Temporal Dynamics of Norovirus GII.4 Variants in Brazil between 2004 and 2012

Julia Monassa Fioretti1*, Gonzalo Bello2, Mônica Simões Rocha1, Matias Victoria3, José Paulo Gagliardi Leite1, Marize Pereira Miagostovich1

1Laboratório de Virologia Comparada e Ambiental, Instituto Oswaldo Cruz – Fiocruz, Rio de Janeiro, Brazil, 2Laboratório de AIDS & Imunologia Molecular, Instituto Oswaldo Cruz – Fiocruz, Rio de Janeiro, Brazil, 3Universidade de La República, Laboratorio de Virologia Molecular – Regional Norte, Ciudad de Salto, Uruguay

Abstract

Noroviruses (NoVs) are the major cause of acute gastroenteritis outbreaks, and, despite a wide genetic diversity, genotype II.4 is the most prevalent strain worldwide. Mutations and homologous recombination have been proposed as mechanisms driving the epochal evolution of the GII.4, with the emergence of new variants in 1–3-year intervals causing global epidemics. There are no data reporting the dynamics of GII.4 variants along a specific period in Brazil. Therefore, to improve the understanding of the comportment of these variants in the country, the aim of this study was to evaluate the circulation of NoV GII.4 variants during a 9-year period in 3 out of 5 Brazilian regions. A total of 147 samples were sequenced, and a phylogenetic analysis of subdomain P2 demonstrated the circulation of six GII.4 variants, Asia_2003, Hunter_2004, Den Haag_2006b, Yerseke_2006a, New Orleans_2009, and Sydney_2012, during this period. The most prevalent variant was Den Haag_2006b, circulating in different Brazilian regions from 2006 to 2011. A Bayesian coalescent analysis was used to calculate the mean evolutionary rate of subdomain P2 as 7.3×10^{-3} (5.85×10^{-3}–8.82×10^{-3}) subst./site/year. These analyses also demonstrated that clad Den Haag_2006b experienced a rapid expansion in 2005 and another in 2008 after a period of decay. The evaluation of the temporal dynamics of NoV GII.4 in Brazil revealed a similar pattern, with few exceptions, to the worldwide observation. These data highlight the importance of surveillance for monitoring the emergence of new strains of NoV GII.4 and its impact on cases of acute gastroenteritis.

Citation: Fioretti JM, Bello G, Rocha MS, Victoria M, Leite JPG, et al. (2014) Temporal Dynamics of Norovirus GII.4 Variants in Brazil between 2004 and 2012. PLoS ONE 9(3): e92988. doi:10.1371/journal.pone.0092988

Editor: Pierre Roques, CEA, France

Received December 17, 2013; Accepted February 27, 2014; Published March 25, 2014

Copyright: © 2014 Fioretti et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by Excellence Program of Research (PROEP-CNPq/IOC) and General Coordination of Laboratories/Secretary of Health Surveillance, Ministry of Health. MP Miagostovich, JPG Leite and G Bello are CNPq fellows. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: juliafior@ioc.fiocruz.br

Introduction

Noroviruses (NoVs) are considered to be the major causative agent of acute gastroenteritis (AGE) outbreaks worldwide [1]. In developing countries, is estimated that it causes annually 1.1 million hospitalizations and 200,000 deaths in children under 5 years of age [2].

The genus Norovirus belongs to the Caliciviridae family, which comprises nonenveloped viruses with an icosahedral symmetry of approximately 27–30 nm in diameter. The NoV genome is composed of a positive-sense single-stranded RNA covalently linked to VPg at the 5’ end and polyadenylated at the 3’ end, 7.5–7.7 kb in length. The genome is organized into three open reading frames (ORF), with ORF1 encoding six non-structural proteins, including the RNA-dependent RNA-polymerase, ORF2 encoding the structural protein VP1 that composes the viral capsid, and ORF3 encoding the structural protein VP2 [3]. VP1 is divided into a highly conserved shell (S) domain and a variable protruding (P) domain, which is further subdivided into subdomains P1 and P2. The protruding P2 domain, located at the most exposed surface of the viral capsid, is the major responsible for the antigenicity of the virus and determines interaction with host cell attachment factors (histo-blood group antigen, HBGA) [4,5]. NoV has been classified into five genogroups on the basis of the VP1 sequence (GI to GV) that can be further subdivided into more than 36 genotypes [6]. Recently, a standardization of NoV nomenclature was proposed based on the classification of the VP1 and ORF1 sequences due to the common event of recombination in the hotspot ORF1/ORF2 [7].

