Virus Evolution, 2024, **10(1)**, veae045
DOI: https://doi.org/10.1093/ve/veae045
Advance Access Publication 22 June 2024
Research Article

Evolution of DS-1-like G8P[8] rotavirus A strains from Vietnamese children with acute gastroenteritis (2014–21): Adaptation and loss of animal rotavirus-derived *genes* during human-to-human spread

Thi Nguyen Hoa-Tran,[1,†,*] Toyoko Nakagomi,[2] Hung Manh Vu,[1] Trang Thu Thi Nguyen,[1] Anh Thi Hai Dao,[1] Anh The Nguyen,[1] Julie E. Bines,[3] Sarah Thomas,[3] Varja Grabovac,[4] Chikako Kataoka-Nakamura,[5] Takemura Taichiro,[6,7] Futoshi Hasebe,[6,7] Toshio Kodama,[8] Miho Kaneko,[9] Huyen Thi Thanh Dang,[10] Hong Thi Duong,[10] Dang Duc Anh,[11] and Osamu Nakagomi[2]

[1]Department of Virology, National Institute of Hygiene and Epidemiology, Hanoi 100000, Vietnam, [2]Department of Hygiene and Molecular Epidemiology, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki 852-8523, Japan, [3]Enteric Diseases Group, Murdoch Childrens Research Institute, Royal Children's Hospital, Parkville, Victoria 3052, Australia, [4]Vaccine-Preventable Diseases and Immunization Unit, Division of Programmes for Disease Control, World Health Organization Regional Office for the Western Pacific, Manila 1000, Philippines, [5]Center Surveillance Division, The Research Foundation for Microbial Diseases of Osaka University, Osaka 768-0065, Japan, [6]Vietnam Research Station, Institute of Tropical Medicine, Nagasaki University, Nagasaki 852-8523, Japan, [7]Vietnam Research Station, National Institute of Hygiene and Epidemiology-Nagasaki University, Hanoi 100000, Vietnam, [8]Department of Bacteriology, Institute of Tropical Medicine, Nagasaki University, Nagasaki 852-8523, Japan, [9]Department of Molecular Microbiology and Immunology, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki 852-8523, Japan, [10]National office for Expanded Program on Immunization, National Institute of Hygiene and Epidemiology, Hanoi 100000, Vietnam and [11]Department of Bacteriology, National Institute of Hygiene and Epidemiology, Hanoi 100000, Vietnam
[†]https://orcid.org/0000-0001-6887-7317
*Corresponding author: E-mail: ttnh@nihe.org.vn; trnghoasuk@gmail.com

Abstract

Animal rotaviruses A (RVAs) are considered the source of emerging, novel RVA strains that have the potential to cause global spread in humans. A case in point was the emergence of G8 bovine RVA consisting of the P[8] VP4 *gene* and the DS-1-like backbone *genes* that appeared to have jumped into humans recently. However, it was not well documented what evolutionary changes occurred on the animal RVA-derived *genes* during circulation in humans. Rotavirus surveillance in Vietnam found that DS-1-like G8P[8] strains emerged in 2014, circulated in two prevalent waves, and disappeared in 2021. This surveillance provided us with a unique opportunity to investigate the whole process of evolutionary changes, which occurred in an animal RVA that had jumped the host species barrier. Of the 843 G8P[8] samples collected from children with acute diarrhoea in Vietnam between 2014 and 2021, fifty-eight strains were selected based on their distinctive electropherotypes of the genomic RNA identified using polyacrylamide gel electrophoresis. Whole-genome sequence analysis of those fifty-eight strains showed that the strains dominant during the first wave of prevalence (2014–17) carried animal RVA-derived VP1, NSP2, and NSP4 *genes*. However, the strains from the second wave of prevalence (2018–21) lost these *genes*, which were replaced with cognate human RVA-derived *genes*, thus creating strain with G8P[8] on a fully DS-1-like human RVA *gene* backbone. The G8 VP7 and P[8] VP4 *genes* underwent some point mutations but the phylogenetic lineages to which they belonged remained unchanged. We, therefore, propose a hypothesis regarding the tendency for the animal RVA-derived *genes* to be expelled from the backbone *genes* of the progeny strains after crossing the host species barrier. This study underlines the importance of long-term surveillance of circulating wild-type strains in order to better understand the adaptation process and the fate of newly emerging, animal-derived RVA among the human population. Further studies are warranted to disclose the molecular mechanisms by which spillover animal RVAs become readily transmissible among humans, and the roles played by the expulsion of animal-derived *genes* and herd immunity formed in the local population.

Keywords: rotavirus; DS-1-like; G8P[8]; evolution; interspecies transmission; reassortment.

1. Introduction

Rotavirus A (RVA), a species within genus *Rotavirus*, family *Sedoreoviridae* (Matthijnssens et al. 2022), is the major cause of acute diarrhoea among children as well as the young of many mammalian and avian species (Crawford et al. 2017). The RVA genome is made up of eleven *gene* segments of double-stranded RNA (dsRNA), each containing a single *gene*, except for segment 11 that possesses two *genes*. These segments collectively encode six structural proteins (VP1–VP4, VP6, and VP7) and six non-structural proteins (NSP1–NSP6) (Crawford et al. 2017). The outer

capsid proteins, VP7 and VP4, contain independent neutralisation epitopes, on which the binary classification of G and P genotypes was established. The vast majority of RVA infections in humans are caused by G1P[8], G2P[4], G3P[8], G4P[8], G9P[8], and G12P[8] (Crawford et al. 2017). This binary classification system was subsequently extended to cover the entire genome segments. Thus, the constellation of the VP7-VP4-VP6-VP1-VP2-VP3-NSP1-NSP2-NSP3-NSP4-NSP5 *genes* is described as G*x*-P[*x*]-I*x*-R*x*-C*x*-M*x*-A*x*-N*x*-T*x*-E*x*-H*x*, where *x* denotes a genotype number (Matthijnssens et al. 2008). Most human RVA strains are classified into three genotype constellations; the Wa-like, DS-1-like, and AU-1-like genotype constellations, which are described as G1/G3/G4/G9-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1, G2P[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2, and G3P[9]-I3-R3-C3-M3-A3-N3-T3-E3-H3, respectively (Heiman et al. 2008).

The diversity of the human rotavirus genome is generated by several mechanisms, including accumulation of point mutations, reassortment, direct transmission of animal RVA, recombination, and *gene* rearrangement (Kirkwood 2010). Of those, accumulation of point mutations and reassortment between two human strains or human and animal strains are major mechanisms (McDonald et al. 2016). Interspecies transmission, in which an animal RVA as whole virions sporadically crosses the species barrier and spread in other host populations, also contributes to the evolution of rotavirus genome (Martella et al. 2010). Many of the spillover events, however, are dead end, but viruses may sometimes adapt to gain human-to-human transmissibility (Webby et al. 2004). A good example of a spillover event of an animal RVA was recovery of a G6P[1] bovine rotavirus strain from an Israeli child, who developed diarrhoea within a few days after his first contact with cows on a farm (Doan et al. 2013). However, there appeared to be no onward transmission of this bovine RVA because sentinel surveillance detected no other RVA of suspected bovine origin (Doan et al. 2013).

The introduction of rotavirus vaccines in the national immunisation programmes in over 100 countries in the world since 2006 has made a marked impact on rotavirus disease burden with a 59 per cent (interquartile range (IQR): 46, 74) reduction of rotavirus hospitalisations, a 36 per cent (IQR: 23, 47) reduction of acute gastroenteritis hospitalisations, and a 36 per cent (IQR: 28, 46) reduction in the mortality attributable to acute diarrhoea among children less than 5 years of age (Burnett et al. 2020). Nevertheless, RVA still accounts for more than one-third of diarrhoeal deaths in children less than 5 years of age living mostly in low-income countries (Tate et al. 2016). It was also hypothesised that the introduction of rotavirus vaccine would increase the immune pressure against wild-type RVA strains, thereby resulting in the selection and spread of novel reassortant strains. Thus, the surveillance of RVA strains circulating in the paediatric population in low-income countries has been coordinated by the World Health Organisation (WHO) to monitor vaccine effectiveness and to detect novel RVA variants that may escape vaccine-acquired immune protection.

In Vietnam, systematic rotavirus surveillance funded by the WHO started in 2012 (Huyen et al. 2018). As part of this surveillance network, we detected a few samples that typed as G8P[8] in the last quarter of 2014, and G8P[8] increased to 26.5 per cent of the genotyped samples during the first half of 2015 (Hoa-Tran et al. 2016). As G8 strains, particularly their outbreaks, were rarely observed in areas other than the African continent, where an average of 12 per cent of RVA was reported to be G8 (Todd et al. 2010), we analysed the G8P[8] samples by genomic RNA electrophoresis, and showed that they had almost identical short RNA patterns, which are suggestive of a clonal origin (Hoa-Tran et al. 2016). We then carried out a whole-genome analysis on the representative strain and showed that it had a bovine-like G8 VP7 *gene*, whereas its remaining *genes* except for the VP1, NSP2, and NSP4 *genes* were more than 98 per cent homologous to the contemporary human RVAs (Hoa-Tran et al., 2016). In further strain surveillance, we continued to observe an increase in the prevalence of G8P[8]; however, during the latter half of 2021 we identified no G8P[8] strains.

Therefore, the availability of 8 years (i.e. between 2014 and 2021) of samples spanning from the time of emergence, expansion, and finally to the disappearance of G8P[8] strains in Vietnam provided us with a unique opportunity to investigate the evolutionary changes that a newly emergent animal RVA-derived strain underwent during circulation in humans. These observations were not comprehensively described in the existing literature. Of particular interest is the fate of animal RVA-derived *genes* that may have been introduced into the genome of the recipient human RVA strain, when an animal G8 RVA strain jumped across the species barrier and underwent reassortment with a concurrently infected human RVA strain. Thus, the aim of this study was to describe the evolutionary processes that emergent G8P[8] strains went through while they circulated among children in Vietnam over 8 years.

2. Materials and methods

2.1 Stool specimens under the WHO surveillance programme

A total of 14,840 stool specimens were collected from children aged less than 5 years seeking treatment for acute diarrhoea in sentinel hospitals in the Central and Northern regions of Vietnam between 2012 and 2021. RVAs were detected by either ProSpecT Rotavirus Microplate Assay (Thermo Scientific™) or RIDASCREEN® Rotavirus (R-Biopharm AG) Enzyme-linked Immunosorbent Assay (ELISA) as per WHO recommendation. Rotavirus-positive samples were randomly selected but with consideration to reflect the time period of sample collection (the minimum time period was one week) for G and P genotyping according to the WHO and the Centers for Disease Control and Prevention (CDC) protocols (Simmonds et al. 2008; World Health Organization 2009; Esona et al. 2015; Fig. 1).

