OPEN ACCESS Freely available online

ADVANCES IN PEDIATRIC RESEARCH

# Spatiotemporal evolutionary dynamics of norovirus GII.4 variants

**Ruta Kulkarni [1], Atul M. Walimbe [2], Shobha D. Chitambar [1]***

**1** Enteric Viruses Group, National Institute of Virology, Pune, India **2** Bioinformatics Group, National Institute of Virology, Pune, India

**Abstract**

***Background:*** Noroviruses (NoVs) are the second most common cause of sporadic childhood gastroenteritis worldwide. NoVs of the GII.4 genotype are predominant globally and undergo continuous evolution in the VP1 gene, encoding the major capsid protein, resulting in the emergence of novel variants, with fourteen GII.4 variants identified to date.

***Methods:*** The present study investigated the spatiotemporal evolutionary dynamics of the globally circulating GII.4 NoVs using Bayesian approach, phylogeographic, and migration pattern analyses on a dataset of complete VP1 sequences representing each of the GII.4 variants.

***Results:*** The estimated mean evolutionary rate for GII.4 VP1 was $5.1x10^{-3}$ nucleotide substitutions per site per year (sub/site/yr), the time to Most Recent Common Ancestor (tMRCA) was ~1971, and the most probable ancestral location was the United States of America (USA). The fourteen known GII.4 variants displayed variable evolutionary rates ($4.5–7.4x10^{-3}$ sub/site/yr) and tMRCA (~1984 to ~2006), with the majority having USA/Asia as their ancestral location. Migration pattern analysis indicated the important role of Australia-New Zealand, India, and Southern Europe in the global dispersal of GII.4 noroviruses.

***Conclusion:*** The study contributes to the understanding of GII.4 norovirus evolutionary dynamics.

**Citation:** Kulkarni R, Walimbe AM, Chitambar SD (2017) Spatiotemporal evolutionary dynamics of norovirus GII.4 variants. Adv Pediatr Res 4:5. doi:10.12715/apr.2017.4.5

**Received:** October 25, 2016; **Accepted:** February 1, 2017; **Published:** May 1, 2017



**Competing interests:** The authors have declared that no competing interests exist.

**Sources of funding:** The project was funded by National Institute of Virology (NIV), Pune and Indian Council of Medical Research (ICMR), New Delhi, India.

* Email: drshobha.niv@gmail.com

## Introduction

Noroviruses (NoVs) are an important cause of acute gastroenteritis worldwide, with children aged <5 years having the highest incidence of the disease. In countries implementing rotavirus vaccination as routine, NoVs are identified as the primary cause of viral gastroenteritis among children [1].

NoVs have a positive sense ssRNA genome consisting of three open-reading frames (ORFs). ORF1 encodes the non-structural proteins (NS1-NS7) including NTPase, Proteinase, and RNA dependent RNA polymerase (RdRp). ORF2 and ORF3 code for the capsid proteins, VP1 and VP2, respectively. The major capsid protein, VP1, carries the antigenic determinants and is involved in interaction with host antibodies and cellular binding ligands. VP1 is divided into the shell (S) and two protruding (P) domains – P1 and P2 [2]. NoVs are classified into six genogroups, GI–GVI [1]. Human infections are mainly caused by GII NoVs which are further classified into 23 genotypes based on the sequence divergence of the capsid (ORF2/VP1) [3, 4].

Genogroup II genotype IV (GII.4) NoVs are predominant globally and account for ~62% of norovirus outbreaks and ~70% of sporadic infections

