Norovirus diversity in children with gastroenteritis in South Africa from 2009 to 2013: GII.4 variants and recombinant strains predominate

J. MANS1*, T. Y. MURRAY1, S. NADAN2, R. NETSHIKWETA2, N. A. PAGE2 AND M. B. TAYLOR1

1Department of Medical Virology, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa
2Virology Division, Centre for Enteric Diseases, National Institute for Communicable Diseases, Sandringham, South Africa

Received 22 July 2015; Final revision 25 August 2015; Accepted 26 August 2015; first published online 16 September 2015

SUMMARY

From 2009 to 2013 the diversity of noroviruses (NoVs) in children (≤5 years) hospitalized with gastroenteritis in South Africa was investigated. NoVs were genotyped based on nucleotide sequence analyses of partial RNA-dependent RNA polymerase (RdRp) and capsid genes. Seventeen RdRp genotypes (GI.P2, GI.P3, GI.P6, GI.P7, GI.P not assigned (NA), GI.Pb, GI.Pf, GII.P2, GII.P4, GII.P7, GII.P13, GII.P16, GII.Pc, GII.Pe, GII.Pg, GII.PNA) and 20 capsid genotypes (GI.1, GI.2, GI.3, GI.5, GI.6, GI.7, GI.NA, GII.1, GII.2, GII.3, GII.4, GII.6, GII.7, GII.10, GII.12, GII.13, GII.14, GII.16, GII.17, GII.21) were identified. The combined RdRp/capsid genotype was determined for 275 GII strains. Fifteen confirmed recombinant NoV strains circulated during the study period. NoV GII.P4/GII.4 (47%) and GII.Pe/GII.4 (18%) predominated, followed by GII.PNA/GII.3 (10%) and GII.P21/GII.3 (7%). Other prevalent strains included GII.Pg/GII.12 (6%) and GII.Pg/GII.1 (3%). Two novel recombinants, GII.Pg/GII.2 and GII.Pg/GII.10 were identified. In 2013 the replacement of GII.4 New Orleans 2009 and GII.P21/GII.3, which predominated during the early part of the study, with GII.4 Sydney 2012 and GII.PNA/GII.3 was observed. This study presents the most comprehensive recent data on NoV diversity in Africa.

Key words: Norovirus diversity, norovirus recombinants, paediatric, South Africa, Sydney 2012.

INTRODUCTION

Norovirus (NoV) is a major cause of severe gastroenteritis in young children [1]. Since the introduction of the rotavirus vaccine, NoV is reported to be the predominant pathogen causing hospitalization of children with gastroenteritis in the United States [2] and Nicaragua [3]. NoV, a small RNA virus from the Caliciviridae family, has a single-stranded, positive-sense genome which is comprised of three open reading frames (ORFs). ORF1 encodes a large polyprotein which is processed in infected cells to produce the non-structural proteins including the RNA-dependent RNA polymerase (RdRp). ORF2 encodes the major capsid protein (VP1) and ORF3 the minor capsid protein (VP2) [4]. Currently NoVs are classified into six genogroups of which GI, GII and GIV infect humans. The genogroups are further subdivided into at least nine (GI), 22 (GII) and two (GIV) genotypes,
respectively based on analysis of the complete capsid gene. The globally dominant genotype causing outbreaks and sporadic cases of gastroenteritis is genogroup II, genotype 4 (GII.4) [5]. NoVs are characterized by high genetic diversity and frequent recombination between genotypes. GII.4 strains evolve rapidly and a new variant arises every 2–3 years which then spreads across the world and replaces the previously dominant GII.4 strain. The most recent GII.4 variant was described in Sydney in 2012 and is currently dominant globally [6].

In 2008, NoVs were detected in 14% of children hospitalized with gastroenteritis at two hospitals in the Gauteng province of South Africa (SA) [7]. The Apeldoorn 2007 and New Orleans 2009 GII.4 variants were then described for the first time in SA. Other studies in Africa have reported a wide range of NoV prevalence from 0.8% in Morocco [8] to 25.5% in Nigeria [9] as well as a predominance of GII.4 strains [7, 10, 11]. To understand the true molecular epidemiology of NoVs, including the prevalence of recombinant strains, characterization of both the RdRp and capsid regions is necessary [12]. Recently Kroneman and co-workers [13] proposed a nomenclature system for NoVs which combines the RdRp and capsid genotypes if both are available, e.g. GII.Pg/GII.12. To date few studies in Africa [10, 14] have provided data on both genomic regions and the prevalence of recombinants. In this study the diversity of NoVs in SA, based on both RdRp and capsid regions, in children aged ≤5 years suffering from severe gastroenteritis was investigated.

