Phylogenetic analysis of partial VP7 gene of the emerging human group A rotavirus G12 strains circulating in Tunisia

Amal Moussa,1,2 Mouna Ben Hadj Fredj,1,3 Meriam BenHamida-Rebaï,1,2 Imene Fodha,1,2 Noureddine Boujaafar2,4 and Abdelhalim Trabelsi1,2,*

Abstract

Purpose. Group A rotavirus (RVA) is the leading cause of severe gastroenteritis in children younger than 5 years. The most common human G-types are G1-4 and G9. G12 genotype is currently emerging worldwide, becoming the sixth most prevalent RVA G-genotype. In Tunisia, an emergence of G12 RVA strains was observed. To understand the evolution and origin of these Tunisian G12 strains, phylogenetic analyses were conducted.

Methodology. A total of 1127 faecal samples were collected from Tunisian children under 5 years consulting for gastroenteritis between 2009 and 2014. Samples were screened by ELISA for the presence of RVA antigen. RVA-positive samples were used for the detection of G12 RVA strains by semi-nested RT-PCR. G12-positive specimens were subjected to VP4 genotyping reaction. PCR products of the G12-positive samples were sequenced and characterized by phylogenetic analysis of partial VP7 gene sequence.

Results. Globally, 270 (24 %) stool specimens were RVA-positive. Fourteen presented the G12 genotype (5.2 %) and were found to be in combination with either the P[6] (50.0 %) or the P[8] (50.0 %) genotype. Phylogenetic analysis revealed that all characterized Tunisian G12 strains clustered in the modern G12 lineage III and appear to form three different subclusters.

Conclusion. Thus, the Tunisian G12 strains may have originated from not a single, but at least three distinct ancestral G12 strains. Detailed molecular characterization of the entire genome of these strains remains essential to help determine the extent of genetic variation and the relatedness of Tunisian G12 RVA strains to G12 strains described worldwide.

INTRODUCTION

Group A rotavirus (RVA), a member of the Reoviridae family, is the leading etiological agent of severe gastroenteritis in the young of humans and many animal species worldwide [1]. In humans, RVA infections are associated with high morbidity and mortality and are estimated to cause between 196 000 and 453 000 deaths every year among children <5 years of age, with more than 90 % of these deaths estimated to occur in developing countries in Asia and Africa [2–4].

The RVA virion is a triple-layered, non-enveloped icosahedron, enclosing a genome of 11 segments of dsRNA that encode six viral structural proteins (VP1–VP4, VP6 and VP7) and six nonstructural proteins (NSP1–NSP6) [5]. Because of the segmented nature of the genome, reassortment between/within human and animal strains is one of the major processes of genetic evolution of this virus. The three main antigenic proteins of the virus are as follows: the middle layer protein VP6, which is used to classify rotaviruses into groups A–H [6], and the two outer capsid proteins VP7 and VP4, which are implicated independently in neutralization. Based on the serotype specificities and the nucleotide sequence diversity of their two outer capsid protein genes, VP7 and VP4, RVA have been classified into G (glycoprotein) and P (protease-sensitive) genotypes, and to date, 27 G- and 37 P-genotypes have been identified [7, 8]. Recently, with

Received 21 October 2016; Accepted 30 December 2016

Author affiliations: 1 LR14SP02, Epidemiology and Immunogenetics of Human Viral Infections, Laboratory of Microbiology, Sahloul University Hospital, 4054 Sousse, Tunisia; 2 Faculty of Pharmacy, University of Monastir, 5000 Monastir, Tunisia; 3 Faculty of Sciences and Techniques, University of Kairouan, 9100 Kairouan, Tunisia; 4 Laboratory of Microbiology, Sahloul University Hospital, 4054 Sousse, Tunisia.

*Correspondence: Abdelhalim Trabelsi, abdelhalim.trabelsi@gmail.com

Keywords: G12 strains; human group A rotavirus; phylogenetic analysis; VP7 gene.

