant NoV genotypes were GII.3 (52; 45.22%) followed by GII.4 (47; 40.87%) (Table 1). In addition, 24 NoV sequences that were submitted to GeneBank previously between 2006 and 2007 were found to be GI.2, GII.3 and GII.4, GII.6 and GII.14, respectively. This genotyping was based on genetically diverse regions C in the capsid protein-encoding gene, which is recommended for routine NoV genotyping. No evidence of recombination was found in both GII.3 (52) and GII.4 (47) NoVs.

Table 1
Genotypes distribution of NoVs in Hebei Province.

Genotype	2011			2012				Constituent (% constituent ratio)
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	
GII.2	5	4	1	0	0	0	1	11 (9.57)
GII.3	32	14	2	1	2	1	0	52 (45.22)
GII.4	26	10	6	4	0	1	0	47 (40.87)
GII.6	0	0	1	0	0	3	0	4 (3.48)
GII.7	1	0	0	0	0	0	0	1 (0.87)
Total	64	28	10	5	2	5	1	115

3.2. The phylogenetic tree of NoVs in Hebei Province

For the phylogenetic analyses, two independent Maximum Likelihood (ML) phylogenetic trees were constructed using the model described above for two different data sets: (1) 11 reference sequences from RIVM and 115 NoV sequences obtained from our surveillance study, and (2) 24 NoV sequences isolated from 2006 to 2007 in Hebei Province and 115 NoV sequences from our surveillance study. The HKY + G model of nucleotide substitution was used and the results were shown in Fig. 1 [the 115 NoV sequences from surveillance study with township information was provided in the Supplemental materials (Fig. S1)] and 2. From Fig. 1 the relatively large evolutionary distance between reference sequences and our sequences was seen, since they located in two separate branches. Only 5 GII.4, 2 GII.3 and 1 GII.2 were located in the branch of the reference sequences, and 10 GII.2, 50 GII.3, 42 GII.4, 4 GII.6 and 1 GII.7 were located in another branch. Similar results were found in Fig. 2, where the tree was also made up of two major branches. The first branch (red) mainly comprised 10 GII.2, 50 GII.3, 42 GII.4, 4 GII.6 and 1 GII.7 NoVs from our study. Another branch (cyan) was mainly composed of 4 GI.2, 6 GII.3 and 14 GII.4 NoVs downloaded from GenBank that were identified previously in Hebei Province from 2006 to 2007 and 5 GII.4, 2 GII.3 and 1 GII.2 NoVs that were collected from 2011 to 2012 in rural Zhengding in Hebei Province. Therefore, the data shown in both Figs. 1 and 2 are consistent each other.

3.3. The temporal network of NoVs in Hebei Province at two different time points

The 24 NoV sequences from GenBank were used along with the 115 NoV sequences obtained from our study to establish relationships among NoVs in Hebei Province (Fig. 3). The network has two layers, the upper layer (cyan) are composed of NoV sequences from 2006 to 2007; the lower layer (red) consists of sequences from our study from 2011 to 2012. Haplotypes are represented by scaled ellipses, so that their areas correspond to the numbers of sequences. The result showed that an identical haplotype existed in GII.4 and no other genotypes overlap at the two time points.

3.4. The nucleotide substitution rates of GII.3 and GII.4 NoVs in Hebei Province

To estimate the evolutionary rates of the dominant NoV strains in Hebei Province, that is, GII.3 and GII.4, a Bayesian coalescent method was used to infer the rates of evolutionary change expressed as nucleotide substitutions per site per year, using a strict-clock model implemented by BEAST software. The most conservative clock, the strict model, estimated that the VP1 gene of GII.3 NoVs evolved at a rate of 6.29×10^{-2} nucleotide substitutions/site/year (HPD95% = 2.238×10^{-2}–1.043×10^{-1}). The same Bayesian approach was used to estimate the evolutionary rate of GII.4 NoVs using the same clock model. The mean nucleotide substitution rate (nucleotide substitutions/site/year) of the VP1 gene of GII.4 was 3.67×10^{-2} with 95% HPD range from to 1.65×10^{-2}–5.78×10^{-2}. The effective sample sizes for GII.3 and GII.4 were 403 and 656, respectively.

