Phylogenetic characterization and evolution of the VP4 gene of P[9] rotaviruses

Shamali Ahmed  
Universiti Malaysia Sabah

Takaaki Yahiro  
Universiti Malaysia Sabah

Nor Amalina Emran  
Universiti Malaysia Sabah

Lia Natasha Amit  
Universiti Malaysia Sabah

Hidekatsu Iha  
Oita Daigaku

Kamruddin Ahmed (ahmed@ums.edu.my)  
Universiti Malaysia Sabah  https://orcid.org/0000-0002-1869-3701

Research note

Keywords: Rotavirus, genotype P[9], phylogenetic analysis, reassortant, lineages

DOI: https://doi.org/10.21203/rs.3.rs-43564/v1

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Objectives: Rotavirus is one of the major causes of gastroenteritis in children under 5 years of age and is responsible for over 200,000 deaths annually. Rotavirus can evolve by reassortment, in which gene segments are exchanged between strains of different origins. Rotavirus strains with the P[9] genotype is an example of reassortment, in which the P[9] genotype is from feline species. A number of outbreaks by P[9] strains have been documented in several countries. However, details regarding the epidemiological relationships between the strains remains largely unknown. Therefore, in the present study, genetic characterization and evolutionary analyses were performed to gain insight into P[9] strains circulating in different parts of the world.

Results: The P[9] strains could be divided into five lineages, and that the common ancestor of currently circulating P[9] strains is around 168 years old. In each lineage, the strains were not only from different countries, but also from different continents. These findings suggest that none of the lineages has a specific region of distribution, and although humans have had interactions with cats for thousands of years, the ancestor of the current P[9] strain is relatively recent.

Introduction

Rotavirus is a major cause of gastroenteritis and is responsible for over 200,000 deaths annually in children under 5 years of age [1]. Rotavirus has a double-stranded RNA genome divided into 11 segments, encoding six structural (VP1–VP4, VP6, and VP7) and six nonstructural proteins (NSP 1–6) [2]. A binary classification system developed on the basis of two outer capsid proteins, VP7 and VP4, using G and P genotypes is used to identify different strains of group A rotavirus. To date 35 G genotypes and 50 P genotypes have been found [3]. Among the numerous G/P-genotype combinations, only a limited number are commonly found in human infections; these are G1P[8], G2P[4], G3P[8], G4P[8], and G9P[8].

One of the mechanisms by which rotavirus evolves is reassortment, where segments of genes are exchanged between strains of different origin. Among the known reassortment strains, G3P[9] is the least studied, but has become prominent, where the P[9] component originated from feline rotavirus [4]. The most common G genotypes in combination with P[9] are G3P[9], G6P[9], G1P[9], and G12P[9] [4–10]. The incidence of P[9] strains causing infection in humans is relatively low, at about 2.5% worldwide [11]. Despite this low incidence, cases with P[9] strains have frequently appeared in different parts of the world [12–22]. In Brazil and Ireland, around 10% and 18% of the patients infected with rotavirus had P[9] strains [20, 23]. The other point of concern is that P[9] strains might become more virulent over time with multiple reassortments have the potential to cause outbreaks in other areas.

Based on limited studies, the evolutionary patterns of the human P[9] rotaviruses appear to be complex [22]. Therefore, the present study was performed to determine the genetic relationships among the P[9] strains circulating in different countries, and their evolutionary timelines.
Materials And Methods

Phylogenetic analyses

A total of 88 full- and partial-length VP4 gene nucleotide sequences of P[9] strains were extracted from GenBank (Additional file 1). Phylogenetic analyses were conducted with the neighbour-joining method using MEGA after aligning the nucleotide sequences using CLUSTAL W [24]. The branching patterns were evaluated based on a bootstrap analysis of 1,000 replicates. In all of the phylogenetic trees, lineages were designated based on significant bootstrap values of 70%. Several phylogenetic trees were constructed using full-length gene and partial-length genes of different lengths (Additional file 2).

Nucleotide identity

The nucleotide identities of the full- and partial-length P[9] genes of different lineages were compared with BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Timeline and evolution

Evolutionary analysis was performed using the full-length nucleotide sequences of VP4 genes. We inferred a maximum clade credibility phylogenetic tree using the Bayesian Markov Chain Monte Carlo method available in BEAST version 1.6.1 [25]. The nucleotide sequences were analyzed by using a relaxed molecular clock (uncorrelated lognormal) and general time-reverse model (GTR + I model). The sequences were run for 20 million generations and sampled at every 2,000 steps. The end result was a sample size of 1,000 Bayesian trees, which was then verified for convergence by Tracer version 1.5.

