Unusual G9P[4] rotavirus emerged after the dynamic changes in rotavirus genotypes from equine-like G3 to typical human G1/G3 in Indonesia

Zayyin Dinana[1], Yen Hai Doan[2], Aussie Tahta Maharani[1], Anisa Lailatul Fitria[1], Laura Navika Yamani[1], * Juniastuti[1], Rury Mega Wahyuni[1], Soegeng Soegijanto[1], Soetjipto Soetjipto[1], Takako Utsumi[3], Chieko Matsui[3], Lin Deng[3], Nobuhiro Takemae[2], Tsutomu Kageyama[2], Kazuhiko Katayama[4], Maria Inge Lusida[1], and Ikuo Shoji[3]

[1]Universitas Airlangga Institute of Tropical Disease
[2]Kokuritsu Kansensho Kenkyujo Murayama Chosha
[3]Kobe Daigaku Daigakuin Igakukei Kenkyuka Igakubu Kansensho Center Kansen Seigyogaku Bunya
[4]Kitasato Daigaku Kansen Seigyo Kenkyu Kiko

August 30, 2024

Abstract

Background: Inter-genogroup reassortment of Rotavirus A (RVA) strains has highlighted the spread of unusual RVA strains worldwide. We previously reported the equine-like G3 RVA as predominant strains in Indonesia in 2015-2016. However, since July 2017, typical human genotypes G1 and G3 have replaced these strains completely. To understand how dynamic changes in RVA occur in Indonesia, we performed a detailed epidemiological study. **Main body:** A total of 356 stool specimens were collected from hospitalized children in Sidoarjo, Indonesia between 2018 and 2022. Whole-genome sequencing was performed for all 26 RVA-positive samples using next-generation sequencing. Twenty-four samples were determined to be the unusual RVA G9P[4], while 2 were G9P[6]. Detailed analysis revealed that seven G9P[4] strains had the typical DS-1-like backbone, while the other strains exhibited a double-reassortant profile (G9-N1) on the DS-1-like backbone. The Bayesian evolutionary analyses suggested that the Indonesian G9P[4] strains share a common ancestor with previously reported G9P[4] strains in the VP7 and VP4 genes. **Conclusions:** G9P[4] DS-1-like strains were identified as the predominant genotype in Indonesia in 2021 for the first time. These results suggest that the G9P[4] strains were generated from the previous G9P[4] strains that had undergone further intra-reassortments with the other circulating strains.

Unusual G9P[4] rotavirus emerged after the dynamic changes in rotavirus genotypes from equine-like G3 to typical human G1/G3 in Indonesia

Zayyin Dinana[1,2], Yen Hai Doan[3], Aussie Tahta Maharani[1], Anisa Lailatul Fitria[1], Laura Navika Yamani[1,2], Juniastuti[1,2], Rury Mega Wahyuni[1], Soegeng Soegijanto[1], Soetjipto[1,2], Takako Utsumi[4], Chieko Matsui[4], Lin Deng[4], Nobuhiro Takemae[3], Tsutomu Kageyama[3], Kazuhiko Katayama[5], Maria Inge Lusida[1,2], Ikuo Shoji[4*]

Author Afiliation:

[1] Laboratory of Viral Diarrhea, Indonesia-Japan Collaborative Research Center for Emerging and Re-emerging Infectious Diseases, Institute of Tropical Disease, Universitas Airlangga, Surabaya, Indonesia

[2] Research Center on Global Emerging and Re-emerging Infectious Diseases, Institute of Tropical Disease, Universitas Airlangga, Surabaya, Indonesia

[3] Office of Laboratory Emergency Preparedness, Center for Emergency Preparedness and Response, National Institute of Infectious Diseases, Tokyo, Japan

[4] Division of Infectious Disease Control, Center for Infectious Diseases, Kobe University Graduate School of Medicine, Hyogo, Japan

[5] Laboratory of Viral Infection I, Department of Infection Control and Immunology, Ōmura Satoshi Memorial Institute, Graduate School of Infection Control Sciences, Kitasato University, Tokyo, Japan

*Correspondence:

Ikuo Shoji,

Centre for Infectious Diseases, Kobe University Graduate School of Medicine, 7-5-1 Kusunoki-Cho, Chuo-ku, Kobe, Hyogo 650- 0017, Japan.

Email: ishoji@med.kobe-u.ac.jp

Keywords: RVA, G9P[4], complete genome, reassortment, evolution, NGS.

Abstract

Background: Inter-genogroup reassortment of Rotavirus A (RVA) strains has highlighted the spread of unusual RVA strains worldwide. We previously reported the equine-like G3 RVA as predominant strains in Indonesia in 2015-2016. However, since July 2017, typical human genotypes G1 and G3 have replaced these strains completely. To understand how dynamic changes in RVA occur in Indonesia, we performed a detailed epidemiological study.

Main body: A total of 356 stool specimens were collected from hospitalized children in Sidoarjo, Indonesia between 2018 and 2022. Whole-genome sequencing was performed for all 26 RVA-positive samples using next-generation sequencing. Twenty-four samples were determined to be the unusual RVA G9P[4], while 2 were G9P[6]. Detailed analysis revealed that seven G9P[4] strains had the typical DS-1-like backbone, while the other strains exhibited a double-**reassortant** profile (G9-N1) on the DS-1-like backbone. The Bayesian evolutionary analyses suggested that the Indonesian G9P[4] strains share a common ancestor with previously reported G9P[4] strains in the VP7 and VP4 genes.

Conclusions: G9P[4] DS-1-like strains were identified as the predominant genotype in Indonesia in 2021 for the first time. These results suggest that the G9P[4] strains were generated from the previous G9P[4] strains that had undergone further intra-reassortments with the other circulating strains.