2.2 Polyacrylamide gel electrophoresis for assignment of electropherotypes

Samples genotyped as G8P[8] were subjected to polyacrylamide gel electrophoresis (PAGE) to assign electropherotypes (Fig. 1) according to the method described by Gauchan et al. (2013) (Supplementary Fig. S1). The selection of samples for PAGE was done by taking into consideration whether there was a sufficient amount of viral RNA present in the sample to repeat PAGE at least three times. The viral RNA was extracted from 20 per cent stool suspension (w/v) using a QIAamp Viral RNA Mini Kit (QIAGEN Sciences, Germantown, MD, USA). The samples (20 μl of extracted RNA/sample) were first run on polyacrylamide gels to screen RNA migration profiles (screening gels). Samples on a single gel that showed similar RNA migration profile were grouped together under a single label of RNA pattern (Supplementary Fig S1.A–B). Samples representative of each RNA pattern from different gels were taken and run on the same gel for comparison (grouping gels). Further running on grouping gels of the samples that were grouped under the same RNA pattern were repeated until it was possible to give them a distinct electropherotype. Thus, an electropherotype in this study is defined as a unique RNA migration profile that distinguishes one or multiple samples from all other co-circulating strains (Supplementary Fig. S1.C–D).

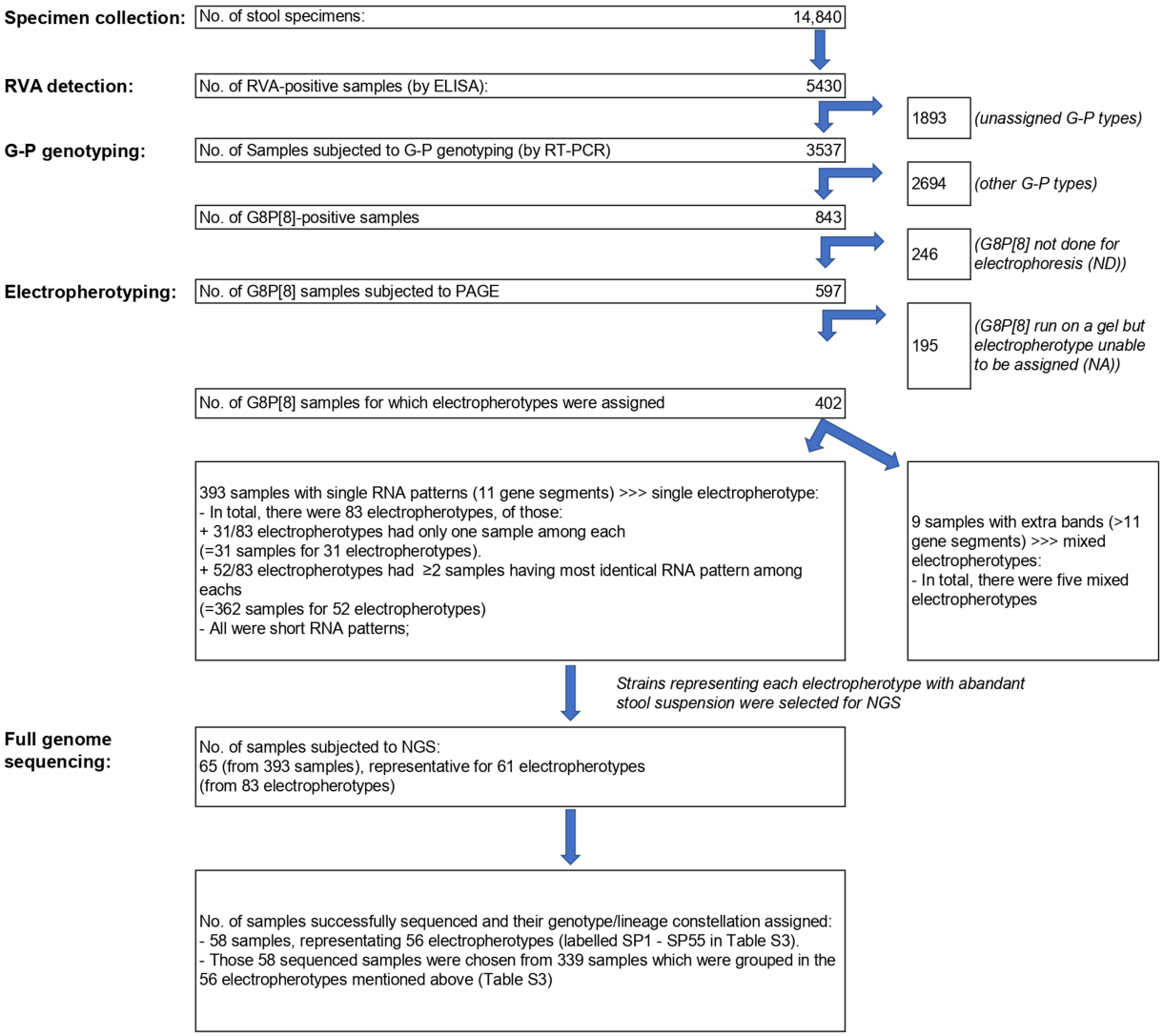

Figure 1. The workflow of the experiments to determine electropherotypes and genotype/lineage constellations of G8P[8] strains.

2.3 Full-genome sequencing using Illumina MiSeq platform

Samples representing each electropherotype were subjected to whole-genome sequencing using previously described methods (Hoa-Tran et al. 2019; Fig. 1). The selection was made by considering the representativeness of electropherotypes and the remaining amount of sample volumes. In addition, to ensure the success in sequencing, the selected sample was chosen because the migration profile of its eleven *gene* segments was clearly visible, when a maximum of 5 μl of extracted RNA of the samples was loaded on the polyacrylamide gel.

Complementary DNA (cDNA) library was constructed from RNA samples pretreated with DNase I (TURBO DNA-free™ Kit/Invitrogen) and normalised to 100 ng; subsequent steps included 200-bp RNA fragmentation, priming, double-stranded cDNA synthesis, end repair_dA-tailing, adaptor ligation, and polymerase chain reaction (PCR) enrichment using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® or the NEBNext® Ultra™ II RNA Library Prep Kit for Illumina® (New England Biolabs, Ipswich, MA, USA), and the NEBNext® Multiplex Oligos for Illumina® (New England Biolabs). After enriched PCR products were purified using the Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA), the quality of the purified cDNA libraries was measured using Qubit™ dsDNA Quantification Assay Kits (Thermo Fisher Scientific) and normalised according to the method of standard normalisation of Illumina. Nucleotide sequencing was performed on an Illumina MiSeq sequencer with a MiSeq Reagent Micro Kit v2 (300 cycles) (Illumina) to generate 150-cycle paired-end reads.

Data analysis was carried out using CLC Genomics Workbench v7.0.3 (CLC Bio, Tokyo, Japan). Contigs were initially assembled from the obtained sequence reads by de novo assembly, and then by reference mapping, using Basic Local Alignment Search Tool (BLAST) against local data in CLC Genomics Workbench with the assembled contigs as query sequences. The recommended average coverage level was ≥100x. For those contigs that were assembled with low-depth read and coverage (30x–50x), Sanger sequencing was used to confirm the sequences using the 5′- and 3′-end consensus primer sets (Doan et al. 2015).

2.4 Assignment of genotype/lineage constellations

To assign the genotype to a *gene* sequence, we used the web-based Rotavirus A Genotyping Tool (version 0.1;

https://www.rivm.nl/mpf/typingtool/rotavirusa/). This automated genotyping tool has an algorithm, developed from RotaC by Maes et al. (2009), in which the submitted sequence is compared with the reference sequences representing all available genotypes according to the new guidelines proposed by Rotavirus Classification Working Group (https://rega.kuleuven.be/cev/viralmetagenomics/virus-classification/rcwg).

To assign a lineage within the genotype for the DS-1-like backbone *genes*, we used the method by Agbemabiese et al. (2019) and followed their standardised numbering system. To assign a lineage within the G8 VP7 and the P[8] VP4 *genes*, we carried out phylogenetic analysis using the coding sequences (CDSs) of the VP7 and VP4 *genes* using the MEGA7 v.7.0.26 (Kumar, Stecher, and Tamura et al. 2016), and followed the numbering systems used by Agbemabiese et al. (2015) and Do et al. (2016) for the VP7 and VP4 *genes*, respectively, which were the same as used widely by rotavirus researchers.

Multiple sequence alignments were performed with ClustalW algorithm and the CDS of each *gene* segments were identified and extracted for further analyses. When the pairwise distance (the p-distance) was calculated, we used the pairwise deletion option. The nucleotide identities were calculated using the formula: % nucleotide identity = 100 × (1 – p-distance). For drawing Maximum likelihood (ML) phylogenetic trees, the nucleotide substitution model testing was carried out and the best-fit evolutionary model for each *gene* was selected based on the lowest Bayesian Information Criterion score. Specifically, the ML trees of VP7 and VP6 were based on the T92 + G + I model; that of VP4 was based on T93 + G + I; those of VP1–VP3 were based on GTR + G + I model; and those of NSP1–NSP5/6 were based on T92 + G model. The trees were analysed by bootstrapping with 1,000 replicates. ML trees were constructed by including representative human and animal RVA strains from different genetic lineages of *gene* segment. The ML tree of VP4 *gene* included RVA/Human-tc/USA/DS-1/1976/G2P[4] and RVA/Human-tc/JPN/AU-1/1982/G3P[9], while those of the other *genes* included RVA/Human-tc/USA/Wa/1974/G1P[8] and RVA/Human-tc/JPN/AU-1/1982/G3P[9] as outgroups to root phylogenetic trees.

An ML tree was constructed for the full-genome sequences of the fifty-three Vietnamese G8P[8] strains together with the twenty-eight DS-1-like strains detected in Thailand, Japan, USA, and China between 2013 and 2022. A full-genome sequence has been made by merging the CDS of the eleven genome segments in the order of VP1, VP2, VP3, VP4, NSP1, VP6, VP7, NSP2, NSP3, NSP5/6, and NSP4. A total of 17,427 nucleotides were found in the final data set of full-length nucleotide sequence when it was merged from eleven complete CDS of eleven *gene* segments (Supplementary Table S2). The evolutionary model used in this tree was based on the T92 + G + I model. Three outgroup strains were used: i.e. RVA/Human-tc/USA/DS-1/1976/G2P[4] (DS-1-like genotype constellation), RVA/Human-tc/USA/Wa/1974/G1P[8] (Wa-1-like genotype constellation), and RVA/Human-wt/JPN/AU-1/1982/G3P[9] (AU-1-like genotype constellation).

Bayesian Evolutionary Analysis by Sampling Trees (BEAST) was conducted using BEAST v1.10.4 (Drummond et al. 2012) according to the guidance described by Degiuseppe et al. (2022). We ran two iterations, each consisting of 30 million generations, and combined the data with LogCombiner v1.10.4 (Bouckaert et al. 2014), annotated with TreeAnnotator v1.10.4 (Helfrich et al. 2018), and drew the Maximum Clade Credibility (MCC) tree with FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

2.5 Estimation of recombination events and positive-selection sites

Putative recombination events were evaluated on separately aligned sequences of all eleven *gene* segments using the RDP4.101 BetaBig (Martin et al. 2015) with a *P*-value cut-off of <0.001 and algorithms of 3Seq, Chimæra, SiScan, MaxChi, Bootscan, Geneconv, and RDP. A recombination event was predicted when the evaluation was shown at least with 3Seq, Chimæra, and MaxChi as the best methods for pinpointing the precise location of recombination break points. BLAST search was used to identify the possible major and minor parents of the putative recombinants.