Despite the wide genetic diversity of NoV, GII.4 has been described as the most prevalent NoV genotype and has been associated with AGE outbreaks and sporadic cases worldwide since the mid-1990s [8,9,10,11]. Two mechanisms have been proposed to allow the persistence of NoV GII.4 in the population: receptor switching and antigenic variation, which facilitate an expansion of the host range and the escape of new strains from herd immunity [12]. Many studies have shown that the rapid rate of evolution and fixation of amino acid changes in the capsid P2 subdomain might result in altered antigenicity, allowing escape from the host immune responses directed toward the previous circulating variant [10,13,14,15]. This continuous diversification allows the emergence of new epidemic NoV GII.4 lineages in incremental periods, a temporal dynamic comparable to that of influenza A virus [16].

The first GII.4 variant associated with epidemic outbreaks was 95/96US detected in Australia, Europe, and the USA [17,18]. Since this time, six pandemic variants named Farmington
Molecular Characterization

A 10% stool suspension was prepared in 0.01 M Tris-HCl and 0.0015 M Ca\(^{2+}\) (pH 7.2). The RNA extraction was performed using a QIAamp viral RNA Mini kit (Qiagen, Valencia, Calif., USA) according to the manufacturer’s instructions, and the synthesis of complementary DNA (cDNA) was performed using a random primer, pd(N)6 (Amersham Biosciences, UK). The amplification was performed using a set of primers that target the partial region of the genome that encodes the main VP1 capsid protein (subdomain P2), a fragment situated in ORF2 [23]. The amplicons obtained were purified using the QIAquick PCR Purification Kit (QIAGEN, Valencia, CA, USA) following the manufacturer’s recommendations and quantified using a spectrophotometer (Qubit, USA). DNA sequencing was performed using the dideoxynucleotide chain termination method with the ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction Kit 1, v. 3.1 and the ABI Prism 3730 Genetic Analyzer (both from Applied Biosystems, Foster City, CA, USA) by the genomic platform of DNA sequencing PDITS/FIOCRUZ.

Phylogenetic Analyses

The Brazilian sequences obtained in this study were aligned with reference strains of each NoV GII.4 variant and three sequences from Taiwan (HQ456344), Sweden (JN183166), and Japan (AB541348) using the tool CLUSTALW in the MEGA5.2 program [45,46]. A Maximum Likelihood (ML) tree was reconstructed with program PhyML [47] using an online web server [48]. A heuristic tree search was performed using the SPR branch-swapping algorithm, and the reliability of the obtained topology was estimated with the approximate likelihood-ratio test (aLRT) [49] based on a Shimodaira-Hasegawa-like procedure. The trees were visualized using the FigTree v1.3.1 program [50].

Evolutionary and Demographic Analyses

The rate of nucleotide substitution per site per year (subs./site/year), the time to the most recent common ancestor (TMRCA), and the demographic history of NoV GII.4 were jointly estimated using the Bayesian Markov chain Monte Carlo (MCMC) approach implemented in the BEAST v1.7.5 package [51]. The temporal scale of the evolutionary process was directly estimated from the sampling dates of the sequences using the GTR+I+F, nucleotide substitution model, selected using the jModeltest program [52], a relaxed uncorrelated lognormal molecular clock model [53], and a Bayesian Skyline coalescent tree prior [54]. The MCMC analysis was performed for 50 million generations, and the convergence of parameters was assessed by calculating the Effective Sample Size (ESS) using the TRACER v1.5 program after excluding the initial 10% of the run. Uncertainty in the parameter estimates was reflected in 95% HPD intervals. The programs TreeAnnotator v1.7.5 and FigTree v1.4.0 were used to summarize the posterior tree distribution and to visualize the annotated Maximum Clade Credibility (MCC) tree, respectively.