(200 words)

1. INTRODUCTION

Rotavirus A (RVA) is a major cause of viral acute gastroenteritis (AGE) in infants and young children worldwide[1-4]. RVA-associated AGE is a global concern, affecting 128,500 children under 5 years of age globally in 2016[5]. Notably, within Southeast Asia, RVA-associated diarrhea accounted for 40.8% of cases between 2008 and 2018, and is the leading cause of illness and mortality among children under 5 years of age[6]. In Indonesia in 2016-2018, the prevalence of RVA infection in patients under five years of age was 31.7%-55.4%, and RVA was a major cause of acute gastroenteritis in hospitalized patients in pediatric hospitals and general hospitals[7].

The first RVA vaccine was approved in the United States in 2006. The World Health Organization (WHO) currently recommends the use of four oral RVA vaccines: Rotarix® (GlaxoSmithKline Biologicals, Rixenstart, Belgium), RotaTeq® (Merck Sharp & Dohme LLC,

Rahway, NJ), RotaSiil® (Serum Institute of India, Pune, India), and Rotavac® (Bharat Biotech, Telangana, India)[8]. Clinical trials and real-world usage of these vaccines have validated the safety and efficacy of all RVA vaccines. However, their effectiveness is lower in developing countries than in developed countries. In low-income countries lacking government RVA vaccination programs, diarrhea caused by RV infection has serious consequences, contributing to the diversity of RVA genotypes[9]. In Indonesia, two RVA vaccines, Rotarix and RotaTeq, have been available commercially since 2013. However, neither of these vaccines has been included in the universal immunization program thus far.

RVA belongs to the family *Reoviridae* . RVA is a**double-stranded RNA virus consisting of 11 segments of the genome encoding six structural proteins** (VP1, VP2, VP3, VP4, VP6, and VP7 genes) **and six non-structural proteins** (NSP1, NSP2, NSP3, NSP4, and NSP5/NSP6)[2,3]. **The outer layer proteins consist** of VP7 glycoprotein (the G genotype) and VP4 protein (the P genotype)**. Up to June 2024, Rotavirus Classification Working Group has proposed at least 42G, 58P, 32I, 28R, 24C, 23M, 39A, 28N, 28T, 32E, and 28H genotypes (https://rega.kuleuven.be/cev/viralmetagenomics/virus-classification/rcwg). The common recombinant genotypes were found in seven G genotypes (G1-G4, G8, G9 and G12) and three P genotypes (P[4], P[6], and P[8]), all of which are usually associated with human infections[2,10,11]. RVA genomic constellations are divided into three groups, Wa-like constellation (G1-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1), DS-1-like (G2-P[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2), AU-1-like (G3-P[9]-I3-R3-C3-M3-A3-N3-T3-E3-H3) a minor genomic constellation[2,4,11].**

The existence of the inter-genogroup reassortants highlights the wide diversity within RVA[13]. Moreover, recent findings on these inter-genogroup reassortant strains underscore the ongoing spread of rare RVA strains across Asia and other regions[12]. Although the **reassortant** RVA strains have been described in many studies, they appeared only sporadically until the emergence of reassortant G1P[8]-DS-1-like reported in Japan in 2014[13]. We previously reported the equine-like G3 RVAs as predominant strains among children in Indonesia in 2015-2016. However, from July 2017 until 2018, the predominant RVA strains were completely replaced by typical human G1 and G3 genotypes, suggesting the dynamic changes in RVA genotypes from equine-like G3 to typical human G1 and G3[12].

The G9 genotype have emerged and spread globally since 1995. It is now acknowledged as the fifth major human RVA genotype[14]. The geographical distribution of G9 strains has contributed to their diversity, leading to the identification of various combinations involving the G9 genotype and the P genotype, including G9P[8], G9P[6], G9P[4], and G9P[11][13,15-17]. It is noteworthy that the majority of G9 RVA reported cases were characterized as a combination of P[8] genotypes[16]. Reassortant G9P[4]-DS-1 like RVA strains exhibited notably high prevalence rates in several countries, including Mexico (80%), Guatemala (66%), Bangladesh (31.8%), India (35.3%), Iran (30.3%), and Ghana (16%)[15,17-19]. In Japan, G9P[4] prevalence has surged, reaching the levels between 11.5% and 21.2% during the years 2010-2012[13]. Although unusual rotavirus strains have been described, they appeared only sporadically until the reassortment G1P[8] genotype on the DS-1 genetic backbone was discovered in several countries[20]. Notably, the emergence of G9 RVA with the P[8] genotype combination has been documented in many studies[4,6,21].

During RVA surveillance conducted in Sidoarjo, Indonesia, we initially identified four cases of G9P[4] RVA infections in the years 2018 and 2020. Surprisingly, in 2021, the majority of RVA-positive specimens (19 out of 21) were identified as G9P[4], while the remaining three strains were G9P[6]. To gain insights into the genetic characteristics of these G9P[4]/P[6] strains, we conducted whole-genome sequencing for all 26 of the detected G9P[4]/P[6] strains using next-generation sequencing (NGS). In addition, we investigated the evolutionary patterns and genetic diversity among Indonesian G9P[4] strains, along with comparisons to RVA strains detected in other countries. We demonstrate that the emergence of G9P[4] was derived from the genetic diversity circulating in Indonesia, whereas the predominant strain survived for about a year and changed rapidly alongside or in connection with other genotypes.

MATERIAL AND METHODS

Specimens

We collected 356 stool samples from children under 5 years old who were hospitalized for acute gastroenteritis (AGE), in Sidoarjo, Indonesia between August 2018 and February 2022. AGE was defined as three times or more occurrence of looser than normal stools within 24 h. All stool samples were screened by an immunochromatography, Dipstick "Eiken" Rota kit (Eiken Chemical, Tokyo). The kit detected a total of 26 RVA-positive samples, which were analyzed by RT-PCR to determine their G and P genotypes. We primarily focused on detecting and characterizing G9P[4]/P[6] strains, which were subsequently subjected to whole-genome sequencing at the National Institute of Infectious Diseases in Tokyo.

