Complete genome analysis of contemporary G12P[8] rotaviruses reveals heterogeneity within Wa-like genomic constellation

Simona De Grazia a,⁎, Renáta Dóró b, Floriana Bonura a, Szilvia Marton b, Antonio Cascio a, Vito Martella c, Krisztián Bányai b, Giovanni M.Giammanco a

a Department of Health Promotion Sciences and Mother and Child Care “G. D’Alessandro”, University of Palermo, Palermo, Italy
b Veterinary Medical Research Institute, Centre for Agricultural Research, Hungarian Academy of Sciences, Budapest, Hungary
c Department of Veterinary Medicine, University Aldo Moro of Bari, Valenzano, Italy

Available online 21 June 2016
Accepted 19 June 2016
Received in revised form 16 June 2016
Received 18 April 2016

Article history:
Article info
Keywords:
G2
Phylogenetic analysis
Full length-genome
Italy

Complete genome analysis of contemporary G12P[8] rotaviruses reveals heterogeneity within Wa-like genomic constellation

A R T I C L E   I N F O

G12 rotaviruses are globally emergent rotaviruses causing severe childhood gastroenteritis. Little is known about the evolution and diversity of G12P[8] rotaviruses and the possible role that widespread vaccine use, globally, has had on their emergence. In Sicily, Italy, surveillance activity for rotaviruses has been conducted uninterruptedly since 1985, thus representing a unique observatory for the study of human rotaviruses in the pre- and post-vaccine era. G12 rotaviruses were first detected only in 2012 and between 2012 and 2014 they accounted for 8.7% of all rotavirus-associated infections among children, with peaks of 27.8% in 2012/2013 and 21% in 2014. We determined and analyzed the full-genome of 22 G12P[8] rotaviruses collected during the 2012-2014. Although all G12P[8] rotaviruses exhibited a typical Wa-like genotype constellation (G12P[8]-I1-R1-M1-A1-N1-E1-H1), phylogenetic analysis allowed distinguishing either two or three (sub)lineages in each genome segment. On the basis of the segregation patterns into lineages/sublineages, 20 G12P[8] rotaviruses could be grouped into three stable major genomic sub-constellations, whilst two strains displayed unique genome architectures, likely due to reassortment with co-circulating strains. Altogether, these findings indicate that the onset and prolonged circulation of G12 rotaviruses was due to repeated introductions of different G12 rotaviruses circulating globally. Importantly, as regional rotavirus vaccination was initiated in 2012 reaching a 45% coverage in newborns in 2014, a correlation between the appearance and spread of G12 rotaviruses and the enacted vaccination program could not be drawn. Constant epidemiologic surveillance remains important to monitor the epidemiological dynamics of human rotaviruses.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Group A rotavirus (RVA) is the leading etiological agent of severe gastroenteritis in the young of humans and of many animals worldwide. The RVA virion is a triple-layered, non-enveloped icosahedron enclosing 11 segments of double stranded RNA encoding six structural (VP) and five/six non-structural proteins (NSP) (Estes and Greenberg, 2013). Accumulation of point mutations and reassortment events between homologous and/or heterologous strains are the major mechanisms driving the evolution of RVAs. The two outer capsid proteins, VP7 (Glycoprotein) and VP4 (Protease-sensitive protein), both eliciting neutralizing immune response, are used for the binary classification of RVAs into at least 27 G and 37 P genotypes (Estes and Greenberg, 2013; Trojnar et al., 2013). Epidemiological studies worldwide have shown that G1-, G3- and G4P[8] and G2P[4] strains are more commonly associated with human infections and those types have been targeted by the rotavirus vaccines. More recently, new G/P combinations have emerged acquiring global epidemiological relevance. G9P[8], G12P[6] and G12P[8] RVAs emerged and spread worldwide at the end of the 1990s and at the beginning of the new millennium, respectively, and they are now considered among the major human genotypes (Matthijnssens et al., 2009). In recent years full genome based phylogenetic analyses have been used to define the genotype constellations of RVA strains (Matthijnssens et al., 2008). Based on the whole genome analyses, most human RVA strains show high sequence similarity in all their genes to either the Wa or DS-1 prototype strains. Usually G1-, G3-, G4-, G9- and G12P[8] strains share a Wa-like genotype constellation (I1-R1-M1-A1-N1-E1-H1), whereas G2P[4] and G12P[4] strains display a DS-1-like backbone (I2-R2-M2-A2-N2-E2-H2) (Matthijnssens et al., 2008). Moreover, several reassortment events between genotype constellations have been...
frequently described (Rahman et al., 2007; Komoto et al., 2014). The recent introduction and widespread circulation of G12 strains into the human population seems to be the result of reassortment events introducing a G12 VP7 gene into a Wa-like genome constellation. The origin of the G12 VP7 segment is unclear, although pigs and bovines are suspected to be the potential host reservoir for this genotype (Rahman et al., 2007; Ndze et al., 2013a, 2013b; Midgley et al., 2012). G12 rotaviruses were first identified in humans in association with the P[4] VP4 genotype (strains L26 and L27) (Taniguchi et al., 1990) and they were found to have a unique genotype constellation (Rahman et al., 2007). However, G12 RVAs started spreading successfully in human population only a decade later, when a novel G12 VP7 variant combined with more stable genome constellations (Wa-like and DS-1-like) emerged (Rahman et al., 2007; Cilla et al., 2013). The Wa-like constellation seems to have the ability to propagate extremely well in human hosts, as it is the most successful backbone of RVAs globally (Desselberger et al., 2006). In Sicily, Italy, a documented uninterrupted RVA surveillance has been conducted for more than three decades since 1985 and RVA infection was detected in 31.4% (2293) of 7293 samples tested until December 2014 (De Grazia et al., 2014). Fluctuations of the predominant G types and strains were observed over the years, but G12 strains were never detected until 2012. G12 RVAs emerged locally in 2012-2014 when they accounted for 8.7% of the strains characterized, with peaks of 27.8% in Ragusa in 2012/2013 and of 21% in Palermo in 2014 (Giammanco et al., 2016). All the detected G12 RVA strains carried a P[8] VP4 type but a unique RVA G12 strain that was P[9] (De Grazia et al., 2015). In this study we determined the full-length genome sequence of a selection of G12 RVAs, in order to investigate the origin and genetic diversity of these emerging enteric pathogens in local population.

