Genetic diversity of norovirus in children under 5 years of age with acute gastroenteritis from Angola

A. Esteves1,*, J. Nordgren2,*, C. Tavares1, F. Fortes3, R. Dimbu3, N. Saraiva3 and C. Istrate1

*These authors contributed equally to this work.


Received: 11 October 2017
Revised: 11 January 2018
Accepted: 5 February 2018

Key words: Acute gastroenteritis; Angola; children; norovirus

Author for correspondence: Claudia Istrate, E-mail: claudiaistratedvm@gmail.com

Norovirus (NoV) is a major cause of acute gastroenteritis (AGE). In this study, we investigated the genetic diversity of NoV strains identified in children under 5 years of age with AGE in four provinces of Angola. Faecal samples from 343 children were screened for NoV by an in house real-time PCR assay and genotyping was performed by partial capsid gene sequencing.

NoV was detected in 17.4% (58/334) of the samples, with high detection rates in children <6 months old (19%) and in children aged 12–24 months (23%). Genotype diversity was large, as demonstrated by the 11 identified genotypes. GII.4 was the predominant genotype (20% of all NoV-positive samples), followed by GII.6 (15%), GI.3 (12%), GII.7 (10%) and by other genotypes to a lesser extent. Two GII.4 variants, New Orleans 2009 and Sydney 2012, were detected and several genetic clusters were observed for genotypes GL3, GIL6 and GIL7. The present study shows high detection rates and genetic diversity of circulating NoV genotypes in paediatric AGE samples from Angola. This information emphasises the importance of continuous assessment of NoV burden and evolution in the target population.

Introduction

Norovirus (NoV) is recognised as the second most common cause of viral diarrhoea in children under-5, accounting for approximately 18% of acute gastroenteritis (AGE) cases worldwide in this age group [1]. Furthermore, in countries having introduced infant vaccination for rotavirus, NoV has become the most common cause of paediatric AGE [2, 3]. In sub-Saharan Africa, a region with a high diarrhoeal disease burden, there is still a lack of comprehensive epidemiological studies regarding NoV [4–6].

Angola is a sub-Saharan African country, located on the South Atlantic coast of West Africa, between Namibia and the Republic of Congo. The country registered 25.8 million inhabitants in 2014, distributed throughout the 18 provinces. According to the recent data, the young population (0–14 years old) represent 47% [7]. The World Health Organization (WHO) estimates in 2015 showed that Angola has one of the highest under-5 death rate in the world (194 deaths per 1000 live births) and a high amount of these deaths is attributed to diarrhoea (20 000 registered deaths in 2010 indicating approximately 0.5% of children under 5 died due to diarrhoea) [8]. The present study, performed during 2012–2013, describe the period prior to rotavirus vaccine introduction in the National Immunization Programme in April 2014.

NoVs are non-enveloped, single-stranded, positive-sense RNA viruses within the Caliciviridae family, with a 7.5-kb genome comprising three open reading frames (ORFs). The first ORF (ORF1) encodes non-structural proteins, including the RNA-dependent RNA polymerase; ORF2 encodes the viral capsid protein (VP1); and ORF3 encodes a minor structural protein (VP2). The NoV genus comprises seven genogroups (GI to GVII), of which GI is the most common genogroup infecting humans, but GI and, to a lesser extent, GIV also infect and cause disease in humans [9]. Through phylogenetic analysis of the capsid gene, 22 genotypes have been recognised within GH and nine within GI [9, 10]. Among all NoV genotypes, GIL4 is the single most common genotype infecting humans worldwide, associated with approximately 60% of all reported NoV outbreaks and 70% of sporadic NoV AGE in children [11, 12]. In addition, the GIL4 genotype has been associated with more severe symptoms [2, 13]. Significant advances in understanding NoV strain evolution and diversity have been observed globally in the last decade. Regional surveillance networks are established on many continents, with exception of the African continent where only a few surveillance studies regarding NoV have been performed till date.
In this context, prior to rotavirus vaccine introduction in Angola, we investigated the detection rate and distribution of circulating NoV strains in children with AGE living in four Angolan provinces.