[5, 6]. GII.4 NoVs undergo continuous evolution leading to the periodic emergence of novel variants which replace the previous circulating variants [7, 8]. Fourteen variants of the GII.4 NoVs have been reported to date [3, 4]. Of these, Bristol_1993 and Camberwell_1994, circulating in the 1980s and 1990s, represent the earliest known GII.4 variants. The US95_96 pandemic variant, circulating during 1995–2002, was the first GII.4 variant reported to have a global prevalence [9]. Since then, the emergence of six pandemic GII.4 variants, Farmington Hills_2002 (2002-2004), Hunter_2004 (2004-2006), Yerseke_2006a (2006-2007), DenHaag_2006b (2006-2009), New Orleans_2009 (2009-2012) and Sydney_2012 (2012 onwards), has been reported to date [6, 10]. Each of these variants predominated for 2–3 years before being replaced by the next emerging variant. The Osaka_2007 and Apeldoorn_2007 variants have also been described to be prevalent worldwide during 2007–2009; however, these variants did not gain global predominance over the pandemic GII.4 variants circulating in this period [6]. In addition to the above variants which showed global presence, other variants, namely, Lanzou_2002, Kaiso_2003 and Asia_2003, with limited circulation, have been reported, with the Asia_2003 variant being an important cause of outbreaks in Asia during 2004–2005 [6].

Previous studies on the evolutionary dynamics of GII.4 NoV variants have reported a rapid rate of evolution of the VP1 gene [11-13]. However, such studies have not been carried out for the recent GII.4 variants (DenHaag_2006b onwards) using a dataset of globally representative complete VP1 sequences. The studies that have included the recent GII.4 variants have been either restricted to a limited geographical region or have analyzed partial VP1 sequences [14-18]. Moreover, phylogeographic studies, important for understanding the global epidemiology and migration patterns of GII.4 NoVs, have not been conducted so far. The present study, therefore, aimed to investigate the spatiotemporal dynamics of evolution and dispersal of the GII.4 NoVs circulating worldwide between 1974 and 2014.

## Materials and Methods

### *Preparation of the sequence dataset*

Complete VP1 nucleotide sequences (~1620 bp) of NoV GII.4 strains, available in GenBank as of November 2014, with known sampling year and country (through GenBank entries or associated literature), were considered for the study. All such sequences (n=1614) were retrieved from GenBank and subjected to the Noronet automated genotyping tool [3] for GII.4 variant assignment. A dataset of 270 sequences was compiled by selection of representative sequences based on the GII.4 variant assignment, year, and country of isolation (1 sequence per GII.4 variant, per year, per country). Another 15 sequences, representing the GII.4 strains from Pune, India, were added to this dataset. Thus, the final dataset was comprised of 285 VP1 sequences, representing the GII.4 strains collected during 1974–2014 from 28 different countries in the 5 different continents of the world. Of these 285 sequences, 279 were distributed into the 14 known GII.4 variants while 6 (5 non-recombinant and 1 recombinant) could not be assigned to any of the GII.4 variants (Supplemental Table S1, Supplemental Fig. S1).

### *Evolutionary analysis using Bayesian approach*

Bayesian phylogenetic analysis was carried out for the estimation of evolutionary rates and ancestral times of the GII.4 NoVs. A Maximum Clade Credibility (MCC) tree was generated using the Bayesian Markov Chain Monte Carlo approach as implemented in BEAST 1.6.2 [19]. The best-fit model of substitution, GTR + G + I (general time-reversible model with gamma-distributed rates of variation among sites and a proportion of invariable sites) was employed. The relaxed uncorrelated lognormal clock model [20] was employed with the Bayesian Skyline tree prior as described previously for studies on GII.4 NoVs [12, 14-18]. Three independent runs of the chain were carried out, with chain length 300 million and a sampling frequency of 5000. The convergence of the output was assessed by Tracer 1.6. The MCC tree was visualized in FigTree 1.4.2. TreeStat program was used to estimate the cladewise rates of substitution.