The alignment schemes of the amino acid sequences of VP7 and VP4 from Vietnamese G8P[8] strains were exported from Sequence Data Explorer into Excel format to locate variable sites, including substitutions in putative neutralisation epitopes (Chan-It, Chanta, and Ushijima et al. 2023).

The non-synonymous (dN) to synonymous (dS) substitutions ratio (dS/dN ratio) was estimated using the Estimate Selection at Codons (via HyPhy) with the statistical method of ML (Sergei et al. 2020) implemented in the MEGA7 package. Additionally, the online Datamonkey server, a web server for the HyPhy package, was also used to determine the site under positive selection (dN > dS) or negative selection (dN < dS) using the Single-Likelihood Ancestor Counting and fixed effects likelihood methods with the *P*-value threshold of 0.05. Mixed Effects Model of Evolution was then used to determine the site under positive selection (Kosakovsky and Frost et al. 2005; Weaver et al. 2018).

2.6 Ethics statement

The study was approved by the institute's ethics committee, with numbers of IRB - VN01057 – 24/2015, NIHE IRB – 41/2021, NIHE IRB – 08/2021, and NIHE IRB – 35/2022.

3. Results

3.1 The biphasic circulation of G8P[8] between 2014 and 2021 in Vietnam

Out of 14,840 stool specimens collected between 2012 and 2021, 5,430 (36.6 per cent) samples were tested positive for RVA. A total of 3,537 RVA-positive samples, which correspond to 65.1 per cent of the RVA-positives, were subjected to G and P genotyping (Fig. 1). The typing identified 843 (23.8 per cent) samples as G8P[8], which was the second most frequently detected genotype after G1P[8] (27.3 per cent) (Supplementary Fig. S2). However, its relative frequency varied considerably from year to year (Fig. 2A). The detection rate of G8P[8] was 0 per cent for the first 2 years of the surveillance (Fig. 2A). The first detection of G8P[8] was in the last quarter of 2014 when three G8P[8] specimens were detected, which accounted for just 1 per cent of the specimens genotyped throughout 2014. The variation in detection rate thereafter was biphasic; the first peak of prevalence occurred in 2016 when eighty-one G8P[8] specimens were detected, which accounted for 41 per cent of the genotyped samples in 2016. The second peak, which looked more like a plateau, occurred in 2019 and 2020 when there were 113 and 67 samples, respectively, and they accounted for 63–65 per cent of the genotyped samples in these years. Subsequently, detection dropped to seventeen samples in 2021, accounting for 19 per cent of the genotyped samples (Fig. 2A). However, no G8P[8] samples were detected in the latter half of 2021, suggesting that G8P[8] likely disappeared from circulation. To summarise, the circulation of G8P[8] strains in Vietnam was divided into two waves of prevalence, with the first between 2014 and 2017, and the second between 2018 and 2021 (Fig. 2A).

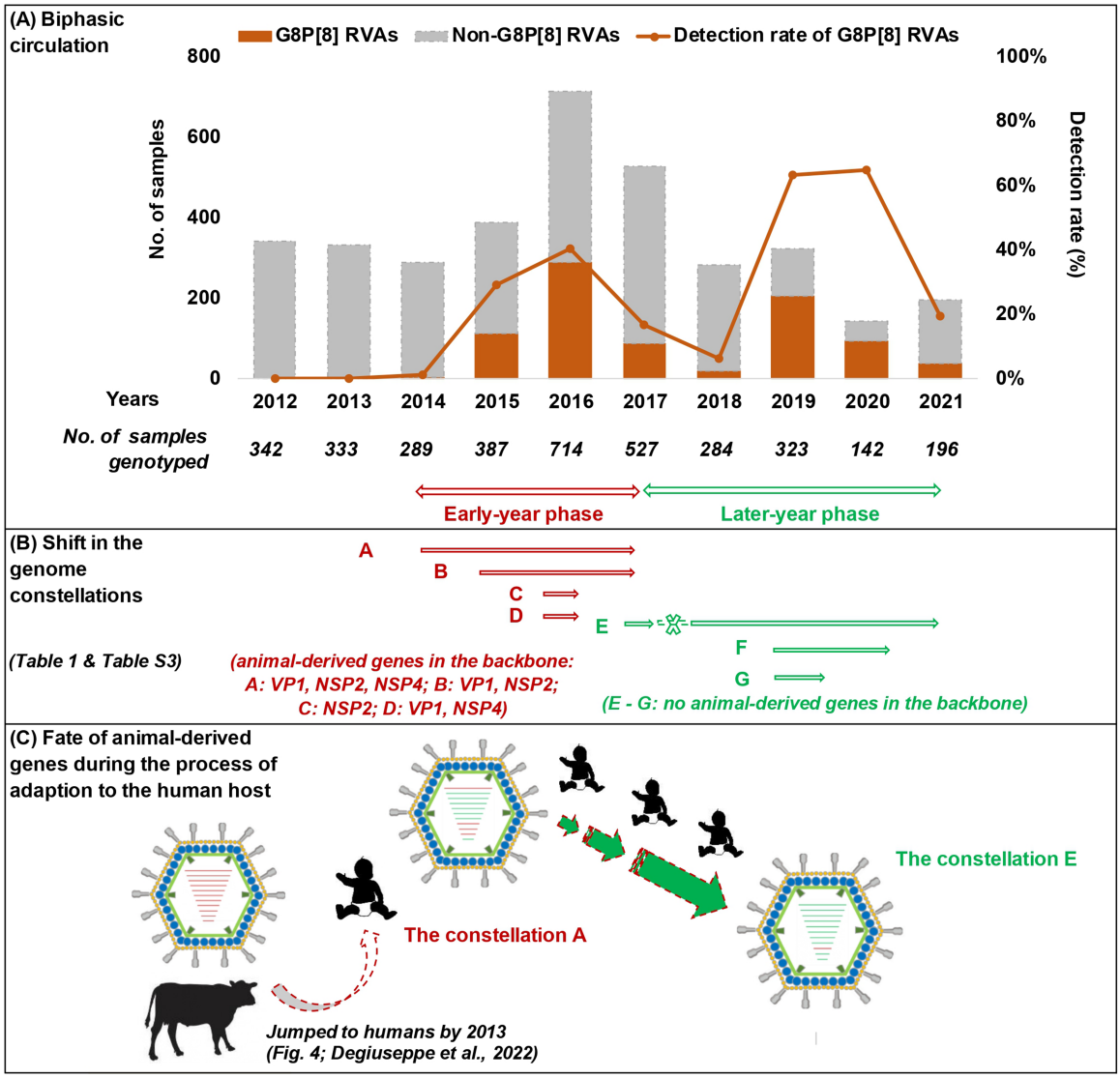

Figure 2. Molecular epidemiology and evolutionary changes of DS-1-like G8P[8] rotavirus A strains, 2014–21: A hypothesis on the fate of the animal RVA-derived *genes* during the process of adaptation to the human host. (A) The circulation was divided into two waves of prevalence, with the first between 2014 and 2017, and the second between 2018 and 2021. The detection rate of G8P[8] is defined as the number of positive samples with G8P[8] divided by the total number of samples genotyped in each year. (B) The shift in the genome constellation agreed well with two waves of G8P[8] prevalence. (C) It was speculated that a bovine-like G8 strain jumped into humans between 2009 and 2012.The strains dominant during the first wave had three animal RVA-derived backbone *genes* but the strains dominant during the second wave expelled all these animal *genes*, comprising fully human RVA *genes* except the G8 VP7 *gene*. Mutational changes were observed in neutralisation epitopes of VP7 and VP8* concurrent with the two waves of G8P[8] prevalence, but none of these sites appeared to be positively selected.

3.2 The shift in the genome constellations concurrent with the two waves of G8P[8] prevalence in Vietnam

To further characterise the samples, 597 (71 per cent) of the 843 G8P[8] samples were subjected to PAGE, and an electropherotype was assigned to each of 402 samples (Fig. 1). Of those, 393 had short RNA patterns which were classified into eighty-three electropherotypes, and nine samples had mixed RNA patterns (i.e. short RNA patterns with some extra bands which were taken as evidence of mixed infection). These samples were classified into five mixed electropherotypes. The rest of the samples (195) did not produce clearly visible bands on the gel, so we were not able to assign an electropherotype to any of them (Fig. 1).

Furthermore, sixty-five samples chosen from the 393 samples and representative of 61 different electropherotypes were subjected to whole-genome sequencing. Seven samples failed to yield meaningful sequences enough for further analysis. Whole-genome sequences were obtained for fifty-three samples; whereas sequences were obtained for all but parts of the genome in the remaining five samples (Fig. 1). The sequencing read information, including the number of reads and the average coverage of *gene* segments of those G8P[8] strains in this study, was described in Supplementary Table S1. The medians of total read number and average coverage were respectively 61,612 (IQR: 21,240–141,092) and 5,097 (IQR: 2,065–10,458) (Supplementary Table S1). Sequences determined in this study were deposited under the following accession numbers LC074731-LC074741 (Hoa-Tran et al., 2016) and LC784433-LC785054 (Supplementary Table S2). Supplementary Table S2 also included information on the length of CDS and breadth of coverage. The total lengths of

full-genome sequences merged from CDS of eleven *gene* segments ranged from 15,404 to 17,427 nucleotides, with an average of 17,065 nucleotides, covering 97.9 per cent (17,065/17,427) of the full-length sequence merged from eleven complete CDS (Supplementary Table S2). Those fifty-eight samples were representative of fifty-five different electropherotypes (Fig. 1) which were numbered from SP1 to SP55 (Supplementary Table S3).

Phylogenetic analysis was carried out to determine the lineages to which each of the *gene* segments of the fifty-eight RVAs belonged (Supplementary Figs S3–S13). Supplementary Table S3 lists the genotype/lineage constellation and the electropherotype of each of the fifty-eight G8P[8] samples and the number of samples that shared the same electropherotype. Phylogenetic analysis on the whole-genome sequences revealed that seven *gene* segments, i.e. the VP7, VP4, VP6, VP2, NSP1, NSP3, and NSP5 *genes* remained in the same lineage throughout the 8 years of G8P[8] strain circulation (Supplementary Table S3). The VP7 *gene* belonged to a bovine-like lineage (G8.IV) (Supplementary Fig. S3) while the other *genes* belonged to human lineages, i.e. P[8].III (VP4), I2.V (VP6), C2.IVa (VP2), A2.IVa (NSP1), T2.V (NSP3), and H2.IVa (NSP5) (Supplementary Figs S4, S5, S7, S9, S11, and S13, respectively). In contrast, there were changes of lineages in the VP1 (from R2.XI to R2.V), VP3 (M2.V to M2.VII), NSP2 (N2.XV to N2.V), and NSP4 (E2.XII to E2.VI) *genes* between the first wave of prevalence and the second wave (Supplementary Table S3 and Supplementary Figs S6, S8, S10, and S12). Of those lineages in which changes occurred, the R2.XI, N2.XV, and E2.XII were bovine (artiodactyl)-RVA derived (Supplementary Figs S6, S10, and S12, respectively) (Hoa-Tran et al. 2016; Tacharoenmuang et al. 2016; Agbemabiese et al. 2019). There were two samples (RVN19-0507: SP25 and RVN19-0909: SP30) that showed some extra bands indicative of mixed infections upon PAGE. They had the T1 genotype in the NSP3 *gene*, suggesting that they were intergenogroup reassortant in the making (Supplementary Table S3 and Fig. S11). There were five samples for which the sequences of one to three genome segments were not obtained (blanks in Supplementary Table S3), precluding complete assignment of a lineage constellation.