**Material and methods**

**Study population and sampling**

The study was conducted in the Angolan provinces of Cabinda, Luanda, Huambo and Zaire, in two periods, between June and August 2012 and between September and October in 2013, as described in a previous report [14]. These periods correspond mostly to the dry season, which usually lasts from May until September. Due to logistic reasons, a full year sampling could not be performed. Considering reports from other sub-Saharan African countries, the dry season has been associated with higher virus-caused diarrhoea burden [15, 16]. The samples from the Huambo province were collected during June 2012 in six locations: the hospitals of Alto Hama (n = 26), Bailundo (n = 43) and Caala (n = 100), and the health care centres of Calenga (n = 9), Casseque III (n = 11) and Mineira (n = 57). The samples from the Zaire province were collected from July to August 2012 in two health care facilities, the hospital (n = 12) and the local health care centre of Mbanza Kongo (n = 13). During September and October 2013, samples were collected from the Luanda province at the ‘Paz’ health care centre (n = 22) and from the northern province of Cabinda at the Chinga hospital (n = 41) (Fig. 1).

During the sampling periods, all children under 5 years of age attended with acute diarrhoea at the above-mentioned health care centres and hospitals were enrolled. Faecal samples from a total of 334 children were collected in specific containers and stored at −20 °C until used for RNA extraction.

**Clinical and epidemiological data**

Information regarding clinical characteristics such as symptoms of AGE (i.e. diarrhoea (n = 334), vomiting (n = 172), body temperature (n = 160), date of symptom onset (n = 334)) and demographics (i.e. gender (n = 256) and date of birth (n = 182)) was collected when possible, using a standard case report. Acute diarrhoea was considered as inclusion criteria and defined as three or more passages of watery/loose stools within a 24 h period.

**Ethics statement**

The study was approved by the National Ethics Committee in Luanda, Angola, and the Ethics Committee of the Institute of Hygiene and Tropical Medicine in Lisbon, Portugal (Process identification: 5-2012-PN). Each parent or legal guardian of the children to be enrolled in the study was informed about the objectives of the study, the type of sampling needed and voluntary nature of the study. Written informed consent was obtained from each parent or legal guardian before participating in the study.

**NoV detection and genotyping**

NoV detection was performed by an in-house real-time PCR assay that allows simultaneous identification of NoV genogroups GI and GII as described previously [16]. Ten per cent suspensions of faecal samples were prepared in phosphate buffer saline and used for RNA extraction with the innuPREP Virus RNA kit (Analytik Jena, Jena). cDNA was synthesised by reverse transcription using random hexamers and the NZY First-Strand cDNA Synthesis kit (NZYTech, Lisbon). Multiplex TaqMan real-time PCR was carried out on the cDNA with the SensiFAST™ Probe Lo-Rox kit (Bioline, London) and two NoV genogroup-specific primer sets and TaqMan® probes [16, 17] on a 7500 Fast Real-Time PCR System (Applied Biosystems, Carlsbad, California). A Ct-value >40 was considered negative.

For NoV genotyping, a fragment of the capsid N/S domain was amplified by PCR using primer sets NVGIf1b/G1SKR and G2SKF/G2SKR for GI and GII strains, respectively [17, 18]. PCR products with the expected size were purified from agarose gels with the Zymoclean™ Gel DNA Recovery kit (Zymo Research, Irvine, California), and sequenced by STABVIDA (Costa da Caparica, Portugal).