METHODS

Study population

From April 2009 to December 2013 stool specimens were collected from children aged ≤5 years, hospitalized with severe gastroenteritis, as part of the South African Rotavirus Sentinel Surveillance Programme (RSSP). Seven hospitals (Chris Hani Baragwanath, Dr George Mukhari, Ngwelazane, Edendale, Mapulaneng, Matikwane, and the Children’s Red Cross Hospital) located in urban, peri-urban and rural areas in four provinces of SA (Gauteng, KwaZulu-Natal, Mpumalanga, and the Western Cape) were included in the study. NoV detection was performed at the Virology Division, Centre for Enteric Diseases, National Institute for Communicable Diseases using real-time RT–PCR (N. A. Page et al., unpublished data).

Ethical approval

Ethical approval for the RSSP was obtained from the ethics committees of the University of the Witwatersrand (M091018), University of Limpopo (MREC/P/10/2009), University of Cape Town (068/2010) and the University of KwaZulu-Natal (BF074/09). The RSSP protocol included testing of stool specimens for other viral gastroenteritis pathogens, such as NoV, as well as subsequent characterization of the identified viruses.

Stool specimen preparation and nucleic acid extraction

Ten percent stool suspensions of the NoV-positive specimens were submitted to the Department of Medical Virology, University of Pretoria for genotyping and stored at −20 °C until nucleic acid extraction. Total nucleic acids were extracted as described previously [15] and stored in aliquots at −70 °C.

Genotyping of NoVs

NoVs were genotyped based on amplification and nucleotide sequencing of the partial RdRp gene (region A) [15] and the partial capsid gene (region C) [16]. The RdRp (320 bp) and capsid (320 bp) gene sequences were submitted to the online Norovirus Genotyping Tool (http://www.rivm.nl/mpf/norovirus/typingtool/) to assign genotypes [17]. All GII.4 strains that could not be assigned a variant type by the typing tool were analysed with BLAST (Basic local alignment search tool) [18]. Virus strains with >99% identity to GII.4 variants in GenBank, over the partial capsid region, were assigned to these variants. The complete capsid gene nucleotide sequence was determined for selected strains as described [15].

Phylogenetic analyses

Nucleotide sequences were aligned using the online MAFFT server (http://mafft.cbrc.jp/alignment/server/index.html). Phylogenetic analyses were conducted in MEGA6 [19] using the Neighbour-joining method with the Kimura two-parameter model and 1000 bootstrap replicates.

Recombination analyses

Discrepant genotypes obtained for RdRp and capsid sequences amplified from a single specimen indicated possible recombination. The polymerase/capsid overlap
region was amplified as described by Mans and co-workers [15]. Putative breakpoint analysis was performed using SimPlot version 3.5.1 (http://sray.med.som.jhmi.edu/SCRoftware/simplot/) and the maximum \( \chi^2 \) test as implemented in RDP version 4.33 [20].

**Statistical analyses**

A \( t \) test was used to determine significant differences in mean Ct values of GII.4 and non-GII.4 infections and association between GII.4 infections and age, gender and HIV status was assessed using 2 \( \times \) 2 tables (Fisher’s exact test) (http://www.openepi.com).

**Nucleotide sequence accession numbers**


**RESULTS**

Stool specimens collected from children hospitalized with gastroenteritis in SA over a 5-year period (2009–2013) from seven surveillance sites were screened for NoV and 14·1% (837/5950) of specimens were positive. Specimens from five of the seven surveillance sites were screened for viruses, bacteria and parasites (N. A. Page et al., unpublished data).

Of the 837 NoV-positive specimens, 54 NoV GI- and 350 NoV GII-positive specimens were available for genotyping. NoV GI-positive specimens, collected between 2011 and 2013, were genotyped whereas NoV GII strains collected between 2009 and 2013 were characterized. In total, seven polymerase-based (16 typed specimens) and seven capsid-based (27 typed specimens) GI genotypes were identified (Fig. 1, Supplementary Table S1). The polymerase and capsid genotype of 11 GI strains could be determined (GI.P2/GI.2, GI.P3/GI.3, GI.P6/GL.6, GI.P7/GL.7, GI.P/GI.3, GI.PNA/GI.NA). GI.3 strains predominated overall, followed by GI.7. Since the overlap region between ORF1 and ORF2 could not be amplified, the GI.PI/GI.3 combination was not confirmed. Of the 32 genotyped NoV GI strains, four were found in single infections. The capsid genotypes in these single infections included GI.5, GI.7 and GI.NA.