GenBank Accession Numbers

The nucleotide sequences obtained in this study were submitted to the National Center for Biotechnology Information (GenBank, http://www.ncbi.nlm.nih.gov/) and received the following accession numbers: JX975499-JX975614 and KF434555-KF434585.

Results

Identification of NoV GII.4 Variants Circulating in Brazil

An ML phylogenetic analysis of the subdomain P2 of gene encoding VP1 from 147 NoV GII.4 Brazilian samples demonstrated the presence of six variants in different states of the country during the 2004–2012 period (Fig. 1). The samples classified as Den Haag_2006b were grouped into two remarkably separate subclusters, here denominated as “O” (oldest), which comprises 40 samples collected during 2006–2008, and “Y” (youngest), which comprises 24 samples from the period 2009–2011 (Fig. 1).

For most of the years analyzed, it was possible to detect at least two different variants co-circulating in Brazil (Table 1 and Fig. 2), and a clear variation in the frequency of the different variants over time was also observed. The most prevalent variants were Asia_2003 in 2004, Hunter_2004 in 2005, Den Haag_2006b in 2006–2010, Yerseke_2006a in 2008–2009, New Orleans_2009 in 2010–2011, and Sydney_2012 in 2012 (Table 1 and Fig. 2). In addition, the variants Den Haag_2006b, Yerseke_2006a, New Orleans_2009, and Sydney_2012 were detected in all three Brazilian regions evaluated.
Evolutionary History of NoV GII.4 Variants

To reconstruct the time-scale of the NoV GII.4 epidemics, we performed a Bayesian coalescent analysis of partial VP1 sequences of Brazilian origin and reference samples isolated between 1974 and 2012. The mean evolutionary rate of the subdomain P2 was calculated as $7.3 \times 10^{-3} \text{ (5.85} \times 10^{-3}, 8.82 \times 10^{-3}) \text{ subst./site/year.}$

Figure 1. ML tree of P2 subdomain sequences (630 pb) from Brazilian (n = 147) and non-Brazilian strains (n = 15). The black circles represent nodes with aLRT $\geq 0.9$. The non-Brazilian samples are marked in black and with an asterisk. Horizontal branch lengths are drawn to scale with the bar at the bottom indicating nucleotide substitutions per site. doi:10.1371/journal.pone.0092988.g001

Table 1. NoV GII.4 variants detected in Brazil during the period 2004–2012.

| Year | Brazilian region | GII.4 variant |
|------|------------------|---------------|
| 2004 | Southeast        | Asia_2003 (n = 2) |
| 2005 | Southeast        | Asia_2003 (n = 1); Hunter_2004 (n = 5) |
| 2006 | Northeast, South | Hunter_2004 (n = 1); Den Haag_2006b (n = 21) |
| 2007 | Northeast, South | Yerseke_2006a (n = 2); Den Haag_2006b (n = 12) |
| 2008 | Northeast, South | Yerseke_2006a (n = 6); Den Haag_2006b (n = 7) |
| 2009 | Northeast, South | Yerseke_2006a (n = 8); Den Haag_2006b (n = 7); New Orleans_2009 (n = 3) |
| 2010 | Northeast, South | Yerseke_2006a (n = 4); Den Haag_2006b (n = 15); New Orleans_2009 (n = 22) |
| 2011 | Northeast, South | Den Haag_2006b (n = 2); New Orleans_2009 (n = 20) |
| 2012 | Northeast, South | Sydney_2012 (n = 9) |

The number in parentheses denotes the quantity of samples sequenced. doi:10.1371/journal.pone.0092988.t001
Considering this substitution rate, the mean TMRCA of the NoV GII.4 variants were estimated as follows: 2002 (95% HPD: 2001–2003) for Asia_2003; 2003 (95% HPD: 2003–2004) for Hunter_2004; 2005 (95% HPD: 2005–2006) for Yerseke_2006a; 2005 (95% HPD: 2004–2006) for Den Haag_2006b; 2008 (95% HPD: 2008–2009) for New Orleans_2009; and 2011 (95% HPD: 2011–2012) for Sydney_2012 (Fig. 3). Within the Den Haag_2006b variant, the TMRCA of subclusters “O” and “Y” was dated to 2005 (95% HPD: 2005–2006) and 2008 (95% HPD: 2007–2008), respectively.