2. Materials and Methods

2.1. Specimen selection

During 2012-2014 routine surveillance for RVA on children (aged 0-14 years old) admitted for acute gastroenteritis to three Sicilian Hospitals (“G. Di Cristina” Children Hospital, Palermo; “G. Di Martino” University Hospital, Messina; and “Civile” Hospital, Ragusa), 33 G12P[8] strains were detected by RT-PCR (Gentsch et al., 1992; Gouvea et al., 1990; Iturria-Gomara et al., 2004). In this study, full-length genome sequencing and phylogenetic analyses were performed on 22 G12P[8] strains representative of the surveillance period and selected on the basis of preliminary sequence analyses of VP7 and VP4 genes (Giammanco et al., 2016). Although the collection of faecal samples was part of the general process of diagnosis of the acute gastroenteritis affecting the minors/children, we obtained verbal informed consent from the family or caretakers and the consent was recorded in the patient’s medical chart.

2.2. Complete genome sequencing and phylogenetic analysis

The complete genome of selected G12P[8] strains was determined as previously described (Dorò et al., 2014). Phylogenetic analyses of the 11 dsRNA segments were performed by MEGA 6 software, using the Kimura 2-parameter substitution and the maximum likelihood method to construct phylogenetic trees (Tamura et al., 2013). The statistical significance of the phylogenies inferred was estimated by bootstrap analysis with 1000 replicate data sets. The results of phylogenetic analyses were confirmed with the neighbor-joining method. The nucleotide sequences of the full-length genome of G12P[8] strains of this study were deposited in GenBank under accession numbers: VP1 (KU048527 - KU048548); VP2 (KU048549 - KU048570); VP3 (KU048571 - KU048592); VP4 (KU048593 - KU048614); VP6 (KU048615 - KU048636); VP7 (KU048637 - KU048658); NSP1 (KU048659 - KU048680); NSP2 (KU048681 - KU048702); NSP3 (KU048703 - KU048724); NSP4 (KU048725 - KU048746); NSP5 (KU048747 - KU048768).