Abbreviations: RVA, group A rotavirus; WHO, World Health Organization.

increasing numbers of complete RV genome sequences becoming available, a standardized RV strain nomenclature system has been proposed by the Rotavirus Classification Working Group; each individual rotavirus strain has to be named as follows: RV group/species of origin/country of identification/common name/year of identification/G- and P-type [7]. Of all possible VP7/VP4 combinations, G1P[8], G2P[4], G3P[8], G4P[8] and G9P[8] are considered the most prevalent in humans worldwide, accounting for approximately 75% of cases globally [9]. However, over the last decade, G12P[8] emerging genotype has been detected with high frequency in many areas of the world [10–13]. After RVA genotype G12 was first detected in children <2 years old with acute diarrhoea in 1987 in the Philippines [14], it was detected only sporadically for more than 13 years [15–18]. In subsequent years, this genotype has re-emerged worldwide, becoming the sixth most prevalent RVA VP7 genotype [11, 12]. The G12 VP7 specificity is predominantly found to be in combination with either the P[6] or P[8] VP4 genotype and less commonly with P[4] or P[9] [11, 12, 17–23]. Since 2004, G12 strains have been increasingly identified in several African countries [20, 21, 24–31], indicating the ongoing expansion of G12 strains in Africa.

In Tunisia, the first case of infection with rotavirus G12 was identified in 2010 [32]. Thereafter, through retrospective and continuous surveillance (2009–2014) for human RVA circulating in Tunisia, an emergence of G12 RVA strains could be observed [33]. In order to better understand the evolution and origin of these emerging strains, phylogenetic analyses were conducted among Tunisian G12 RVA isolated strains.

METHODS

Specimen collection

During ongoing surveillance of RVA circulating strains in Tunisia, a total of 1127 faecal specimens were collected from children younger than 5 years. Patients were either hospitalized in a paediatric unit or consulting for gastroenteritis, between January 2009 and December 2014. Fourteen distinct hospitals were included in this multi-centre study realized through 11 Tunisian cities: Sousse (Sahloul University Hospital; Farhat Hached University Hospital; Hospital of M’Saken), Monastir (Fattouma Bourguiba University Hospital; Hadj Ali Soua, Ksar Helal), Mahdia (Tahar Sfar University Hospital), Sfax (Hedi Chaker University Hospital), Gabes (Hospital of Gabes), Nabeul (Hospital of Nabeul), Tunis (Bechir Hamza Children Hospital), Kairouan (Ibn El Jazzar Hospital), Jendouba (Hospital of Jendouba), Beja (Hospital of Beja) and Bizerte (Hospital of Bizerte).

A standardized questionnaire including the duration of diarrhoea, maximum number of stools passed per day, duration and peak frequency of vomiting and degree of fever was filled in during the medical visit for each child.

All samples were immediately screened for RVA detection. RVA-positive samples were stored at −20 °C until their molecular analysis.

Detection of RVA antigens

All samples were processed for rapid diagnosis by direct sandwich ELISA (IDEIA Rotavirus; Dako) according to the manufacturer’s instructions.

RNA extraction

Viral RNA was extracted and purified from 10% facal suspensions in PBS using the TRizol method (Gibco BRL, Invitrogen).

Detection of G12 RVA strains

Detection of G12 RVA was performed as described previously by Das et al. [34]. The extracted RNA was used for RT-PCR, using 9con1 and VP7R as specific VP7 primer pair to amplify the full-length segment coding for VP7 [34]. Afterwards, a semi-nested PCR was performed using G12 specific primer (5’-TAACGCTAATGAAATTGCTAGT3’; nt498–nt475) [18] to yield an amplicon of 464 bp.

VP4 genotyping

All G12-positive specimens were subjected to VP4 genotyping reaction. VP4 genotyping of RVA was performed as described previously by Gentsch et al. [35]. The extracted RNA was used for RT-PCR using primers VP4F and VP4R to amplify a partial VP4 gene [36]. For the VP4 specificity, VP4 genotypes were determined by a semi-nested multiplex PCR using a cocktail of specific primers designed for common P[4], P[6], P[8] and unconventional P[9], P[10] and P[11] human VP4 genotypes [35–37].

All amplified products were examined by gel electrophoresis in 2% agarose gels stained with ethidium bromide and viewed under UV illumination.

Sequencing reaction

PCR products obtained after the PCR aiming to detect the G12 VP7 specificity were sequenced. Amplicons were first purified using ExoSAP-IT™ reagent (USB Corporation) [38]. Sequencing reactions were performed in the DNA Engine Tetrad 2 Peltier Thermal Cycler (Bio-Rad) using the ABI BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems), on an automated sequencer (ABI 3730xl DNA Analyzer; Applied Biosystems).