### *Phylogeographic and migration pattern analyses*

To perform the phylogeographic analysis, the sampling locations (n=28) were classified into 12 geographical regions: United States of America (USA), Canada (Can), South America (S.Am), Northern Europe (N.Eur), Southern Europe (S.Eur), Eastern Asia (E.As), South-Eastern Asia (SE.As), India (Ind), Bangladesh (Ban), Northern Africa (N.Af), Southern Africa (S.Af), and Australia-New Zealand (AuNZ). The association between phylogeny and geographical locations of the sequence data and the suitability of the phylogeographic analysis was evaluated using the phylogeny-trait association tests (AI, Association Index and PS, Parsimony Score) available in software BaTS [21]. Since a strong association was indicated between phylogeny and geographic locations, phylogeographic analysis was conducted using the spatial information of the GII.4 strains. The geographic spread patterns of the virus were inferred by fitting a standard continuous-time Markov chain (CTMC) model with the Bayesian stochastic search variable selection (BSSVS) [22] in BEAST. The output of the Bayesian phylogeographic analyses, as generated by BEAST, was summarized using MCC tree. The output of BEAST was also input to the program SPREAD 1.0.3 [23] to visualize and analyze the dispersion pathways. The well-supported migration links (BF>3) between the different geographical regions were identified using Bayes Factor (BF) test available in SPREAD.

## Results

### *Bayesian phylogenetic analysis*

According to the MCC tree, the NoV GII.4 strains were distributed into 14 clusters corresponding to the 14 GII.4 variants identified to date (Fig. 1). This clustering was in agreement with the GII.4 variant assignment carried out by the Noronet automated genotyping tool [3], with two exceptions. First, the AF145896/Camberwell-101922 strain, identified as Camberwell_1994 variant according to Noronet typing, did not cluster with the rest of the strains (n=3) of this variant in the MCC tree. Secondly, seven of the Noronet classified US95_96 strains clustered separately from the remaining (n=18) strains of this variant in the MCC tree. A maximum likelihood tree of the GII.4 strains, constructed to confirm the topology of the MCC tree, also showed separate clustering within the Camberwell_1994 and US95_96 variants (data not shown), thus indicating sequence divergence within these variants.

Further examination of the topology of the MCC tree revealed that with the exception of the Kaiso_2003 variant, all of the variants post-year 2000 clustered close to each other, representing a group of "contemporary GII.4 NoVs," separate from the previous variants (Bristol_1993, Camberwell_1994 and US95_96) and the earliest GII.4 strains (1970s–1980s). The Kaiso_2003 variant clustered close to the older Bristol_1993 variant.

### *Estimation of the evolutionary rate and ancestral time of the GII.4 NoVs*

The information regarding the year of isolation of the GII.4 strains in the dataset was used to estimate the evolutionary rate of the GII.4 NoVs. According to the relaxed lognormal clock model, the mean rate of nucleotide substitution for the VP1 gene of GII.4 NoVs was estimated to be $5.1x10^{-3}$ substitutions per site per year (highest posterior density (HPD) limits: $4.6–5.7x10^{-3}$). The evolutionary rate for the contemporary GII.4 NoVs was estimated to be $5.5x10^{-3}$ sub/site/yr. The evolutionary rates estimated individually for each of the 14 GII.4 variants ranged between $4.5x10^{-3}$ and $7.4x10^{-3}$ substitutions per site per year (sub/site/yr) with Bristol_1993 and Camberwell_1994 showing the lowest and highest rates respectively (Table 1).

Based on the mean evolutionary rate, the mean root age for the GII.4 NoVs was inferred to be 42.5 years (HPD: 40.0–46.0 years). The tMRCA was calculated to be around the year 1971 (HPD: 1968–1974) (Table 1). For the contemporary GII.4 NoVs, the tMRCA was found to be ~1993. For the individual GII.4 variants, the tMRCA varied between ~2006 and ~1984 for the youngest (Sydney_2012) to oldest (Camberwell_1994) variants.


**Figure 1.** Maximum clade credibility (MCC) tree of GII.4 NoV VP1 sequences.

The distribution of the GII.4 strains into 14 GII.4 variants is shown as collapsed clusters. The posterior supports for each variant cluster are shown. One strain of the Camberwell_1994 variant and seven of the US95_96 variant cluster separately from the remaining strains (collapsed branches) of the respective variants. The GII.4 strains are represented as follows: Accession number_Strain name_Sampling location (ISO country code)_Geographical region code_Year of sampling.