This study was approved by the Research Ethics Board of both hospitals, Universitas Airlangga (ethics approval number 2054/ UN3.14/LT/2015) in Indonesia and Kobe University (ethics approval number 1857) in Japan. Written informed consent was obtained from the children's parents or guardians.

RNA extraction and RT-PCR genotyping

To prepare a 10% (wt/vol) stool suspension, we utilized phosphate-buffered saline (PBS; pH 7.4) and clarified it by centrifugation at 21,130×g for 10 min. The supernatant was collected and stored at −80°C until further use. The viral RNA was extracted using the Direct-zol-96 MagBead RNA (with DNase treatment) kit (Zymo Research, Irvine, CA) . In brief, 100 µl of stool suspension in PBS was mixed with 300 µl of Trizol LS reagent (Thermo Fisher Scientific, Waltham, MA). Further RNA extraction steps were then carried out using the automated KingFisher Flex System (Thermo Fisher Scientific). The extracted viral RNAs were eluted using DNase/RNase-free water and were employed for RT-PCR genotyping and whole-genome analysis by NGS.

All of the positive samples were subjected to genotyping in the VP7 (G typing) and VP4 (P typing) genes by multiplex RT-PCR. The primer sets of VP7 and VP4 were described previously [12], which enabled us to genotype the equine-like G3 strain along with other strains (G1, G2, typical human G3, G4, G8, G9, and G12)[13]. The RT-PCR products were subjected to 2% agarose gel electrophoresis.

Complementary DNA (cDNA) library building and Illumina MiSeq sequencing

The cDNA library preparation and the Illumina Miseq sequencing were conducted as previously described[15]. Briefly, a 300-bp fragment library ligated with bar-coded adapters was constructed for individual strains using an NEBNext Ultra RNA library Prep Kit for Illumina ver. 1.2 (New England Biolabs, Ipswich, MA). Library purification was performed with Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA). The quality of the final DNA libraries was assessed using the High Sensitivity D1000 ScreenTape assay on 4150 TapeStation (Agilent Technologies, Santa Clara, CA) and quantified with Qubit 1X dsDNA High Sensitivity (HS) using the Qubit Flex Fluorometer (Thermo Fisher Scientific). Sequencing was performed on an Illumina MiSeq sequencer (Illumina, San Diego, CA) using the Miseq Reagent Kit ver. 2 (Illumina) to generate 151 paired end reads. Data analysis was performed using QIAGEN CLC Genomics Workbench v22 (CLC Bio, Tokyo, Japan). The nucleotide sequences of each gene segment of 26 Indonesian G9P[4]/P[6] strains was obtained by de novo assembly.

Phylogenetic analysis

Nucleotide sequence comparisons were carried out with the references retrieved from GenBank by using the BLAST program (http://blast.ncbi.nlm.nih.gov/) with the sequences of our G9P[4] and G9P[6] strains as the query sequences for the VP7 (G9 genotype), VP4 (P[4] and P[6] genotype) and NSP2 (N1 and N2 genotype) genome segments. Phylogenetic trees were constructed using Molecular Evolutionary Genetic Analysis (MEGA) ver. 11.0 with reference sequences retrieved from the DDBJ/EMBL/GenBank databases using the Maximum Likelihood method. Alignments were performed using the CLUSTAL W in MEGA 11 and phylogenetic trees were constructed by the neighbor-joining method. To confirm the reliability of phylogenetic tree analysis, bootstrap resampling and reconstruction were carried out 1,000 times[22]. The full-length genome sequences determined in this study were subjected to multiple alignment with the sequences obtained from the GenBank database using the MAFFT multiple sequence alignment, ver. 7.0[23].

Bayesian evolutionary analysis using BEAST

Estimation time of Most Recent Common Ancestor (tMRCA) was determined for the VP7, VP4 and NSP2 genes of the G9P[4] and G9P[6] strains by the Bayesian Markov chain Monte Carlo (MCMC) method in BEAST ver.1.8.1[24]. The models used for BEAST analyses of G9P[4] and G9P[6] were GTR + G (VP7), T92 + 1 (VP4) and GTR + G (NSP2). Strict clock, relaxed clock and coalescent exponential growth models were used. MCMC runs were carried out for 100 million generations and evaluated using Tracer ver.1.6[24]. Only parameters with an effective sample size of >200 were accepted. The maximum clade credibility (MCC) trees were annotated with the Tree annotator and viewed with FigTree ver.1.4[24].

Nucleotide sequence accession numbers

All the nucleotide sequences of the 11 gene segments of G9P[4] and G9P[6] have been deposited in the DDJB/GenBank/EMBL database under the accession numbers PP948908 - PP949193.

RESULTS

Nucleotide sequence and genotype constellation of Indonesian G9 strains

From 2018 to 2021, a total of 356 stool samples were collected from hospitalized children with AGE. In our previous surveillance conducted in Sidoarjo, Indonesia, from 2015 to 2017, the G9P[4] strain was not detected. Interestingly, in this study, we first detected the G9P[4] strain in 2018, and G9P[4] strain continued to be detected in 2019 and 2020, albeit with low prevalence (2 out of 44, 4.5% in 2019 and 1 out of 7, 14% in 2020). Remarkably, in 2021, the majority of RVA-positive specimens (19 out of 22, 86.4%) were identified as G9P[4], while the remaining three strains (13.6%) were G9P[6].