4. Discussion

NoVs are genetically and antigenically diverse, evolving in pattern similar to influenza virus, therefore, its epidemiology outcome is often referred as “Gastric Flu” (Lopman et al., 2008). The more prevalent GII.4 strain had a 5 to 36-fold higher mutation rate compared with the less frequently detected GII.b/GII.3 and GII.7 strains, ultimately leading to on average a 1.7-fold higher rate of evolution within the capsid sequence (Bull et al., 2010). The latest research suggests that the evolutionary mechanism of GII.4 NoVs may include altering the carbohydrate-binding domains over time in response to human susceptibility (Tan and Jiang,

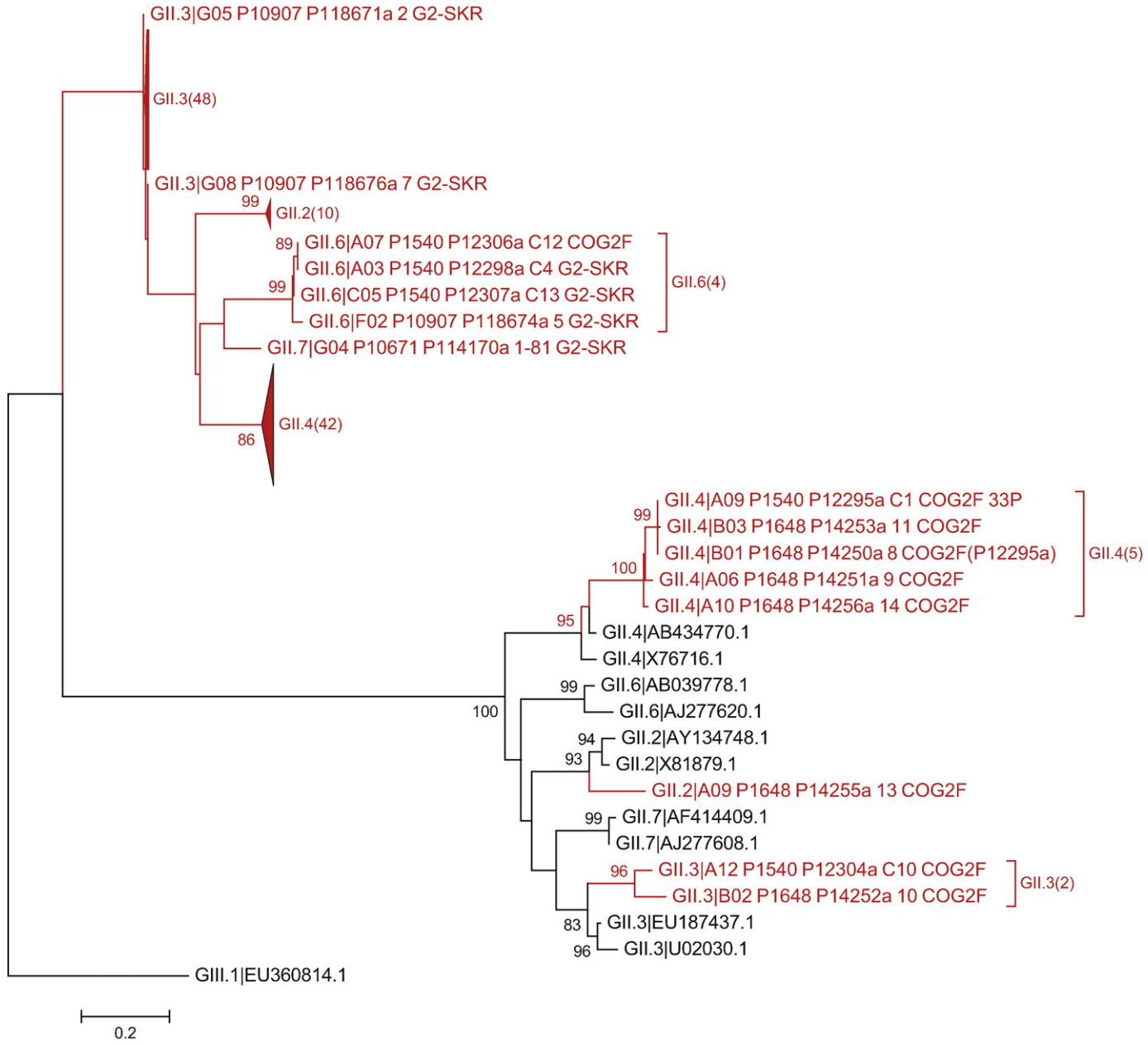

Fig. 1. Phylogenetic Tree of NoVs in Hebei Province. Maximum Likelihood (ML) phylogenetic trees were constructed using 11 reference sequences from RIVM and 115 NoV sequences obtained from our surveillance study. The HKY + G model of nucleotide substitution was shown. GIII.1|EU360814.1 was used as an out group to determine the root of the tree. Numbers in the brackets indicate the numbers of samples of the corresponding NoV genotypes. The relatively large evolutionary distance between reference sequences and our sequences was seen. Only 5 GII.4, 2 GII.3 and 1 GII.2 were located in the branch of the reference sequences, and 10 GII.2, 50 GII.3, 42 GII.4, 4 GII.6 and 1 GII.7 were located in another branch.

2011; Tan and Jiang, 2014; Tan and Jiang, 2010) as well as antigenic drift in the receptor-binding regions of the P2 subdomain in the face of human herd immunity (Lindesmith et al., 2008). However, sequences form different geographic locations may have confounding effects on the evolutionary mechanism (Lindesmith et al., 2008). Therefore, the above hypothesis needs to be further confirmed (Lindesmith et al., 2008).