Sequence and phylogenetic analysis

The nucleotide sequences obtained were inspected using Chromas 2.3 (Technelysium) and compared to sequences deposited in the National Center for Biotechnology Information (NCBI, National Institutes of Health) BLAST® server on the GenBank database. Phylogenetic analysis was conducted using the neighbor-joining method in the MEGA, version 7 (www.megasoftware.net) [39].

RESULTS

Rate of detection of G12 RVA strains

Among 1127 stool specimens screened for RVA, 270 (24%) were positive by ELISA. Among these 270 RVA-positive samples, 14 strains (5.2%) presented the G12 VP7
specificity: one case was detected in 2009 and 13 in 2010. Clinical features of children infected by G12 RVA are described in Table 1.

**VP4 genotyping**

G12 RVA strains were found to be in combination with either the P[6] or the P[8] VP4 genotype: indeed, seven G12P[6] (50.0%) and seven G12P[8] (50.0%) strains could be detected.

**Partial VP7 gene sequencing**

The presence of highly positive PCR products, visualized by a specific band of great intensity on agarose gel, is necessary to allow a successful sequencing. Unfortunately, PCR products of the entire VP7 gene did not reveal a sufficient intensity for sequencing. Therefore, amplicons obtained after the semi-nested PCR were used for the sequencing and VP7 genes of G12 strains were only partially analysed.

A total of 11 PCR products were successfully sequenced and presented interpretable results. The failure of the remaining sequencing reactions (3 out of 14) can be explained by a too low viral load in PCR products.

**Nucleotide sequence accession numbers**

The nucleotide sequences of G12 RVA strains characterized in the present study have been deposited in the GenBank database (accession numbers KX451284–KX451293).

The strain RVA/Human-wt/TUN/34333/2010/G12P8 (accession number JX986579) was already sequenced and published by Mouna et al. [32].

**Phylogenetic analysis**

The 11 Tunisian G12 sequences were compared with relevant RVA sequences available in GenBank (Fig. 1). The partial VP7 gene sequence analysis showed high to maximum nucleotides and amino acid similarity between all Tunisian G12 RVA strains. Multiple sequence alignment and phylogenetic analysis revealed that all characterized Tunisian G12 RVA strains clustered in the modern G12 lineage III (Fig. 1).

Human Tunisian G12 strains clustered separately from the porcine G12 strain, RU172, and from the first G12 strain, L26, detected in the Philippines. Furthermore,

---

**Table 1. Clinical features of Tunisian children infected by G12 RVA**

<table>
<thead>
<tr>
<th>Strain designation</th>
<th>Date of isolation (mo/yr)</th>
<th>Age (months)</th>
<th>Gender</th>
<th>City of detection</th>
<th>Diarrhoea</th>
<th>Temperature (°C)</th>
<th>Vomiting</th>
<th>GenBank accession numbers</th>
</tr>
</thead>
</table>

ND, Data not determined.

*RVA strain described by Mouna et al. [32].
Tunisian G12 strains appear to form three different subclusters. Most Tunisian strains (9 among 11) were very closely related to one another with maximum similarity among themselves and clustered with G12 strains from Saudi Arabia, South Korea, Pakistan, Belgium and France with maximum homology (100% at the nucleotide level). On the other hand, each one of the two remaining strains formed a distinct subcluster, indicating its distinct origin: Tunisian strain 33572 closely related to G12 strains isolated in Spain and Cameroon, while strain 34333 was closely associated with G12 strains described in Bangladesh.

DISCUSSION

Genetic characterization of the partial sequence of gene encoding the VP7 protein was carried out in order to study the relationship existing between Tunisian G12 RVA strains and G12 strains reported elsewhere.

The global spread of G12 rotavirus strains from the Philippines to USA, South America, South Asia, Africa and finally to Europe appeared to be similar to that of G9 which became the fifth most important global genotype in the late 1990s [40, 41]. First, both genotypes were identified in humans about two decades ago and thereafter remained...
rarely detected for many years despite intensified surveillance activity and use of sensitive and specific molecular typing methods. Second, sequence analysis and molecular studies demonstrated significant genetic and antigenic differences between early and modern variants within both genotypes, suggesting a new or separate introduction into humans. Third, of several coexisting genetic lineages, only one, the modern lineage of each serotype, has been recognized as being capable of spreading globally. It will be interesting to determine whether, as with the spread of G9 strains [42], emergence of G12 rotavirus will occur predominantly through reassortment of a single gene, VP7, into a background of globally common human Wa genogroup strains [43].

