Amino Acid Sequence Analysis of the Two Major Outer Capsid Proteins (VP7 and VP4) from Human-Derived Canine G3P[3] Rotavirus Strain Detected in Brazil


Adriana Luchs¹, Maria do Carmo Sampaio Tavares Timenetsky¹

1. Instituto Adolfo Lutz, Centro de Virologia, Núcleo de Doenças Entéricas, São Paulo, SP, Brasil

Abstract

Introduction: A close look at the rotavirus group A (RVA) genotypes in Brazil revealed the detection of a rare G3P[3] strain close related to canine strains. The aim of this study was to add to the already known genetic analysis by the description of the G3P[3] (IAL-R2638 strain) amino acid characteristics.

Methods: Amino acid sequence analysis and protein based trees were conducted using BioEdit and MEGA 4.0.

Results: The VP7 and VP4 protein of the IAL-R2638 strain displayed the highest amino acid identity to the canine-derived human strain HCR3A (99.2%), and to the canine strain RV52/96 (96.4%), respectively. IAL-R2638 strain did not possess an extra VP7 N-linked glycosylation site at amino acid 238 recently described for some G3 strains, as well as RotaTeq™ G3 vaccine strain. The topology exhibited by phylogenetic trees in previous analysis were maintained in the present amino acid-based trees, reinforcing a stable relationship between G3P[3] strains.

Conclusions: Amino acid analysis data were consistent with the previous sequence of data obtained for the IAL-R2638 strain, supporting its possible canine origin. Theoretically, RotaTeq™ vaccine could efficiently protect against G3P[3] infections based on the lack of the extra VP7 N-linked glycosylation site at amino acid 238. Phylogenetic analysis hypothesizes that all features undergo evolution independently of each other; however, unfavorable effects of nucleotide substitutions may be compensated by substitutions in other positions. The present study raises the question as to whether the amino acid-based trees could be applied as an approach to the study of RVA evolution, avoiding incorrect phylogenetic reconstructions.


Resumo


Correspondence: Adriana Luchs, Maria do Carmo Sampaio Tavares Timenetsky. Instituto Adolfo Lutz, Centro de Virologia, Núcleo de Doenças Entéricas. Address: Av. Dr. Arnaldo, nº 355, São Paulo, SP, Brasil, 01246-902. Fax: 55-11-3088 3753; phone: 55-11-3068 2909. E-mail: driluchs@gmail.com; timenetsky.m@gmail.com

Conflicts of interest: None declared.

Funding: None.

Ethical Approval: Previous Ethics Committee approval was granted by Adolfo Lutz Institute – São Paulo – Brazil (Ref. 14/05; Ref. 53/05). This was an anonymous unlinked study and informed consent was not required according to the resolution 196/96 concerning research evolving humans beings – Conselho Nacional de Saúde (CNS)/Ministério da Saúde (MS), Brasília, 1996.

Received 13 Oct 2013; Revised 14 Nov 2013; Accepted 27 Nov 2013.
INTRODUCTION

Group A rotavirus (RVA) are an important cause of severe gastroenteritis among infants and young children worldwide, as well as in animals of a wide variety of species.1 RVA is classified into different P and G genotypes based upon the main neutralization antigens, namely, the spike protein (VP4) and the major outer capsid glycoprotein (VP7). The most prevalent genotypes, G1P[8], G2P[4], G3P[8], G4P[8], and G9P[8], are responsible for approximately 90% of worldwide human RVA infections.2 However, RVA epidemiology is constantly changing, mainly due to reassortment events, and interspecies transmission of animal RVA to humans.3

With regard to the high incidence of RVA infections worldwide, two oral live-attenuated vaccines have been licensed. Rotarix™ is a monovalent vaccine derived from a human G1P[8] strain, and has been introduced in Brazilian Immunization Program since 2006.4,6 Rotarix™ is a human-bovine reassortant RVA vaccine that contains the human genotypes G1, G2, G3, G4, and P[8],6 and is available at private clinics. The post-marketing surveillance of circulating RVA genotypes is crucial for vaccine efficacy studies. In addition, selective vaccine pressure could also increase the circulation of uncommon animal-like strains.7 In fact, a close look at RVA genotypes in Brazil has already revealed the detection of uncommon G3P[3] strain with similarity to canine strains in 2012.8

Sequence similarities have been used as evidence for evolutionary relationships between strains. However, phylogenetic trees based on the analyses of DNA sequence may be misleading, especially when G+C content differs widely among lineages; therefore protein-based trees from amino acid sequences may be more reliable.9,10,11 Moreover; simulations on phylogenetic trees based on the analyses of DNA sequence of circulating RVA genotypes is crucial for vaccine efficacy studies. In addition, selective vaccine pressure could also increase the circulation of uncommon animal-like strains.7