Despite many studies about NoV evolution, our understanding remains limited regarding the impact of recombination on NoV adaptation and the relative position and function of substitution and recombination in the process of virus evolution. Notably, previous evolutionary studies were mostly based on hospitals, and the research objects were patients with medium to severe symptoms. Besides, it is difficult to study the population with mild infection, which may affect the accuracy of virus evolution analysis. Therefore, immunity induced by NoV natural infection as well as the impact of NoV evolutionary pattern, rate, and mechanism on the immune persistence are urgent key issues in the development and application of NoV vaccines (Robilotti et al., 2015).

The evolutionary rates of both GII.3 and GII.4 NoVs in Hebei Province were about ten times higher than that reported in the literature (Boon et al., 2011; Fernández et al., 2012; Siebenga et al., 2010; Mahar et al., 2013), and our results also showed that the VP1 gene evolutionary rate of GII.4 NoVs was supported by a previous study (Victoria et al., 2009). Our results also demonstrated that the evolutionary rate of GII.3 NoVs is about two times higher than GII.4 in Hebei Province. High evolutionary rates may contribute to the antigenic diversity of NoVs in human populations. It is noted that our surveillance covered only 6 months from October 2, 2011 to April 2, 2012, and most of our stool samples were collected in October (55.65%), November (24.35%), December (8.70%) 2011 (see Table 1). Therefore, future study to expand the sampling time range is necessary.