Demographic History of the NoV GII.4 Den Haag_2006b Variant

Estimations of the demographic history of the NoV GII.4 Den Haag_2006b lineage was next performed using the Bayesian skyline plot method (Fig. 4). This analysis demonstrated that clade Den Haag_2006b experienced a rapid expansion in 2005 after which the population reached a maximum effective population size in early 2006 and then began to decay. It is possible to observe a new and sharp expansion of the effective population size circa mid-2008, which coincides with the TMRCA of subcluster “Y”, and a later stabilization during 2009 and 2010.

Discussion

NoV GII.4 is the most prevalent genotype of NoV is associated with worldwide acute gastroenteritis outbreaks [10,11,14,42,55]. A phylogenetic analysis of subdomain P2 of the VP1 gene allowed the identification of six different variants of NoV GII.4 that were sequentially detected in Brazil during the 2004–2012 periods. In most of the years analyzed, two or three different GII.4 variants were detected simultaneously in Brazil, a pattern comparable to that observed in Japan during the years 2005–2006 [22].

The first variant detected in Brazil in 2004 was Asia_2003, which was described as a recombinant strain, with an ORF1 sequence belonging to GI.12 (Wortley-like) and ORF2 and –3 sequences belonging to GI.4. However, this variant does not feature pandemic characteristics and was demonstrated as only causing outbreaks in Asia, Oceania, and the United Kingdom between 2002 and 2006 [14,30,56,57]. The lineage Hunter_2004 was the next to become dominant in Brazil in 2005 and was associated with global gastroenteritis outbreaks between 2004 and 2005 [58]. The first description of this variant in Latin America was in Paraguayan children also during the years 2004 and 2005 [58].

Two different variants most likely emerged in 2005 and were detected globally in 2006: the Yerseke_2006a and Den Haag_2006b variants that share a common ancestor with Hunter_2004 and Fargminton Hills_2002, respectively. Both variants were detected in all Brazilian regions analyzed, but they spread and died out with very different dynamics. The Den Haag_2006b variant replaced Hunter_2004 in Brazil in 2006 and was the most prevalent lineage during the period of 2006–2008. The Yerseke_2006a variant was first detected in Brazil in 2007 and did not replace the previous variant Den Haag_2006b, though both variants co-circulated at equal proportions during 2008–2009. The detection rate of Yerseke_2006a decreased by 2010, whereas Den Haag_2006b remained detectable until 2011. Another different temporal dynamic was observed in Europe and Oceania, where Yerseke_2006a was the dominant clade in the 2005/2006 season and was replaced by Den Haag_2006b in summer/fall of 2007 [60,61,62].

The variant New Orleans_2009 was first detected to be circulating in Brazil in December 2009. By 2010, this variant reached a frequency that was similar to variants Yerseke_2006a and Den Haag_2006b together, becoming the dominant clade by 2011. In the USA, this variant was first detected during the winter of 2009/2010, causing 60% of all outbreaks associated with NoV and replacing the Den Haag_2006b variant as the predominant GII.4 lineage [23,24].
Figure 3. MCC tree of NoV GII.4 subdomain P2 sequences (630 pb) from Brazilian (n = 147) and non-Brazilian strains (n = 12). The black circles represent nodes with posterior $\geq 0.9$. The non-Brazilian samples are marked in black and with an asterisk.

doi:10.1371/journal.pone.0092988.g003

Figure 4. Bayesian skyline plot of variant Den Haag_2006b. The black line represents the median posterior value, and the blue lines indicate the 95% Highest Probability Density (HPD) intervals. The Y-axis depicts the value of Net on a logarithmic scale.

doi:10.1371/journal.pone.0092988.g004
The variant Sydney_2012 was first detected in 2012 in all of the Brazilian regions evaluated - Northeast, Southeast, and South - replacing New Orleans_2009. A recent report demonstrated this variant as circulating in Northern Brazil, causing sporadic cases in hospitalized children [41]. This strain also became the predominant NoV variant in the USA in 2012/2013, but its emergence was not associated with an increase in the activity of outbreaks when compared to previous seasons [63].