Phylogenetic analysis showed that all Tunisian G12 RVA strains clustered within the modern lineage III, as well as strains isolated in South Korea, Saudi Arabia, Nepal, USA and Bangladesh, indicating the derivation of those strains from a common origin. On the other hand, within the lineage III, Tunisian G12 strains appear to form three different subclusters. Thus, the G12 strains that have emerged in Tunisia may have originated from not a single, but at least three distinct ancestral G12 strains.

According to literature data, based on phylogenetic and phylodynamic analyses, G12 strains are subdivided into four lineages [11, 12]. Lineage I contains the prototype L26, the first G12P[4] strain, which was identified in the Philippines in 1987 [14, 44]. Lineage II consists of G12P[9] strains isolated in South America and Asia. Lineage IV comprises the only porcine G12 strain known to date: the Indian strain RU172 detected in 2002, found to be in combination with P[7] VP4 specificity [45]. Lineage III contains the majority of the currently known G12 strains (more than 90%) which are associated with P[6] or P[8] VP4 genotypes [11]. The exact relationship existing between these four distinct lineages could become apparent only after detailed molecular analysis.

To date, the exact origin of G12 strains remains unclear. The diversity found among the G12 strains described worldwide may suggest either accidental human infection by animal rotavirus strains or reassortant rotaviruses generated in nature between animal and human RVA strains. Since animal rotaviruses and human-animal reassortant rotaviruses circulate simultaneously in the environment, more intensified investigation of animal rotavirus strains should be conducted to evaluate the origin of emerging genotypes. Indeed, such survey may identify the possible animal ancestor of each new genetic lineage.

In 2009, two RVA vaccines, Rotarix® and RotaTeq®, have been recommended by the World Health Organization (WHO) for routine immunization of all infants: Rotarix® (GlaxoSmithKline Biologicals) is a monovalent live attenuated vaccine derived from a G1P1A[8] human strain [46] and RotaTeq® (Merck) is a pentavalent live vaccine composed of a bovine (strain WC3) background into which human G1-G4 genotypes and the P1A[8] genotype were inserted by reassortment [47, 48].

In Tunisia, Rotarix® was introduced on the private market in 2009 according to the WHO recommendation but has not yet been included in the national childhood vaccination program due to its high cost per dose (82 TND/dose; 32.8 EUR/dose) [32].

G12 RVA strains might escape recognition by host immune system which recognizes the common G1–G4 genotypes and may not be adequately covered by existing vaccines [18]. Supporting such a hypothesis, simultaneously to the wide adoption of RVA vaccination program in different parts of the world, an unusually high prevalence of G12 strains was observed [13, 49, 50]. Nevertheless, since G12 genotype is usually associated with P[8] genotype which is present in both licensed vaccines, these vaccines should be effective against G12P[8] RVA strains [51]. Thus, in the USA, the high efficacy of the RotaTeq vaccine (83% [95% confidence interval: 57 to 93%]) against G12 RVA strains has been observed [52] and in Africa, Rotarix vaccine showed 51.5% (95% confidence interval: −6.5 to 77.9%) efficacy against severe gastroenteritis caused by G12 RVA strains [53]. In fact, efficacy trials of the two licensed RVA vaccines (Rotarix, GlaxoSmithKline and RotaTeq, Merck) against G12 RVA strains remain quite limited and must be continued [51].

The present study documents the phylogenetic analysis of a partial sequence of the VP7 gene of the emerging G12 genotype in Tunisia. Detailed and complete molecular characterization of the entire genome of these strains remains essential to help determine the extent of genetic variation and the relatedness of Tunisian G12 strains to others. As more sequence data become available currently, the evolutionary history of rotaviruses should become clearer soon and may yield insight into future changes of the rotavirus genome [54]. Moreover, continued rotavirus surveillance has to be performed in order to track the emergence of new strains, document possible shifts in strain prevalence and assess the ability of current vaccines to protect children effectively against new unconventional genotypes.

Funding information
This study was funded by the WHO (GL.GLO.IVD.646.XC.04.2.999.00).

Acknowledgements
We thank the staff of laboratories of Microbiology for their technical help and we are grateful to the nursing staff of the Paediatric Units of all hospitals included in Tunisian rotavirus surveillance group for assisting in sample collection. This study was supported by the Ministry of Higher Education and Scientific Research of Tunisia.

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
This article does not contain any studies with human participants or animals performed by any of the authors.
Reference


---

**Five reasons to publish your next article with a Microbiology Society journal**

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.