---

**Fig. 1.** Geographical location of the study sites, hospitals (Chinga, Mbanza Kongo, Alto Hama, Bailundo and Caala) and health care centres (Mbanza Kongo, Paz, Calenga, Caseque III, Mineira) pointed out by filled triangles within the four provinces (Cabinda, Zaire, Luanda and Huambo) of the country participating in the study and highlighted with light grey. The number of samples collected in each study site was indicated.
**Sequence analysis**

Nucleotide sequences were manually edited with the BioEdit Sequence Alignment Editor, version 7.2.5 (Ibis Biosciences, Carlsbad, USA) and submitted to BLAST sequence similarity search (http://blast.ncbi.nlm.nih.gov/Blast.cgi). NoV strains were further analysed using the NoV typing tool available through Noronet (http://www.rivm.nl/mpf/norovirus/typingtool), where they were assigned to a genotype or GI.4 variant. Phylogenies of NoV strains, as well as assignment of GI.4 variants that was not possible using the NoV typing tool, were inferred from the analysis of multiple sequence alignments of Angolan strains and reference virus genotypes selected from the GenBank database (http://www.ncbi.nlm.nih.gov) generated with the ClustalW algorithm from the MEGA 6.0 package. Maximum Likelihood (ML) trees were constructed with MEGA 6.0 using Tamura–Nei model. The reliability of the inferred topologies was evaluated by bootstrap analysis of 1000 pseudo-replicates of the original sequence data.

**Nucleotide sequence accession number**

Sequences obtained in this study were deposited in the GenBank database under accession numbers KT382965 to KT382998 and KX816577 to KX816585.

**Statistical analysis**

Data were analysed using IBM SPSS software, version 22 (IBM Corp, Armonk, New York). Fischer’s exact test was used to analyse differences between categorical data. A two-tailed P-value <0.05 was considered significant.

**Results**

**NoV was detected in nearly one-fifth of the diarrhoeal samples**

Among the 334 stool samples from children with AGE, NoV RNA was detected in 58 samples, corresponding to a detection rate of 17.4%. The detection range varied between the four provinces of the study: Luanda (22.7%), Huambo (18.4%), Cabinda (14.6%) and Zaire (8.0%). The epidemiological profile of paediatric NoV diarrhoea in Angola was assessed only in children with filled in epidemiological inquiry. Fever and vomiting were reported for 64.3% of the NoV-positive children, i.e. 18 out of 28 for whom this clinical information was available. The frequency of detection of NoV was high both in children <6 months of age and children over 1 year of age (Table 1), with children aged 7–12 months significantly less likely to be infected with NoV compared with the other age groups (P < 0.05, Fischer’s exact test). Rotavirus coinfection was detected in 36.2% (21/58) of the NoV cases with rotavirus detection rate previously observed (exact test). Rotavirus coinfection was detected in 36.2% (21/58) of the NoV cases with rotavirus detection rate previously observed (exact test). Rotavirus coinfection was detected in 36.2% (21/58) of the NoV cases with rotavirus detection rate previously observed (exact test). Rotavirus coinfection was detected in 36.2% (21/58) of the NoV cases with rotavirus detection rate previously observed (exact test). Rotavirus coinfection was detected in 36.2% (21/58) of the NoV cases with rotavirus detection rate previously observed (exact test). Rotavirus coinfection was detected in 36.2% (21/58) of the NoV cases with rotavirus detection rate previously observed (exact test).

**High diversity of circulating NoVs**

Concerning the capsid region analysed, a large genetic diversity of circulating NoV strains was observed (Table 2). Overall, 46 of the 58 NoV-positive samples could be genotyped, being assigned to 11 genotypes: GI.3, GI.5, GI.1, GI.2, GI.3, GI.4, GI.6, GI.7, GI.9, GI.14 and GI.16. The majority of the NoV strains were classified as genogroup II (80%). The GI.4 was the predominant genotype (20% of all NoV-positive samples), followed by GI.6 (15%), GI.3 (12%), GI.7 (10%) and the other genotypes to a lesser extent.