For NoV GII, 10 polymerase-based (292/350 typed) and 13 capsid-based (312/350 typed) genotypes were characterized (Fig. 2a, b, Supplementary Table S1). The strains from 312 specimens that could be typed in the capsid region comprised of 46 single NoV GII infections, nine mixed GI + GII + bacteria and/or parasite infections and 257 infections with NoV GII + bacteria and/or parasites. The capsid genotypes in the single GII infections were 69·6% (32/46) GI.4, 17·4% (8/46) GII.3, 6·5% (3/46) GII.2, 4·3% (2/46) GII.12 and 2·2% (1/46) GII.13. Combined NoV GII polymerase and capsid genotyping data was obtained for 275 specimens (Fig. 2c, Supplementary Table S2). NoV GII.4 strains represented 61% of the typed GII strains. Based on the partial capsid gene the GII.4 strains comprised of Osaka 2007 (10/191), New Orleans 2009 (96/191), Sydney 2012 (57/191) and several unassigned GII.4 variants (28/191). The monthly distribution of the GII.4 variants between 2009 and 2013 in SA is shown in Figure 3a. From 2009 to 2012 the New Orleans 2009 strain predominated, with unassigned GII.4 variants circulating at low frequencies. In 2011, the Osaka 2007 variant was detected in 6 months of the year. The Sydney 2012 variant was first detected in May and July of 2010, then again in November of 2011, after which it was
Fig. 2. Prevalence of norovirus GII (a) RdRp and (b) capsid genotypes detected in children with gastroenteritis from 2009 to 2013 in South Africa. (c) Prevalence of polymerase-capsid combinations detected in 275 specimens from children with gastroenteritis from 2009 to 2013 in South Africa.
Fig. 3. (a) Monthly distribution of norovirus GII.4 variants, based on partial capsid sequence, from 2009 to 2013. (b) Neighbour-joining phylogenetic analysis of the complete nucleotide sequence (1623 bp) of 39 representative GII.4 variants detected from 2009 to 2013 in South Africa. Significant bootstrap support (>70%) is indicated. Norovirus reference strains are represented by their GenBank accession numbers. # Strain 11939 has a GII.P4 RdRp. * Pre-epidemic Sydney 2012 strain. The following symbols indicate the year of detection: 2009 (●), 2010 (▲), 2011 (■), 2012 (◇), 2013 (○).
detected in all months of 2012 except March and July until it essentially replaced the New Orleans 2009 variant in 2013 (Fig. 3a). The GII.4 strains predominated across all surveillance sites (Supplementary Fig. S1).

The 1623 bp capsid gene sequence of 39 GII.4 variants detected over the course of the study was determined. The phylogenetic relationship between the identified variants and GII.4 reference strains is shown in Figure 3b. The New Orleans 2009 strains formed two distinct clusters, with the majority of earlier strains detected from 2009 to 2010 clustering separately from the GU445325 type strain, whereas strains detected in 2011 and 2012 were more closely related to the type strain. One strain detected during 2013 is located on a separate branch indicating significant divergence from the original New Orleans type strain.

The Sydney 2012 variant circulated in two forms in SA during the study period. The first strains detected in 2010 (Johannesburg_5520, Johannesburg_5935) and at the end of 2011 (Johannesburg_8429), were GII.P4 New Orleans 2009/GII.4 Sydney 2012 recombinants. The GII.Pe/GII.4 Sydney 2012 strain, first detected in February 2012, co-circulated with the New Orleans 2009/GII.4 Sydney 2012 recombinant during 2012 and predominated from September 2012 onwards. Interestingly, in the phylogenetic tree the Sydney 2012 capsid genes clustered according to their combination with either the New Orleans 2009 or GII.Pe polymerase types, except for a GII.P4 NA/GII.4 Sydney 2012 strain (Empangeni_11939) detected in 2013, which clustered with the GII.Pe recombinants (Fig. 3b). The unassigned RdRp gene of Empangeni_11939 is 97% identical over 324 nucleotides to a New Orleans 2009 variant (KF429778) detected in the United States in 2012. Two types of unassigned GII.4 variants, grouping in the Hunter 2004 or Apeldoorn 2007 lineages, were identified in the study (Fig. 3b).

