Complete genome analysis of a rare group A rotavirus, G11P[25], isolated from a child in Mumbai, India, reveals interspecies transmission and reassortment with human rotavirus strains

Sushmitha A. Shetty,1 Meenakshi Mathur2 and Jagadish M. Deshpande1

1Enterovirus Research Centre, ICMR, Mumbai, India
2Lokmanya Tilak Medical College & Hospital, Mumbai, India

Hospital-based rotavirus surveillance was carried out in Mumbai during 2005–2009. An isolate (B08299) with a rare genotype combination (G11P[25]) was detected. The present study was undertaken to characterize the complete genome of the isolate. B08299 exhibited a G11–P[25]–I2–R1–C1–M1–A1–N1–T1–E1–H1 genotype constellation. Phylogenetic analysis of the 11 gene segments of B08299 revealed that the VP2 and NSP5 genes of B08299 had a human origin, while the VP6 gene represented an I12 genotype of obscure origin. The remaining six genes formed a lineage distinct from human and porcine rotaviruses within genotype 1. Analysis of the structural and non-structural genes suggested that B08299 has evolved by gene reassortment. Our findings provide further evidence that interspecies transmission is an important mechanism involved in the evolution and genetic diversity of human rotaviruses in nature.

INTRODUCTION

Group A rotaviruses of the family Reoviridae are the major causative agents of acute gastroenteritis in children and animals. The rotavirus genome, consisting of 11 segments of dsRNA, is enclosed in a triple-layered icosahedral capsid. It encodes six structural (VP1–4, VP6, VP7) and six non-structural (NSP1–6) proteins. Rotaviruses are classified into G (glycoprotein) and P (protease-sensitive) genotypes on the basis of the outer-layer proteins VP7 and VP4, which elicit neutralizing antibodies (Estes & Kapikian, 2007). To date, 27 G and 37 P genotypes have been identified globally, with various G–P combinations (Matthijnssens et al., 2011; Trojnars et al., 2013). G1, G2, G3, G4 and G9 in combination with P[4] and P[8] are the most common VP7 and VP4 genotypes associated with human rotavirus infections globally (Santos & Hoshino, 2005). In addition, the P[6] genotype in combination with the common G genotypes has been found to be highly prevalent in Africa, India, Brazil and Argentina (Ramachandran et al., 1996; Araújo et al., 2001; Bok et al., 2001; Cunliffe et al., 2001). Epidemiological studies have revealed an upsurge in the prevalence of genotype G12 in Asia, Argentina, Brazil and the USA (Das et al., 2003; Castello et al., 2006; Uchida et al., 2006; Rahman et al., 2007; Paul et al., 2008). However, genotypes commonly associated with animal infections have also been detected in humans in certain locations. These include G5 (porcine) in Brazil, G6 (bovine) in India and Australia, G8 (bovine) in Africa, G10 (bovine) in Brazil and India, P[9] (feline) in Brazil, P[11] (cattle) in India and P[14] (lapine) in Australia, South Africa and Egypt (Santos et al., 1998; Cooney et al., 2001; Santos & Hoshino, 2005).

The segmented nature of the virus genome provides a unique mechanism for the generation of novel or atypical genotypes by reassortment during co-infection (Palombo, 2002). Interspecies transmission of rotaviruses as whole-genome constellations or individual gene segments has been frequently reported in the literature (Adah et al., 2001; Duan et al., 2007; Ghosh et al., 2007; Khamrin et al., 2007; Chittambar et al., 2009; Matthijnssens et al., 2010a; Mukherjee et al., 2013). Therefore, classification on the basis of the VP7 and VP4 genes only may be insufficient to represent the genetic diversity of rotaviruses. In 2008, the Rotavirus Classification Working Group proposed a full-genome classification, in which genotyping is based on specific nucleotide cut-off values for each segment (Matthijnssens et al., 2008a). Full-genome analyses have provided a complete picture of rotavirus diversity, and have been useful in investigating interspecies transmission and reassortment events (Ghosh & Kobayashi, 2011). Based on this classification, the 27G (VP7), 37P (VP4), 18I (VP6), 9R (VP1), 9C (VP2), 8M (VP3), 18A (NSP1), 10N (NSP2), 12T (NSP3), 15E (NSP4) and 11H (NSP5) genotypes of...
rotaviruses have been described to date (Matthijnssens et al., 2011; Guo et al., 2012; Papp et al., 2012; Trojnar et al., 2013; Jere et al., 2014).