In fact, a close look at RVA genotypes in Brazil has already revealed the detection of uncommon G3P[3] strain with similarity to canine strains in 2012.8

Sequence similarities have been used as evidence for evolutionary relationships between strains. However, phylogenetic trees based on the analyses of DNA sequence may be misleading, especially when G+C content differs widely among lineages; therefore protein-based trees from amino acid sequences may be more reliable.9,10,11 Moreover; simulations on phylogenetic trees based on the analyses of DNA sequence of circulating RVA genotypes is crucial for vaccine efficacy studies. In addition, selective vaccine pressure could also increase the circulation of uncommon animal-like strains.7

In fact, a close look at RVA genotypes in Brazil has already revealed the detection of uncommon G3P[3] strain with similarity to canine strains in 2012.8

MATERIAL AND METHODS

Strain

RVA IAL-R2638 strain (G3P[3]) was isolated from a 1 year old child with gastroenteritis symptoms in São Paulo, Brazil in 2011.8 Partial VP7 and VP4 sequences suggested a common origin between this human strain and canine strains.8 The amino acid characteristic of IAL-R2638 strain was determined in this study.

RNA extraction and reverse transcription-PCR (RT-PCR)

Viral RNA was extracted using QIAamp® Viral RNA Mini kit (Qiagen, Inc., Valencia, CA) according to the manufacturer’s instructions. RT-PCR for VP7 and VP4 genes was performed according to the protocol previously described.13,14

Nucleotide sequencing

The G3P[3] (IAL-R2638 strain) PCR amplicons were sequenced using the BigDye™ kit v3.1 (Applied Biosystems, Inc., Foster City, CA) with primers Beg9 and End9 for the VP7 gene; Con2 and Con3 for the VP4 gene. Dye-labeled products were sequenced using an ABI 3130 sequencer (Applied Biosystems, Inc., Foster city, CA, USA).

Protein sequence analysis

The sequences obtained for the VP7 and VP4 (accession number JN848803 and JN848804, respectively) genes, and a set of cognate sequences of human and animal RVA available in the GenBank database were assembled with the BioEdit Sequence Alignment Editor (version 7.0.5.2) program. Protein sequence analysis was conducted using BioEdit and MEGA software version 4.0.15 Amino acid trees were constructed using the neighbor-joining method based on Poisson correction model.

Results

The deduced amino acid sequence of the VP7 gene from human strain IAL-R2638 was determined and compared to those of reference RVA strains belonging to the G3 genotype (Figure 1). The VP7 amino acid sequence of strain IAL-R2638 was 90.5 to 99.2% identical to those of RVA strains exhibiting G3 genotype specificity, and the highest amino acid identity (99.2% and 98.8%) was found to human RVA strain HCR3A (USA) and Ro1845 (Israel), respectively. Strain IAL-R2638 also exhibited high amino acid identity to canine strains A79-10 (98.8%), CU-1 (98.8%), RV198-95 (98.8%), K9 (98.4%), and to feline strain Cat97 (98.8%) (Figure 2). The lowest amino acid identity was observed between strain IAL-R2638 and murine strain MelMuRV (90.5%) (Figure 2).

The VP7 protein of strain IAL-R2638 had a potential N-linked glycosylation site located at amino acid 69 (Asn) (Figure 1). The antigenic regions A-F of strain IAL-R2638 clearly support its classification as genotype G3. Within the VP7 hyper variable regions A-F, the VP7 of human strain IAL-R2638 was completely identical to the canine strains A79-10, CU-1, and RV198-95; to the human strains HCR3A and Ro1845; and to the feline strain Cat97 (Figure 1). VP7 hyper-variable region B (amino acids 87 to 101) seems to be the most conserved region considering all strains analyzed, while region A (amino acid 39 to 50) seems to be the most variable (Figure 1).
The alignment of the deduced amino acid sequences from VP7 gene revealed amino acid substitutions inside the variable region D (amino acids 143 to 152) at position 147A→T; region E (amino acids 207 to 220) at positions 212A/T→V and 213A/T/N→V; and region F (amino acids 233 to 242) at position 242A/N/S→V. Amino acid substitutions were also observed outside VP7 hyper-variable regions in IAL-R2638 strain: 66A→P, 76F→L, 221A→T and 278A→V (Figure 1).

Among RVA strains with G3 genotype specificity, two distinct VP7 branches or groups were observed in amino acid-based tree, and designated A and B (Figure 2). The Brazilian RVA strain IAL-R2638 clustered into group A, along with human and animal RVA G3 genotypes from different countries, including the reference simian strains RRV (EU636932) and SA11-H96 (DQ838620) both from USA (Figure 2). Group B comprised most of the human G3 strains, excluding the porcine strain A131 (L35055) and the feline strain Cat2 (EU708961) (Figure 2).