The second most prevalent capsid type identified was GII.3. This capsid was found in combination with an unassigned polymerase (28/54), GII.P21 (18/54) and GII.P16 (2/54) and 11% (6/54) of GII.3 specimens could not be typed in the RdRp region. In the early part of the study, from 2009 to February of 2011, the GII.P21/GII.3 combination predominated. Thereafter only a few cases of this combination were detected up to November 2012, after which the GII.PNA/GII.3 combination increased in prevalence during 2013. BLAST analysis of the partial capsid region of the GII.3 strains show that they exhibit high nucleotide identity (97–99%) to GII.3 strains detected globally (results not shown). Phylogenetic analysis indicated that diverse GII.3 capsids circulated in SA, with the GII.PNA/GII.3 strain capsid regions clustering separately from the GII.3 sequences associated with GII.P21 (Fig. 4).
The GII.Pg/GII.12 recombinant was the next most prominent strain detected, representing 5.8% (16/275) of the NoV strains with combined typing data (Fig. 2c). In total, 19 NoV GII RdRp/capsid combinations with different genotypes in the RdRp and capsid regions were identified. Fifteen of these combinations have been confirmed by amplification of the overlap region, four in the present study and 11 in a previous study [15]. The GII.P2/GII.2, GII.P4/GII.4 and GII.P21/GII.21 were the only non-recombinant strains characterized in both genomic regions. The number of genotypes detected per site varied from four (Pietermaritzburg) to 11 (Bushbuckridge and Cape Town) (Supplementary Fig. S1).

Analysis of specimens collected during 2013 identified an additional four recombinant types within GII, not identified previously in SA. Phylogenetic analysis grouped the partial RdRp and capsid regions of each strain into different genotypes, suggesting recombination (data not shown). Subsequent maximum \( \chi^2 \) and SimPlot analysis of a 1090 bp region, spanning the polymerase and capsid typing regions, indicated that the recombination breakpoints \( (P < 0.05) \) were at the ORF1/2 junction (Fig. 5). The newly identified combinations were GII.Pg/GII.2, GII.Pg/GII.10, GII.P7/GII.14 and GII.P21/GII.13. SimPlot and RDP analysis could not be performed for GII.P14 polymerase sequence data associated with the GII.14 capsid are available. Several additional RdRp/capsid combinations, namely GII.Pc/GII.2, GII.PNA/GII.16, GII.P13/GII.17 and GII.P16/GII.3 were detected but remain putative because the overlap region could not be amplified from these strains.

A statistically significant difference was observed between the mean real-time RT–PCR \( C_t \) values of GII.4 strains [22.4, 95% confidence interval (CI) 21.6–23.2, interquartile range (IQR) 18.0–26.1] compared to non-GII.4 strains (25.5, 95% CI 24.5–26.6, IQR 20.4–30.2) \( (P < 0.001) \). This indicates a nearly tenfold higher viral load in GII.4-positive stool specimens compared to other genotypes. No association was found between GII.4 and age, gender or HIV status compared to non-GII.4 genotypes \( (P > 0.05) \).

**DISCUSSION**

In this study, which spanned 5 years and included four provinces of SA, a great diversity of NoV genotypes was observed in paediatric patients with severe gastroenteritis. Overall 17 RdRp types and 20 capsid types were characterized. Despite this remarkable diversity, the GII.4 and GII.3 capsid genotypes predominated and the replacement of one variant (GII.4) or cluster...
(GII.3) by another was observed over time. Both GII.3 and GII.4 have been described as prevalent NoV genotypes in children [5].

The GII.3 capsid genotype is quite diverse and three sub-clusters (GII.3a, b, c) have been defined [12]. The diversity of GII.3 strains is enhanced by recombination with a variety of RdRp genotypes such as GII.P4, GII.P12, GII.P16, GII.P21, GII.P22, GII.Pa and GII.Pe [21–24]. The GII.P21/GII.3 strain (formerly known as GII.b/GII.3) has been reported widely [25, 26]. The majority of these SA strains cluster with the GII.3c reference strain in the capsid region. The GII.P16/GII.3 recombinant was first detected in infants in Bangladesh [27] and subsequently in children in Italy [28]. Therefore, although the overlap region of the SA GII.P16/GII.3 strains could not be amplified, it is likely that they are recombinant. A novel recombinant between an unassigned RdRp and GII.3b-like capsid strain (Bushbuckridge_6387) was identified in SA recently [15]. In 2011, this type of RdRp was first identified in three SA provinces, a related but not identical strain subsequently emerged and in 2013 was detected in Gauteng, KwaZulu-Natal and Mpumalanga. It remains to be seen whether this recombinant is circulating in other parts of Africa or on other continents.

