Infection, Genetics and Evolution 35 (2015) 184–193

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Whole genomic constellation of the first human G8 rotavirus strain detected in Japan

Chantal Ama Agbemabiese, Toyoko Nakagomi, Yen Hai Doan [1], Osamu Nakagomi *

Department of Hygiene and Molecular Epidemiology, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan

ARTICLE INFO

Article history:
Received 7 July 2015
Received in revised form 28 July 2015
Accepted 30 July 2015
Available online 11 August 2015

Keywords:
Rotavirus
G8
Genotype constellation
Ruminant rotavirus
Phylogenetic analysis
Reassortment

ABSTRACT

Human G8 *Rotavirus A* (RVA) strains are commonly detected in Africa but are rarely detected in Japan and elsewhere in the world. In this study, the whole genome sequence of the first human G8 RVA strain designated AU109 isolated in a child with acute gastroenteritis in 1994 was determined in order to understand how the strain was generated including the host species origin of its genes. The genotype constellation of AU109 was G8-P[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2. Phylogenetic analyses of the 11 genome segments revealed that its VP7 and VP1 genes were closely related to those of a Hungarian human G8P[14] RVA strain and these genes shared the most recent common ancestors in 1988 and 1982, respectively. AU109 possessed an NSP2 gene closely related to those of Chinese sheep and goat RVA strains. The remaining eight genome segments were closely related to Japanese human G2P[4] strains which circulated around 1985–1990. Bayesian evolutionary analyses revealed that the NSP2 gene of AU109 and those of the Chinese sheep and goat RVA strains diverged from a common ancestor around 1937. In conclusion, AU109 was generated through genetic reassortment event where Japanese DS-1-like G2P[4] strains circulating around 1985–1990 obtained the VP7, VP1 and NSP2 genes from unknown ruminant G8 RVA strains. These observations highlight the need for comprehensive examination of the whole genomes of RVA strains of less explored host species.

1. Introduction

Rotavirus A (RVA), a non-enveloped virus belonging to the family *Reoviridae* and genus *Rotavirus*, is an important aetiological agent of viral diarrhoea in children and the young of animals. It is responsible for an estimated mortality of 453,000 globally (Parashar et al., 2009; Tate et al., 2012). Thus in 2009 the World Health Organization (World Health Organisation, 2009) recommended the inclusion of rotavirus vaccines in all national immunization programmes worldwide to help reduce the high morbidity and mortality caused by RVA. Currently, two rotavirus vaccines, Rotarix (GlaxoSmithKline Biologicals) and RotaTeq (Merck) both of which were reported in phase III clinical trials to be safe and efficacious (Ruiz-Palacios et al., 2006; Vesikari et al., 2006) have been licensed and incorporated into the national immunization programmes of many countries. Recently, India introduced Rotavac, a locally manufactured low cost vaccine that was reported to be effective and well tolerated in children (Bhandari et al., 2014).

The rotavirus genome is composed of 11 segments of double-stranded RNA encoding six structural and six non-structural viral proteins (VP1-VP4, VP6, VP7; NSP1-NSP6) (Estes and Greenberg, 2013). Based on the two outermost capsid proteins, VP4 and VP7, RVA strains were dually classified into G and P genotypes. At least 27 G-genotypes and 37 P-genotypes have been identified (Matthijnssens et al., 2011, 2008b; Trojnar et al., 2013) and in humans, RVA strains with G1, G2, G3, G4 or G9 in combination with P[4], P[6] or P[8] have been identified as the commonest strains globally (Banyai et al., 2012; Gentsch et al., 2005; Santos and Hoshino, 2005). G12 RVA strains also increased globally as one of the important causes of diarrhoea in children (Castello

Abbreviations: RVA, *rotavirus A*; I, **I**ntermediate capsid shell; R, **R**NA polymerase; C, **C**ore shell; M, RNA-capping **M**ethyltransferase; A, interferon **A**ntagonist; N, octameric **N**TPase; T, **T**ranslation regulation; E, **E**nterotoxin; H, p**H**osphoprotein; VP, viral protein; NSP, non-structural protein; MEGA, molecular evolutionary genetics analysis; BIC, bayesian information criterion; BEAST, bayesian evolutionary analysis sampling trees; tMRCA, time of most recent common ancestor; MCMC, Markov chain Monte Carlo.