After excluding those five incomplete samples, there were seven genotype/lineage constellations, namely A to G (A–G), observed in the fifty-three samples for which a complete genotype constellation was determined (Table 1). Constellations A–D comprised the samples which were detected in the first wave (2014–17), while constellation E–G were in the second wave (2018–21), except for only two samples detected in 2017.

Of the seven genotype/lineage constellations, there were two dominant ones, constellations A and E, accounting for 79.2 per cent of the G8P[8] samples for which complete genotype constellations were determined. In particular, the samples with constellation A accounted for 22.6 per cent (12/53) and those with constellation E accounted for 56.6 per cent of the actual sequences (Table 1). Assuming that the samples having the same electropherotype possessed the same genotype/lineage constellation, the percentage accounted for by constellation A would be 36.1 per cent (103/285) and that accounted for by constellation E would be 36.8 per cent (105/285) (Table 1). So, it is clear that the shift in the genome constellations from A to E occurred concurrently with the two waves of G8P[8] prevalence whether the number was counted by the actually sequenced strains or by the estimates extrapolated by the assumption based on the electropherotypes. Furthermore, the shift in the genome constellation agreed well with two waves of G8P[8] prevalence determined by G–P typing alone (Fig. 2B).

Constellation A as represented by RVA/Human-wt/VNM/RVN1149/2014/G8P8 which possessed three animal-derived *genes*, R2.XI (VP1), N2.XV (NSP2), and E2.XII (NSP4), in the backbone dominated during the first wave of G8P[8] prevalence (2014–17). Constellation E as represented by RVN19-0608 which possessed no animal-derived *genes* in the backbone dominated during the second wave of G8P[8] prevalence (2018–21) (Table 1, Fig. 2B). The other two constellations of the second wave, i.e. F and G, also had no animal-derived backbone *gene* (Table 1). In short, the strains dominant during the first wave had three animal RVA-derived backbone *genes* but the strains dominant during the second wave expelled all these animal *genes*, comprising fully human RVA *genes* except G8 VP7 *genes* (Fig. 2C).

The ML tree on the full-genome sequences merged from eleven CDS of eleven *gene* segments of those fifty-three DS-1-like G8P[8] RVA strains, for which a complete genotype constellation was determined (Supplementary Tables S2 and S3), also significantly showed two distinct genetic clusters of constellations A and E with the nucleotide differences of those merged CDSs up to 4.7 per cent (IQR: 4.7–4.9 per cent) (Fig. 3). The average length of the merged CDSs from those fifty-three G8P[8] samples was 17,154 nucleotides, covering up to 98.4 per cent of the full-length sequence merged from eleven complete CDS (i.e. 17,427 nucleotides).

BLAST was employed by using the sequences of eleven *gene* segments from two representative strains RVA/Human-wt/VNM/RVN1149/2014/G8P8 and RVN19-0608 to search the GenBank database for similar sequences of DS-1-like G8P[8] strains. Phylogenetic analysis of eleven *gene* segments revealed that RVA/Human-wt/VNM/RVN1149/2014/G8P8 shared similar genotype/lineage constellation to G8P[8] strains detected in other countries in the period 2014–19, while the genotype/lineage constellation of RVN19-0608 was similar with those detected in 2019–22. In particular, constellation A was found from G8P[8] strains in Thailand (in 2014), USA (2016–17), and Japan (2014, 2016, 2018–19); constellation E was observed in Thailand (in 2019) and China (2020–22) (Supplementary Figs S3–S13; Fig. 3). Moreover, the nucleotide identities of 11 CDS from eleven *gene* segments among RVA/Human-wt/VNM/RVN1149/2014/G8P8 with some worldwide G8P[8] strains with constellation A and among RVN19-0608 with those with were as high as 99.2100 per cent (Table 2). The results suggested similar shift in the genome constellations among G8P[8] strains circulating worldwide, from the animal-derived to the full human RVA *gene* backbones.

3.3. BEAST analysis on the G8 lineage IV VP7 genes

To provide a detailed, time-resolved view of viral evolutionary dynamics, BEAST analysis on the G8 lineage IV VP7 *genes* primarily of Vietnamese origin was further incorporated. The mean evolutionary rate of the G8 lineage IV VP7 *genes* was calculated to be 2.3×10^{-3} substitutions per site per year with the 95 per cent highest posterior density interval being 1.6×10^{-3}–3.3×10^{-3}, which is in good agreement with yet narrower than that reported by Degiuseppe et al. (2022). While the tree was almost exclusively composed of Vietnamese strains collected in this study, it also included a few Thai strains collected in 2013 and Indian artiodactyl strains, whose VP7 *genes* belong to G8 lineage IV, in order to address the question of when an artiodactyl strain of G8 lineage IV crossed the species barrier to get into humans. The tree shows that the interspecies transmission occurred sometime between 2009

Table 1. Genotype/lineage constellations of G8P[8] strains detected in Vietnam between 2014 and 2021.

	Genotype/lineage constealltions													Circulation period of the extraplorated samples	
Constellation	VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5	Representative strain	Number of samples with the same constellation (actual)	Number of samples with the same constellation (extraploration)	Years	Phases
A	**G8.IV**[a]	P[8].III	I2.V	**R2.XI**[a]	C2.IVa	M2.V	A2.IVa	**N2.XV**[a]	T2.V	**E2.XII**[a]	H2.IVa	RVN14-1149[b]	12 (22.6%)	103 (36.1%)	2014–17	First wave
B	**G8.IV**[a]	P[8].III	I2.V	**R2.XI**[a]	C2.IVa	M2.V	A2.IVa	**N2.XV**[a]	T2.V	E2.VI	H2.IVa	RVN15-0632	2 (3.8%)	13 (4.6%)	2015–17	
C	**G8.IV**[a]	P[8].III	I2.V	R2.V	C2.IVa	M2.V	A2.IVa	**N2.XV**[a]	T2.V	E2.VI	H2.IVa	RVN16-0459	1 (1.9%)	2 (0.7%)	2016	
D	**G8.IV**[a]	P[8].III	I2.V	**R2.XI**[a]	C2.IVa	M2.V	A2.IVa	N2.V	T2.V	**E2.XII**[a]	H2.IVa	RVN16-0614	1 (1.9%)	23 (8.1%)	2016	
E	**G8.IV**[a]	P[8].III	I2.V	R2.V	C2.IVa	M2.V	A2.IVa	N2.V	T2.V	E2.VI	H2.IVa	RVN19-0608	30 (56.6%)	105 (36.8%)	2017[c]–21	Second wave
F	**G8.IV**[a]	P[8].III	I2.V	R2.V	C2.IVa	M2.VII	A2.IVa	N2.V	T2.V	E2.VI	H2.IVa	RVN19-0064	5 (9.4%)	35 (12.3%)	2019–20	
G	**G8.IV**[a]	P[8].III	I2.V	R2.V	C2.IVa	M2.V	A2.IVa	N2.V	T1	E2.VI	H2.IVa	RVN19-0507	2 (3.8%)	4 (1.4%)	2019	
Total													53 (100%)	285 (100%)	2014–21	

[a]Lineages of animal-RVA origin (Hoa-Tran et al., 2016; Tacharoenmuang et al. 2016; Agbemabiese et al. 2019).
[b]RVA/Human-wt/VNM/RVN1149/2014/G8P8.
[c]Only two samples were detected in 2017 and all the rest (103) were detected between 2018 and 2021.

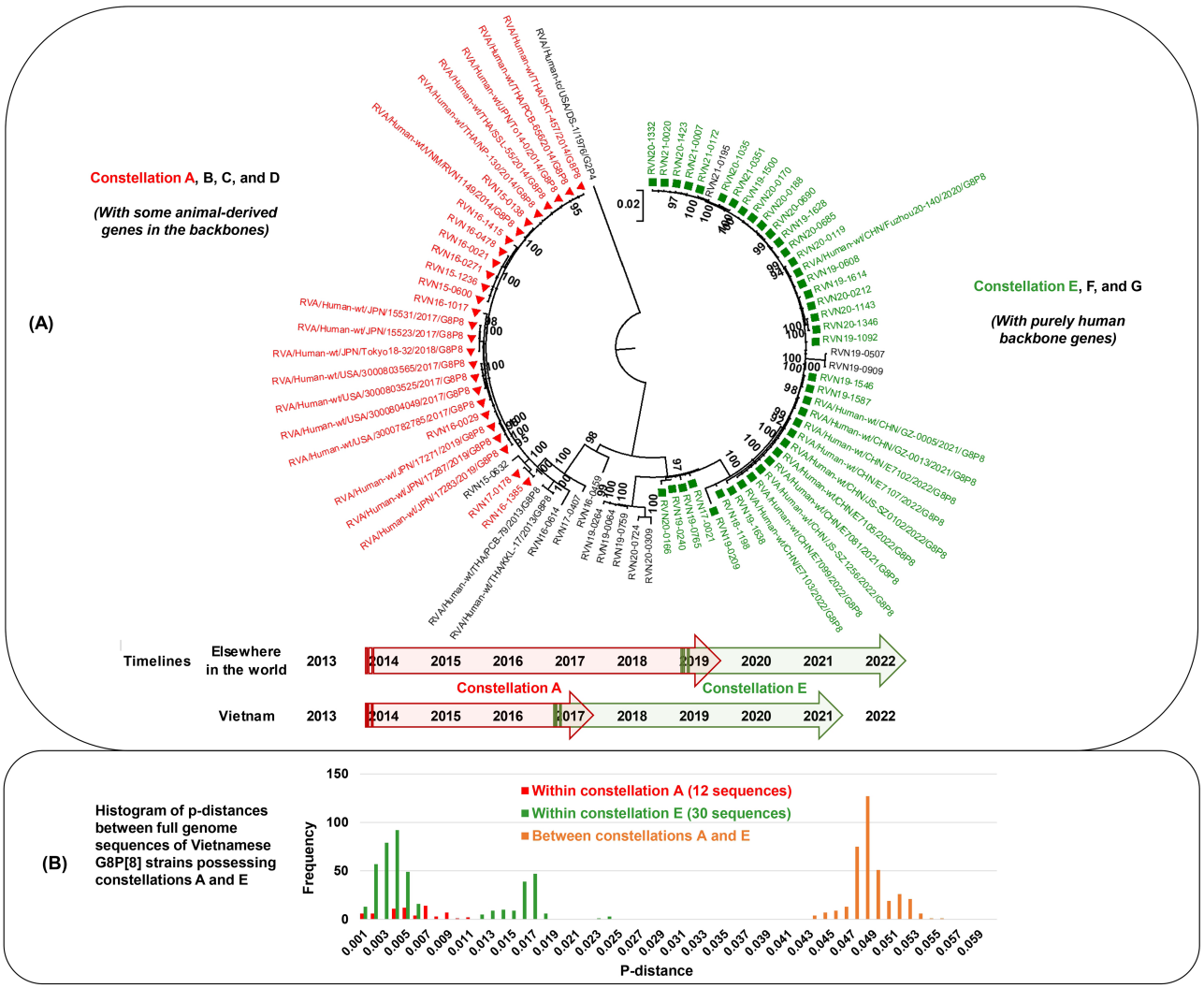

Figure 3. Genetic analyses on full-genome sequences that were made by merging the eleven CDS of the eleven *gene* segments of DS-1-like G8P[8] strains from Vietnam and elsewhere in the world. (A) A ML subtree was constructed from which two outgroups (Wa and AU-1) were excluded. The bootstrap values less than 90 per cent were hidden. The branch lengths were scaled to the genetic distance defined as the number of substitutions per site. The G8P[8] strains possessing constellation A were displayed with triangles and those possessing constellation E with squares. (B) Histogram of pairwise-distances (p-distances) between the full-genome sequences of Vietnamese G8P[8] strains possessing constellations A and E. The p-distances were significantly separated into two categories: (1) within constellation A (0.1–1 per cent) and E (0–2.4 per cent), respectively; (2) between constellations A and E (4.3–5.5 per cent).

and 2012, and that there were Thai G8P[8] strains that preceded the similar strains detected in Vietnam in 2014 (Fig. 4).