G11 strains were isolated from pigs in Mexico in 1988 and Venezuela in 1994 (Ruiz et al., 1988; Ciarlet et al., 1994). More than a decade later, G11 was detected in a 21-year-old male patient (Dhaka6) in Bangladesh (Rahman et al., 2005). The VP4 gene of the Dhaka6 strain shared less than 76% amino acid identity with known rotavirus reference strains and was therefore assigned to a new genotype, P[25]. The VP4 protein of Dhaka6 contains an extra amino acid at position 135, as compared with the VP4 protein of human rotavirus strains. This feature has previously been demonstrated in animal rotavirus strains only, suggesting that Dhaka6 could be of zoonotic origin (Taniguchi et al., 1989; Rahman et al., 2005). Since then, only five human G11P[25] strains have been reported: KTM368 (Nepal), CRI 10795 and N-38/2009 (India), and G0703034 and CAU12-2 (Korea) (Uchida et al., 2006; Banerjee et al., 2007; Mullick et al., 2013; Than et al., 2013).

The present study pertains to the molecular characterization of a G11P[25] strain identified through rotavirus surveillance in Mumbai, India. This is the seventh G11P[25] isolate reported globally.

METHODS

Rotavirus surveillance. Rotavirus surveillance was carried out in Mumbai between December 2005 and June 2009. Faecal specimens were collected from children aged 5 years or younger who were hospitalized with acute gastroenteritis. Institutional ethics committee approval was obtained. Informed consent of the parent/caregiver was obtained for sample collection. The algorithm described by Kang et al. (2009) was followed for detection and genotypic characterization of rotaviruses.

Rotavirus detection. Faecal specimens were screened for group A rotavirus using antigen-capture ELISA targeting the VP6 antigen (Rota IDEIA; Dako) as per the manufacturer’s instructions. The G and P genotypes of ELISA-positive specimens were determined by reverse transcription (RT)-PCR followed by semi-nested multiplex PCR using genotype-specific primers (Gouvea et al., 1990; Gentsch et al., 1992).

RNA extraction and RT-PCR genotyping. Faecal suspensions (30% w/v) were prepared in 0.01 M PBS. Viral RNA was extracted using TRIzol (Invitrogen) as per the manufacturer’s instructions. cDNA was synthesized at 37 °C using random primers (dN)6 (Roche Diagnostics) and 100 U M-MLV reverse transcriptase (Invitrogen) (Iturriza-Gómez et al., 1999). Con3/Con2 and BgE/End9 primer pairs were used for PCR amplification of the VP4 and VP7 genes, respectively (Gouvea et al., 1990; Gentsch et al., 1992). The multiplex PCR for G genotyping was performed using genotype-specific primers for G1–6, G8–10 and G12 (Gouvea et al., 1990). P genotyping was performed using a pool of genotype-specific primers P[4], P[6] and P[8]–[11] (Gentsch et al., 1992). Non-typable specimens were characterized by sequencing the first-round products of the VP7 and VP4 genes. Novel genotypes detected were further characterized by sequencing the remaining nine gene segments (VP1–3, VP6 and NSP1–5). Amplification of the nine gene segments was carried out using primer pairs and conditions that have been described previously (Cunliffe et al., 1997; Matthijnssens et al., 2006, 2008b; Bányai et al., 2009). PCR products were electrophoresed in 2% agarose gels containing ethidium bromide and visualized under a UV transilluminator.