* Corresponding author at: Department of Hygiene and Molecular Epidemiology, Graduate School of Biomedical Sciences, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan.

E-mail address: onakagom@nagasaki-u.ac.jp (O. Nakagomi).

[1] Present address: Department of Virology 2, National Institute of Infectious Diseases, Tokyo, Japan

http://dx.doi.org/10.1016/j.meegid.2015.07.033
1567-1348/

et al., 2006; Cunliffe et al., 2009; Matthijnssens et al., 2010; Pun et al., 2007; Rahman et al., 2007; Uchida et al., 2006).

The dual classification system of RVA strains was extended to include the other nine genome segments; hence, the whole genome VP7-VP4-VP6-VP1-VP2-VP3-NSP1-NSP2-NSP3-NSP4-NSP5/6 of RVA strains is respectively denoted by the descriptor G*x*-P[*x*]-I*x*-R*x*-C*x*-M*x*-A*x*-N*x*-T*x*-E*x*-H*x* (*x* = Arabic numerals $\geqslant 1$). In this regard, human RVA strains were grouped into three genotype constellations namely: the Wa-like which includes G1/G3/G4/G9-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1; the DS-1-like which includes G2/G8-P-[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2 and the AU-1-like which includes G3-P[9]-I3-R3-C3-M3-A3-N3-T3-E3-H3 (Matthijnssens et al., 2008a,b, 2011).

In cattle, RVA genotype G8 accounts for about 3.5% of rotavirus diarrhoeal infections (Papp et al., 2013) and have also been detected in other less explored animal host species such as guanacos, sheep, goats and camels (Ciarlet et al., 2008; Jere et al., 2014; Louge Uriarte et al., 2014; Sieg et al., 2015). In humans, the first G8 RVA strain 69 M bearing a DS-1-like genotype constellation was detected in an Indonesian child in 1980 (Matsuno et al., 1985). Since then, a large number of G8 RVA strains bearing either DS-1-like or Wa-like genotype constellations were detected in children on the African continent some of which possessed genes of rotaviruses of animal host species origin (Dennis et al., 2014; Esona et al., 2009; Ghosh et al., 2011; Matthijnssens et al., 2006; Nakagomi et al., 2013). While G8 strains bearing the DS-1-like genotype constellation are abundant and more frequently reported than those bearing the Wa-like genotype constellation, both genotype constellations are detected across the African continent (Esona et al., 2009). However, elsewhere in the world, human G8 strains, almost exclusively those bearing the DS-1-like genotype constellation, are infrequently detected (Banyai et al., 2010; Delogu et al., 2013; Gautam et al., 2015; Ianiro et al., 2014; Pietsch et al., 2009).

In Japan, the G8 RVA strain is one of the commonly detected bovine rotaviruses (Fukai et al., 2004a,b); meanwhile, there are so far no reports of human G8 RVA strains. A single G8P[4] RVA strain designated AU109 was for the first time detected during the 1993–1994 rotavirus season in a hospital-based study conducted between 1987 and 1996 in Akita prefecture (Nakagomi et al., 2009). In 2010, Banyai and colleagues published a Hungarian human G8P[14] RVA strain whose VP7 gene was closely related to that of AU109 and whose genotype constellation was identical to that of a Spanish sheep G8 RVA strain described by Ciarlet et al. (2008). In this regard, the whole genome sequence of AU109 was determined to obtain insight into how it relates to the previously reported bovine G8 strains in Japan as well as those reported from humans and other host species elsewhere in the world.

2. Materials and methods

2.1. Rotavirus strain

During a ten-year retrospective survey between 1987 and 1996 in Akita Prefecture located in the northern part of Japan (Nakagomi et al., 2009), a single G8P[4] strain was detected during the 1993–1994 rotavirus season and was found to possess a short electropherotype. This G8P[4] strain was subsequently adapted to MA104 cells, plaque-purified thrice and designated AU109.