The GI strains, assigned to two genotypes GI.3 and GI.5, were detected in two provinces, Huambo and Luanda. Seven of the GI.4 genotypes were classified as the New Orleans 2009 variant and four as the Sydney 2012 variant, whereas one GI.4 genotype could not be assigned a variant due to short sequencing reads (<220 bp). The GI.4 genotypes generally infected younger children and were more associated with fever and vomiting than other genotypes (data not shown), but the number of GI.4 infections was low for a reliable statistical analysis. The GI.4 New Orleans 2009 viruses were only detected in Huambo province during 2012, while the GI.4 Sydney 2012 strains were present in all districts during both years of study. The GI.6 and GI.7 strains, the second and third most abundant, originated in

**Table 1. Epidemiological profile of paediatric norovirus diarrhoea in children ≤5 years of age from different provinces of Angola (June–August, 2012 and September–October, 2013)**

<table>
<thead>
<tr>
<th>Age range (months)</th>
<th>Total diarrhoea cases*</th>
<th>NoV-positive (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–6</td>
<td>62</td>
<td>12 (19)</td>
</tr>
<tr>
<td>7–12</td>
<td>65</td>
<td>7 (11)</td>
</tr>
<tr>
<td>13–24</td>
<td>43</td>
<td>10 (23)</td>
</tr>
<tr>
<td>&gt;24</td>
<td>12</td>
<td>3 (25)</td>
</tr>
</tbody>
</table>

*Only considering the cases with epidemiological inquiries filled in.

**Table 2. Norovirus strain diversity in children with AGE in four provinces of Angola**

<table>
<thead>
<tr>
<th>NoV type</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genogroup I*</td>
<td>12 (20)</td>
</tr>
<tr>
<td>GI.3*</td>
<td>7 (12)</td>
</tr>
<tr>
<td>GI.5</td>
<td>1 (2)</td>
</tr>
<tr>
<td>GI.6</td>
<td>4 (7)</td>
</tr>
<tr>
<td>Genogroup II*</td>
<td>47 (80)</td>
</tr>
<tr>
<td>GI.1</td>
<td>3 (5)</td>
</tr>
<tr>
<td>GI.2</td>
<td>1 (2)</td>
</tr>
<tr>
<td>GI.3*</td>
<td>3 (5)</td>
</tr>
<tr>
<td>GI.4*</td>
<td>12 (20)</td>
</tr>
<tr>
<td>GI.6</td>
<td>9 (15)</td>
</tr>
<tr>
<td>GI.7</td>
<td>6 (10)</td>
</tr>
<tr>
<td>GI.9</td>
<td>1 (2)</td>
</tr>
<tr>
<td>GI.14</td>
<td>2 (3)</td>
</tr>
<tr>
<td>GI.16</td>
<td>1 (2)</td>
</tr>
<tr>
<td>GI.8</td>
<td>9 (15)</td>
</tr>
</tbody>
</table>

*aIncluding one mix GI.3/GI.3 infection observed in one child.

*bNot genotyped.

*Belonging to two variants, New Orleans and Sydney.
Huong and Cabinda. The highest number of samples, also presenting the largest diversity of circulating GII genotypes originated from Huambo province. Genotypes GII.1, GII.2, GII.14 and GII.16 were only detected in this province. GII.3 genotypes were detected in Huambo and Luanda provinces, and the only genotype GII.9 identified in this study was detected in Luanda.

Nine NoVs detected as genogroup II strains and four as genogroup I could not be genotyped, possibly due to sample degradation, low viral load (Ct>31.6 for GI and Ct>36.5 for GII) and/or primer annealing failure. One individual presented a mixed infection (GI and GII).

**Phylogenetic analysis of the NoV strains**

Nucleotide sequences of the partial VP1 gene (N/S domain) were used for phylogenetic analysis of NoV strains as shown in Figure 2a and b. Strains belonging to GI.3 genotype were grouped into two distinct clusters each including other African strains (Fig. 2a). The Angolan sequences in one cluster showed higher similarity with a strain from Central African Republic detected in 1977 (94.0–94.6% nt identity), and the Angolan strains in the second cluster were similar to a strain from Burkina Faso detected in 2010 (96.0–96.4% nt identity). The only Angolan GI.5 sequence was related to the Hungarian Siklos strain detected in 2014 (96.4% nt identity) and a strain detected in Uganda in 1975 (94.9% nt identity).

Seven of the GI.4 sequences were similar to a reference sequence of the New Orleans variant (97.6–98.1% nt identity), and four GI.4 sequences were similar to a reference sequence of the Sydney variant (99.0–99.5% nt identity) Figure 2b.