Nucleotide sequencing and phylogenetic analysis. PCR amplimers were excised from the agarose gel and purified using the QIAquick Gel extraction kit (Qiagen). Nucleotide sequencing was carried out in both directions using the dideoxynucleotide chain-termination method with the BigDye Terminator Cycle Sequencing Reaction kit v3.1 (Applied Biosystems) as per the manufacturer’s instructions. The primers used for PCR amplification were used as the sequencing primers. Sequences were resolved on an ABI 3130xl Genetic Analyzer (Applied Biosystems). Sequences were edited with Sequencher v.4.10.1. The genotype assignment for each gene was accomplished using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and RotaC v2.0 (http://rotac.regatools.be) (Maes et al., 2009). Phylogenetic trees were reconstructed using the neighbour-joining algorithm implemented in MEGA v.4 (http://www.megasoftware.net) (Saitou & Nei, 1987; Tamura et al., 2007). Genetic distances were calculated using the Kimura (1980) two-parameter method.

RESULTS

Non-typable group A rotavirus strain B08299

Faecal samples from 1502 cases of acute gastroenteritis were studied from December 2005 to June 2009. Group A rotavirus was detected in 34.75% of faecal specimens (522/1502). Both the G and P genotypes could be determined for 454 specimens (87%). The G/P combination most commonly found was G1P[8] (151/454, 33.3%), followed by G2P[4] (121/454, 26.7%) and G12 in combination with P[6] (65/454, 14.3%), P[4] (4/454, 0.9%) and P[8] (16/454, 3.5%). Mixed infections were detected in 9.3% specimens. A total of 24/522 specimens (4.6%) were non-typable for both VP4 and VP7 genotyping. Here, we report the characterization of a non-typable rotavirus strain (B08299) which was of rare genotype specificity.

The rotavirus strain B08299 was isolated from a 4-month-old boy suffering from acute gastroenteritis with onset on 20 August 2008. The faecal specimen was positive for group A rotavirus by ELISA and RT-PCR with consensus primers for the VP7 and VP4 genes. However, the G and P genotypes could not be determined using a multiplex PCR.
assay which included the genotype-specific primers G1–4, G6, G8–10 and G12 for G genotyping and P[4], P[6] and P[8–11] for P genotyping. Therefore, genotype identification of B08299 was carried out using nucleotide sequencing of the first-round products of the VP7 and VP4 genes. Since preliminary analysis of the VP7 and VP4 genes showed that isolate B08299 possessed a rare genotype, G11P[25], related to porcine rotaviruses, complete genome characterization was carried out to investigate the origin and exact genetic makeup of this strain. The complete genome was amplified for each gene segment except for VP4, which was partially amplified (Table 1). Each gene segment was subjected to both the RotaC online rotavirus genotyping tool and BLAST, which provided global dendrograms representing all rotavirus genotypes. Although B08299 was compared with all rotavirus genotypes, only those relevant to the present study are shown in Figs 1 and 2.

**Analysis of structural genes of B08299**

Analysis of the VP7 gene (nt 72–952) revealed close genetic relatedness to the 13 human and two porcine G11 strains reported to date. The phylogenetic tree presented in Fig. 1(a) includes the VP7 gene sequences of all known G11 strains of human and animal origin. The G11 strains reported from Vellore (CRI 10795) and Bangladesh (Dhaka13-06) were also compared but are not included in the phylogenetic tree as the nucleotide sequences deposited in GenBank are comparatively short (730 bp). The G11 rotaviruses isolated to date cluster into four lineages (Bányai et al., 2009). Lineages I and II are represented by the two porcine G11 strains isolated from Mexico (YM) and Venezuela (A253), respectively. The G11 human strains isolated from Asia cluster in lineage III, while the Ecuadorian G11 strain forms a separate lineage (IV). Strain B08299 from Mumbai clustered with the Asia G11 strains (N-38/2009, KTM368, CAU-1, CUK1, GJ0703034, CAU12-2, Dhaka6 and Maltab36) and showed close genetic relatedness to CRI 10795 isolated from Vellore, India (Table 1).