***Analysis of the ancestral location and migration pattern of the GII.4 NoVs***

The most probable ancestral geographical location for the GII.4 NoVs was found to be USA with a probability of 0.55. Analysis carried out to determine the ancestral locations of individual GII.4 variants revealed Northern Europe for Bristol_1993, Asia for Kaiso_2003, and the USA for Camberwell_1994 and US95_96 variants (Table 1). For the contemporary GII.4 NoVs, the ancestral location was identified to be the USA for Lanzou_2002, Farmington Hills_2002, and Osaka_2007; Eastern Asia for Asia_2003; Southern Europe for Hunter_2004; and Asia for DenHaag_2006b variants. The most probable ancestral locations were identified with a relatively lower probability for Yerseke_2006a (Southern Europe), Apeldoorn_2007 (Asia), New Orleans_2009 (Asia), and Sydney_2012 (Asia) variants.

Analysis of the migration patterns of the GII.4 NoVs (Fig. 2) revealed that Australia-New Zealand showed the maximum number (six) of significant transmission links with other geographical regions (Southern Europe, USA, Canada, Eastern Asia, South-Eastern Asia, and India) in three different continents. The migration links were strongest between Australia-New Zealand and Southern Europe (BF:135.4). India had significant transmission links with three geographical regions: Southern Europe, Southern Africa, and Australia-New Zealand. Southern Europe also showed links with three regions: Australia-New Zealand, India, and Northern Africa. The USA, Canada, Northern Europe, Northern Africa, Southern Africa, Eastern Asia, and South-Eastern Asia each showed significant migration links with 1–2 geographical regions, while South America and Bangladesh showed none.

## Discussion

The GII.4 NoVs are known to be predominant globally. The persistence of these NoVs in the human population has been attributed to their continuous evolution in the VP1 gene resulting in the emergence of novel variants [7, 8]. The emerging variants have been described to escape pre-existing host immunity through variations in the surface exposed P2 domain of VP1, particularly in the epitopic regions [24], which have been reported to be under positive selection pressure [8,11-12,14-15]. In addition, positive selection has also been detected in the S and P1 domains suggesting the importance of these regions in GII.4 NoV evolution [8,11-12,14].

In the present study, the spatiotemporal evolutionary dynamics of the 14 GII.4 variants, identified to date, was examined using a dataset of complete VP1 sequences representing the GII.4 NoV strains circulating worldwide between 1974 and 2014. The VP1 gene of GII.4 NoV was estimated to evolve at a rate of $5.1x10^{-3}$ nucleotide sub/site/yr with the evolutionary rate varying between $4.5x10^{-3}$ and $7.4x10^{-3}$ for the individual GII.4 variants (Table 1). Previous studies have described evolutionary rates in the range of $3.9–7.2x10^{-3}$ sub/site/yr for the GII.4 NoVs and have reported this rate to be higher than that estimated for other NoV genotypes ($1.9–2.4x10^{-3}$) [11-15]. The rapid rate of VP1 evolution of the GII.4 NoVs is consistent with their higher RdRp mutation rate [13] and may partly explain the diversity within the GII.4 genotype and the periodic emergence of novel variants. In this context, it is worth noting that NoVs of the GII.3 genotype reportedly evolve at a rate ($4.2–7.4x10^{-3}$ nucleotide sub/site/yr) comparable to that of GII.4 NoVs; however, GII.3 NoVs show lower diversity in amino acid sequence as compared to the GII.4 NoVs suggesting differences in the host selection pressures on these two NoV genotypes [25].

The tMRCA for the GII.4 NoVs was estimated to be around the year 1971 in the present study (Table 1) and thus falls within the range (1967–1982) suggested by previous studies [11,12]. The tMRCA for the contemporary GII.4 NoVs was more recent (~1993). The tMRCA for the different GII.4 variants was variable (~1984 to ~2006), reflecting the periodic emergence of novel variants.

The most probable ancestral geographical location for the GII.4 genotype was inferred to be the USA. However, this observation may be biased by the fact that the USA was the sampling location of the earliest known GII.4 strains included in the dataset and that the placement of these oldest strains was closer to the root in the MCC tree.