The GII.6 sequences were separated into three clusters (Fig. 2b). A first cluster showed similarity to a GII.6 strain detected in Burkina Faso in 2009 (KF434305, 94.2–95.6% nt identity), a second cluster showed similarity to a strain from Kenya detected in 2013 (KFT79386, 95.6–96.1% nt identity), while in a third cluster the Angolan GII.6 sequence was similar to GII.6 sequences from Brazil (JX898890, 99.0% nt identity) and from the USA (KJ415786, 98.5% nt identity) identified in 2009 and 2008, respectively.

The GII.1, GII.2, GII.14 and GII.16 sequences were closely related with recently reported NoV sequences from South Africa, Kenya, Ethiopia and Burkina Faso, respectively.

**Discussion**

We have performed a molecular epidemiology study of NoV in paediatric diarrhoea in four different provinces of Angola. Since a higher number of NoV diarrhoeal cases have been associated with the dry season in several African countries [6, 15, 16], sample collection was carried out mostly during this period. The NoV detection rate was 17.4% in children under 5, showing NoV to be an important pathogen in paediatric diarrhoea in Angola. Differences in the sampling periods considered in this study may have biased the results obtained and, at least in part, justify this discrepancy.

Moreover, young children (<6 months of age) in this study were highly affected by NoV, with a detection rate of 19%, the youngest NoV-positive child being only 2 months of age. Such high rate of NoV detection in very young children is rare in literature, but has been previously reported in hospitalised children with AGE from Brazil and Spain [19, 20]. A factor explaining the high detection rate of NoV in infants could be the poor level of hygiene and sanitation found in this low-income settings. Another reason could be the not exclusive breast feeding (only 50.9% of the mothers exclusively breast feed in Angola [21]).

Interestingly, in this study, children in the age group 7–12 months were significantly less likely to be infected with NoV compared with the other age groups (P < 0.05); although number of NoV positives in each age group were relatively few warranting cautionary interpretation. Other studies have shown that globally NoV diarrhoea peaks at this age group [22, 23]. The lower detection rate of NoV in the 7–12 months old group could partly reflect short time protection from previous infection at younger age as noted above, as well as high prevalence of rotavirus infection with AGE from Brazil and Spain [19, 20]. A factor explaining the high detection rate of NoV in infants could be the poor level of hygiene and sanitation found in this low-income settings. Another reason could be the not exclusive breast feeding (only 50.9% of the mothers exclusively breast feed in Angola [21]).

Interestingly, in this study, children in the age group 7–12 months were significantly less likely to be infected with NoV compared with the other age groups (P < 0.05); although number of NoV positives in each age group were relatively few warranting cautionary interpretation. Other studies have shown that globally NoV diarrhoea peaks at this age group [22, 23]. The lower detection rate of NoV in the 7–12 months old group could partly reflect short time protection from previous infection at younger age as noted above, as well as high prevalence of rotavirus infection with AGE from Brazil and Spain [19, 20]. A factor explaining the high detection rate of NoV in infants could be the poor level of hygiene and sanitation found in this low-income settings. Another reason could be the not exclusive breast feeding (only 50.9% of the mothers exclusively breast feed in Angola [21]).
infection in this age group as shown in an earlier study \[14\]. NoV infection at a young age is important in light of recent efforts towards developing and implementation of a NoV vaccine as suggested by Steele et al. \[24\], and our data suggest that vaccination should be accomplished in the first months of life to prevent NoV paediatric gastroenteritis in Angola. Presently several

---

**Fig. 2.** Continued.
vaccines against NoV are in the pipeline [25]. Last year, Steele et al. used a model-based analysis of mass vaccination in the US population demonstrated that the potential public health value of a NoV vaccine is likely greatest with paediatric immunisation [24]. Altogether, these findings argue for a clinical development plan for a NoV vaccine with a safety and efficacy profile suitable for use in young children.