3. Results

Preliminary sequence analyses of the VP7 and VP4 genes revealed temporal clustering patterns among the G12P[8] strains detected from 2012 to 2014 in Sicily (Desselberger et al., 2006). This information was used to select 22 G12P[8] RVA strains for further investigations. Sequencing and phylogenetic analyses was extended to the whole genome of the 22 G12P[8] strains to investigate their evolution pathways. Full-length genome analysis allowed classifying all the G12P[8] strains as I1-R1-C1-M1-A1-N1-T1-E1-H1, according to the current classification system. Phylogenetic and pairwise identity analyses of all eleven genome segments were performed to determine the genetic relationships of the G12 strains with each other and with RVA complete genomes retrieved from GenBank. Sequence comparison and phylogenetic analysis allowed distinguishing either two or three lineages/sublineages in all the genome segments. Therefore, for the limited purposes of the present study, we defined with letters (a to c) the various lineages/sublineages according to the patterns of segregation for each gene (Fig. 1). On the basis of the combination of segregation patterns observed in the 11 genome segments, 20 G12P[8]s could be grouped into three major sub-constellations (A-C), while two strains displayed unique genome sub-constellations (D and E). Strains belonging to the same sub-constellation shared a high nucleotide (nt) identity (id), ranging from 99 to 100% in each gene (Table 1). The nt id among different sub-constellations ranged from 89.9% to 98.6%, while a 99.2% of nt id was observed in the NSP5 gene. BLAST analyses were performed on GenBank database in order to identify the RVA strains with the highest identity to the Italian G12P[8] RVAs over the whole genome constellation (Table 2). In detail, G12 RVA strains of sub-constellation A shared seven gene segments with the strain GER126/08/2008/G12P[8], while in the other segments (VP2, VP3, NSP1) they were more similar to other G12P[8] and G1P[8] strains. G12 RVA strains of sub-constellation B showed high identity to the strain GER172/08/2008/G12P[6] detected in a child hospitalized with gastroenteritis in Leipzig (Germany) in 2008 and to the strain USA2013774166/2013/G12P[8], identified in an outbreak of gastroenteritis in adults in USA in 2013 and reported by the Center for Disease Control (CDC, Atlanta, USA). G12 RVA strains of sub-constellation C showed the highest sequence identity in nine gene segments to the African strain KEN/KDH651/2010/G12P[8], identified from a child with diarrhea in Kiambu district, Kenya, in 2010, with the NSP4 gene being related to the strains ZAF/3133WC/2009/G12P[4] and GER172/08/2008/G12P[6]. Other two strains, PA417/14 and ME659/14, showed a complex genetic structure, sharing gene segments with strains of the main sub-constellations A to C and they were named herewith as sub-constellation D and E, respectively (Table 2). As shown in Table 2, the NSP5/6 gene was genetically highly conserved across all the G12P[8] strains, although it was still possible, phylogenetically, to identify segregation patterns. Upon BLAST analysis, the NSP5 gene of the sub-constellations A, B-C, D and E resembled, respectively, the NSP4 of the G12 RVA strains ZAF/3137/WC/2009/G12P[6], KEN/KDH651/10/2010/G12P[8], USA/VU08-09-39/2008/G12P[8] and GER172/08/2008/G12P[6] (Table 2). To investigate whether amino acid (aa) differences might exist within in G12P[8] strains of different sub-constellations, the antigenic epitopes (7-1a, 7b-1b and 7-2) and variable regions (VR1-9) of the VP7 were analyzed. Detailed inspection of the aa sequence alignment revealed a unique aa residue (I44V) in the VP7 variable region 3 (VR3) that changed across the VP7 lineages. Inspection of the antigenic epitopes on the VP4 multimeric (VP5* and VP8*) did not reveal any peculiar aa substitutions among the analyzed G12P[8] strains and the reference strains available in GenBank. With respect to reference strain Wa and to vaccine strains, only the G12P[8] strains of cluster a showed two
peculiar aa substitutions (N87S and L194D) in the epitopes 8-1 and 8-4 of the VP8* (Table 3).

4. Discussion

The prevalence of G12 RVAs has remained relevant over the past few years across Europe, suggesting that this genotype should be now considered stably as one of the most common RVA VP7 type in humans (Matthijnssens et al., 2009). In Sicily, surveillance activity for RVAs has been conducted since 1985, but G12 RVAs were first detected only in 2012. Between 2012 and 2014, G12 RVAs accounted on average for 8.7% of the detected RVA strains, although we registered high peaks of prevalence (27.8% in Ragusa in 2013 and 21% in Palermo in 2014) in some prefectures. In the last decade, many studies documented the emergence of G12 RVAs, predominantly in combination with P[8] and P[6] VP4 types, and more rarely with P[4] and P[9] types. Full-length genome analysis of selected G12P[8] RVAs has suggested that this G-genotype is mostly associated with a Wa-like backbone (Delogu et al., 2015; Gomez et al., 2014; Ide et al., 2015; Komoto et al., 2014; Pietsch and Liebert, 2009; Wangchuk et al., 2014). This stable genetic constellation...
might have conferred to G12 rotaviruses an evolutionary advantage to settle and spread into human population (Rahman et al., 2007). P[4] and P[8] strains seem to use histo-blood group antigens (HBGAs) like the H-type 1 antigen and Lewis b antigen for attachment to cells. Different affinity for receptors/co-receptors, i.e. HBGAs, among human RVAs could explain the epidemiological predominance of some RVA strains.
Table 1
Percentage of identity of single gene segments in the three different Rotavirus G12P[8] sub-constellations described in the study