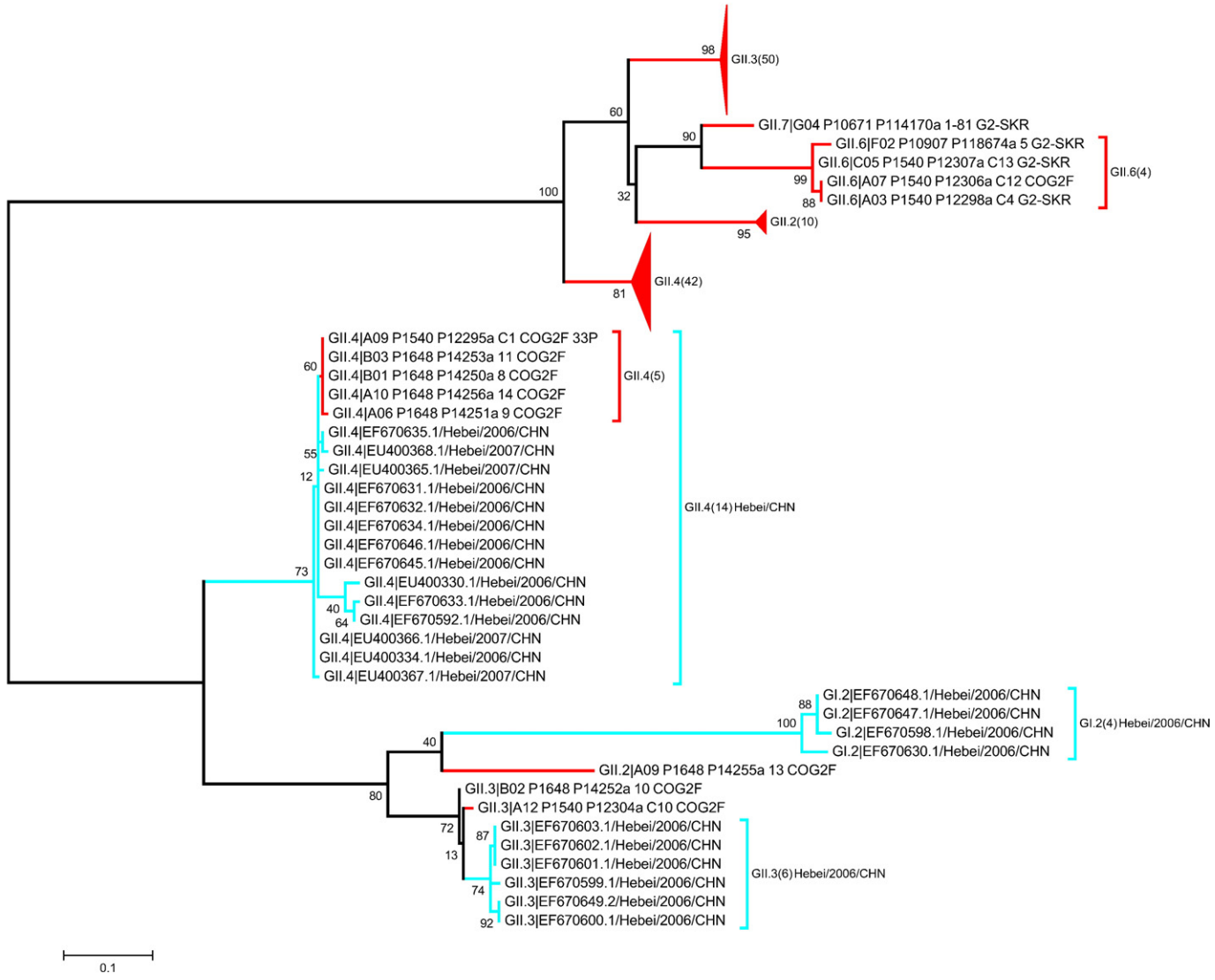

Fig. 2. Phylogenetic tree of NoVs in Hebei Province at two time points. Maximum Likelihood (ML) phylogenetic trees were constructed using 24 NoV sequences isolated from 2006 to 2007 in Hebei Province and 115 NoV sequences from our surveillance study. The HKY + G model of nucleotide substitution was shown. GIII.1|EU360814.1 was used as an out group to determine the root of the tree. Numbers in the brackets indicate the numbers of samples of the corresponding NoV genotypes. The tree contains two major branches. The first branch (red) mainly comprised 10 GII.2, 50 GII.3, 42 GII.4, 4 GII.6 and 1 GII.7 NoVs from our study. Another branch (cyan) was mainly composed of reference sequences downloaded from GenBank that were identified previously in Hebei Province from 2006 to 2007 and 5 GII.4, 2 GII.3 and 1 GII.2 NoVs that were collected from 2011 to 2012 in rural Zhengding in Hebei Province.

The NoV detection rate among children < 5 yrs. in rural Hebei Province was 10.54% (115/1091), which was consistent with that summarized in a systematic review (Patel et al., 2008; Glass et al., 2009). The genotypes distribution of NoVs in Hebei Province was roughly similar to that reported in the literatures (Ahmed et al., 2014; Siebenga et al., 2009; Lindesmith et al., 2008; Boon et al., 2011; Barreira et al., 2010), which is mainly GII NoVs with GII.3 (45.22%) and GII.4 (40.87%) as the two predominant genotypes. This indicated that certain genotypes, such as GII.3 and GII.4, remained predominant in Hebei Province, China regardless of the evolutionary mechanism of different NoVs genotypes. The detection rate of GII.3 NoVs was higher than GII.4 NoVs in children <5 years of age in Hebei Province in our analysis, probably because GII.3 NoVs were often identified as the dominant genotype associated with sporadic NoVs infection, particularly in children (Siebenga et al., 2009; Boon et al., 2011; Barreira et al., 2010).