Results

Phylogenetic analyses

In the phylogenetic tree constructed using 39 full-length gene sequences of P[9] strains, five lineages could be identified (Fig. 1). Lineage I consisted of G3P[9], G1P[9], G3G4P[9], and G12P[9] strain from Paraguay, Japan, China, Thailand and Korea. Paraguayan strains were of genotypes G1P[9], G3P[9], G3G4P[9], and G12P[9]. Strains from other countries were of genotype G3P[9] (Additional file 3). Lineage II consisted of strains from Italy, Australia, and Japan. All these strains were of genotype G3P[9]. Lineage III consisted of G3P[9] and G6P[9] strains from Russia, Japan, Italy, Korea, and Tunisia. Among these, strains from Russia, Korea, and Italy were G3P[9], while strains from Tunisia and Japan were G6P[9]. Lineage IV consisted of strains from the USA, Italy, and Russia; all were G3P[9]. Lineage V consisted of strains from Paraguay, Brazil, Italy and Thailand. All of the strains in lineage V were of the G12P[9] combination.

A total of 11 phylogenetic trees were constructed using different partial-length nucleotide sequences (Additional file 2) of P[9] strains. All of these could be classified into five lineages. The distribution of
strains in each lineage was consistent with the phylogenetic tree constructed using full- and partial-length gene sequences.

Overall, five lineages of P[9] strains could be identified in the phylogenetic tree constructed using 88 partial gene sequences (187–741 nt) of P[9] strains (Fig. 2). Lineage I consisted of G3P[9], G1P[9], G9P[9], GxP[9], G12P[9], and G3G4P[9] strain from Brazil, Japan, Paraguay, Thailand, China, the USA, and Korea. The strains from Brazil, were of the G3P[9], G1P[9], and G9P[9] combination (Additional file 3). G3P[9] and GxP[9] combination were from Japan. Strains from Paraguay were G3P[9], G1P[9], G12P[9], and G3G4P[9]. All strains from Thailand, China, the USA, and Korea were G3P[9]. Lineage II consisted of G3P[9] and G6P[9] strains from Italy, Australia and Japan. Genotype G3P[9] and G6P[9] were from Italy. The strains from Australia and Japan were G3P[9]. Lineage III consisted of genotype G3P[9], G6P[9], and G9P[9] from Bulgaria, Russia, Hungary, Italy, Tunisia, France, Japan, and Korea. The strains from Bulgaria were of genotype G6P[9], and G3P[9]. Russian strains were of the G3P[9] combination. The strains from Hungary were of genotype G6P[9] and G9P[9]. Italian strains were of genotype G3P[9] and G6P[9]. The strains from Tunisia and Japan were G6P[9], and from France and Korea were G3P[9]. Lineage IV consisted of genotypes G3P[9] and GxP[9] from Hungary, the USA, Italy, Russia and Israel. The strains from Hungary, the USA, Italy, and Russia were G3P[9], while the strain from Israel was GxP[9]. Lineage V consisted of strains from Paraguay, Brazil, Argentina, Japan, Thailand and Italy. All strains were of the G12P[9] combination.

**Nucleotide identity**

Comparision of nucleotide identities of 39 full-length and 88 partial-length nucleotide sequences of the outer capsid protein VP4 gene among the five lineages of rotavirus P[9] are shown in additional file 4 and 5, respectively. P[9] strains of the same lineage shared close nucleotide identity among themselves. The strains of different lineages showed a decrease in nucleotide identity, except for lineage V, which shared relatively low nucleotide identity with strains of other lineages.

**Timeline of evolution**

The phylogenetic tree constructed using the Bayesian method also showed five lineages of P[9] strains (Fig. 3), which corresponded with the lineages as determined using the neighbour-joining method. Approximately 168 years ago the current P[9] strains of rotavirus emerged (95% highest posterior density [HPD] years) and evolved into different lineages (Fig. 3). Lineage I derived from lineages II, III, and IV approximately 79 years ago (95% HPD 56.07–99.69 years) and different strains of lineage I evolved approximately 48 years ago (95% HPD 37.00–54.96 years). Lineage II diverged from lineage III approximately 70 years ago (95% HPD 49.45–87.93 years), and the strains in lineage II started evolving around 47 years ago (95% HPD 35.61–55.07 years). Lineage III diverged from lineage II approximately 70 years ago (95% HPD 49.45–87.93yrs) ago, and the strains in lineage III started evolving around 40 years ago (95% HPD 26.65–50.03 years). Lineage IV diverged from lineages II and III approximately 74 years ago (95% HPD 52.49–83.44 years), and the strains in lineage IV evolved around 47 years ago.
(95% HPD 43.09–83.44 years). Lineage V diverged from lineages I, II, III, and IV around 168 years ago, at the very beginning of the emergence of the P[9] strain.