3.4. Putative recombination event on the VP7 *gene*

Analysis of all eleven *gene* segments of Vietnamese G8P[8] strains suggested that there was only one recombination event occurring on the VP7 of strain RVN17-0178 (constellation A). The event was predicted to occur in the positions 189–220 (in alignment) or 163–230 (95 per cent CI) of the VP7 CDS, in which no neutralisation epitope was located, by the five methods such as RDP, Geneconv, MaxChi, Chimæra, and 3Seq (KA P-Val: 3.38 E-19). The potential major parent was strain RVN16-0021 (constellation A) (99.6 per cent similarity). The potential minor parent was unknown, although BLAST search using the region of breaking points suggested the closely related sequence of the VP7 *gene* from G3 strains, including RVA/Human-wt/VNM/VE3425/2018/G3P[8] (100 per cent similarity) (Fig. 5).

3.5 Nucleotide difference and amino acid variation of VP7 and VP4 *genes*: Mutational changes on neutralisation epitopes concurrent with the two waves of G8P[8] prevalence in Vietnam

More precise examination was conducted for the nucleotide sequences and amino acid variations of the two outer capsid protein *genes* VP7 and VP4.

Excluding the putative recombinant RVN17-0178, whose VP7 nucleotide sequence differed by 2.1–3.5 per cent from that of the other fifty-seven G8P[8] strains, the nucleotide differences (p-distances) of the CDS of VP7 of fifty-seven G8P[8] strains detected over 8 years of circulation ranged from 0 per cent to 1.6 per cent.

Using the VP7 sequence of RVA/Human-wt/VNM/RVN1149/2014/G8P8 as a query, amino acid substitutions on the VP7 were determined at eight residues, comprising of Y32C, V42A, A65T/V, A125S, N147D, A225T/V, Y235H, and T281N (Supplementary Fig. S14). Of those, the substitutions of A125S and N147D were

Table 2. G8P[8] strains detected elsewhere in the world that were most closely related in their nucleotide sequences to Vietnamese G8P[8] strains of constellations A and E.

				Nucleotide identities (%)[a]											
Queries	**Representative strains**	**Country**	**Period of circulation**	**VP7**	**VP4**	**VP6**	**VP1**	**VP2**	**VP3**	**NSP1**	**NSP2**	**NSP3**	**NSP4**	**NSP5**	**Reference**
Comparision based on complete CDS sequences															
RVA/Human-wt/VNM/RVN1149/2014/G8P8 (Constellation A) (2014–17)	RVA/Human-wt/JPN/To14-0/2014/G8P8	Japan	2014	99.6%	99.8%	99.8%	99.8%	99.8%	99.8%	100%	99.9%	99.7%	99.8%	99.7%	Kondo et al., 2017
	RVA/Human-wt/THA/NP-130/2014/G8P8	Thailand		99.7%	99.8%	99.9%	99.9%	99.7%	99.8%	99.9%	99.8%	99.7%	99.8%	99.3%	Tacharoen-muang et al. 2016
	RVA/Human-wt/THA/PCB-656/2014/G8P8			99.8%	99.7%	99.8%	99.8%	99.8%	99.8%	99.9%	99.8%	99.8%	99.6%	99.8%	
	RVA/Human-wt/THA/SKT-457/2014/G8P8			99.8%	99.8%	99.8%	99.9%	99.8%	99.8%	99.9%	99.9%	99.8%	99.8%	99.2%	
	RVA/Human-wt/THA/SSL-55/2014/G8P8			99.8%	99.9%	99.8%	99.8%	99.8%	99.8%	99.9%	99.9%	99.7%	99.8%	99.8%	
	RVA/Human-wt/JPN/15,523/2017/G8P8		2017	99.6%	99.3%	99.8%	99.8%	99.8%	99.8%	99.4%	99.9%	100%	99.8%	99.7%	Phan et al. 2022
	RVA/Human-wt/JPN/15,531/2017/G8P8			99.6%	99.3%	99.8%	99.8%	99.8%	99.7%	99.4%	99.9%	100%	99.8%	99.7%	
	RVA/Human-wt/JPN/Tokyo18-32/2018/G8P8	Japan	2018	99.7%	99.3%	99.8%	99.9%	98.5%	99.7%	99.3%	99.9%	99.9%	99.6%	99.5%	Fujii et al. 2019
	RVA/Human-wt/JPN/17,271/2019/G8P8		2019	99.7%	99.3%	99.9%	99.8%	98.9%	99.8%	99.3%	99.9%	99.9%	99.4%	99.8%	Phan et al. 2022
	RVA/Human-wt/JPN/17,283/2019/G8P8			99.8%	99.3%	99.8%	99.8%	98.9%	99.8%	99.4%	99.9%	99.9%	99.4%	99.8%	
	RVA/Human-wt/JPN/17,287/2019/G8P8			99.8%	99.3%	99.8%	99.8%	98.9%	99.8%	99.4%	99.9%	99.9%	99.4%	99.8%	
RVN19-0608 (Constellation E)(2018–21)	RVA/Human-wt/CHN/GZ-0005/2021/G8P		2021	99.9%	99.5%	99.8%	99.8%	99.8%	99.6%	99.8%	99.7%	99.9%	100%	100%	Wang et al. 2022
	RVA/Human-wt/CHN/GZ-0013/2021/G8P8			99.7%	99.6%	99.8%	99.8%	99.9%	99.7%	99.7%	99.7%	99.9%	99.8%	99.8%	
	RVA/Human-wt/CHN/E7081/2021/G8P8			100%	99.8%	99.8%	99.8%	99.8%	99.6%	99.7%	99.6%	100%	99.6%	100%	Zhou et al. 2023
	RVA/Human-wt/CHN/E7099/2022/G8P8			100%	99.7%	99.7%	99.8%	99.8%	99.7%	98.1%	99.4%	99.9%	99.8%	100%	Zhou et al. 2023
	RVA/Human-wt/CHN/E7102/2022/G8P8			99.9%	99.8%	99.8%	99.8%	99.8%	99.6%	99.6%	99.6%	99.9%	99.8%	100%	
	RVA/Human-wt/CHN/E7103/2022/G8P8			99.9%	99.8%	99.6%	99.8%	99.7%	99.6%	99.6%	99.6%	100%	99.8%	100%	
	RVA/Human-wt/CHN/E7105/2022/G8P8			99.9%	99.7%	99.8%	99.8%	99.7%	99.7%	99.6%	99.6%	100%	99.8%	100%	
	RVA/Human-wt/CHN/E7107/2022/G8P8			100%	99.8%	99.8%	99.8%	99.8%	99.6%	99.7%	99.7%	99.9%	99.6%	100%	
Comparision based on partial nucleotide sequences (471–804 nucleotides) with Thai strains which shared similar genotype/lineage constellation[b]															
RVN19-0608 (Constellation E)(2018–21)	RVA/Human-wt/THA/5CR11/2019/G8P8	Thailand	2019	99.9%	99.6%	100%	96.8%	99.7%	99.5%	100%	99.5%	96.6%	96.7%	97.6%	Chan-It, Chanta, and Ushijima et al. 2023
	RVA/Human-wt/THA/5CR23/2019/G8P8			99.9%	99.6%	99.8%	96.6%	99.7%	99.5%	100%	99.5%	96.6%	96.7%	97.9%	
	RVA/Human-wt/THA/5CR33/2019/G8P8			99.9%	99.6%	100%	96.6%	99.7%	99.4%	99.7%	98.6%	96.5%	97.0%	97.0%	
	RVA/Human-wt/THA/5CR56/2019/G8P8			99.9%	99.5%	100%	96.6%	99.7%	99.5%	99.7%	98.6%	96.6%	97.0%	99.5%	

The italic values: Nucleotide identitites were less than 99 per cent but the sequences belonged to the same lineage as the queries, i.e. RVA/Human-wt/VNM/RVN1149/2014/G8P8 or RVN19.0608.

[a]Nucleotide identities (%) = 100 × (1 – p-distance).

[b]Those four Thai strains with constellation E were included in the 11 ML trees of 11 *gene* segments (Supplementary Figs S3–S13) but they were not included in ML tree of full-genome sequence (Fig. 3) because the length of sequence merged from 11 CDS was 7037 nucleotides, covering only 40.1 per cent of the full-length genome (i.e. 17,427 nucleotides).