Moreover, the phylogeny of the NoV sequences in Hebei Province showed two main independent lineages that were composed of the reference NoV sequences form the GeneBank, and our NoV sequences, respectively. Haplotypes presenting in the two layers, each consisting of NoVs isolated from 2006 to 2007 and 2011 to 2012 showed that an identical haplotype existed in GII.4 NoVs at the two time points, which means that it may have a common origin. The observation that there was no GII.3 genotypes overlap illustrated that the evolution distance of GII.3 may be relatively large at the two time points. The reasons for two separate branches and no other genotypes overlap, might be due to NoV evolution rate was high. Alternatively, the sequences from our study and those from the GeneBank were actually from different ancestors.

Viral genome recombination affects virus classification, alters phylogenetic groupings, changes viral virulence and impacts vaccine strategy. Literatures (Bull et al., 2007; Bull et al., 2005) that NoV recombination often occurred at the joint region of ORF1 and ORF2. In this study genotyping was based on genome sequences spanning from position 4954 and 5390 of the genome of NoV GII (GenBank accession No. U07611) containing a joint region between the 3′ end of ORF1 and the 5′ end of ORF2. No evidences of recombination were found in both GII.3 (52) and GII.4 (47) NoVs in our analysis, indicating that these NoVs evolved through mutations instead of homologous recombination.