**Table 1.** Summary of the evolutionary and phylogeographic analysis of the GII.4 NoVs

| GII.4 variant | Evolutionary rate (95% HPD) X$10^{-3}$ nucleotide sub/site/yr | Node age (95% HPD) | tMRCA (95% HPD) | Most probable ancestral location (probability) |
|---|---|---|---|---|
| Root | 5.1 (4.6–5.7) | 42.5 (40.0–46.0) | 1971.5 (1968.0–1974.0) | USA (0.55) |
| Bristol_1993[a] | 4.5 (2.4–7.1) | 22.7 (22.2–23.5) | 1991.3 (1990.5–1991.8) | N.Eur (0.85) |
| Camberwell_1994[a] | 7.4 (4.6–10.7) | 29.3 (27.8–30.9) | 1984.7 (1983.1–1986.2) | USA (0.93) |
| US95_96 | 4.9 (3.9–6.1) | 20.7 (20.0–21.6) | 1993.3 (1992.4–1994.0) | USA (0.64) |
| Lanzou_2002[a] | 6.5 (4.4–8.8) | 14.8 (14.1–15.7) | 1999.2 (1998.3–1999.9) | USA (0.94) |
| Farmington Hills_2002 | 5.4 (4.3–6.6) | 16.3 (15.1–17.6) | 1997.7 (1996.4–1998.9) | USA (0.77) |
| Kaiso_2003[a] | 5.6 (3.6–7.9) | 15.0 (13.0–17.3) | 1999.0 (1996.7–2001.0) | Asia[b] (0.51) |
| Asia_2003[a] | 5.3 (3.8–6.8) | 11.9 (11.1–12.8) | 2002.1 (2001.2–2002.9) | E.As (0.91) |
| Hunter_2004 | 5.0 (4.1–5.9) | 12.9 (12.3–13.6) | 2001.1 (2000.4–2001.7) | S.Eur (0.62) |
| Yerseke_2006a | 5.4 (4.2–6.5) | 13.0 (11.4–14.6) | 2001.1 (1999.4–2002.6) | S.Eur (0.32) |
| DenHaag_2006b | 5.2 (4.5–6.1) | 10.7 (9.6–11.7) | 2003.3 (2002.3–2004.4) | Asia (0.42) |
| Osaka_2007 | 6.4 (5.1–7.7) | 15.2 (12.8–17.6) | 1998.8 (1996.4–2001.2) | USA (0.44) |
| Apeldoorn_2007 | 5.7 (4.8–6.7) | 9.4 (8.5–10.3) | 2004.6 (2003.7–2005.5) | Asia (0.36) |
| New Orleans_2009 | 5.1 (4.3–6.0) | 8.2 (7.3–9.2) | 2005.8 (2004.8–2006.7) | Asia (0.31) |
| Sydney_2012 | 5.6 (4.5–6.9) | 8.0 (6.6–9.4) | 2006.0 (2004.6–2007.5) | Asia (0.38) |
| Contemporary GII.4 | 5.5 (4.8–6.2) | 20.2 (19.0–21.5) | 1993.8 (1992.5–1995.1) | - |

For the phylogeographic analysis, the sampling locations were classified into 12 geographical regions: United States of America (USA), Canada (Can), South America (S.Am), Northern Europe (N.Eur), Southern Europe (S.Eur), Eastern Asia (E.As), South-Eastern Asia (SE.As), India (Ind), Bangladesh (Ban), Northern Africa (N.Af), Southern Africa (S.Af), and Australia-New Zealand (AuNZ). HPD: highest posterior density; tMRCA: time to Most Recent Common Ancestor.

[a] The evolutionary analysis for this variant was based on a limited number (< 10) of sequences available in GenBank.