**Discussion**

All currently circulating P[9] strains were divided into five lineages, and the strains in each lineage were from multiple countries on different continents. This suggests that strains in a lineage do not belong to a specific geographical area. This might exemplify the role of human migration in the spread of strains in different countries. Humans tend to bring their accompanying animals, which might include domestic cats, during migration, and as human migration is a continuous process, this trend is expected to continue in future.

The results of this study confirmed sequences of the VP4 gene as short as 555 nt were adequate for lineage designation. For new strains, it is still recommended where possible to use full-length nucleotide sequences for phylogenetic analysis to provide a more robust and comprehensive for analysis of the evolution, spread, and genome-wide heterogeneity of a given virus. In the present study, full-length nucleotide sequences were used during timeline evolutionary analysis because different portions of the gene have different rates of evolution, which might affect the determination of lineage age.

Lineage I contained the AU-1 strain (G3P[9]), which was the first detected P[9] strain in humans [4]. The Brazilian strains were from several outbreaks [20]; these strains were also clustered in this lineage and shared high nucleotide identity with AU-1 [26]. All Paraguayan P[9] strains of lineage I were detected in the same year; however, the P[9] genotype was in combination with different G genotypes. By contrast, the Paraguayan strains of lineage V were all G12P[9]. These were considered emerging strains [18], suggesting the possibility of outbreak. The results of the present study support the finding that Paraguayan strains share 99% nucleotide identity with T152, another G12P[9] rotavirus strain discovered in Thailand [9]. However, the underlying mechanism for the specific combination of the P[9] from lineage V with G12 rather than G3 requires further study.

When full nucleotide sequences were compared, P[9] strains of the same lineage shared close nucleotide identity among themselves. When strains of different lineages were compared, a decrease in nucleotide identity was observed, except for lineage V, which shared relatively low nucleotide identity with strains of other lineages. All of the strains in lineage V were G12P[9]; the significance of this genotype combination on nucleotide identity requires further study. Also, eight of 12 strains in lineage V were from South America, which suggests the possibility of a single clone spreading across the continent. Few differences were seen when the identities of partial-length sequences were compared with those of full-length sequences. Such differences in nucleotide identity using partial-length nucleotide sequences might be acceptable when no full-length sequences are available.

The information obtained in this study indicates that the origin of the common ancestor of currently circulating P[9] rotavirus strains was approximately 168 years ago, which might be too recent, as humans have interacted with cats for several thousand years. We postulate that there could have been several
rotavirus transmission events from cats to humans; however, older strains might have been wiped out by evolutionary constraints, and the currently circulating strains that evolved from 168-year-old ancestor could have survived and dispersed in different places with further local evolution. The Bayesian tree showed that the rate of evolution among the strains was significantly higher in the last few decades. This is possibly because of increased human–animal interaction in recent decades. It is possible that the current strains might evolve further and give rise to virulent strains of rotavirus.

Human–animal interaction has increased in recent years for several reasons, such as more humans having pets, a loss of animal habitats because of deforestation, and increase in large-scale farming and ecotourism. As a result, the potential for zoonotic transmission of viruses has increased, which could lay the foundation for the emergence of reassorted strains. The development of a common P[9] vaccine for humans and cats might help control rotavirus infection by this reassorted strain.

We conclude that the available P[9] strains could be divided into five lineages. VP4 gene as short as 555 nt were adequate for lineage designation. Although humans have had interactions with cats for thousands of years, the common ancestor of the current P[9] strain is relatively recent, around 168 years old. The rate of evolution among the strains was significantly higher in the last few decades. This is possibly because of increased human–animal interaction in recent decades. In support of this, we also found that none of the lineages has a specific region of distribution.

Limitations

- Further study is needed particularly using neocleotide sequences of P[9] strains from cats to evaluate the time-line of evolution of P[9] strains.

Declarations

Ethics approval and consent to participate

Not applicable

Consent to publish

Not applicable

Availability of data and materials

The datasets used and/or analyzed during the current study are available in the GenBank. These are also available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.
Funding

This study was supported in part by a grant, FRGS-0457-2017, from the Ministry of Higher Education, Malaysia.