**Figure 1** Deduced amino acid sequence of the VP7 protein of human strain IAL-R2638, and of a selection of G3 rotaviruses. The VP7 antigenic regions A-F are indicated. The glycosylation site NST (amino acids 69 to 71) is indicated by asterisks. Species and isolate of each strain are indicated.
Proteins (VP7 and VP4) from Canine-Related Human Rotavirus G3P[3]

Figure 2 Neighbor-joining tree of the partial VP7 deduced amino acid sequence generated with MEGA 4.0 software of the IAL-R2638 strain (arrow). Reference G3 strains were obtained from GenBank database. Accession number, species, isolates, countries and year of each strain are indicated. Group A and B represent two clusters genetically distinct. The scale indicates the number of divergent amino acid residues. Percentages of bootstrap values are shown at the branch node.

Figure 3 shows the deduced amino acid sequence of the VP4 (subunit VP8*) of the human strain IAL-R2638 and representative VP4 amino acid sequences of RVA P[3] genotype. The potential cleavage sites, arginine (R) at positions 231 and 241, were maintained in strain IAL-R2638. The third arginine at position 247 in strain IAL-R2638 was substituted by a lysine (K). The same substitution was also observed in feline strains Cat97 and FRV64; in canine strains A79-10, K9, CU-1, and RV52/96; and in human strains HCR3A, Ro1845, 6212, and 6235. The highly conserved cysteines (C) at residues 203 and 216, and proline (P) at residues 68, 71, 225 and 226 were maintained in strain IAL-R2638 (Figure 3).

Comparative analysis of the deduced amino acid sequences of the strain IAL-R2638 VP4 fragment (VP8* subunit) showed that the variable region between amino acid 71 and 204 was fairly conserved among all strains analyzed, confirming the classification of strain IAL-R2638 as P[3] genotype. Within the VP8* subunit variable region, a unique and remarkable substitution occurred in strain IAL-R2638 at positions 148Q→L and 149N→S (Figure 3).
Still inside the VP8* variable region, a substitution at residue 155^{Y→H} was observed in strain IAL-R2638, which was also shared by strains K9 (canine), HCR3A (human) and 6212 (human). Another two amino acid substitutions inside variable region were observed in strain IAL-R2638: 92^{I→V} and 201^{A→T} (Figure 3). Other amino acid substitutions were observed outside VP8* hyper-variable region in strain IAL-R2638: (i) an inversion in amino acid residues at positions 233^{V→I} and 234^{I→V}, and (ii) amino acid substitution at position 245^{S→P}. This last substitution was shared exclusively with human strain Ro1845 (Figure 3).

The overall amino acid sequence identity between the VP4 gene of strain IAL-R2638 and those cognate P[3] sequences ranged from 84.2 to 96.8%. The VP4 deduced amino acid sequence of IAL-R2638 showed the highest identity to P[3] canine RVA strain RV52/96 (96.8%). Strain IAL-R2638 also exhibited high amino acid identity to feline strain FRV64 (96.4%), to human strain HCR3A (96.4%), and to canine strain K9 (96.4%). The lowest amino acid identity was observed between strain IAL-R2638 and simian strain RRV (84.2%) (Figure 4). In addition, the VP4 deduced amino acid sequence of P[3] strains could be discriminated into two distinct clusters, designated C and D. The Brazilian RVA strain IAL-R2638 clustered into group C, a part of the reference simian strains RRV (EU636927) (Figure 4).
Proteins (VP7 and VP4) from Canine-Related Human Rotavirus G3P[3]

Figure 4 Neighbor-joining tree of the partial VP4 deduced amino acid sequence generate with MEGA 4.0 software of the IAL-R2638 strain (arrow). Reference P[3] strains were obtained from GenBank database. Accession number, species, isolates, countries and year of each strain are indicated. Group C and D represent two clusters genetically distinct. The scale indicates the number of divergent amino acid residues. Percentages of bootstrap values are shown at the branch node.

DISCUSSION

RVA bearing G3P[3] specificities are common in both cats and dogs.\textsuperscript{16} However, there also have been few reports describing the detection of G3P[3] strains in humans,\textsuperscript{7,17,18} including in Brazil.\textsuperscript{8,19,20} The detection of G3P[3] in humans was presumably due to the interspecies transmission from animals to humans,\textsuperscript{19} and the analysis of these genomes give general insights into the diversity and evolution of the RVA strains.\textsuperscript{3,7,8,17,22,23}

Previous G3P[3] nucleotide analysis studies have demonstrated the existence of a particular canine-feline genogroup comprising animal (i.e. Cat97, CU-1, K9, A79-10) and animal-derived human (i.e. HCR3A, Ro1845) strains\textsuperscript{16,22,23} that were distinct from genomes of human, simian, bovine or porcine origin.\textsuperscript{21} The human G3P[3] strain IAL-R2638 was also shown to belong to this genogroup,\textsuperscript{8} a finding which is confirmed in the present amino acid analysis.