<table>
<thead>
<tr>
<th>Sub-constellation</th>
<th>VP7 (% id nt)</th>
<th>VP4 (% id nt)</th>
<th>VP1 (% id nt)</th>
<th>VP2 (% id nt)</th>
<th>VP3 (% id nt)</th>
<th>VP6 (% id nt)</th>
<th>NSP1 (% id nt)</th>
<th>NSP2 (% id nt)</th>
<th>NSP3 (% id nt)</th>
<th>NSP4 (% id nt)</th>
<th>NSP5 (% id nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>99.9-100</td>
<td>99.9-100</td>
<td>100</td>
<td>99.4-100</td>
<td>99.1-100</td>
<td>99.9-100</td>
<td>100</td>
<td>100</td>
<td>99.3-100</td>
<td>99.7-100</td>
<td>99.3-100</td>
</tr>
<tr>
<td>B</td>
<td>99.9-100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99.9-100</td>
<td>100</td>
<td>100</td>
<td>99.7-100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>C</td>
<td>99.7-100</td>
<td>99.6-100</td>
<td>99.4-100</td>
<td>99.6-100</td>
<td>99.4-100</td>
<td>99.8-100</td>
<td>99.3-100</td>
<td>99.7-100</td>
<td>99.4-100</td>
<td>99.1-100</td>
<td>99.3-100</td>
</tr>
</tbody>
</table>
Table 2
Comparison of the genomic constellations of 22 Sicilian G12P[8] RVA strains analysed in this study (grouped according to phylogenetic sub-constellation) and other RVA strains recovered in GenBank. Percentage of identity of the most related complete genome nucleotide sequences of rotavirus strains are indicated. The lineages/sublineages (a-c) defined according to phylogenetic analyses are indicated for each gene segment, with the exception of the NSP5 gene. Intragenotype similarities are indicated in shades of grey.

<table>
<thead>
<tr>
<th>Sub-constellation</th>
<th>N°</th>
<th>Strains</th>
<th>VP7 (% id nt)</th>
<th>VP4 (% id nt)</th>
<th>VP6 (% id nt)</th>
<th>VP1 (% id nt)</th>
<th>VP2 (% id nt)</th>
<th>VP3 (% id nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-constellation</td>
<td>NSP1 (%) id nt</td>
<td>NSP2 (%) id nt</td>
<td>NSP3 (%) id nt</td>
<td>NSP4 (%) id nt</td>
<td>NSP5 (%) id nt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
and account for the preferential adoption of the Wa-like backbone by G12 RVAs (Liu et al., 2012).

In this study, we determined the complete genome of 22 G12P[8] RVAs detected over three years of surveillance in a settled population and compared this information with the sequence data available in the databases. All the analyzed G12P[8] strains exhibited a typical Wa-like genotype constellation, G12-P[8]-II1-C1-M1-A1-N1-T1-E1-H1, although upon sequence comparison and phylogenetic analysis a discrete genetic heterogeneity was observed. In all the genome segments, but the NSP5/6, either two or three different lineages could be identified, which were named, for the limited purposes of the present study, with the letters “a” to “c”. Based on the segregations patterns of the genome segments, the 22 Italian G12P[8] RVAs could be distinguished into at least three different major sub-constellations, herewith indicated as A, B and C (Fig. 2). G12 RVA strains of sub-constellation A and B were phylogenetically linked, respectively, to the strains GER126-08/2008/G12P[8] and GER172-08/2008/G12P[6], identified in Germany in 2008 during post-marketing monitoring of rotavirus vaccine efficacy. The two German strains were not related to each other or to local animal G12 RVA strains. Interestingly, strain GER172-08 (sub-constellation A) was closely related to Southeast Asian RVA strains, while the origin of strain GER126-08 (sub-constellation B) remained unclear (Pietsch and Liebert, 2009). The G12 strains of sub-constellation C were genetically linked to the Italian strain KEN/KDH651/2010/G12P[8], which in turn was closely related to several human G1 and G12 strains detected in the African continent (Komoto et al., 2014). Although both the German and African G12 RVA strains carried a Wa-like backbone, phylogenetic analysis suggest that these strains were likely generated by separate reassortment events with other co-circulating RVA genotypes, i.e. G1[P[8]] and G9[P[8]] (Komoto et al., 2014; Pietsch and Liebert, 2009; Ianiro et al., 2013) that are currently the predominant RVA types worldwide (Banyai et al., 2012).