Phylogenetic analysis based on the partial S’ sequence of the NoV capsid gene demonstrated an unusually large genetic diversity of circulating strains consisting of two genogroups (GI and GII) comprising 11 genotypes. This variety of genotypes including some uncommon ones is a feature of the NoV epidemiology in many sub-Saharan countries [22, 26–30]. GI.4 was the most prevalent genotype (20%), with two different GI.4 circulating variants, New Orleans 2009 and Sydney 2012. Moreover, each of the three other highly prevalent genotypes in this study, i.e. GI.6 (15%), GI.3 (12%) and GI.7 (10%), presented several phylogenetically well-supported sub-clusters (Fig. 2a and b). This observation is in agreement with a recent publication of Parra et al. [31] showing that more than one genetic cluster of emerging genotypes (e.g. GI.6 and GI.3) are co-circulating in the same time and space. This is in contrast with the previous evolutionary trends for the GI.4 genotype, where newly emerging variants replace older ones (e.g. GI.4-New-Orleans and GI.4-Sydney); a process likely driven by the immune response of the host [31]. A factor explaining the high NoV diversity in this study could be that sampling was carried out in different regions, which differ in their socio-economic/hygiene status. Moreover, a process likely driven by the immune response of the host [31]. A factor explaining the high NoV diversity in this study could be that sampling was carried out in different regions, which differ in their socio-economic/hygiene status. Moreover, susceptibility to NoV is strongly dependent on host genetics related to histo blood group antigens (HBGAs) in a NoV genotype-dependent manner [32]. The NoV genotype variability observed in this study could thus also be dependent on differences in population-specific HBGAs distribution between the regions.

Recent reviews on NoV genotype distribution in Africa and worldwide described GI.4 as the most frequently identified and widely distributed genotype, representing roughly 60% of the strains, followed by GI.3 and GI.3, while GI.6 and GI.7 showed a relative low prevalence [22, 30, 33]. Newly emerging genotypes, such as GII.17, have recently been predominant in some studies, particularly in Asian countries [34]. Although this specific genotype, with pre-pandemic potential, was shown to originate in Africa, within this study, no GII.17 strains were detected. Differences in the study design, e.g. sampling period and inclusion criteria (outpatients not hospitalised below five), may account for this distinct pattern of circulating genotypes in Angola.

Additionally, a putative reason for the relatively low detection rate of GI.4 viruses in our study could be differences on the expression of host genetic factors associated with susceptibility to NoV infection, namely concerning the ‘secretor status’. A birth cohort study carried out in Ecuador described GI.4 infection only among secretor-positive children and higher infection rates with non-GI.4 strains in secretor-negative individuals [35]. Additional longitudinal studies covering a longer period of time and a greater number of individuals, including asymptomatic ones, and determination of corresponding HBGAs expression might clarify these issues and evaluate the NoV infection burden in these populations.

Conclusions

We observed a high detection rate and high genetic diversity of NoV in paediatric diarrhoea during the dry season of Angola. NoV was found in children of all ages, with a high detection rate in children <6 months of age. The results of this study can be important in order to address the burden of NoV in children with diarrhoea in Angola.

Acknowledgements. The authors are grateful to all the children who participated and all parents/caregivers who collaborated in this study. The authors would also like to thank all the clinical and laboratory staff in all the health units enrolled in this study for their support and collaboration. They acknowledge Cristina Mendes, Filipa Esteves and Angela Mendes for participation in NoV genotyping. They also acknowledge Professor Dr Virgilio do Rosario who initiated the collaboration with the Ministry of Health in Angola, Dr Antónia Ribeiro for local logistics, Dr Maria Cecília de Almeida (MD) and Marina Oliveira for sample collection in Huambo province.

Financial support. The study was supported by Ministry of Health and the ‘Departamento do Controlo de Doenças da Direção Nacional de Saúde Pública’ from Luanda, Angola and by the Portuguese Ministry of Education and Science (via Fundação para a Ciência e a Tecnologia), through the GHTM-UID/Multi/04413/2013 funding. The study sponsors had no role in the design of the study.

Declaration of Interest. None.

References