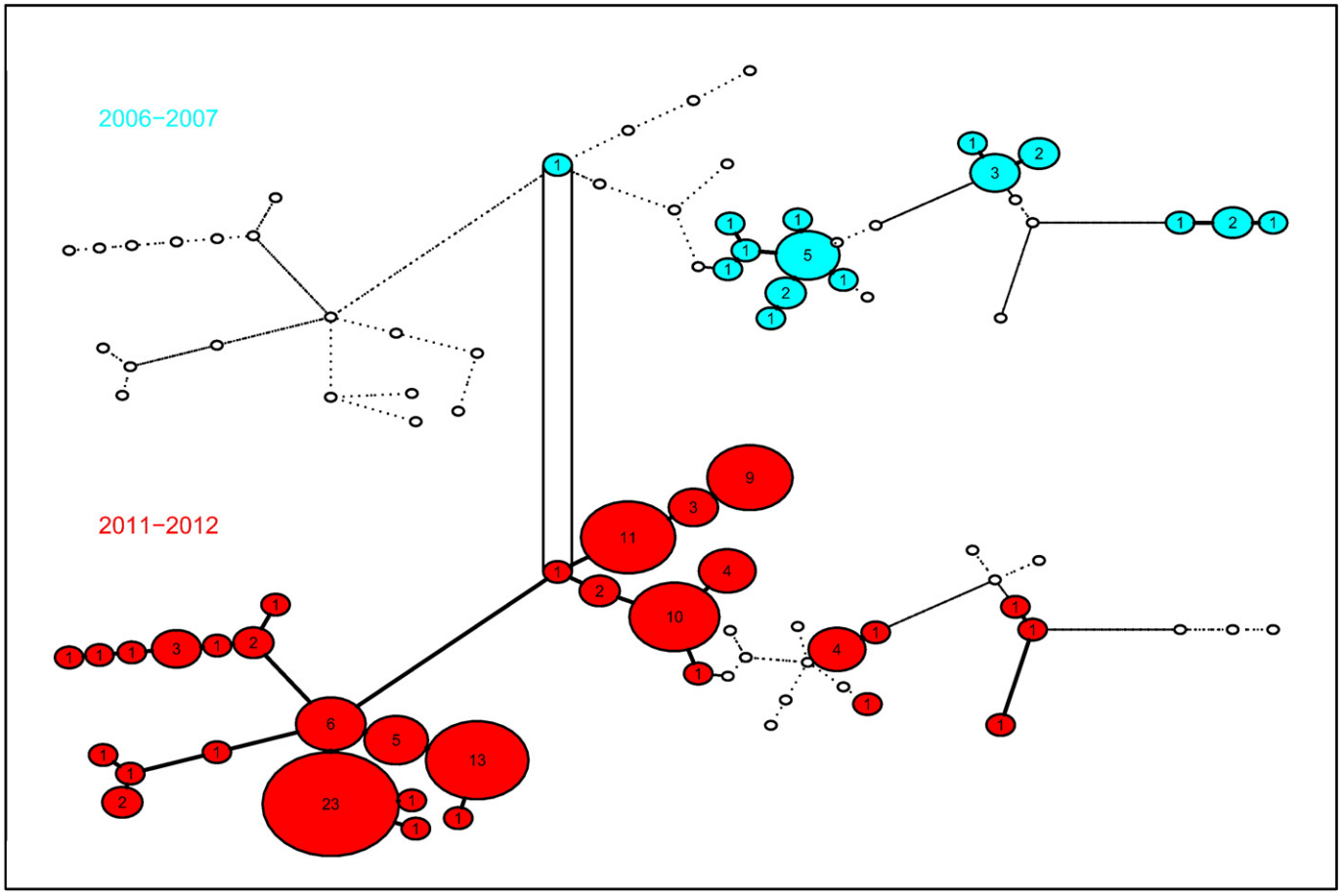

Fig. 3. Phylogenetic tree of NoVs in Hebei Province at two different time points (White ellipse indicates that a haplotype is not found in a particular layer but is found elsewhere in the network. Extant haplotypes are connected by solid lines, whereas dotted lines connect at least one unsampled haplotype. Haplotypes separated by more than one mutation are indicated by one small black circle for each additional mutation. Haplotypes present in consecutive layers are connected by vertical lines.)

In conclusion, all identified NoV sequences from Hebei Province were analyzed for their genogroup distribution and GII.3 and GII.4 were found to be predominant NoV genotypes during the study period. By constructing a haplotype network, the NoV genetic relationships in Hebei Province at two different time points were illustrated. Bayesian inference was also performed using MCMC sampling to calculate the nucleotide substitution rate of the two dominant NoV genotypes and demonstrated that the GII.3 NoVs evolved at a faster rate than GII.4 NoVs. However, due to the limitation number of specimens, sampling locations, and the relative short sample collection time, the results may not fully represent the evolutionary rate of GII.3 and GII.4 NoVs in China. To this end, future study with increased numbers, broader locations, and extent timing of sample collection is warranted. Moreover, in-depth study of the evolutionary pattern, evolutionary rate, genetic diversity, and epidemic cycles of NoVs is essential to monitor how the viruses are changing. Considering the notable disease burden, highly contagiousness of virus, and persistence immunity induced by natural infection, improvement of sanitation approaches and development of effective vaccines are of high significance to control and prevent against NoV associated diseases.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.meegid.2017.06.007.

Declarations

Ethics approval and consent to participate

This study was reviewed and approved by the Institutional Review Board of the Hebei Center for Disease Control and Prevention, and the Institutional Review Board of the Institutes of Biomedical Sciences, Fudan University. Written inform consent was obtained from the parent/guardian of each child.

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National High Technology Research and Development Program (2015AA020104) and “Precision Medical Research” Key Project (2106YFC0901900).

Authors' contributions

XYW, SMW participated in conception, design and acquisition of data, NQ, JXW, LL participated in design, analysis and interpretation of data and drafting the manuscript. BK, SSZ, JXW, XJZ, ZYH, JCM, CQ and YLZ participated in analysis and sequencing of data. All authors read and approved the final manuscript.

References

Ahmed, S.M., Hall, A.J., Robinson, A.E., Verhoef, L., Premkumar, P., Parashar, U.D., Koopmans, M., Lopman, B.A., 2014. Global prevalence of norovirus in cases of gastroenteritis: a systematic review and meta-analysis. Lancet Infect. Dis. 14 (8), 725–730.

Ando, T., Noel, J.S., Fankhauser, R.L., 2000. Genetic classification of Norwalk-like viruses. J Infect Dis 181 (Suppl. 2), S336–S348.

Ao, Y.Y., Yu, J.M., Li, L.L., Jin, M., Duan, Z.J., 2014. Detection of human norovirus GIV.1 in China: a case report. J. Clin. Virol. 61 (2), 298–301.