Analysis of the VP4 gene (nt 26–1853) revealed close genetic relatedness to genotype P[25]. The VP4 gene of isolate B08299 showed close genetic relatedness to the CRI 10795 strain isolated from Vellore, India (Table 1). A phylogenetic tree of the P[25] strains isolated globally was reconstructed (Fig. 1b). CRI 10795 was not included in the phylogenetic tree as the nucleotide sequences deposited in GenBank are only 514 bp. Evolutionary analysis of P[25] rotaviruses revealed the presence of two distinct lineages (Fig. 1b). Lineage I was represented by Dhaka6 (Bangladesh), KTM368 (Nepal), CAU12-2 (Korea) and GJ0703034 (Korea), while N-38/2009 isolated from Kolkata, India, formed a separate lineage II. The P[25] strain from Mumbai clustered into lineage I. The genetic distance between the Kolkata strain and other P[25] strains was 8.2–11.3 %.

The VP1, VP2 and VP3 genes of B08299 belonged to genotypes R1, C1 and M1, respectively. The VP2 gene of B08299 clustered with the human rotavirus strains within the C1 genotype (Fig. 1e). The VP1 and VP3 genes formed distinct clades and showed equal genetic distances from the human and porcine rotaviruses within the R1 and M1 genotypes (Fig. 1d, f). The VP6 gene of B08299 belonged to an H1 genotype of unknown origin (Fig. 1c). Thus, B08299 contains VP1, VP3 and VP6 genes that have evolved from rotaviruses of unknown origin.

**Analysis of non-structural genes of B08299**

The NSP1–5 genes of B08299 were found to belong to genotypes A1, N1, T1, E1 and H1, respectively. The NSP1–4 genes of B08299 formed clades distinct from all known human and porcine rotaviruses within genotype 1 (Fig. 2a–d). The NSP5 gene clustered with the human strains within the H1 genotype (Fig. 2e).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence length obtained for B08299 (bp)</th>
<th>KTM368</th>
<th>N-38/2009</th>
<th>CRI 10795</th>
<th>Dhaka6</th>
<th>CAU12-2</th>
<th>GJ0703034</th>
</tr>
</thead>
<tbody>
<tr>
<td>VP1</td>
<td>3222</td>
<td>0.013</td>
<td>0.005</td>
<td>NA</td>
<td>0.145</td>
<td>0.150</td>
<td>NA</td>
</tr>
<tr>
<td>VP2</td>
<td>2696</td>
<td>0.020</td>
<td>0.017</td>
<td>NA</td>
<td>0.018</td>
<td>0.026</td>
<td>NA</td>
</tr>
<tr>
<td>VP3</td>
<td>2524</td>
<td>0.014</td>
<td>0.020</td>
<td>NA</td>
<td>0.022</td>
<td>0.019</td>
<td>NA</td>
</tr>
<tr>
<td>VP4</td>
<td>1828</td>
<td>0.028</td>
<td>0.113</td>
<td><strong>0.004</strong></td>
<td>0.032</td>
<td>0.034</td>
<td>0.028</td>
</tr>
<tr>
<td>VP6</td>
<td>1331</td>
<td>0.012</td>
<td>0.064</td>
<td>*</td>
<td>0.189</td>
<td>0.024</td>
<td>NA</td>
</tr>
<tr>
<td>VP7</td>
<td>881</td>
<td>0.017</td>
<td>0.017</td>
<td><strong>0.010</strong></td>
<td>0.018</td>
<td>0.023</td>
<td>0.026</td>
</tr>
<tr>
<td>NSP1</td>
<td>1460</td>
<td>0.181</td>
<td><strong>0.032</strong></td>
<td>NA</td>
<td>0.183</td>
<td><strong>0.026</strong></td>
<td>NA</td>
</tr>
<tr>
<td>NSP2</td>
<td>953</td>
<td>0.007</td>
<td>0.008</td>
<td>NA</td>
<td>0.141</td>
<td>0.017</td>
<td>NA</td>
</tr>
<tr>
<td>NSP3</td>
<td>1002</td>
<td>0.012</td>
<td>0.039</td>
<td>NA</td>
<td>0.094</td>
<td>0.011</td>
<td>NA</td>
</tr>
<tr>
<td>NSP4</td>
<td>730</td>
<td>0.016</td>
<td>0.018</td>
<td><strong>0.010</strong></td>
<td>0.147</td>
<td>0.045</td>
<td>NA</td>
</tr>
<tr>
<td>NSP5</td>
<td>642</td>
<td>0.053</td>
<td>0.057</td>
<td>NA</td>
<td><strong>0.012</strong></td>
<td>0.060</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA, No sequence data available in GenBank.