The mean estimated rate of evolution (7.3 x 10^-3 subst./site/year) of subdomain P2 of VP1 in our study was similar to the rates reported in previous studies for the entire VP1 gene (1.9 to 9.0 x 10^-3 subst./site/year) [15,64,65,66]. The mean TMRCA of the six variants detected in Brazil [Asia_2003 (2002), Hunter_2004 (2003), Yerseke_2006a (2005), Den Haag_2006b (2005), New Orleans_2009 (2008), and Sydney_2012 (2011)] demonstrates that the diversification process starts at least one year before the variant becomes a dominant strain and causing epidemics worldwide. Many studies have proposed that a strain circulates at low levels, accumulating mutations for some time before becoming pandemic [62].

The reconstruction of the demographic history of the Den Haag_2006b clade demonstrates that, after a rapid increase in 2005, the population reached a maximum effective size early in 2006 (when this variant was first detected) and then decayed until mid-2008. Since mid-2008, the Den Haag_2006b lineage sharply increased again, reaching a new maximum effective population size in 2009. A recent Argentinian study evaluated the temporal dynamics of this variant and found that the effective population size increased in mid-2004 and started to decrease in 2008 [67]. Such complex demographic patterns of this variant might be related to the existence of two distinct sublineages of Den Haag_2006b: the “Y” sublineage that emerged circa 2008 is a descendant from the “O” sublineage; from most likely emerged circa 2005. The estimated emergence of subcluster “Y” coincides with the second expansion phase of the Den Haag_2006b variant. The identification of a few samples from Taiwan, Sweden, and Japan that grouped with subcluster “Y” suggests that this sublineage is not exclusive to Brazil. It is interesting to note that subcluster “Y” displayed a large number of substitutions at residues 255, 357, 368 and 425 in relation to subcluster “O” (data not shown). The residue 368 is located in putative epitope “A” of the viral capsid that is most likely responsible for the difference in antigenicity among NoV GI.4 variants [68,69], thus suggesting that subclusters “O” and “Y” of the Den Haag_2006b lineage may represent distinct antigenic variants of NoV GI.4.

As suggested by de Rougemont et al. (2011) [70], the greater prevalence of variant Den Haag_2006b could correlate with the binding capability to secretor and nonsecretor populations when compared with Yerseke_2006a strains. In Brazil, there is only one study that characterizes the genetic diversity of the HBGAs types from a small population of descendants of slaves living in semi-isolated rural areas [71]. Further studies including structure modeling of these variants and characterization of HBGAs polymorphisms of Brazilian population to relate the binding properties of NoV should be conducted for clear understanding of predominance of variants GI.4 in the country.

In developed countries, the emergence and replacement of NoV GI.4 variants over the years have pointed to the establishment of NoV surveillance networks as an approach to a better understanding of the epidemiology of this virus and to guide preventive measures against outbreaks [23,32]. In the present study, the temporal dynamics of NoV GI.4, with few exceptions, were similar worldwide, stressing the need for further studies on the continuous epidemiological surveillance of the circulation of NoV genotypes and GI.4 variants in Brazil.

Acknowledgments

The authors thank the staff of the Central Laboratories of the States (LACEN) for sending the clinical specimens.

Author Contributions

Conceived and designed the experiments: JMF MSR MPM. Performed the experiments: JMF MSR. Analyzed the data: JMF GB MV JPGL MPM. Contributed reagents/materials/analysis tools: JPGL MPM. Wrote the paper: JMF GB JGPLM. MPM.