Among them, a total of 26 samples, collected between 2018 and 2021 (Table 1) and **tested positive for rotavirus kit, were analyzed by whole genome sequencing using NGS** Illumina MiSeq technology**. The most predominant genotype was G9P[4] (24/26, 92.3%)** followed by G9P[6] (2/26, 7.7%), respectively. The genome constellation of G9P[4] (6/19, 31.5%) was G9-P[4] (G9-P[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2), i.e., the DS-1-like backbone. The reassortant strain found in G9P[4] (14/24, 58.4%), including the G9-VP7 and N1-NSP2 genes with a DS-1-like genotype constellation (G9-P[4]-I2-R2-C2-M2-A2-N1-T2-E2-H2) (Table 2). The reassortant strain found in all G9P[6] genotypes, including the G9-VP7 and N1-NSP2 genes with a DS-1-like genotype constellation (G9-P[6]-I2-R2-C2-M2-A2-N1-T2-E2-H2) (Table 2).

Phylogenetic tree analysis

To assess the evolutionary dynamics of Indonesian reassortant G9 RVA strains at the lineage level, our phylogenetic analysis incorporated sequences from the published G9P[4] RVA strains available in the DNA database. Additionally, we included a set of reference strains for genotype 2 RVA for lineage and sub-lineage classification[21]. Furthermore, to elucidate the relationship between the DS-1-like backbone of the newly detected reassortant G9 RVA strains in this study and the equine-like G3P[8] strains found in Indonesia during 2015-2016, we analyzed the nucleotide sequences carrying genotypes I2, R2, C2, M2, A2, N2, T2, and H2 from these equine G3P[8] DS-1-like strains (Supplementary Fig. 1).

Based on the recently proposed lineage classification for globally circulating DS-1-like strains and G9 VP7 genes[13], the Indonesian reassortant G9P[4]/P[6] strains were assigned to the following significant lineages: major sub-lineage III for VP7; lineages IV for VP4, VP2, NSP1, and NSP5; lineages V for VP1, VP6, NSP2-N2, and NSP3; and lineage VI for VP3 and NSP4.

The phylogenetic trees showed that the branch topology of most segments of the Indonesian reassortant G9 RVA strains detected during 2018 to 2021 fell within the same monophyletic lineage, except for the NSP3 and NSP4 genes of three G9 strains detected in 2021 and 2022 (SOEP-733, SOEP-735 and SOEP-740). These results indicate a high nucleotide identity of 99.2% among the sequences of the Indonesian G9P[4]/P[6] strains

(Fig. 1). The genetic background remains consistent among G9P[4] strains found sporadically between 2018 and 2021, when they circulated at a high prevalence.

In the VP7 phylogenetic tree, Indonesian reassortant G9 RVA strains were**clustered** within the major lineage sub-lineage-III, comprising 103 G9 strains retrieved from the GenBank database spanning from 2006 to 2020 **(Fig. 1A)** . While the major sub-lineage III includes the G9P[4] strains identified in India, Denmark, and the USA between 2011 and 2015 as well as G9P[4] strains detected in 2018 and 2020, along with G9P[8] strains detected in Indonesia in 2018 (DSA48, DSA50, and DSA62), the newly identified G9 strains in this study exhibits a closer genetic affinity with G9P[4] strains detected in the Czech Republic in 2018 (CZE/H186/2018 and CZE/H187/2018), Russia in 2018, and Italy from 2019 to 2020. The Indonesian reassortant G9 DS-1-like strains formed a monophyletic lineage with G9P[4] strains detected between 2018 and 2020, sharing approximately 99.6% nucleotide identity within their VP7 genes **(Fig. 1).**

In the VP4 phylogenetic tree, all of the 23 Indonesian G9P[4] strains were clustered within lineage IV, alongside VP4 genes from previously identified G9P[4]-DS-1-like strains **(Fig. 2)** . Similar to the VP7 genes, the VP4 gene of Indonesian G9P[4] strains exhibited the closest relationship to the VP4 genes of G9P[4] DS-1-like strains detected in the Czech Republic in 2018 and in Italy in 2020 (ITA/PA374-20/2020, CZE/H186/2018, and CZE/H187/2018). The nucleotide similarities among them ranged from 99.6% to 99.7%. However, the VP4 gene of the reassortant G9P[6] strain detected in Indonesia in this study exhibited the highest similarity to VP4 genes from equine-like G3P[6] DS-1-like strains identified in Indonesia during 2015-2016 (Fig. 2).

In the NSP2 phylogenetic tree, 14 samples exhibited reassortment in N1-NSP2, with a 99.6% sequence identity shared with CZE/H186/2018 and CZE/H187/2018, and were therefore clustered into lineage-III**(Suppl. Fig. S1)** . This result indicates that Czech G9P[4] strains contain Wa-like segment in the N1-NSP2 region, a trait reminiscent of G9P[4] strains found in Indonesia. To elucidate the origin of N1-NSP2, we collected the N1 sequences available from the GenBank database **(Suppl. Fig. S1)** .

The remaining 8 phylogenetic trees are shown in supplementary Fig. S3-S10 associated with the VP1-VP3, VP6, and NSP1-NSP5 genes. Indonesian reassortant G9P[4]/P[6]-DS-1-like strains exhibited the most significant genetic closeness to the G9P[4] DS-1-like strains detected in the Czech Republic in 2018, along with other contemporary DS-1-like RVA strains. Notably, these Indonesian strains differed slightly from G9P[4]-DS-1-like strains identified during the period from 2011 to 2015. Furthermore, the VP1-VP3, VP6, and NSP1-NSP5 genes in most of the Indonesian reassortant G9 DS-1-like strains showed genetic variations different from those observed in the equine-like G3P[8] strains found in Indonesia during 2015-2016. Notably, three strains detected in 2021-2022 (SOEP-733, SOEP-735, and SOEP-740) in this study showed similar genetic distances in the NSP3 and NSP4 genes to the equine-like G3P[8] strains.