*Sequences <500 nt of the ORF were not included in the analysis.
Fig. 1. Phylogenetic analysis based on the nucleotide sequences of structural genes: (a) VP7/G11, (b) VP4/P[25], (c) VP6/I12, (d) VP1/R1, (e) VP2/C1 and (f) VP3/M1 of group A rotaviruses and B08299. The scale bar indicates the fractional genetic distance. The numbers at the nodes indicate the frequency (%) at which the particular tree topology was found in 1000 bootstrap replicates; only values ≥80% are shown. Strains in the dendrogram are represented as accession no./host species/specimen ID/ geographical area/year of isolation/G–P genotype. B08299 is highlighted by ●.
Fig. 2. Evolutionary analysis based on the nucleotide sequences of non-structural genes: (a) NSP1/A1, (b) NSP2/N1, (c) NSP3/T1, (d) NSP4/E1 and (e) NSP5/H1 of group A rotaviruses and B08299. The scale bar indicates the fractional genetic distance. The numbers at the nodes indicate the frequency (%) at which the particular tree topology was found in 1000 bootstrap replicates; only values ≥ 80 % are shown. Strains in the dendrogram are represented as accession no./host species/specimen ID/geographical area/year of isolation/G–P genotype. B08299 is highlighted by ◆.
Thus, nucleotide sequencing and phylogenetic analysis of all 11 gene segments of B08299 showed that it belongs to genotype G11–P[25]–I12–R1–C1–M1–A1–N1–T1–E1–H1.

**DISCUSSION**

A rotavirus isolate (Dhaka6) detected in Bangladesh in 2001 was the first strain with the G11P[25] genotype combination (Rahman et al., 2005). The remaining nine gene segments of Dhaka6 represent genotypes similar to Wa-like human rotaviruses (I1–R1–C1–M1–A1–N1–T1–E1–H1) (Matthijnssens et al., 2010a). G11P[25] was then detected from an asymptomatic patient (CRI 10795) in a community-based birth cohort study in Vellore, India, in 2003 (Banerjee et al., 2007). Unfortunately, only partial VP7, VP4, VP6 and NSP4 gene sequences of this strain are available in GenBank. The VP6 gene of CRI 10795 exhibited an I1 genotype similar to that of Dhaka6. But, unlike Dhaka6, the NSP4 gene of CRI 10795 formed a lineage distinct from all known human and porcine rotaviruses within genotype E1 (Banerjee et al., 2007). A third G11P[25] strain (KTM368) was isolated in Nepal in 2004. The VP6 gene sequence of KTM368 did not belong to any of the established VP6 genotypes; it was therefore assigned to a new genotype, I12, by the Rotavirus Classification Working Group (Matthijnssens et al., 2010a). In 2007, yet another G11P[25] human rotavirus strain (GJ0703034) was isolated in Korea, but only the VP7 and VP4 gene sequences of this strain are available in GenBank (accession no. EU259894 and EU259896, respectively). This strain is not discussed in the current paper. Two more G11–P[25]–I12 rotavirus strains, one each in India (N-38/2009) and South Korea (CAU12-2), were reported in 2009 and 2012, respectively (Mullick et al., 2013; Than et al., 2013). B08299 is the fourth strain isolated from a human source with the G11–P[25]–I12 genotype combination.