2.2. Whole genome amplification and sequencing

Viral RNA extracted from 140 μL of infected cell culture fluid was recovered into 60 μL of elution buffer using the QIAamp Viral RNA Mini Kit (Qiagen) following the manufacturer's protocol. Complementary DNA of the double-stranded RNA was generated by reverse transcription using the SuperScript™ III first-strand synthesis system for RT-PCR (Invitrogen) following the manufacturer's protocol.

Briefly, an initial reaction mixture consisting of viral RNA and 50 mM random primers was denatured at 97 °C for 5 min and quickly chilled on ice for 5 min. After addition of a reverse transcription reaction mix containing 20 μM dNTP and 20 U/μL SuperScript™ III reverse transcriptase to the reaction mix to make up a final volume of 20 μL, cDNA was synthesized under the following condition: 25 °C for 5 min; 42 °C for 60 min; 70 °C for 15 min. With the exception of the VP7 gene whose sequence was already deposited in the GenBank (Nakagomi et al., 2009) under the accession number AB272753, the remaining 10 genome segments were amplified by PCR from 2 μL of cDNA using gene specific primers (Supplementary Table 1) (Doan et al., 2012; Gentsch et al., 1992; Gouvea et al., 1990; Matthijnssens et al., 2008a) and GoTaq® Green Master Mix System (Promega) under the following conditions: 95 °C/5 min followed by 40 cycles of PCR at 94 °C/1 min; 45 °C/1 min; 72 °C/2–6 min depending on the size of the gene and final extension at 72 °C/8 min.

PCR products were purified using Exosap-IT purification system (USB products) following the manufacturer's protocol and sequenced in both forward and reverse directions by the fluorescent dideoxy chain termination chemistry using Big Dye Terminator Cycle Sequencing Ready Reaction Kit v3.1 (Applied Biosystems). Nucleotide sequence reads were obtained with the aid of the ABI-PRISM 3730 Genetic Analyzer (Applied Biosystems).

2.3. Sequence and phylogenetic analyses

Nucleotide sequences for each genome segment were assembled into contigs using the SeqMan program in DNAstar Lasergene core suite software v11 (DNAstar Inc.) and the genotypes were determined using the RotaC v.2.0 automated online genotyping tool for RVA strains (Maes et al., 2009). The Basic Local Alignment Search Tool on the NCBI website was employed to retrieve sequences similar to each of the 11 genome segments of AU109. Multiple sequence alignment files were constructed using the online version of Multiple Alignment using Fast Fourier Transform (MAFFT version 7). The nucleotide and amino acid identity matrices were calculated for the aligned sequences of each of the 11 genes using MEGA v6.06. Maximum likelihood phylogenetic trees were constructed using the datasets for each genome segment; the best-fit nucleotide substitution models selected for the datasets using MEGA v6.06 were based on the lowest Bayesian Information Criterion (BIC) scores (Schwarz, 1978) and were as follows: T92+G (VP7, VP4, VP6, VP2, NSP1-NSP5), TN93+G (VP1) and GTR+G+I (VP3). The trees were constructed using 1000 pseudo-replicate datasets.

2.4. Bayesian evolutionary analysis using BEAST

The time of most recent common ancestor (tMRCA) was determined for the VP7, VP1 and NSP2 genes of AU109. The sequence datasets were the same as those used for the maximum likelihood phylogenetic analysis except that sequences of strains whose year of detection were not available were omitted. The final datasets for the VP7, VP1 and NSP2 genes consisted of 187, 207 and 214 dated sequences respectively (Table 3). The Bayesian Markov chain Monte Carlo (MCMC) framework implemented in BEAST v1.8.1 (Drummond et al., 2012) was employed. Briefly, using the following substitution models: VP7 and NSP2 (TN93+G), VP1 (GTR+G), a lognormal relaxed clock (Drummond et al., 2006) and a coalescent constant size (Drummond et al., 2002) model, two