[b] Asia = E.As + SE.As + Ind + Ban


**Figure 2.** World map displaying significantly non-zero transition rates.

For the phylogeographic and migration pattern analysis, the sampling locations were classified into 12 geographical regions: United States of America (USA), Canada (Can), South America (S.Am), Northern Europe (N.Eur), Southern Europe (S.Eur), Eastern Asia (E.As), South-Eastern Asia (SE.As), India (Ind), Bangladesh (Ban), Northern Africa (N.Af), Southern Africa (S.Af), and Australia-New Zealand (AuNZ). The figure shows ten well-supported (Bayes Factor, BF > 3) migration links (shown by purple colored lines): AuNZ–S.Eur (BF=135.4), AuNZ–E.As (BF=50.3), AuNZ–Can (BF=25.7), AuNZ–SE.As (BF=19.4), Can–N.Eur (BF=13.4), AuNZ–USA (BF=12.7), AuNZ–Ind (BF=5.0), Ind–S.Af (BF=3.7), Ind–S.Eur (BF=3.5) and S.Eur–N.Af (BF=3.5). The darker the colour gradient, the stronger the support is for the corresponding rate.

The ancestral locations for the different GII.4 variants varied; however, the USA or the Asian continent was identified as the most probable location for the majority of the GII.4 variants (Table 1).

Further analysis of the spatial and temporal information of the GII.4 variants (Fig. S1) revealed that in the cases of the seven pandemic GII.4 variants and the globally prevalent Osaka_2007 and Apeldoorn_2007 variants, the GII.4 variant was detected within 1–2 years in multiple locations in three or more different continents of the world indicating global dispersal of these GII.4 variants. Analysis of the migration patterns of the GII.4 NoVs (Fig. 2) revealed an important role for Australia-New Zealand, India, and Southern Europe, all of which showed significant transmission links with other geographical regions around the world. The global prevalence of the GII.4 NoVs indicates that apart from the statistically significant transmission links noted in the present study, other migration patterns may also contribute to the dispersal of GII.4 NoVs.

The present study, thus, contributes to the understanding of evolutionary dynamics of the NoV GII.4 variants. However, the spatiotemporal estimates made in the present study are limited by the low number (<10) of sequences available for some of the GII.4 variants, namely, Bristol_1993, Camberwell_1994, Lanzou_2002, Kaiso_2003, and

Asia_2003, as well as a lack of older sequences from different countries/geographical regions and the under-sampling of a few of the geographical regions due to sequence unavailability. The increase in nucleotide sequencing in recent years, as evidenced by the availability of a large number of sequences from different geographical regions for the recent GII.4 variants, indicate that future studies may be possible using a more exhaustive sequence dataset. Such studies would further elucidate the spatiotemporal evolutionary dynamics of the GII.4 NoVs that pose a challenge for infection control due to their continuous evolution, diversity and global dispersal.

## Acknowledgements

The authors thank the University Grants Commission (UGC), New Delhi, for supporting the first author (R.K.) with a Senior Research Fellowship and Dr. D.T. Mourya, Director, National Institute of Virology, Pune for his constant support.