In contrast with reports from other parts of the world [29, 30] the number of NoV-positive cases in SA did not increase in 2012 (14%) compared to 2011 (16%) (N. A. Page et al., unpublished data). In SA, the GII.Pe/GII.4 Sydney 2012 variant started predominating in January 2013 and completely replaced New Orleans 2009 by October 2013. This replacement was delayed compared to other studies where Sydney 2012 replaced New Orleans 2009 as the predominant strain causing outbreaks or sporadic gastroenteritis towards the end of 2012 and in the first few months of 2013 in Oceania [6], parts of Europe [31, 32] and China [29]. Analysis of the complete capsid gene sequences of the SA Sydney 2012 strains revealed that two capsid types circulated in the population. The strains detected in the early part of the study cluster separate from the Sydney 2012 type strain (JX459908), and all have a GII.P4 RdRp gene. Auckland 2010 (KF060124), a strain detected in New Zealand possibly representing a pre-epidemic version of Sydney 2012 [6], is 95–96% identical to the capsid gene of these SA Sydney 2012 strains identified between 2010 and 2012 and is the most closely related strain in GenBank. Thus the SA capsid sequences may represent another pre-epidemic form of Sydney 2012. Although the majority of these strains circulated from 2010 to 2012, one GII.P4 New Orleans 2009/GII.4 Sydney 2012 strain (Cape Town_12544) was detected in October 2013, indicating continuing circulation of the strain at low levels. Essentially, three subgroups of Sydney 2012 circulated in SA, the type strain-related GII.Pe/GII.4 Sydney 2012, GII.P4/GII.4 pre-epidemic Sydney 2012 and GII.P4 New Orleans 2009/GII.4 Sydney 2012.

The unassigned Hunter 2004-like SA strains are related to NoV strains (95–96% identity in the capsid gene) circulating in Tunisia from 2003 to 2004. The Apeldoorn 2007-like SA strains exhibit 96% nucleotide identity in the capsid gene to NoVs circulating in Asia (AB541322) [33, 34] during 2008. These strains appear to be minor GII.4 variants unique to SA and circulated at low levels.

By 2007, 20 NoV recombinant types had been described [35], and since then many other RdRp/capsid combinations have been reported [23, 27, 28, 36, 37]. Fifteen confirmed NoV recombinants have circulated in SA between 2009 and 2013 (the present study and [15]). The recombination breakpoints of all recombinant types were located at the ORF1/2 junction, as has been described for the majority of NoV recombinants [35]. Two of the recombinant combinations characterized in the present study, GII.Pg/GII.2 (Bushbuckridge_12257) and GII.Pg/GII.10 (Johannesburg_12243), have not been reported to date and the most closely related NoV strains in GenBank exhibited 93% nucleotide identity over the 1090 bp RdRp/capsid overlap region. The GII.P2/GII.2, GII.P21/GII.2, GII.Pg/GII.1 and GII.Pg/GII.12 strains were some of the more prevalent strains detected during the study, presenting frequent opportunities for co-infection and subsequent recombination. The RdRp genotypes encountered most frequently in the SA NoV recombinants were GII.Pg (four different types), GII.P21 (three different types) and GII.P7 (two different types). The GII.Pg and GII.P21 RdRp genotypes have been reported in combination with multiple capsid genotypes and therefore appear to recombine more frequently than many other NoV RdRp genotypes.

The majority of NoV infections were detected in combination with other enteric pathogens, with no difference in genotype distribution between mixed and single infections. Therefore it is difficult to assess the contribution of NoVs to the gastroenteritis symptoms observed in the children. GII.4 strains, however, were shed at higher concentrations in stool compared to non-GII.4 strains, as measured by real-time detection Ct values, in agreement with a NoV study in children in Nicaragua [38].
These data provide the most comprehensive analysis of NoV diversity in paediatric patients with severe diarrhoea in the post-rotavirus vaccination era in SA and underscores the importance of GII.4 and GII.3 NoV strains in causing disease in this population. In addition, the ability of several polymerase types to recombine with a variety of capsid types, thereby creating novel recombinants, was highlighted. Since this study focused on children with severe diarrhoea, it is likely that even greater NoV diversity is present in the broader population.

SUPPLEMENTARY MATERIAL
For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0950268815002150.

ACKNOWLEDGEMENTS
The authors acknowledge the staff and participants of the South African Rotavirus Sentinel Surveillance Programme. This work was supported by the Poliomyelitis Research Foundation of South Africa (Research Grant 12/13). Dr J Mans was funded by a Post-doctoral Fellowship (2011–2012) and Research Fellowship (2013–2014) from the University of Pretoria and Dr T. Y. Murray was funded by the Poliomyelitis Research Foundation of South Africa (PhD Bursary). The South African Rotavirus Sentinel Surveillance Programme was funded through a grant from GlaxoSmithKline.

DECLARATION OF INTEREST
None.

REFERENCES