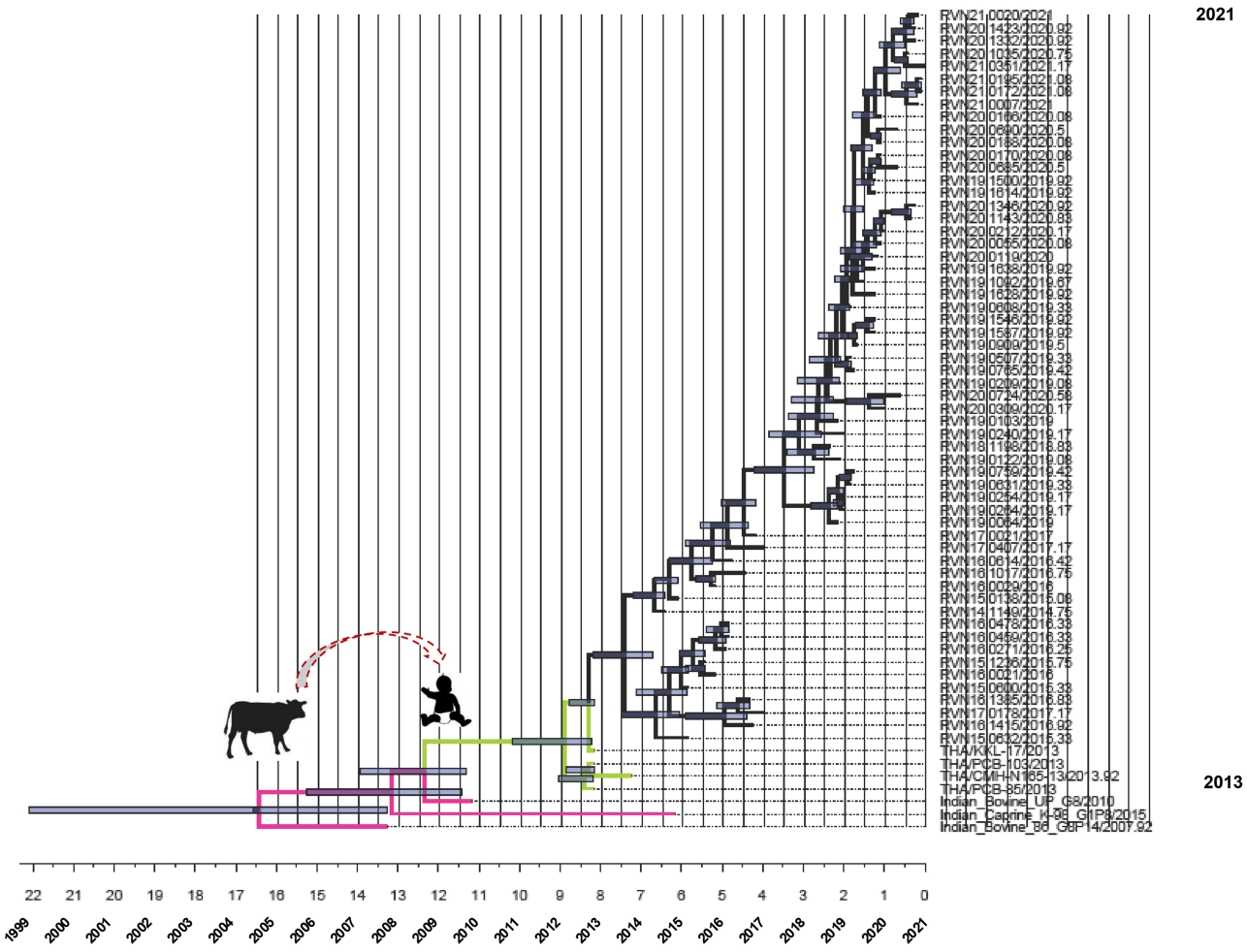

Figure 4. The time-scaled phylogenetic tree for the VP7 *gene*, estimating the time of emergence, and the spread of bovine-like G8 RVAs into human population. The interspecies transmission of bovine-like RVAs into human occurred sometime between 2009 and 2012, and that there were Thai G8P[8] strains that preceded the similar strains detected in Vietnam in 2014. The 22-year timescale (i.e. 1999–2021) was indicated below the tree.

located on the neutralisation epitopes 7–1a and 7–2, respectively (Chan-It, Chanta, and Ushijima et al. 2023). Residue 125A was found on the 7–1a regions of G8P[8] strains circulating in the first wave (up to 2017, including strain RVN17-0021 (constellation E)) while 125S was detected from the strains in the second wave (since 2018). Similarly, residue 147 N was found on the 7–2 regions among the G8P[8] strains circulating in the first wave (up to 2017, including strain RVN17-0021 (constellation E)), while 147D was found among most of strains circulating in the second wave (since 2018). In all, 147 N was re-detected in some strains of constellation E, circulating in 2020 (Supplementary Fig. S14). Notably, an exception was the VP7 of strain RVN17-0021, a 'transitional' strain, although assigned to constellation E, shared the VP7 amino acid sequences more closely related to those of constellations A–D.

The nucleotide differences (p-distances) of the VP4 of fifty-eight G8P[8] strains detected over the 8 years of circulation ranged from 0 per cent to 2.5 per cent. Amino acid substitutions in VP4 identified using RVA/Human-wt/VNM/RVN1149/2014/G8P8 as a query occured at seventeen residues. Of those ten substitutions occurred on the VP8*, comprising of K7R/S, V35I, D113N, R131K, D135N, S146N/G, S149N, D195G, S236P, and K246R (Supplementary Fig. S15). The substitutions D113N, R131K, and D135N were located on the neutralisation epitopes 8–3, while the substitutions S146N/G and D195G/N were located on the epitope 8–1 of the VP8* (Chan-It, Chanta, and Ushijima et al. 2023). Notably, 146S and 195D (epitope 8–1) were detected among the G8P[8] strains with constellations A–D, circulating in the first wave (2014–17), while 146 G and 195 G were detected among the strains with constellations E–G, circulating in the second wave (2017Q1 (RVN17-0021) and 2018–21) (Supplementary Fig. S15). Notably, unlike VP7, the 'transitional' strain RVN17-0021 possessed the VP4 amino acid sequences typical of the second-wave strains.

Analysis of the dN/dS ratios for the VP7 and VP4 CDS of the DS-1-like G8P[8] failed to identify any positive selection site, including the sites having substitutions located on the putative neutralisation epitopes (Supplementary Fig. S14–S15) ($P \leq 0.05$). Furthermore, the substitutions in VP7 and VP4 among G8P[8] strains were suggested to be under neutral evolution (Supplementary Table S4).

4. Discussion

Interspecies transmission contributes to the evolution of the rotavirus genome (Martella et al. 2010). However, the reality is not as simplistic as an animal RVA strain jumping to humans and causing an epidemic. When an animal rotavirus crosses the host species barrier, infection occurs as whole virions. Here, a vast majority of them are considered to terminate in dead-end infection (McDonald et al. 2016; Warren and Sawyer et al. 2019).

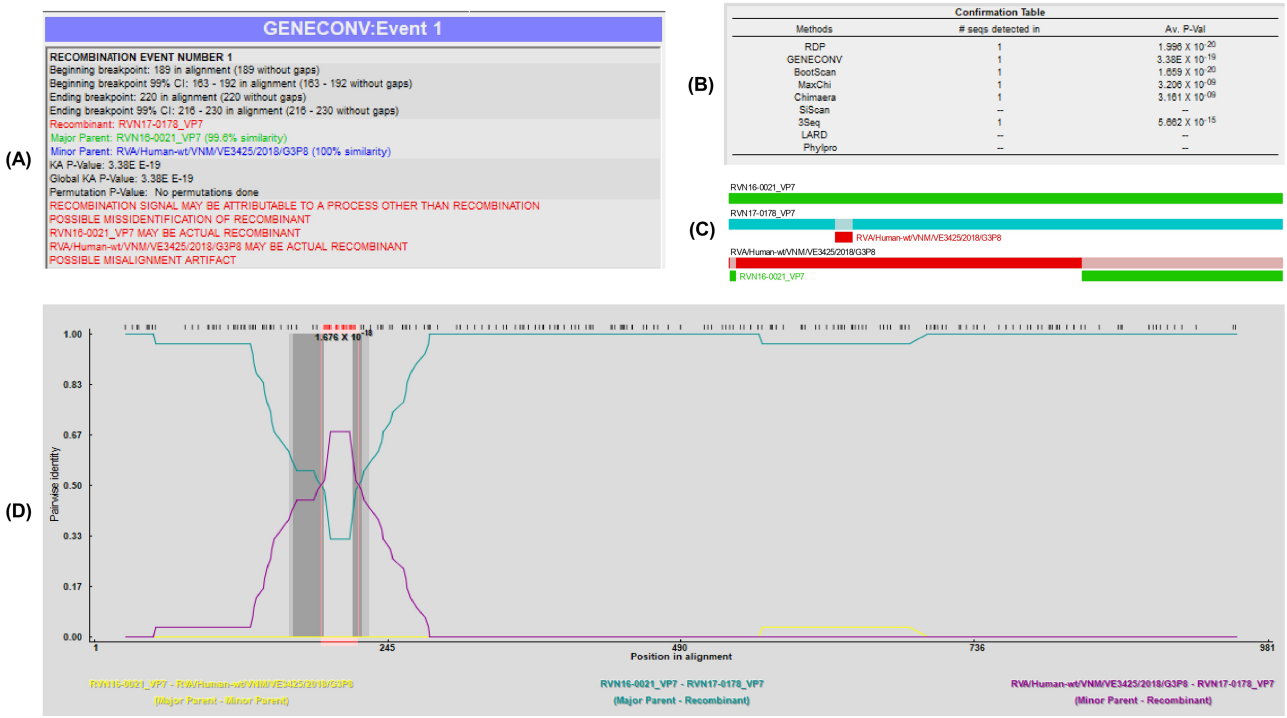

Figure 5. Evaluation of recombination event occurring on the CDS of VP7 of strain RVN17-0178, using the RDP4.101 BetaBig: (A) recombination information; (B) the recombination event was confirmed by five detection methods: RDP, Geneconv, MaxChi, Chimæra, and 3Seq. (C) Recombination graphic; (D) visualisation of evidence of recombination using RDP4.101 BetaBig tool.

Therefore, in theory, spillover infections of.. animal RVAs can make little contribution to rotavirus evolution, unless they succeed in establishing a human-to-human transmission chain. Only very rarely, RVAs of animal origin likely establish such onward transmission among the new host (Webby, Hoffmann, and Webster et al. 2004; Mukherjee. et al. 2013; Malik et al. 2020). No evidence thus far exists that the animal-like RVA in its entirety caused efficient onward transmission among humans. RVA strains of suspected zoonotic origin that caused an identifiable spread were almost exclusively reassortants between human and animal strains (Kirkwood 2010). Indeed, exemplary reassortant RVA strains that caused global spread were almost limited to equine-like G3P[8] RVA that caused an outbreak for the first time in Australia (Cowley et al., 2016), and bovine-like G8P[8] that was reported for the first time in Vietnam (Hoa-Tran et al. 2016). While a zoonotic virus is considered to establish and spread among humans undergoing genetic reassortment to enhance its fitness within the human population (McDonald et al. 2016; Warren and Sawyer et al. 2019), there was no study in which the fate of the animal RVA-derived *genes* was pursued over the course of spread among humans. As a zoonotic RVA strain, by definition, had a genome entirely of animal RVA at the time of infecting a human host, we hypothesised that such zoonotic reassortant capable of human-to-human transmission would have lost its animal RVA-derived *genes* and gained their cognate human RVA *genes* over time (Fig. 6).

In this context, the detection of the bovine-like G8 strains, which were speculated to have jumped into humans between 2009 and 2012 (Fig. 4; Degiuseppe et al. 2022) in Vietnam at an early phase of emergence provided a unique opportunity to monitor the evolutionary changes that the ancestral animal strain having jumped to humans would undergo. This study showed that the G8P[8] strains that suddenly emerged in Vietnam in 2014 subsequently circulated in two waves of predominance during the following 8 years before they went extinct. Electropherotyping coupled with whole-genome sequencing of selected G8P[8] samples revealed that this biphasic predominance of G8P[8] strains was accompanied, if not caused, by the concurrently occurring shift of the predominant strains with two different types of genotype/lineage constellations, which we termed constellation A and constellation E (Fig. 2B, Table 1, Supplementary Table S3). The most significant difference between the strains with two constellations was the number of bovine RVA-derived *genes* in their genetic backbones (i.e. the internal capsid protein *genes* and non-structural protein *genes*). The earlier strains with constellation A possessed three bovine RVA *genes* in their VP1, NSP2, and NSP4 *genes*, whereas the later strains with constellation E replaced all of the three bovine RVA *genes* with cognate human RVA-derived *genes* (Fig. 2 and Table 1).