Authors’ contributions

KA conceived and designed the experiments. NAE and KA performed general supervision and guidance in the research process. SA TY LNA KA performed the experiments. SA TY NAI LNA HI analyzed and interpreted the data. SA TY NAI LNA HI drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgments

Not applicable

Author details

1 Department of Pathobiology and Medical Diagnostics, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia. 2 Department of Microbiology, Faculty of Medicine, Oita University, Yufu-shi, Oita, Japan. 3 Borneo Medical and Health Research Centre, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia.

Abbreviations

RNA: Ribonucleic acid; MEGA: Molecular evolutionary genetics analysis; BEAST: Bayesian evolutionary analysis by sampling trees; GTR: General time reverse; USA: United States of America; HPD: Highest posterior density; Nt: Nucleotide.

References

1. Tate JE, Burton AH, Boschi-Pinto C, Parashar UD. Global, Regional, and National Estimates of Rotavirus Mortality in Children < 5 Years of Age, 2000–2013. Clin Infect Dis. 2016;62(Suppl 2):96–105.

2. Matthijnssens J, Ciarlet M, Rahman M, Attoui H, Banyai K, Estes MK, Gentsch JR, Iturriza-Gomara M, Kirkwood CD, Martella V, et al. Recommendations for the classification of group A rotaviruses using all 11 genomic RNA segments. Arch Virol. 2008;153:1621–9.

3. Agbemabiese CA, Nakagomi T, Gauchan P, Sherchand JB, Pandey BD, Cunliffe NA, Nakagomi O. Whole genome characterisation of a porcine-like human reassortant G26P[19] Rotavirus A strain detected in a child hospitalised for diarrhoea in Nepal, 2007. Infect Genet Evol. 2017;54:164–9.

4. Nakagomi O, Ohshima A, Aboudy Y, Shif I, Mochizuki M, Nakagomi T, Gotlieb-Stematsky T. Molecular identification by RNA-RNA hybridization of a human rotavirus that is closely related to rotaviruses of
feline and canine origin. J Clin Microbiol. 1990;28:1198–203.

5. De Grazia S, Giammanco GM, Doro R, Bonura F, Marton S, Cascio A, Martella V, Banyai K. Identification of a multi-reassortant G12P[9] rotavirus with novel VP1, VP2, VP3 and NSP2 genotypes in a child with acute gastroenteritis. Infect Genet Evol. 2015;35:34–7.

6. Isegawa Y, Nakagomi O, Nakagomi T, Ueda S. A VP4 sequence highly conserved in human rotavirus strain AU-1 and feline rotavirus strain FRV-1. J Gen Virol. 1992;73:1939–46.

7. Mladenova Z, Nawaz S, Ganesh B, Iturriza-Gomara M. Increased detection of G3P[9] and G6P[9] rotavirus strains in hospitalized children with acute diarrhea in Bulgaria. Infect Genet Evol. 2015;29:118–26.

8. Nakagomi O, Nakagomi T. Interspecies transmission of rotaviruses studied from the perspective of genogroup. Microbiol Immunol. 1993;37:337–48.

9. Pongsuwanna Y, Guntapon R, Chiwakul M, Tacharoenmuang R, Onvimala N, Wakuda M, Kobayashi N, Taniguchi K. Detection of a human rotavirus with G12 and P[9] specificity in Thailand. J Clin Microbiol. 2002;40:1390–4.

10. Yamamoto D, Kawaguchiya M, Ghosh S, Ichikawa M, Numazaki K, Kobayashi N. Detection and full genomic analysis of G6P[9] human rotavirus in Japan. Virus Genes. 2011;43:215–23.

11. Khamrin P, Maneekarn N, Peerakome S, Toniusin S, Phan TG, Okitsu S, Ushijima H. Molecular characterization of rare G3P[9] rotavirus strains isolated from children hospitalized with acute gastroenteritis. J Med Virol. 2007;79:843–51.

12. Ben Hadj Fredj M, Heylen E, Zeller M, Fodha I, Benhamida-Rebai M, Van Ranst M, Matthijnssens J, Trabelsi A. Feline origin of rotavirus strain, Tunisia, 2008. Emerg Infect Dis. 2013;19:630–4.

13. De Grazia S, Giammanco GM, Martella V, Ramirez S, Colomba C, Cascio A, Arista S. Rare AU-1-like G3P[9] human rotaviruses with a Kun-like NSP4 gene detected in children with diarrhea in Italy. J Clin Microbiol. 2008;46:357–60.