Ayres, D.L., Darling, A., Zwickl, D.J., Beerli, P., Holder, M.T., Lewis, P.O., Huelsenbeck, J.P., Ronquist, F., Swofford, D.L., Cummings, M.P., et al., 2011. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. Syst. Biol. 61 (1), 170–173.

Baehner, F., Bogaerts, H., Goodwin, R., 2016. Vaccines against norovirus: state of the art trials in children and adults. Clin. Microbiol. Infect. 22 (Suppl. 5), S136–S139.

Barreira, D.M.P.G., Ferreira, M.N.S.E., Túlio Machado, F., Rita, C., Sadovsky, A.D.I.D., Leite, J.P.G., Marize Pereira, M., Liliana Cruz, S., 2010. Viral load and genotypes of noroviruses in symptomatic and asymptomatic children in Southeastern Brazil. J. Clin. Virol. 47 (1), 60–64.

Beth, S., Andrew, R., Drummond, A.J., 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol. Biol. Evol. 23 (1), 7–9.

Boon, D., Mahar, J.E., Abente, E.J., Kirkwood, C.D., Purcell, R.H., Kapikian, A.Z., Green, K.Y., Bok, K., 2011. Comparative evolution of GII.3 and GII.4 norovirus over a 31-year period. J. Virol. 85 (17), 8656–8666.

Bull, R.A., White, P.A., 2011. Mechanisms of GII.4 norovirus evolution. Trends Microbiol. 19 (5), 233–240.

Bull, R.A., Hansman, G.S., Clancy, L.E., Tanaka, M.M., Rawlinson, W.D., White, P.A., 2005. Norovirus recombination in ORF1/ORF2 overlap. Emerg. Infect. Dis. 11 (7), 1079–1085.

Bull, R., Tanaka, M., White, P., 2007. Norovirus recombination. J. Gen. Virol. 88 (12), 3347–3359.

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 (3), e1000831.

Delport, W., Poon, A.F.Y., Frost, S.D.W., Kosakovsky Pond, S.L., 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26 (19), 2455–2457.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29 (8), 1969–1973.

Eden, J.S., Hewitt, J., Lim, K.L., Boni, M.F., Merif, J., Greening, G., Ratcliff, R.M., Holmes, E.C., Tanaka, M.M., Rawlinson, W.D., et al., 2014. The emergence and evolution of the novel epidemic norovirus GII.4 variant Sydney 2012. Virology 450-451, 106–113.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32 (5), 1792–1797.

Fernández, M.D.B., Torres, C., Poma, H.R., Riviello-López, G., Martínez, L.C., Cisterna, D.M., Rajal, V.B., Nates, S.V., Mbayed, V.A., 2012. Environmental surveillance of norovirus in Argentina revealed distinct viral diversity patterns, seasonality and spatio-temporal diffusion processes. Sci. Total Environ. 437, 262–269.

Glass, R.I., Parashar, U.D., Estes, M.K., 2009. Norovirus gastroenteritis. N. Engl. J. Med. 361 (18), 1776–1785.

Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82 (4), 711–732.

Hall, B.G., 2013. Building phylogenetic trees from molecular data with MEGA. Mol. Biol. Evol. 30 (5), 1229–1235.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22 (2), 160–174.

Kirsten, M., Elsie, G., Brian, A., Hugues, C., Drews, S.J., Angela, F., Nicole, G., Stephen, H., Alain, H., Lee, B.E., 2009. Multicenter comparison of two norovirus ORF2-based genotyping protocols. J. Clin. Microbiol. 47 (12), 3927–3932.

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 (s 1–2), 107–114.

Kosakovsky Pond, S.L., 2006. Automated phylogenetic detection of recombination using a genetic algorithm. Mol. Biol. Evol. 23 (10), 1891–1901.

Kosakovsky Pond, S.L., Posada, D., Gravenor, M.B., Woelk, C.H., Frost, S.D.W., 2006. GARD: a genetic algorithm for recombination detection. Bioinformatics 22 (24), 3096–3098.

Kroneman, A., Vennema, H., Deforche, K., Avoort, H.V.D., Pe Aranda, S., Oberste, M.S., Vinjé, J., Koopmans, M., 2011. An automated genotyping tool for enteroviruses and noroviruses. J. Clin. Virol. 51 (2), 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. 158 (10), 2059–2068.