The main hypothesis on the basis of many methods of phylogenetic analysis is that all features undergo evolution independently of each other.\textsuperscript{14} However, this supposition could not be completely applied for polypeptides because potentially unfavorable effect of some nucleotide substitutions may be compensated by one or several substitutions in other positions of the sequence (i.e. coordinated or correlated substitutions).\textsuperscript{24,25} Relatively high proportion of coordinated substitutions may result in significant errors in phylogenetic analysis, and an incorrect topology of a phylogenetic tree.\textsuperscript{12} It should be noted that the same topology exhibited by phylogenetic trees in Luchs et al\textsuperscript{8} study was maintained in the present amino acid-based trees, reinforcing the rather stable relationship between G3P[3] strains.

The VP7 hyper variable regions A-F was completely identical between IAL-R2638 and the strains A79-10 (canine), CU-1 (canine), RV198-95 (human), HCR3A (human), Ro1845 (human), and Cat97 (feline); with amino acid sequence identity ranging from 98.8 to 99.2%. Therefore, the amino acid divergence was located outside the variable regions. However, the evolutionary value of this data remains unclear. Interestingly, the amino acid identity of VP7 protein was higher (99.2%) between canine-derived human HCR3A (USA) and IAL-R2638 (Brazil) strains. Taking the risks, it could be suggested that canine RVA might acquire mutations during replication in a heterologous host (i.e. humans). Geographic location may also play a role in this high amino acid identity observed, since both strains were detected in American Continent.
In addition, IAL-R2638 strain shared the amino acid substitutions at positions 147A→T, 212A/T→V, 213A/T/N→V, and 242V/N/A→V with the canine-feline genogroup strains, including the canine-feline-derived human RVA strains (i.e. HCR3A and Ro1845). The amino acid substitution at position 147A→T occurred inside of the major antigenic site, region D (amino acids 143 to 152). The amino acid substitution at positions 212A/T→V and 213A/T/N→V occurred inside the antigenic site E (amino acid 207 to 220), which is spatially very close to site D. Some authors have been described region E ranging from amino acid 208 to 223/224; therefore, the amino acid substitution at position 221A→T, could also be considered inside antigenic site E. Bearing in mind all data together, those substitutions could modify the antigenicity of the corresponding region, and maybe distort three-dimensional structure. It is known that there is a diversity of VP4 sequences in RVA. Animal RVA strains, including canine and feline P[3] strains, were found to have different amino acids in the positions corresponding to those in common human RVA. The previously identified cysteine residue at position 215, proline residues at position 224 and 225, and arginine residues at positions 230, 240, and 246 in human VP4 RVA are found at position 216 (C), 225 (P), 226 (P), 231 (R), 241 (R), and 247 (R), respectively, in animal strains. These differences at amino acids positions were observed in the IAL-R2638 analysis. In fact, a comparison of the VP8* fragment amino acid sequences of seven representative P genotype strains (Wa and W61 for genotype P[8], D5-1 for genotype P[4], Gottfried and M37 for genotype P[6], K9 and HCR3A for genotype P[3]) indicated a insertion of three nucleotides in P[3] strains: two timines at positions 399 and 400, and one citocine at position 409 (data not shown). Probably, the insertion of these three nucleotides resulted in the addition of an amino acid after position 132.

The present data indicate that the VP4 protein of four P[3] strains (simian RRV, human CMH222, caprine GVR, and buffalo 10733) displayed only 84.2-85% amino acid identities (77.5-77.7% nucleotide identities) with the canine-feline genogroup. Previous phylogenetic analysis of relationship among G3P[3] RVA of VP4 gene have already revealed two major lineages within P[3] genotype, keeping these 4 strains together in a separate group.

In summary, data from amino acid analysis were consistent with the sequence data of VP7 and VP4 gene segments of the IAL-R2638 strain, supporting its possible canine origin. However, the choice of the variable regions of VP7 and VP4 to serve as evidence for evolutionary relationships between G3P[3] strains was arbitrary; and other combinations would probably give similar results. It is worth mentioning that this study analyzed only two RVA proteins from IAL-R2638 strain, and it is also possible that one or more of its other segments were derived from human RVA strains making it a human-animal reassortant. In addition, the present study raises the question as to whether the amino acid-based trees could be applied as an approach to the study of RVA evolution, avoiding incorrect phylogenetic reconstructions.

ACKNOWLEDGMENTS

We thank the Enteric Diseases Laboratory of Adolfo Lutz Institute staff: Rita de Cássia C. Carmona for laboratory supervision; Audrey Cilli and Simone G. Morillo for laboratorial analysis assistance; Antonio Erculiani Junior and Sirlene Henrique Rodrigues Silva for technical assistance.
REFERENCES


Como citar este artículo / How to cite this article: