Virus Research 224 (2016) 38–45

Contents lists available at ScienceDirect

Virus Research

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

Complete genome analysis of a rare G12P[6] rotavirus isolated in Thailand in 2012 reveals a prototype strain of DS-1-like constellation

Wilaiporn Saikruang[a], Pattara Khamrin[a], Rungnapa Malasao[b], Kattareeya Kumthip[a], Hiroshi Ushijima[c], Niwat Maneekarn[a,*]

[a] Department of Microbiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand
[b] Department of Virology, Tohoku University Graduate School of Medicine, Sendai, Japan
[c] Division of Microbiology, Department of Pathology and Microbiology, Nihon University School of Medicine, Tokyo, Japan

ARTICLE INFO

Article history:
Received 15 June 2016
Received in revised form 29 July 2016
Accepted 2 August 2016
Available online 24 August 2016

Keywords:
Species A rotaviruses
Human
DS-1-like strain
G12P[6]
Thailand

ABSTRACT

Species A rotaviruses (RVAs) are a major cause of severe diarrhea in children worldwide. G12 RVA detection is currently increasing and has been reported from many countries around the world. However, few studies have reported whole genome sequences of G12 RVAs. In the present study, the complete genome sequence of a G12P[6] RVA strain (RVA/Human-wt/THA/CMHN49-12/2012/G12P[6]) detected in a stool sample from a child with acute gastroenteritis in 2012 in Thailand was analyzed. In the CMHN49-12 strain, all genome segments had a DS-1-like backbone: G12-P[6]-I2-R2-C2-M2-A2-N2-T2-E2-H2 indicates that it is most likely the prototype strain of G12P[6] with a DS-1-like genotype constellation. Based on a Bayesian evolutionary analysis of VP7 nucleotide sequence, G12 RVA strains reported previously from Thailand during the period of 2007–2012 could be divided into 3 clusters, indicating that they originated from at least 3 different ancestral G12 strains. The evolutionary rate of G12 calculated by Bayesian Markov Chain Monte Carlo analysis indicated that the nucleotide substitution rate of G12 was 1.11×10^{-3} mutations/site/year. The finding of a G12P[6] RVA possessing a DS-1-like backbone provides insights into the evolution of global G12 RVAs.

1. Introduction

Species A rotaviruses (RVAs) are a major cause of severe diarrhea in human infants and in young animals of many mammalian and avian species worldwide. Annually, RVA causes approximately 215,000 deaths of children <5 years of age, mainly in countries of sub-Saharan Africa (Tate et al., 2016). In Thailand, RVA is the most common pathogen causing acute gastroenteritis in children admitted to the hospitals with the prevalence ranging from 28.4% to 46% (Chaimongkol et al., 2012; Khamrin et al., 2010, 2007, 2006; Khananurak et al., 2010; Maiklang et al., 2012; Theamboonlers et al., 2008).

RVA is a triple-layered non-enveloped virus of icosahedral symmetry belonging to the genus *Rotavirus*, family *Reoviridae*. The RVA virion encloses a genome of 11 double-stranded RNA segments encoding six structural (VP1-VP4, VP6, VP7) and six non-structural viral proteins (NSP1-NSP5/6) (Estes and Greenberg, 2013). The RVA outer layer capsid consists of two proteins, VP7 (glycoprotein) and VP4 (protease-cleaved protein), both of which are implicated in neutralization and define the G and P genotypes, respectively. To date, at least 27G genotypes and 38P genotypes have been identified in humans and in a variety of animal species (Fujii et al., 2016; Matthijnssens et al., 2011). In humans, five G genotypes (G1–G4, G9) in combination with three P genotypes (P[4], P[6] and P[8]) are commonly associated with human rotavirus infection worldwide (Leshem et al., 2014; Santos and Hoshino, 2005). However, a few new genotypes, such as G12, have emerged and spread through many countries in several continents. G12 RVAs bearing the Wa-like genetic backbone are more frequently detected in the G12 strain worldwide than those bearing the DS-1-like genetic backbone (Ide et al., 2015; Leshem et al., 2014; Mijatovic-Rustempasic et al., 2014; Rahman et al., 2007a; Theamboonlers et al., 2008).

A whole genome-based genotyping system was recently proposed for RVAs based on the genotype assignment of all 11

Abbreviations: RVA, species A rotaviruses; VP, viral protein; NSP, non-structural protein; BLAST, basic local alignment search tool; RT-PCR, reverse transcription polymerase chain reaction; ML, maximum likelihood; AICc, corrected Akaike information criterion; tMRCA, time of the most recent common ancestor; ESS, effective sample size; MCMC, Bayesian Markov Chain Monte Carlo; HPD, highest posterior density.

* Corresponding author at: Department of Microbiology, Faculty of Medicine, Chiang Mai University, Chiang Mai 50200, Thailand.
E-mail address: niwat.m@cmu.ac.th (N. Maneekarn).

http://dx.doi.org/10.1016/j.virusres.2016.08.002
0168-1702/

gene segments (Matthijnssens et al., 2008). In this genotyping system, Gx-P[x]-Ix-Rx-Cx-Mx-Ax-Nx-Tx-Ex-Hx represent the genotypes of the VP7-VP4-VP6-VP1-VP2-VP3-NSP1-NSP2-NSP3-NSP4-NSP5/6 encoding RNA segments, respectively. RVAs were grouped into three genotype constellations; the Wa-like which includes G1-, G3-, G4-, G9-, G12-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1; the DS-1-like which includes G2-, G8-, G12-P[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2; the AU-1-like carries G3-P[9]-I3-R3-C3-M3-A3-N3-T3-E3-H3 (Matthijnssens et al., 2011; Matthijnssens and Van Ranst, 2012).

The first G12 strain, L26, was isolated from a child in the Philippines in 1987 (Taniguchi et al., 1990; Urasawa et al., 1990). After the first report, it was detected sporadically (Rahman et al., 2007a). Over the last decade, the G12 RVA was recognized to be the sixth most important global genotype, mirroring the rise of G9 in the late 1990s. Based on phylogenetic and phylodynamic analyses, G12 genotype have been subdivided into four lineages (Matthijnssens et al., 2010). Lineage I and IV each contains only one strain, the prototype G12P[4] strain L26 (Kobayashi et al., 1989) and a porcine G12P[7] strain RU172 (Ghosh et al., 2006), respectively. Lineage II contains G12P[9] strains from South America and Asia (Matthijnssens et al., 2009). Lineage III contains the majority of the current G12 strains which are associated with P[6] or P[8] genotype and are detected worldwide (Matthijnssens et al., 2009).

Up to now, only a few G12 RVAs have been fully sequenced and characterized, particularly in Thailand and all of those G12 RVAs possessing Wa-like genotype constellation (Theamboonlers et al., 2014). The overall analysis of genotype constellation and evolutionary pattern of G12 strains circulating in Asia will provide more understanding of G12 dynamics. In the present study, the whole genome of CMHN49-12 (G12P[6]) strain was analyzed.

2. Materials and methods

2.1. RVA strains

A total of 186 fecal specimens were collected in 2012 from children hospitalized with acute gastroenteritis at Nakornping Hospital, Chiang Mai province, Thailand. Stool samples were screened for the presence of RVA and then identified for their G and P genotypes by RT- multiplex PCR using the protocol described previously (Chaimongkol et al., 2012). Among these, 35 RVA strains were identified. Twenty-three strains were classified as G1P[8], eleven were G9P[8], and only one was G12P[6] genotype. The G12P[6] strain detected in this study was further characterized its entire genome sequence.

The study was conducted with the approval of the ethical committee for human rights related to human experimentation, Faculty of Medicine, Chiang Mai University (No. 181/2554). All written informed consent for testing of stool samples for RVAs were obtained from children's parents/guardians.

2.2. Whole genome amplification and sequencing

The fecal sample was prepared as a 10% suspension in phosphate-buffered saline (PBS), centrifuged at 2300 *g* for 5 min at room temperature. Viral RNA was extracted from the supernatant using the Geneaid Viral Nucleic Acid Extraction Kit II (Geneaid, Taipei, Taiwan), according to the manufacturer's protocol. The extracted viral RNA was denatured in 50% dimethyl sulfoxide (DMSO) at 95 °C for 5 min and then reverse transcribed to generate cDNA using random hexamer primers (Takara, Shiga, Japan) (Iturriza-Gomara et al., 1999) and RevertAid™ reverse transcriptase (Thermo Fisher Scientific Inc., Waltham, MA, USA) according to the manufacturer's protocol. The full length of the genes coding for VP1-VP4, VP6-VP7, NSP1-NSP5/6 were amplified using specific primers as described previously (Matthijnssens et al., 2008). The PCR products were identified using 1.5% agarose gel electrophoresis, followed by staining with RedSafe (iNtRON Biotechnology, Kyungki-Do, Korea) and observation under UV light. The PCR products were purified using a Gel/PCR DNA Fragment Extraction Kit (Geneaid, Taipei, Taiwan) according to the manufacturer's protocol. Nucleotide sequences of 11 gene segments of the genome were determined by direct sequencing using the Big-Dye terminator cycle sequencing kit and an ABI Prism 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). The sequences of VP6, VP7 and NSP1 to NSP5/6 genes were determined in both directions while the genes encoding VP1 to VP4 were sequenced by primer walking method. All primers used in the present study are listed in Supplementary Table 1. The genotypes were determined using the Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and RotaC 2.0 (Maes et al., 2009).

The nucleotide sequences presented in this paper have been deposited in the GenBank under accession numbers: VP1-4, VP6-7 (KT936626 – KT936631) and NSP1-NSP5/6 (KT936621 – KT936625).

2.3. Sequences and phylogenetic analyses

Nucleotide sequence of each genome segment was assembled and edited using SeaView software (Gouy et al., 2010). The BLAST on NCBI website was employed to retrieve the nearest relative sequences of the 11 genome segments of CMHN49-12 strain and multiple sequence alignments were constructed using Clustal X2 (Larkin et al., 2007). Phylogenetic trees of the nearly full-length nucleotide sequences of the VP1-VP4, VP6, VP7 and NSP1-NSP5/6 genes of strain CMHN49-12 were constructed by the maximum likelihood (ML) method after selecting the best fit evolutionary model for the data set, which was based on the value of the lowest corrected Akaike information criterion (AICc) using MEGA 6.06 software (Tamura et al., 2013). Models used in this study were GTR + I (VP1, VP7), GTR + G + I (VP3, VP4), T92 + G (VP6, NSP1, NSP2, NSP5), T93 + G (NSP4), T93 + I (NSP3), T93 + G + I (VP2). Statistical analysis was performed using bootstrap method with 1000 replicates.

Before estimating evolutionary rates with a molecular clock, it was important to check that there was an actual molecular clock signal in the dataset. We then estimated the clock-like behavior of the data by performing a regression between root-to-tip distance in the ML tree and the date of sampling of the sequence using the software Path-O-Gen v1.4 (http://tree.bio.ed.ac.uk/software/pathogen). This analysis estimated the amount of variation in genetic distances that could be explained by the sampling time. This provided a correlative measure of the goodness-of-fit of the data to a molecular clock (Alizon and Fraser, 2013).

2.4. Bayesian evolutionary analysis

To estimate the rate of evolution (substitutions per site per year) and the time of the most recent common ancestor (tMRCA) of the G12 genotype, the entire open reading frame (ORF) sequences of 143 VP7 genes retrieved from GenBank together with that of the CMHN49-12 strain were used in a Bayesian phylogenetic reconstruction based on the Markov chain Monte Carlo analysis available in the BEAST package v1.6.1 (http://beast.bio.ed.ac.uk). The Bayesian Markov Chain Monte Carlo (MCMC) analysis implemented in the BEAST v.1.8.2 software package (http://beast/bio/ed.ac.th) was used to estimate the rate of evolution (substitutions per site per year) and the time of the most recent common ancestor (tMRCA) (Drummond et al., 2012). A Bayesian tree was constructed using the best-fit nucleotide substitution models selected for the