When we analysed one of the G8P[8] strains that Tacharoenmuang et al. (2016) reported, specifically RVA/Human-wt/THA/PCB-79/2013/G8P[8], and other six strains, they had a genotype/lineage constellation of G8-P[8]-I2.VII-R2.XI-C2.IVa-M2.V-A2.IVa-N2.XV-T2.V-E2.XII-H2.IVa, which uniquely differed from constellation A in the VP6 lineage (I2.VII) (Supplementary Fig. S5; Agbemabiese et al. 2019). The lineage I2.VII was carried by RVA strains detected from cow and pig (Tacharoenmuang et al. 2016). Therefore, these strains, which emerged a year earlier than the earliest identified Vietnamese strain, contained an additional bovine RVA-derived *gene*, totalling five such *genes* including the G8 VP7 *gene*.

There are a few examples that are in good agreement with our observation on the shift of the predominant strains with two different types of genotype/lineage constellations. Tacharoenmuang et al. (2016) and Kondo et al. (2017) reported the emergence and spread of G8P[8] in Thailand (in 2013–14) and Japan (during the 2014 rotavirus season). The Thai and Japanese G8P[8] strains detected in 2014 were 99.3–100 per cent identical in nucleotides

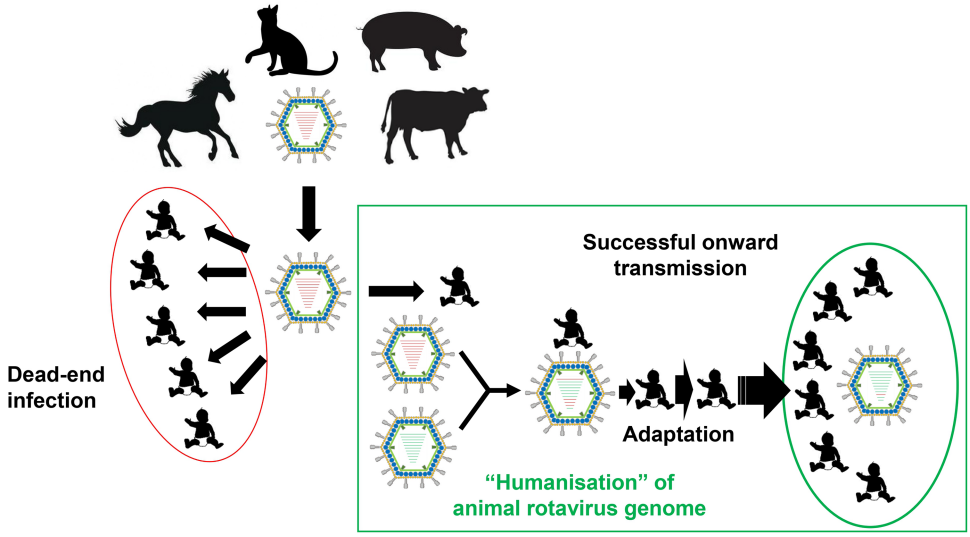

Figure 6. Hypothesis on the fate of animal RVA-derived *genes* of spillover animal RVAs during human-to-human spread. The vast majority of cases result in dead-end infections without further implication for rotavirus evolution. In much smaller proportion of cases, there may be some degree of onward, human-to-human transmission, during which the spillover virus must undergo reassortment with a co-infecting human virus and exchange of cognate *genes* to achieve 'humanisation' of the animal rotavirus genome. During the 'adaptation' process to the human host occurs the expulsion of animal RVA-derived *genes* from the genome of the progeny of the original spillover virus.

in every genome segment to Vietnamese strain with constellation A (Table 2); thus, they also had the same constellation A. G8P[8] strains possessing constellation A were continuously detected in Japan until 2019 (Table 2; Fig. 3).

Chan-It, Chanta, and Ushijima et al. (2023) reported that from 2018 to 2020 in Thailand, the predominate G8P[8] strains were novel reassortant. Their VP1, NSP2, and NSP4 *genes*, distinct from those in RVA/Human-wt/VNM/RVN1149/2014/G8P8 (constellation A), belonged to different lineages. The genetic analysis on those G8P[8] strains, i.e. RVA/Human-wt/THA/5CR11/2019/G8P8, RVA/Human-wt/THA/5CR23/2019/G8P8, RVA/Human-wt/THA/5CR33/2019/G8P8, and RVA/Human-wt/THA/5CR56/2019/G8P8, of which partial nucleotide sequences of the eleven *gene* segments were available in GenBank, suggested that they had a genotype/lineage constellation identical to constellation E (Supplementary Figs S3–13; Table 2). This constellation was that of the most prevalent G8P[8] strains in Vietnam from 2018 to 2021 that lost animal RVA-derived *genes*, as represented by RVN19-0608.

Other examples are from Guangzhou, China, where two out of five RVA-positive samples collected during the winter of 2021–22 had G8P[8] (Wang et al. 2022), and from Wuhan, China, where G8P[8] emerged and was dominant in the 2021–22 epidemic season (Zhou et al. 2023). When we analysed full-genome sequences of the eight G8P[8] strains published in these Chinese studies, all had the same constellation E, and seven out of the eight strains were very closely related in nucleotide identities to Vietnamese RVN19-0608 strain (Table 2). Moreover, a BLAST search revealed that other Chinese G8P[8] strains detected between 2020 and 2022 were 99.6–100 per cent identical in nucleotide sequence to RVN19-0608 (Table 2). Interestingly, during the same 2021–22 epidemic season in China, a high proportion of infants with severe acute gastroenteritis in a multicentre study were found to be infected with G8P[8] strains, although the number of RVA-positive specimens collected in each study site was small: Huizhou (25.0 per cent 1/4), Shunde (55.6 per cent 5/9), Shenzhen (42.1 per cent 8/19) in Guangdong Province, Mianyan (36.4 per cent 8/22) in Sichuan Province, and Xiamen (11.1 per cent 3/27) in Fujian Province (Wang et al. 2022). Taken together, the evolutionary change from constellations A to E in the DS-1-like G8P[8] strains was probably more widely occurring in Asian countries.

Considering the data available to date, we believe it is reasonable to postulate a general principle; that the more *genes* of animal origin in a strain of suspected animal RVA contains, the shorter time has elapsed since its progenitor animal strain crossed the species barrier to infect humans. Reminiscent in this regard is G8P[8] strains prevalent in the African continent. Nakagomi et al. (2013) examined the whole genome of twenty-eight G8 strains with the DS-1-like backbone *genes* circulating in Malawi over a decade (1997–2007), and found no evidence for the introduction of contemporary bovine rotavirus *genes* into any of the G8 strains examined. In addition, the backbone *genes* of these Malawian G8 strains were indistinguishable from those of human DS-1like strains circulating in Africa. Their observation can be interpreted to indicate that a long time elapsed since the bovine RVA strain ancestral to the contemporary Malawian G8 strains had jumped the species barrier to infect humans occurred, thus giving the spillover strain and its progeny a space of time long enough for all bovine-derived *genes* to be replaced with cognate human RVA-derived *genes*.

It is also clear that the two waves of G8P[8] prevalence were accompanied with the point mutations on antigenic epitopes of both VP7 and VP4. Those substitutions, i.e. A125S and N147D in the VP7, and S146G and D195G in the VP8*, were located on the surface-exposed area of the structures of VP7 and VP8* by using homology modelling analysis (Chan-It, Chanta, and Ushijima et al. 2023). Nevertheless, the dN/dS index analysis suggested a concurrent occurrence of random mutational changes, i.e. substitutions were estimated under neutral evolution (Supplementary Table S4), among strains circulating in two waves of prevalence (i.e. constellations A–D vs E–H). This random genetic change was unlikely to cause dominant strain shift between two waves of prevalence.

In addition, the predominance of mutations among G8P[8] strains, as quasispecies, may be explained by change during persistence in environment which may be critical for faecal–oral transmission (Omatola et al., 2022a). A case in point was the occurrence of the T281I substitution that was previously found in rhesus rotavirus populations less sensitive to a disinfectant (Kadoya et al. 2022).

Homologous recombination is rare in rotaviruses as is commonly perceived (McDonald et al. 2016). Out of fifty-eight G8P[8] strains tested, the intragenic recombination was found in only the VP7 *gene* of strain RVN17-0178 (1.7 per cent). In addition, recombination was estimated to have occurred outside the epitope-coded region. It is thus likely that the recombination played little role in contributing to diversity and evolution or effecting on lineage relationships (Hoxie and Dennehy et al. 2020) of Vietnamese DS-1-like G8P[8] strains. However, it deserves mention that out of the 339 samples for which representative electropherotype strains were assigned (Supplementary Table S3), strain RVN17-0178 represented electropherotype SP15, and 20 (5.9 per cent) samples had electropherotype SP5, thereby suggesting that the frequency of intragenic recombination on the VP7 *gene* could be as high as 5.9 per cent. Thus, the real frequency of intragenic recombination events could possibly be higher than previously perceived.

In conclusion, there seems to be evolutionary changes in the backbone *gene* constellation, with the potential for the animal RVA-derived *genes* to be expelled from the genomic backbone of the progeny of an animal RVA that had once crossed the host species barrier during the process of adaptation to the human host (Fig. 6).

Important questions remain unanswered; firstly, were RVAs with constellation E better fitted to replication and transmissibility in humans than RVAs with constellation A? Secondly, what caused disappearance of the unusual yet dominant reassortant strains, particularly with regard to the loss of their animal-derived *genes* during circulation among humans? Possible causes may include effects of increasing use of rotavirus vaccines that may be effective against heterotypic genotypes, the presence of herd immunity induced by vaccination, and decrease in spillover infection of animal RVAs as the supplementary source of pathogens (John and Samuel et al. 2000; Gentsch et al. 2005; Justino et al. 2011; Pollard 2015; Omatola et al., 2022b; Hermann et al. 2024). Additionally, it is conceivable that the bovine-like G8P[8] strains could exchange their animal-derived *genes* to other co-circulating strains possessing genotypes other than G8P[8], thereby escaping the detection in this study. Thus, addressing the cause of disappearance of animal-derived *genes* from the human population will require comprehensive investigations, including whole-genome analysis of co-circulating strains detected in Vietnam and nearby counties, the vaccine effectiveness against animal-derived reassortants, and strain surveillance in animals and environment.

Supplementary data

Supplementary data are available at *VEVOLU* Journal online.

Acknowledgements

All samples used in this study were from the WHO-supported rotavirus surveillance in Vietnam. This research was supported by Japan Agency for Medical Research and Development (AMED) under Grant Numbers (JP20wm0125006, JP223fa627004, JP22wm0225023, and JP23fk0108663) (Japan Initiative for Global Research Network on Infectious Diseases (J-GRID)).

Conflict of interest: The authors have no conflict of interest to disclose.

References

Agbemabiese, C. A. et al. (2015) 'Whole Genomic Constellation of the First Human G8 Rotavirus Strain Detected in Japan', *Infection Genetics & Evolution*, 35: 184–93.

Agbemabiese, C. A. et al. (2019) 'Sub-genotype Phylogeny of the non-G, non-P Genes of Genotype 2 Rotavirus A Strains', *PLoS One*, 14: e0217422.