Evolutionary analysis of the 11 gene segments of the four strains with the G11–P[25]–I12 genotype constellation (KTM368, N-38/2009, CAU12-2 and B08299) revealed that the VP2 gene of all of these strains belongs to a typical human cluster within the C1 genotype, whereas the VP5 and NSP2–4 genes of these strains form a single clade different from all known human and porcine rotaviruses within each respective genotype (Fig. 1e, f and Fig. 2b–d).

Interestingly, for the VP1 and NSP1 genes, at least one of the four strains was an outlier. VP1 of CAU12-2 originated from porcine rotavirus strains, and VP1 of KTM368, N-38/2009 and B08299 formed a new clade within the R1 genotype (Fig. 1d). The NSP1 gene of KTM368 clustered with the human strains, whereas N-38/2009, CAU12-2 and B08299 formed a clade distinct from both human and porcine rotaviruses within genotype A1 (Fig. 2a). The NSP5 gene of B08299 showed a close genetic relationship with NSP5 of commonly circulating G1P[8] human strains in India, whereas the NSP5 gene of KTM368, N-38/2009, formed a clade distinct from both human and porcine rotaviruses, representing the H1 genotype (Fig. 2e).

Strains N-38/2009 and B08299 were isolated in India. Complete genome analysis determined that eight gene segments (VP1–3, VP7 and NSP1–4) of isolate B08299 showed close genetic relatedness to the respective genes of N-38/2009 (Table 1). However, the VP4 (P[25]) and VP6 (I12) genes showed a genetic distance of 11.3 % and 6.4 %, respectively (Table 1). Rotaviruses evolve at a rate of $10^{-3}$ nucleotide substitutions per site per year, or 0.1 % nucleotide changes per year (Jenkins et al., 2002; Matthijnssens et al., 2010b). The genetic distances of the VP4 and VP6 genes of B08299 and N-38/2009 clearly indicate that they may have evolved from independent ancestors.

Complete genome analysis of these four strains with the G11–P[25]–I12 genotype combination revealed that each strain has evolved independently by gene reassortment. In summary, the VP2 and NSP5 genes of B08299 belong to a typical human subcluster within genotype 1. All the remaining gene segments either represent genotypes of animal/unknown origin (VP4; P[25], VP7; G11, VP6; I12) or form a separate lineage within the genotype.

In India, the prevalence of rotavirus infection in pigs can be as high as 25.7 % (Kusumakar et al., 2010). In addition, unusual rotavirus strains of porcine origin (G1P[19], G9P[6], G9P[19] G4P[4] and G4P[6]) have been reported in India in several studies (Varghese et al., 2004; Chitambar et al., 2009; Mukherjee et al., 2009, 2011). The emergence of strains with animal origin or animal–human reassortants in India could be a result of humans and animals sharing living spaces, leading to interspecies transmission and mixed infections. G11 has been reported with a variety of VP4 constellations, such as P[25], P[4], P[6], P[8] and P non-typable (Rahman et al., 2005; Uchida et al., 2006; Banerjee et al., 2007; Hong et al., 2007; Bányai et al., 2009; Matthijnssens et al., 2010a; Mullick et al., 2013; Than et al., 2013). Similar reassortment events have preceded the emergence of G9 and G12 strains globally (Rahman et al., 2007; Phan et al., 2007). It is important to note that the origins of several gene segments of strain B08299 are obscure.

The present study provides further evidence that interspecies transmission and genetic reassortment are the major mechanisms involved in the genetic diversity and evolution of rotaviruses.

**ACKNOWLEDGEMENTS**

This work was performed with support from the Indian Council of Medical Research (grant RFC no. ECD/NTF/3/2005-06). S.A.S. was supported by a senior research fellowship from the Indian Council of Medical Research.

**REFERENCES**

Molecular Journal of Medical Microbiology