Two G12 strains (PA417/14 and ME659/14), detected in Palermo and Messina in 2014, exhibited a diverse and more complex generic sub-constellation. These strains shared some genetic segments with co-circulating G12 RVAs and could have arisen by multiple reassortment events with RVA strains of different G/P-types circulating in the same period. In addition, a multi-reassortant G12[P[9]] strain was detected in the same study period, which was likely derived from multiple reassortment events with human and animal rotaviruses (De Grazia et al., 2015). As a whole, these findings indicate repeated introductions of G12 RVA strains in Sicily in a limited time span (2012-2014), and indicate the possible interaction of G12 RVAs with different ancestral strains, i.e. G12[P8], G12[P6]/P[4], G1[P8] and G9[P8]. These dynamic forces may explain the genetic/antigenic plasticity exhibited by G12 RVAs. Interestingly, peculiar aa substitutions in the neutralizing epitopes of VP7 and VP4 proteins were found to correlate with the various genetic sub-constellations identified during our study, however more studies are needed to better understand how these substitutions would influence the efficacy of the vaccines. The ability of G12 RVAs to spread across human populations has been hypothesized to derive from this peculiar genomic plasticity, driven by multiple reassortment events (Rahman et al., 2007). In addition, G12 strains show an error-prone nature in the VP7 gene, where point mutations tend to accumulate at a high evolutionary rate, as observed for the common human genotype G9P[8] (Matthijnssens et al., 2010). Many studies have shown that the majority of the currently known G12[P8]/P[6] strains clustered into the same lineage (lineage III) further divided into different clusters. These data suggested the co-existence of different genotypic clones during the spread of G12[P8] strains in settled populations using a limited number of full-length genomes (Delogu et al., 2015; Gomez et al., 2014; Ide et al., 2015; Komoto et al., 2014; Pietsch and Liebert, 2009; Wangchuk et al., 2014; Ndze et al., 2013a, 2013b).

This hypothesis seems to be confirmed in our study, in which a much larger set of complete genome sequence of G12[P8] strains was used.

In countries where universal vaccination for RVA has been introduced, changes in circulating RVA genotypes have been observed (Gurgel et al., 2007, Zeller et al., 2012). Since G12 specificity is not included in the vaccine formulations currently available in the market, G12 RVAs have the potential, along with G9 RVAs, to affect the efficacy of the vaccines, when heterotypic immunity is not adequate to protect against RVA disease (Estes and Greenberg, 2013). Continual epidemiologic surveillance is important to monitor the efficacy of the vaccines and the local/global emergence of novel/unusual RVA strains.

**Funding Sources**

Funding was received from the Hungarian Scientific Research Fund (OTKA, T100727).

**Acknowledgments**

Contributors for samples collection: Francesca Di Bernardo, Piera Dones, Antonina Collura, ARNAS Ospedale Civico e Di Cristina, Palermo, Italy; Giuseppe Ferrera, Service of Epidemiology and Preventive Medicine, ASP7, Ragusa, Italy; Maria T. Allò Service of Epidemiology and Preventive Medicine, ASP7, Ragusa, Italy; Mariella Valenzize, Department of Pediatrics, Gynecology, Microbiological and Biomedical Sciences, University of Messina, Italy.

**References**


---

**Table 3**

Point mutations involving the antigenic epitopes of VP8* portion of VP4. The amino acid residues of G12P[8] strains of clusters a and b in relation with the vaccine, Wa and reference strains (20131774166, GER126 and KDH651) are shown. Conserved amino acids are indicated by periods.

<table>
<thead>
<tr>
<th>VP8* Residues</th>
<th>8-1</th>
<th>8-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RVA/Human-tc/USA/1974/G1P[8]</td>
<td>L</td>
<td>N</td>
</tr>
<tr>
<td>RVA/Human-tc/USA/20131774166/G12P[8]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RVA/Human-tc/GER/GER126-08/2008/G12P[8]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 G12P[8]-II Cluster a</td>
<td>D</td>
<td>S</td>
</tr>
<tr>
<td>10 G12P[8]-III Cluster b</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