14. De Grazia S, Giammanco GM, Potgieter CA, Matthijnssens J, Banyai K, Platia MA, Colomba C, Martella V. Unusual assortment of segments in 2 rare human rotavirus genomes. Emerg Infect Dis. 2010;16:859–62.

15. Gomez MM, Resque HR, Volotao Ede M, Rose TL, da Silva MF, Heylen E, Zeller M, Matthijnssens J, Leite JP. Distinct evolutionary origins of G12P[8] and G12P[9] group A rotavirus strains circulating in Brazil. Infect Genet Evol. 2014;28:385–8.

16. Grant L, Esona M, Gentsch J, Watt J, Reid R, Weatherholtz R, Santosham M, Parashar U, O’Brien K. Detection of G3P[3] and G3P[9] rotavirus strains in American Indian children with evidence of gene reassortment between human and animal rotaviruses. J Med Virol. 2011;83:1288–99.

17. Jeong S, Than VT, Lim I, Kim W. Whole-genome analysis of a rare human Korean G3P rotavirus strain suggests a complex evolutionary origin potentially involving reassortment events between feline and bovine rotaviruses. PLoS One. 2014;9:e97127.

18. Martinez M, Amarilla AA, Galeano ME, Aquino VH, Farina N, Russomando G, Parra GI. Predominance of rotavirus G2P[4] and emergence of G12P[9] strains in Asuncion, Paraguay, 2006–2007. Arch Virol.
2010;155:525–33.

19. Nguyen TH, Than VT, Thanh HD, Kim W. Evidence of multiple reassortment events of feline-to-human rotaviruses based on a rare human G3P[9] rotavirus isolated from a patient with acute gastroenteritis. Compar Immunol Microbiol Infect Dis. 2016;46:53–9.

20. Santos N, Volotao EM, Soares CC, Albuquerque MC, da Silva FM, de Carvalho TR, Pereira CF, Chizhikov V, Hoshino Y. Rotavirus strains bearing genotype G9 or P[9] recovered from Brazilian children with diarrhea from 1997 to 1999. J Clin Microbiol. 2001;39:1157–60.

21. Theamboonlers A, Maiklang O, Thongmee T, Chieochansin T, Vuthitanachot V, Poovorawan Y. Complete genome analysis of a rare human G3P[9] rotavirus posing as an AU-1 like strain. Springer Plus. 2013;2:569.

22. Wang YH, Pang BB, Zhou X, Ghosh S, Tang WF, Peng JS, Hu Q, Zhou DJ, Kobayashi N. Complex evolutionary patterns of two rare human G3P[9] rotavirus strains possessing a feline/canine-like H6 genotype on an AU-1-like genotype constellation. Infect Genet Evol. 2013;16:103–12.

23. Lennon G, Reidy N, Cryan B, Fanning S, O’Shea H. Changing profile of rotavirus in Ireland: predominance of P[8] and emergence of P[6] and P[9] in mixed infections. J Med Virol. 2008;80:524–30.

24. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011;28:2731–9.

25. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007;7:214.

26. Tsugawa T, Rainwater-Lovett K, Tsutsumi H. Human G3P[9] rotavirus strains possessing an identical genotype constellation to AU-1 isolated at high prevalence in Brazil, 1997–1999. J Gen Virol. 2015;96:590–600.

Figures
Figure 1
Phylogenetic tree constructed based on the full nucleotide sequences of the outer capsid protein VP4 genes of P[9] rotavirus strains. The numbers adjacent to the nodes represent the bootstrap values; values <70% are not shown. The scale bar shows genetic distance, which is expressed as nucleotide substitution per site.
Figure 2

Phylogenetic tree constructed based on the partial nucleotide sequences of the outer capsid protein VP4 genes of P[9] rotavirus strains, with a length of 555 nucleotides. The numbers adjacent to the nodes represent the bootstrap values; values <70% are not shown. The scale bar shows genetic distance, which is expressed as nucleotide substitution per site.
Figure 3

Bayesian maximum credibility tree showing the ancestry of rotavirus obtained by analyzing full sequences of the outer capsid protein VP4 gene of P[9] rotavirus strains. The nodes correspond to mean age at which the lineages were separated from their most common recent ancestor; blue horizontal bars at the nodes represent the 95% highest posterior density of the most common recent ancestor. The numbers at the main nodes represent posterior values. The horizontal-axis at the bottom represents the time scale in years.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile5.docx
- Additionalfile4.docx
- Additionalfile3.docx
- Additionalfile2.docx
- Additionalfile1.docx