Bayesian evolutionary analysis

To understand the origin and emergence timeline for unusual G9P[4] strains in Indonesia, we performed Bayesian evolutionary analysis on the G9-VP7, P[4]-VP4 and N1-NSP2 genes (Fig.3, Fig. 4, Suppl. Fig. S2). This analysis of the sequence data over time from the GenBank database enabled the estimation of substitution rates and the time of tMRCA for the G9P[4] genotype. Maximum clade credibility (MCC) trees were constructed using the Bayesian MCMC framework. The estimated time of tMRCA for the VP7 MCC tree including G9 was in 2004 with a 95% highest probability density (HPD) interval of 2001-2021.**The estimated evolution rate was 6.43 $\times$ 10^{-3}(4.12-8.74 $\times$ 10^{-3})** (nt substitutions/site/year)**(Fig. 3). The tMRCA for the P[4]-VP4 strains was 1994 (95% HPD interval 1990-2021) and the estimated evolution rate was 2.53 $\times$ 10^{-3} (1.54-3.52 $\times$ 10^{-3}) (Fig. 4). The tMRCA of the P[6]-VP4 strains was 2008 (95% HPD interval 2000-2021), and the estimated evolution rate was 2.11 $\times$ 10^{-3} (1.35-2.86 $\times$ 10^{-3}) (Fig. 4). Within the G9P[4] strains, the NSP2 gene of the reassortant G9-N1 strains diverged from the typical DS-1-like G9-N2 strains in 2003 (95% HPD interval 2000-2021) with the estimated evolution rate of 7.05 $\times$ 10^{-3} (4.22-9.88 $\times$ 10^{-3}) and in 1998 (95% HPD interval**

1996-2021) with the estimated evolution rate of 1.85 × 10^{-3} (1.48-2.22 × 10^{-3}) (Suppl. Fig. S2). Based on the NSP2 MCC tree, the G9-N1 reassortment started in 2011 and was found in several countries such as the Czech Republic in 2018 and in India in 2013. The 11 segments of the unusual G9P[4] Indonesian strain had high nucleotide identities with all the segments from the Czech Republic (VP1-4, VP6, VP7, NSP1-5) ranging from 99.6-99.8% and were clustered into the same lineage.

Amino acid (aa) substitution of VP7 and VP6 Genes

To determine the predominance of the unusual G9P[4] circulating in Indonesia, we performed aa analyses of RVA interaction sites on host cells, focusing on the outer and inner capsid proteins G9-VP7 and I2-VP6. Aa substitutions were deduced from 9 VP7-G9 and VP6-I2 reference strains from the GenBank database (2009-2015). The aa substitutions of T78I and T108I were determined in the VP7 gene (Table 3). Additionally, a single aa substitution of L291S was detected in the VP6 gene (Table 4).

DISCUSSION

Previously, we reported dynamic changes in RVA genotypes from equine-like G3 to human G1/G3[12,25]. Complex RVA genotype diversity is more common in developing countries than in developed countries[21,26,27]. Inter-genogroup reassortant strains highlight the ongoing spread of the unusual RVA strains throughout Asia and other countries[28]. It has been suggested that evolution occurs primarily through the selection of point mutants in the antigenic site or by the reassortant of each segment of the RVA genome[3]. The mutation rate of RVA was estimated to be around the value of 5×10^{-5} per nucleotide during genome synthesis, suggesting that approximately one mutation emerges in each new copy[10].

In this study, we examined the whole genome sequences of the unusual G9P[4] RVA detected in 2021 in Indonesia. This is the first report of the whole genome sequence of the unusual G9P[4] RVA in Indonesia. The newly identified unusual Indonesian G9P[4] strains exhibit genotypes within the G9P[4] constellation distinct from the sporadic G9P[4] strains detected in India, Bangladesh, Japan, and South Korea from 2007 to 2012. The more prevalent G9 genotype in the Americas (18%) and Europe (13%) underwent recombination with the P[8] genotype in 2003 –2004[10]. Human RVA G9P[4] was first identified in Brazil in 1999 as an uncommon genotype[14]. Reassortant in the G9P[4] genotype variation was found in Paraguay with mixed infection of G2G9P[4] from Bangladesh in 2005 and from India in 2008[1,18]. The unusual G9P[4] with a genomic constellation backbone of DS-1-like was reported in Indonesia and was detected in several countries such as the USA, Japan, India, Italy, Mexico, and Bangladesh[21,27-31]. Human RVA reassortant with genomic constellation backbone of Wa-like or DS-1-like were found more frequently[16]. The double-reassortant strain (G9-P[4]-I2-R2-C2-M2-A2-N2/N1-T2-E6-H2) detected in India and in the Czech Republic have the same genotypic constellation as the Indonesian strain[15,32]. In 2013, triple-reassortant strains (G9-P[4]-I2-R2-C2-M2-A2-N1-T2-E6-H2) were found, following the detection of unusual G9P[4] strains in India during 2011-2013[15].

A time-scaled Bayesian phylogenetic tree was constructed to examine the evolution of the RVA genotype[14,32]. The evolutionary rate of VP7-G9 in this study was estimated to be 6.43×10^{-3} nucleotide substitutions/site/year, which was similar to those of the VP7-G9 genotype with 1.38×10^{-3}, $1.87\text{x}10^{-3}$, and $1.609\text{x}10^{-3}$ nucleotide substitutions/site/year in China and Ghana[3,16,33]. The evolutionary rate of the VP4-P[4] in this study was $2.53\text{x}10^{-3}$ nucleotide substitutions/site/year, comparable to the Chinese strain's evolutionary rate of $1.172\text{x}10^{-3}$ nucleotide substitutions/site/year[33]. Conversely, the VP4 gene in Ghana showed a higher evolutionary rate at 8 $\text{x}10^{-4}$nucleotide substitutions/sites/year[21]. These results align with the evolutionary rates typical of many RNA viruses, which evolve at a rate of approximately $1\text{x}10^{-3}$ nucleotide substitutions/site/year[4]. The parent viruses spread regionally within 1-3 years, whereas the ancestral virus takes 2-6 years to propagate[13]. In the present study, the evolutionary rates of both the VP7/VP4 G9P[4] genes of RVA show similar evolutionary speed to those of G9 from China and Ghana. However, the evolution of the P[4] genotype happened more rapidly than previously observed in Ghana. It is essential to monitor the evolutionary rates of RVA genotypes to identify trends in RVA genotype evolution.