Bouckaert, R. et al. (2014) 'BEAST 2: A Software Platform for Bayesian Evolutionary Analysis', *PLoS Computational Biology*, 10: e1003537.

Burnett, E. et al. (2020) 'Global Impact of Rotavirus Vaccination on Diarrhea Hospitalizations and Deaths among Children <5 Years Old: 2006-2019', *Journal of Infectious Diseases*, 222: 1731–9.

Chan-It, W. et al. (2023) 'Predominance of DS-1-like G8P[8] Rotavirus Reassortant Strains in Children Hospitalized with Acute Gastroenteritis in Thailand, 2018-2020', *Journal of Medical Virology*, 95: e28870.

Cowley, D. et al. (2016) 'Emergence of a novel equine-like G3P[8] inter-genogroup reassortant rotavirus strain associated with gastroenteritis in Australian children', *Journal of General Virology*, 97: 403–10.

Crawford, S. et al. (2017) 'Rotavirus Infection', *Nature Reviews Disease Primers*, 3: 17083.

Degiuseppe, J. I. et al. (2022) 'Phylogeography of Rotavirus G8P[8] Detected in Argentina: Evidence of Transpacific Dissemination', *Viruses*, 14: 2223.

Do, L. P. et al. (2016) 'Molecular Characterisation of Wildtype G1P[8] and G3P[8] Rotaviruses Isolated in Vietnam 2008 during a Vaccine Trial', *Archives of Virology*, 161: 833–50.

Do, L. P. et al. (2017) 'Molecular Epidemiology of Rotavirus A, Causing Acute Gastroenteritis Hospitalizations among Children in Nha Trang, Vietnam, 2007-2008: Identification of Rare G9P[19] and G10P[14] Strains', *Journal of Medical Virology*, 89: 621–31.

Doan, Y. H. et al. (2013) 'Identification by Full-Genome Analysis of a Bovine Rotavirus Transmitted Directly to and Causing Diarrhea in a Human Child', *Journal of Clinical Microbiology*, 51: 182–9.

Doan, Y. H. et al. (2015) 'Changes in the Distribution of Lineage Constellations of G2P[4] Rotavirus A Strains Detected in Japan over 32 years (1980–2011)', *Infection Genetics & Evolution*, 34: 423–33.

Drummond, A. J. et al. (2012) 'Bayesian Phylogenetics with BEAUti and the BEAST 1.7', *Molecular Biology and Evolution*, 29: 1969–73.

Esona, M. D. et al. (2015) 'Multiplexed One-Step RT-PCR VP7 and VP4 Genotyping Assays for Rotaviruses Using Updated Primers', *Journal of Medical Virology*, 223: 96–104.

Fujii, Y. et al. (2019) 'Study of Complete Genome Sequences of Rotavirus A Epidemics and Evolution in Japan in 2012–2014', *Frontiers in Microbiology*, 10: 38.

Fujii, Y. et al. (2020) 'Molecular Characteristics of Novel Mono-reassortant G9P[8] Rotavirus A Strains Possessing the NSP4 Gene of the E2 Genotype Detected in Tokyo, Japan', *Japanese Journal of Infectious Diseases*, 73: 26–35.

Gauchan, P. et al. (2013) 'Continued Circulation of G12P[6] Rotaviruses over 28 Months in Nepal: Successive Replacement of Predominant Strains', *Tropical Medicine and Health.*, 41: 7–12.

Gentsch, J. R. et al. (2005) 'Serotype Diversity and Reassortment between Human and Animal Rotavirus Strains: Implications for Rotavirus Vaccine Programs', *The Journal of Infectious Diseases*, 192: S146–59.

Heiman, E. M. et al. (2008) 'Group A Human Rotavirus Genomics: Evidence that Gene Constellations are Influenced by Viral Protein Interactions', *Journal of Virology*, 82: 11106–16.

Helfrich, P. et al. (2018) 'TreeAnnotator: Versatile Visual Annotation of Hierarchical Text Relations', *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, France: European Language Resources Association.

Hermann, L. M. N. et al. (2024) 'Epidemiology of Rotavirus in Humans, Animals, and the Environment in Africa: A Systematic Review and Meta-analysis', *The Journal of Infectious Diseases* 229: 1470–80.

Hoa-Tran, T. N. et al. (2016) 'Abrupt Emergence and Predominance in Vietnam of Rotavirus A Strains Possessing A Bovine-like G8 on A DS-1-like Background', *Archives of Virology*, 161: 479–82.

Hoa-Tran, T. N. et al. (2019) 'Whole Genome Characterization of Feline-like G3P[8] Reassortant Rotavirus A Strains Bearing the DS-1-like Backbone Genes Detected in Vietnam, 2016', *Infection Genetics & Evolution*, 73: 1–6.

Hoxie, I. et al. (2020) 'Intragenic Recombination Influences Rotavirus Diversity and Evolution', *Virus Evolution*, 6: vez059.

Huyen, D. T. T. et al. (2018) 'Epidemiology of Acute Diarrhea Caused by Rotavirus in Sentinel Surveillance Sites of Vietnam, 2012-2015', *Vaccine*, 36: 7894–900.

John, T. J. et al. (2000) 'Herd Immunity and Herd Effect: New Insights and Definitions', *European Journal of Epidemiology*, 16: 601–6.

Justino, M. C. et al. (2011) 'Effectiveness of the Monovalent G1P[8] Human Rotavirus Vaccine against Hospitalization for Severe G2P[4] Rotavirus Gastroenteritis in Belém, Brazil', *The Pediatric Infectious Disease Journal*, 30: 396–401.

Kadoya, S. S. et al. (2022) 'The Intrapopulation Genetic Diversity of RNA Virus May Influence the Sensitivity of Chlorine Disinfection', *Frontiers in Microbiology*, 13: 839513.

Kirkwood, C. D. (2010) 'Genetic and Antigenic Diversity of Human Rotaviruses: Potential Impact on Vaccination Programs', *The Journal of Infectious Diseases*, 202: S43–S48.

Kondo, K. et al. (2017) 'Clinical and Molecular Characteristics of Human Rotavirus G8P[8] Outbreak Strain, Japan, 2014', *Emerging Infectious Diseases*, 23: 968–72.

Kosakovsky, P. S. L. et al. (2005) 'Not so Different after All: A Comparison of Methods for Detecting Amino Acid Sites under Selection', *Molecular Biology and Evolution*, 22: 1208–22.

Kumar, S. et al. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 For Bigger Datasets', *Molecular Biology and Evolution*, 33: 1870–4.

Maes, P. et al. (2009) 'RotaC: A Web-based Tool for the Complete Genome Classification of Group A Rotaviruses', *BMC Microbiology*, 9: 238.

Malik, Y. S. et al. (2020) 'Evolving Rotaviruses, Interspecies Transmission and Zoonoses', *The Open Virology Journal*, 14: 1–6.

Martella, V. et al. (2010) 'Zoonotic Aspects of Rotaviruses', *Veterinary Microbiology*, 140: 246–55.

Martin, D. P. et al. (2015) 'RDP4: Detection and Analysis of Recombination Patterns in Virus Genomes', *Virus Evolution*, 1: vev003.

Matthijnssens, J. et al. (2008) '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*, 82: 3204–19.

Matthijnssens, J. et al. (2022) 'ICTV Virus Taxonomy Profile: Sedoreoviridae', *Journal of General Virology*, 103: 001782.

McDonald, S. M. et al. (2016) 'Reassortment in Segmented RNA Viruses: Mechanisms and Outcomes', *Nature Reviews, Microbiology*, 14: 448–60.

Mukherjee., A. et al. (2013) 'First Report of Human Rotavirus G8P[4] Gastroenteritis in India: Evidence of Ruminants-to-human Zoonotic Transmission', *Journal of Medical Virology*, 85: 537–45.

Nakagomi, T. et al. (2013) 'G8 Rotaviruses with Conserved Genotype Constellations Detected in Malawi over 10 Years (1997-2007) Display Frequent Gene Reassortment among Strains Co-circulating in Humans', *Journal of General Virology*, 94: 1273–95.

Omatola, C. A. et al. (2022a) 'Rotaviruses: From Pathogenesis to Disease Control-A Critical Review', *Viruses*, 14: 875.

Omatola, C. A. et al. (2022b) 'Genetic Heterogeneity of Group A Rotaviruses: A Review of the Evolutionary Dynamics and Implication on Vaccination', *Expert Review of Anti-infective Therapy*, 20: 1587–602.

Phan, T. et al. (2022) 'Whole Genome Sequencing and Evolutionary Analysis of G8P[8] Rotaviruses Emerging in Japan', *VirusDisease*, 33: 215–8.

Pollard, S. L. (2015) 'Estimating the Herd Immunity Effect of Rotavirus Vaccine', *Vaccine*, 33: 3795–800.

Sergei, L. K. P. et al. (2020) 'HyPhy 2.5—A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies', *Molecular Biology and Evolution*, 37: 295–9.

Simmonds, M. K. et al. (2008) 'New Oligonucleotide Primers for P-typing of Rotavirus Strains: Strategies for Typing Previously Untypeable Strains', *Journal of Clinical Virology*, 42: 368–73.

Tacharoenmuang, R. et al. (2016) 'Full Genome Characterization of Novel DS-1-like G8P[8] Rotavirus Strains that Have Emerged in Thailand: Reassortment of Bovine and Human Rotavirus Gene Segments in Emerging DS-1-like Intergenogroup Reassortant Strains', *PLoS One*, 11: e0165826.

Tate, J. E. et al. (2016) 'Global, Regional, and National Estimates of Rotavirus Mortality in Children <5 Years of Age, 2000-2013.', *Clinical Infectious Diseases*, 62: S96–S105.

Todd, S. et al. (2010) 'Rotavirus Strain Types Circulating in Africa: Review of Studies Published during 1997–2006', *The Journal of Infectious Diseases*, 202: S34–42.

Wang, S. J. et al. (2022) 'Genetic Characterization of Two G8P[8] Rotavirus Strains Isolated in Guangzhou, China, in 2020/21: Evidence of Genome Reassortment', *BMC Infectious Diseases*, 22: 579.

Warren, C. J. et al. (2019) 'How Host Genetics Dictates Successful Viral Zoonosis', *PLoS Biology*, 17: e3000217.

Weaver, S. et al. (2018) 'Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes', *Molecular Biology and Evolution*, 35: 773–7.

Webby, R. et al. (2004) 'Molecular Constraints to Interspecies Transmission of Viral Pathogens', *Nature Medicine*, 10: S77–S81.

World Health Organization (2009) *Manual of Rotavirus Detection and Characterization Methods* WHO/IVB/08.17 <https://iris.who.int/handle/10665/70122> accessed 11 Jun 2013.

Zhou, X. et al. (2023) 'Surveillance of Human Rotaviruses in Wuhan, China (2019–2022): Whole-Genome Analysis of Emerging DS-1-like G8P[8] Rotavirus', *International Journal of Molecular Sciences*, 24: 12189.

Virus Evolution, 2024, 10(1), veae045 , DOI: https://doi.org/10.1093/ve/veae045, Advance Access Publication 22 June 2024Research Article