References

1. Glass RI, Parashar UD, Estes MK (2009) Norovirus gastroenteritis. N Engl J Med 361(14): 1776–85.
2. Patel MM, Hall AJ, Vinje J, Parashar UD (2009) Noroviruses: a comprehensive review. J Clin Virol 44(1): 1–8.
3. Prasad BV, Hardy ME, Dondel T, Bella J, Rossmann MG, et al. (1999) X-ray crystallographic structure of the Norwalk virus capsid. Science 286: 287–90.
4. Chen R, Neill JD, Estes MK, Prasad BV (2006) X-ray structure of a native calicivirus: structural insights into antigenic diversity and host specificity. Proc Natl Acad Sci USA 103(21): 8948–53.
5. Cao S, Lou Z, Tan M, Chen Y, Liu Y, et al. (2007) Structural basis for the recognition of blood group trisaccharides by norovirus. J Virol 81(11): 5949–57.
6. Zheng DP, Ando T, Fankhauser RL, Beard RS, Glass RI, et al. (2006) Norovirus classification and proposed strain nomenclature. Virolology 15; 346(2): 312–23.
7. Kroneman A, Veiga E, Vennema H, Vinje J, White PA, et al. (2013) Proposal for a unified norovirus nomenclature and genotyping. Arch Virol 158(10): 2059–68.
8. Noel JS, Faulkner RL, Ando T, Monroe SS, Glass RI (1999) Identification of a distinct common strain of “Norwalk-like viruses” having a global distribution. J Infect Dis 1; 180(5): 1339–44.
9. Lopman B, Vennema H, Kohli E, Pothier P, Sanchez A, et al. (2004) Increase in viral gastroenteritis outbreaks in Europe and epidemic spread of new norovirus variant. Lancet 363(9410): 602–8.
10. Lindemuth LC, Donaldson EF, Lobue AD, Cannon JL, Zheng DP, et al. (2008) Mechanisms of GI.4 norovirus persistence in human populations. PLoS Med 5(2): e31.
11. Ferreira MS, Victoria M, Carvalho-Costa FA, Vieira CB, Xavier MP, et al. (2010) Surveillance of norovirus infections in the state of Rio De Janeiro, Brazil 2005–2008. J Med Virol 82(3): 1442–8.
12. Donaldson EF, Lindemuth LC, Lobue AD, Baric RS (2008) Norovirus pathogenic mechanisms of persistence and immune evasion in human populations. Immunol Rev 225: 190–211.
13. Siebenga JJ, Vennema H, Renckens B, de Bruin E, van der Vreer B, et al. (2007) Epochal evolution of GGII.4 norovirus capsid proteins from 1993 to 2006. J Virol 81(18): 9932–41.
14. Siebenga JJ, Vennema H, Zheng DP, Vinje J, Lee BE, et al. (2009) Norovirus illness is a global problem: emergence and spread of norovirus GI.4 variants, 2001–2007. J Infect Dis 1; 200(5): 802–12.
15. Bull RA, Emsen JS, Rawlinson WD, White PA (2010) Rapid evolution of pandemic noroviruses of the GI.4 lineage. PLoS Pathog 6: 26(5).
16. Koelle K, Coby S, Greenfi L, Pascual M (2006) Epochal evolution shapes the phylogeny of interpandemic influenza A (H1N2) in humans. Science 314: 1898–1903.
17. Wright PJ, Guneserke IC, Doubtje JC, Marshall JA (1998) Small round-structured (Norwalk-like) viruses and classical human caliciviruses in southeast-ern Australia, 1989–1996. J Med Virol 55(4): 312–20.
18. Faulkner RL, Noel JS, Monroe SS, Ando T, Glass RJ (1998) Molecular epidemiology of “Norwalk-like viruses” in outbreaks of gastroenteritis in the United States. J Infect Dis 178(6): 1571–8.
19. Dingle KE, Lambden PR, Caud EO, Clarke IN (1995) Human enteric Caliciviridae: The complete genome sequence and expression of virus-like particles from a genetic group II small round structured virus. J Gen Virol 76: 2349–2355.
20. Widdowson MA, Cramer EH, Hadley L, Bresle JS, Beard RS, et al. (2004) Outbreaks of acute gastroenteritis on cruise ships and on land: identification of a predominant circulating strain of norovirus–United States, 2002. J Infect Dis 1; 190(1): 27–36.