The tMRCA analysis predicts the rate of transmission of RVA genotypes[13,16]. The tMRCA of the G9-VP7 of the unusual G9P[4] strains in lineage-III was estimated to be in 2004 (95% HPD interval 2001-2021) (Fig. 3), showing similar results to the unusual G9P[4] RVA strain in India[15]. Monitoring of genotype diversity revealed that rotavirus G9P[4]-VP7 lineage-III was detected in surveillance networks worldwide in the beginning of the millennium[2]. The tMRCA of the G9P[4]-VP4 in this study was estimated to be in 1994 (95% HPD interval 1990-2021) (Fig.4). The prevalent P[4]-VP4 genotype, G2P[4], emerged in 1959 (95% HPD interval 1923-1984)[16,33]. This finding suggests that the unusual G9P[4] strain began spreading with the onset of reassortment in the VP4 gene in the 1990s. The NSP2 gene was generated by reassortant between N1 and N2. The ancestor of NSP2-N1 emerged in 1986 (95% HPD interval 1981-2021), while NSP2-N2 appeared in 1998 (95% HPD interval 1996-2021). NSP2-N1 appeared 10 years earlier than NSP2-N2, indicating that the Wa-like strain appeared earlier than the DS-1-like strain and that both strains are still circulating worldwide.

The outer membrane protein of the VP7 gene contains major antigenic epitopes that induce specific neutralizing antibody responses[21]. In the present study, the deduced aa substitutions in the VP7 gene were T78I and T108I, conserved in G9P[4] RVA in India[15]. Interestingly, the Indian G9P[4] strain presents a mutation E154K in VP7, whereas the Indonesian G9P[4] strain lacks this mutation. The aa sequence in the VP7 has been deduced and reported[2]. One of the strains was linked to the segment within aa 271-342 of VP6, which is necessary for interaction with VP7[34]. In addition, a pair of van der Waals interactions between aa 279-281 on VP7 and aa 313 on VP6, as well as a side chain hydrogen bond between aa 305 on VP7 and aa 310 on VP6 were reported as interaction sites between VP7 and VP6[13]. In viral self-defense, the virus was capable of mutating at their interaction sites of the outer capsid protein VP7 and the inner capsid VP6[15]. The aa substitution S291L on the VP6 protein is linked to the interaction between VP7 and VP6 proteins. In this context, the altered relationship between G9 and I2, with a substitution at the VP6 protein interaction site, may be more sustainable than the unusual relationship without substitutions[15]. In the present study, aa 291 did not receive any substitutions as previously described. This phenomenon is consistent with the notion that the unusual G9P[4] strains circulating in Indonesia can rapidly change into G9P[6] strains.

CONCLUSIONS

The recent discovery of the unusual G9P[4] RVA strains emphasizes the ongoing spread of RVA in Asia and other regions. In particular, the G9P[4]-DS-1-like strain, which had been reported in several countries outside Southeast Asia since 2011, was first detected in Indonesia in 2018. By 2021, the majority of RVA-positive cases (19 out of 21) were identified as G9P[4], with the remaining 3 as G9P[6]. These Indonesian G9P[4]/P[6]-DS-1-like strains, discovered between 2018 and 2022, show genetic variations from strains reported in other countries between 2011 and 2015, particularly in the NSP4 gene (which has E2 genotype instead of E6 genotype). These strains share a common ancestry in VP7 and VP4 genes with previously reported G9P[4] DS-1-like strains but exhibit slight genetic differences in other genes. These results suggest that there were multiple intra- reassortant events between the original G9P[4] strains and the DS-1-like RVA strains that coexist in Indonesia and elsewhere. Continuous monitoring of RVA genotype dynamics is crucial given that G9P[4] strains emerged as the predominant genotype in Indonesia in 2021. This monitoring is essential to assess the prevalence and genetic diversity of circulating RVA strains and to evaluate the efficacy of RVA vaccines.

AUTHOR CONTRIBUTION

Zayyin Dinana, Yen Hai Doan, Takako Utsumi, Maria Inge Lusida, and Ikuo Shoji: designed the study and wrote the manuscript. Laura Navika Yamani, Soetjipto, Juniastuti, Rury Mega Wahyuni, Soegeng Soegijanto : collected samples. Zayyin Dinana, Aussie Tahta Maharani, Anisa Lilatul Fitria: carried out RVA detection and genome analysis. Zayyin Dinana, Aussie Tahta Maharani, Yen Hai Doan: performed NGS sequencing. Zayyin Dinana, Aussie Tahta Maharani, Yen Hai Doan, Ikuo Shoji: conducted sequence data analysis. Chieko Matsui, Lin Deng, Nobuhiko Takemae, Tsutomu Kageyama, Kazuhiko Katayama, Maria Inge Lusida, Ikuo Shoji: gave critical revision of the article. All the authors contributed to the interpretation of the data, writing of the manuscript, and approved the final manuscript.

DATA AVAILABILITY STATEMENT

The datasets presented in this study will be available at Kernel-Kobe University Repository (https://da.lib.kobe-u.ac.jp/da/kernel/?lang=1). The sequence results were deposited in the DDBJ/GenBank database with accession numbers PP948908 - PP949193.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Airlangga University, Indonesia Kobe University Graduate School of Medicine, Japan. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

FUNDING

This research was supported by the Japan Initiative for Global Research Network on Infectious Diseases (J-GRID) from Ministry of Education, Culture, Sport, Science & Technology in Japan, Japan Agency for Medical Research and Development (AMED: 23wm0125009; 23fk0108667). This research was also supported by the Japan Society for the Promotion of Science (JSPS) (KAKENHI: 19H04354).