## References

1. Ramani S, Atmar RL, Estes MK. Epidemiology of human noroviruses and updates on vaccine development. Curr Opin Gastroenterol. 2014;30(1):25-33.
2. Green KY. Caliciviridae: the noroviruses. In: Knipe DM, Howley PM, Cohen JI, Griffin DE, Lamb RA, Martin MA, Racaniello VR, Roizman B, editors. Fields Virology, 6th edn. Philadelphia, PA: Lippincott Williams & Wilkins; 2013. p. 582-608.
3. Kroneman A, Vennema H, Deforche K, v d Avoort H, Peñaranda S, Oberste MS, et al. An automated genotyping tool for enteroviruses and noroviruses. J Clin Virol. 2011;51:121-5.
4. Kroneman A, Vega E, Vennema H, Vinjé J, White PA, Hansman G, et al. Proposal for a unified norovirus nomenclature and genotyping. Arch Virol. 2013;158:2059-68.
5. Siebenga JJ, Vennema H, Zheng DP, Vinje J, Lee BE, Pang XL, et al. Norovirus illness is a global problem: Emergence and spread of norovirus GII.4 variants, 2001–2007. J Infect Dis. 2009;200:802-12.
6. Hoa Tran TN, Trainor E, Nakagomi T, Cunliffe NA, Nakagomi O. Molecular epidemiology of noroviruses associated with acute sporadic gastroenteritis in children: global distribution of genogroups, genotypes and GII.4 variants. J Clin Virol. 2013;56:185-93.
7. Siebenga JJ, Vennema H, Renckens B, de Bruin E, van der Veer B, Siezen RJ, et al. Epochal evolution of GGII.4 norovirus capsid proteins from 1995 to 2006. J Virol. 2007; 81:9932-41.
8. Lindesmith LC, Donaldson EF, Lobue AD, Cannon JL, Zheng DP, Vinje J, et al. Mechanisms of GII.4 norovirus persistence in human populations. PLoS Med. 2008;5:e31.
9. Noel JS, Fankhauser RL, Ando T, Monroe SS, Glass RI. Identification of a distinct common strain of "Norwalk-like Viruses" having a global distribution. J Infect Dis. 1999;179:1334-44.
10. van Beek J, Ambert-Balay K, Botteldoorn N, Eden JS, Fonager J, Hewitt 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.
11. Bok K, Abente EJ, Realpe-Quintero M, Mitra T, Sosnovtsev SV, Kapikian AZ, et al. Evolutionary dynamics of GII.4 noroviruses over a 34-year period. J Virol. 2009;83:11890-901.
12. Siebenga JJ, Lemey P, Kosakovsky Pond SL, Rambaut A, Vennema H, Koopmans M. Phylodynamic reconstruction reveals norovirus GII.4 epidemic expansions and their molecular determinants. PLoS Pathog. 2010;6:e1000884.
13. Bull RA, Eden JS, Rawlinson WD, White PA. Rapid evolution of pandemic noroviruses of the GII.4 Lineage. PLoS Pathog. 2010;6:e1000831.
14. Cotten M, Petrova V, Phan MV, Rabaa MA, Watson SJ, Ong SH, et al. Deep sequencing of norovirus genomes defines evolutionary patterns in an urban tropical setting. J Virol. 2014;88:11056-69.
15. Eden JS, Hewitt J, Lim KL, Boni MF, Merif J, Greening G, et al. The emergence and evolution of the novel epidemic norovirus GII.4 variant Sydney 2012. Virology 2014;450-451:106-13.
16. Fernández MD, Torres C, Poma HR, Riviello-López G, Martínez LC, Cisterna DM, et al. Environmental surveillance of norovirus in Argentina revealed distinct viral diversity patterns, seasonality and spatio-temporal diffusion processes. Sci Total Environ. 2012;437:262-9.
17. Fioretti JM, Bello G, Rocha MS, Victoria M, Leite JPG, Miagostovich MP. Temporal dynamics of norovirus GII.4 variants in Brazil between 2004 and 2012. PLoS ONE 2014; 9:e92988.
18. Tra My PV, Lam HM, Thompson CN, Phuc HL, Tuyet PT, Vinh H, et al. The dynamics of GII.4 norovirus in Ho Chi Minh City, Vietnam. Infect Genet Evol. 2013;18:335-43.
19. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007;7:214.

20. Drummond AJ, Ho SY, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. PLoS Biol. 2006;4:e88.

21. Parker J, Rambaut A, Pybus OG. Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty. Infect Genet Evol. 2008;8:239-46.

22. Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. PLoS Comput Biol. 2009;5:e1000520.

23. Bielejec F, Rambaut A, Suchard MA, Lemey P. SPREAD: Spatial Phylogenetic Reconstruction of Evolutionary Dynamics. Bioinformatics. 2011;27:2910-12.

24. Lindesmith LC, Beltramello M, Donaldson EF, Corti D, Swanstrom J, Debbink K, et al. Immunogenetic mechanisms driving norovirus GII.4 antigenic variation. PLoS Pathog. 2012;8:e1002705.

25. Boon D, Mahar JE, Abente EJ, Kirkwood CD, Purcell RH, Kapikian AZ, et al. Comparative evolution of GII.3 and GII.4 norovirus over a 31-year period. J Virol. 2011; 85:8656-66.