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 (2), e31.

Lopman, B., Zambon, M., Brown, D.W., 2008. The evolution of norovirus, the "gastric flu". PLoS Med. 5 (2), e42.

Mahar, J.E., Bok, K., Green, K.Y., Kirkwood, C.D., 2013. The importance of intergenic recombination in norovirus GII.3 evolution. J. Virol. 87 (7), 3687–3698.

Martella, V., Medici, M.C., De Grazia, S., Tummolo, F., Calderaro, A., Bonura, F., Saporito, L., Terio, V., Catella, C., Lanave, G., et al., 2013. Evidence for recombination between pandemic GII.4 norovirus strains New Orleans 2009 and Sydney 2012. J. Clin. Microbiol. 51 (11), 3855–3857.

Mesquita, J.R., Barclay, L., Nascimento, M.S.J., Vinjé, J., 2010. Novel norovirus in dogs with diarrhea. Emerg. Infect. Dis. 16 (6), 980–982.

Minin, V.N., Bloomquist, E.W., Suchard, M.A., 2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Mol. Biol. Evol. 25 (7), 1459–1471.

Patel, M.M., Widdowson, M., Glass, R.I., Akazawa, K., Vinjé, J., Parashar, U.D., 2008. Systematic literature review of role of noroviruses in sporadic gastroenteritis. Emerg. Infect. Dis. 14 (8), 1224–1231.

Pevsner, J., 2013. Basic local alignment search tool (BLAST). J. Mol. Biol. 215 (3), 87–125.

Pond, S.L.K., Frost, S.D.W., 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21 (10), 2531–2533.

Prost, S., Anderson, C.N.K., 2011. TempNet: a method to display statistical parsimony networks for heterochronous DNA sequence data. Methods Ecol. Evol. 2 (6), 663–667.

Rambaut, A., Suchard, M., Xie, D., Drummond, A.J., 2014. Tracer v1.6. (In.).

Robilotti, E., Deresinski, S., Pinsky, B.A., 2015. Norovirus. Clin. Microbiol. Rev. 28 (1), 134–164.

Siebenga, J.J., Vennema, H., Zheng, D.P., Vinjé, J., Lee, B.E., Pang, X.L., Ho, E.C.M., Lim, W., Choudekar, A., Broor, S., et al., 2009. Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001–2007. J Infect Dis 200 (5), 802–812.

Siebenga, J.J., Lemey, P., Kosakovsky, P.S., Rambaut, A., Vennema, H., Koopmans, M., 2010. Phylodynamic reconstruction reveals norovirus GII.4 epidemic expansions and their molecular determinants. PLoS Pathog. 6 (5), e1000884.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30 (12), 2725–2729.

Tan, M., Jiang, X., 2010. Virus-host interaction and cellular receptors of caliciviruses. In: Hansman, G., Jiang, X., Green, K. (Eds.), Caliciruses. Caister Academic Press, pp. 111–130.

Tan, M., Jiang, X., 2011. Norovirus-host interaction: multi-selections by human histo-blood group antigens. Trends Microbiol. 19 (8), 382–388.

Tan, M., Jiang, X., 2014. Histo-blood group antigens: a common niche for norovirus and rotavirus. Expert Rev. Mol. Med. 16, e5.

Victoria, M., Miagostovich, M.P., Ferreira, M.S.R., Vieira, C.B., Fioretti, J.M., Leite, J.P.G., Colina, R., Cristina, J., 2009. Bayesian coalescent inference reveals high evolutionary rates and expansion of norovirus populations. Infect. Genet. Evol. 9 (5), 927–932.

Zhen, S.S., Li, Y., Wang, S.M., Zhang, X.J., Hao, Z.Y., Chen, Y., Wang, D., Zhang, Y.H., Zhang, Z.Y., Ma, J.C., 2015. Effectiveness of the live attenuated rotavirus vaccine produced by a domestic manufacturer in China studied using a population-based case–control design. Emerg. Microbes Infect.

Zheng, D., Ando, T., Fankhauser, R.L., Beard, R.S., Glass, R.I., Monroe, S.S., 2006. Norovirus classification and proposed strain nomenclature. Virology 346 (2), 312–323.