CONFLICT OF INTEREST

The authors have no conflict of interest to declare regarding this study.

ACKNOWLEDGEMENTS

We sincerely thank the Ministry of Research, Technology and Higher Education of Indonesia for their advice to carry out this study in Indonesia. We sincerely thank the Ministry of Research, Technology/National Research and Innovation Agency of Republic Indonesia for their advice to carry out this study in Indonesia.

ORCHID

Zayyin Dinana http://orcid.org/0000-0003-1189-4153

Yen Hai Doan http://orcid.org/0000-0002-6782-3541

Aussie Tahta Maharani http://orcid.org/0000-0001-6880-5292

Anisa Lailatul Fitria http://orcid.org/0000-002-9620-4002

Laura Navika Yamani http://orcid.org/0000-0002-2900-908X

Juniastuti http://orcid.org/0000-0002-1437-640X

Chieko Matsui http://orcid.org/0000-0003-3428-5334

Lin Deng http://orcid.org/0000-0002-5993-6597

Nobuhiro Takemae https://orcid.org/0000-0001-5152-8938

Tsutomu Kageyama

Soegeng Soegijanto https://orcod.org/0000-0002-9470-8223

Rury Mega Wahyuni https://orcid.org/0000-0003-2101-1270

Soetjipto https://orcid.org/0000-0003-2203-5565

Takako Utsumi https://orcid.org/0000-0001-6539-5064

Kazuhiko Katayama http://orcid.org/0000-0002-7692-1151

Maria Inge Lusida https://orcid.org/0000-0001-5039-0674

Ikuo Shoji http://orcid.org/0000-0002-0730-4379

Figure legends

Figure 1. Phylogenetic tree analysis of VP7 gene sequences. A total of 24 G9P[4] strains and 2 G9P[6] strains were detected in this study (indicated by black bold font). Additionally, 3 G9P[8] strains were detected in 2018 (indicated by blue bold font). The VP7 genes of 94 G8 RVAs were retrieved from the GenBank database. The genetic distance is indicated at the bottom, and the percent bootstrap support is displayed at each node when the value was 70% or larger.

Figure 2. Phylogenetic tree analysis of VP4 gene sequences. A total of 24 G9P[4] strains and 2 G9P[6] strains were detected in this study (indicated by black bold font). Additionally, 3 G9P[8] strains were detected in 2018 (indicated by blue bold font). The VP7 genes of 94 G8 RVAs were retrieved from the GenBank database. The genetic distance is indicated at the bottom, and the percent bootstrap support is indicated by the value at each node when the value was 70% or larger.

Figure 3. Simplified maximum clade credibility (MCC) drawn from representative sequences of VP7 genes. The MCC tree was constructed using a Bayesian MCMC analysis framework with a strict clock model. Reference sequences were obtained from the GenBank database (www.ncbi.nlm.nih.gov). The years of divergence of each lineage are indicated at each node. Major lineage have been collapsed for simplicity.

Figure 4. Simplified maximum clade credibility (MCC) drawn from representative sequences of VP4 genes. The MCC tree was constructed using a Bayesian MCMC analysis framework with a strict clock model. Reference sequences were obtained from the GenBank database (www.ncbi.nlm.nih.gov). The years of divergence of each lineage are indicated at each node. Major lineage have been collapsed for simplicity.

REFERENCES

1. Babji S, Arumugam R, Sarvanabhavan A, et al. Multi-center surveillance of rotavirus diarrhea in hospitalized children <5 years of age in India, 2009–2012. *Vaccine.* 2014;32:A10-A12.2. Matthijnssens J, Ciarlet M, Heiman E, et al. Full genome-based classification of rotaviruses reveals a common origin between human Wa-Like and porcine rotavirus strains and human DS-1-like and bovine rotavirus strains. *Journal of virology.* 2008;82(7):3204-3219.3. Matthijnssens J, Ciarlet M, McDonald SM, et al. Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). *Archives of virology.* 2011;156(8):1397-1413.4. Matthijnssens J, Heylen E, Zeller M, Rahman M, Lemey P, Van Ranst M. Phylodynamic Analyses of Rotavirus Genotypes G9 and G12 Underscore Their Potential for Swift Global Spread. *Molecular Biology and Evolution.* 2010;27(10):2431-2436.5. Troeger C, Khalil IA, Rao PC, et al. Rotavirus Vaccination and the Global Burden of Rotavirus Diarrhea Among Children Younger Than 5 Years. *JAMA Pediatr.* 2018;172(10):958-965.6. Lestari FB, Vongpunsawad S, Wanlapakorn N, Poovorawan Y. Rotavirus infection in children in Southeast Asia 2008-2018: disease burden, genotype distribution, seasonality, and vaccination. *J Biomed Sci.* 2020;27(1):66-66.7. Wahyuni RM, Utsumi T, Dinana Z, et al. Prevalence and Distribution of Rotavirus Genotypes Among Children With Acute Gastroenteritis in Areas Other Than Java Island, Indonesia, 2016-2018. *Front Microbiol.* 2021;12:672837-672837.8. Folorunso OS, Sebolai OM. Overview of the Development, Impacts, and Challenges of Live-Attenuated Oral Rotavirus Vaccines. *Vaccines (Basel).* 2020;8(3):341.9. Kirkwood CD, Boniface K, Barnes GL, Bishop RF. Distribution of Rotavirus Genotypes After Introduction of Rotavirus Vaccines, Rotarix(r) and RotaTeq(r), into the National Immunization Program of Australia. *Pediatric Infectious Disease Journal.* 2011;30(1):S48-S53.10. Banyai K, Pitzer VE. Molecular Epidemiology and Evolution of Rotaviruses. In. *Viral Gastroenteritis* : Elsevier; 2016:279-299.11. Matthijnssens J, Ciarlet M, Rahman M, et al. Recommendations for the classification of group A rotaviruses using all 11 genomic RNA segments. *Archives of virology.* 2008;153(8):1621-1629.12. Athiyyah AF, Utsumi T, Wahyuni RM, et al. Molecular Epidemiology and Clinical Features of Rotavirus Infection Among Pediatric Patients in East Java, Indonesia During 2015-2018: Dynamic Changes in Rotavirus Genotypes From Equine-Like G3 to Typical Human G1/G3. *Front Microbiol.* 2019;10:940-940.13. Fujii Y, Doan YH, Suzuki Y, Nakagomi T, Nakagomi O, Katayama K. Study of Complete Genome Sequences of Rotavirus A Epidemics and Evolution in Japan in 2012-2014. *Front Microbiol.* 2019;10:38-38.14. Santos N, Volotao EM, Soares CC, et al. Rotavirus strains bearing genotype G9 or P[9] recovered from Brazilian chil-

dren with diarrhea from 1997 to 1999. *Journal of clinical microbiology.* 2001;39(3):1157-1160.15. Doan YH, Suzuki Y, Fujii Y, et al. Complex reassortment events of unusual G9P[4] rotavirus strains in India between 2011 and 2013.*Infection, Genetics and Evolution.* 2017;54:417-428.16. Liu X, Wang M, Li S, et al. Genomic and evolutionary characteristics of G9P[8], the dominant group a rotavirus in China (2016-2018).*Front Microbiol.* 2022;13:997957-997957.17. Mangayarkarasi V, Prema A, Gunasekaran P, Babu BVS, Kaveri K. A unique human rotavirus (non vaccine) G9P4 genotype infection in a child with gastroenteritis.*Indian Pediatrics.* 2012;49(7):569-571.18. Afrad MH, Rahman MZ, Matthijnssens J, et al. High incidence of reassortant G9P[4] rotavirus strain in Bangladesh: Fully heterotypic from vaccine strains.*Journal of Clinical Virology.* 2013;58(4):755-756.19. Felix-Valenzuela L, Cooley-Garcia DP, Cano-Rangel MA, Durazo-Arvizu MdlA, Mata-Haro V. Predominance of G9P[4] Rotavirus from Children with Acute Gastroenteritis in Northwestern Mexico. *Intervirology.* 2016;59(4):228-233.20. Pasittungkul S, Lestari FB, Puenpa J, et al. High prevalence of circulating DS-1-like human rotavirus A and genotype diversity in children with acute gastroenteritis in Thailand from 2016 to 2019. *PeerJ.* 2021;9:e10954-e10954.21. Damanka SA, Agbemabiese CA, Dennis FE, et al. Genetic analysis of Ghanaian G1P[8] and G9P[8] rotavirus A strains reveals the impact of P[8] VP4 gene polymorphism on P-genotyping. *PLoS One.* 2019;14(6):e0218790-e0218790.22. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol.* 2021;38(7):3022-3027.23. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30(4):772-780.24. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol.* 2012;29(8):1969-1973.25. Utsumi T, Wahyuni RM, Doan YH, et al. Equine-like G3 rotavirus strains as predominant strains among children in Indonesia in 2015–2016.*Infection, Genetics and Evolution.* 2018;61:224-228.26. Chitambar SD, Ranshing SS, Pradhan GN, Kalrao VR, Dhongde RK, Bavdekar AR. Changing trends in circulating rotavirus strains in Pune, western India in 2009–2012: Emergence of a rare G9P[4] rotavirus strain. *Vaccine.* 2014;32:A29-A32.27. Rasebotsa S, Uwimana J, Mogotsi MT, et al. Whole-Genome Analyses Identifies Multiple Reassortant Rotavirus Strains in Rwanda Post-Vaccine Introduction. *Viruses.* 2021;13(1):95.28. Lewis J, Roy S, Esona MD, et al. Full Genome Sequence of a Reassortant Human G9P[4] Rotavirus Strain. *Genome Announc.* 2014;2(6):e01284-01214.29. Ianiro G, Recanatini C, D'Errico MM, Monini M. Uncommon G9P[4] group A rotavirus strains causing dehydrating diarrhea in young children in Italy. *Infection, Genetics and Evolution.* 2018;64:57-64.30. Phan T, Hatazawa R, Komoto S, et al. Whole genome sequence of an uncommon G9P[4] species A rotavirus containing DS-1-like (genotype 2) genes in Japan.*Archives of Virology.* 2022;167(7):1603-1606.31. Yamamoto SP, Kaida A, Ono A, Kubo H, Iritani N. Detection and characterization of a human G9P[4] rotavirus strain in Japan. *Journal of Medical Virology.* 2015;87(8):1311-1318.32. Moutelikova R, Sauer P, Prodělalová J. Whole-genome sequence of a reassortant G9P[4] rotavirus A strain from two children in the Czech Republic. *Archives of Virology.* 2020;165(7):1703-1706.33. Zhang T, Li J, Jiang Y-Z, et al. Genotype distribution and evolutionary analysis of rotavirus associated with acute diarrhea outpatients in Hubei, China, 2013-2016. *Virol Sin.* 2022;37(4):503-512.34. Gilbert JM, Feng N, Patton JT, Greenberg HB. Rotavirus assembly - interaction of surface protein VP7 with middle layer protein VP6. *Archives of Virology.* 2001;146(6):1155-1171.

Hosted file

`Unusual G9P[4]_Figures_IS.pptx` available at https://authorea.com/users/824357/articles/1220693-unusual-g9p-4-rotavirus-emerged-after-the-dynamic-changes-in-rotavirus-genotypes-from-equine-like-g3-to-typical-human-g1-g3-in-indonesia

Hosted file

`Unusual G9P[4]_Table_IS.pptx` available at https://authorea.com/users/824357/articles/1220693-unusual-g9p-4-rotavirus-emerged-after-the-dynamic-changes-in-rotavirus-genotypes-from-equine-like-g3-to-typical-human-g1-g3-in-indonesia