Infection, Genetics and Evolution 45 (2016) 122–131

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research paper

# Genomic constellation and evolution of Ghanaian G2P[4] rotavirus strains from a global perspective

Chantal Ama Agbemabiese [a,b,c], Toyoko Nakagomi [b], Yen Hai Doan [b,1], Loan Phuong Do [b,2], Susan Damanka [c], George E. Armah [c], Osamu Nakagomi [a,b,*]

[a] Leading Program, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan
[b] Department of Hygiene and Molecular Epidemiology, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan
[c] Noguchi Memorial Institute for Medical Research, College of Health Sciences, University of Ghana, Legon, Ghana

ARTICLE INFO

*Article history:*
Received 24 May 2016
Received in revised form 3 July 2016
Accepted 20 August 2016
Available online 26 August 2016

*Keywords:*
Ghana
Rotavirus
G2P[4]
Whole genome evolution
NSP4
Reassortment

ABSTRACT

Understanding of the genetic diversity and evolution of *Rotavirus A* (RVA) strains, a common cause of severe diarrhoea in children, needs to be based on the analysis at the whole genome level in the vaccine era. This study sequenced the whole genomes of six representative G2P[4] strains detected in Ghana from 2008 to 2013, and analysed them phylogenetically with a global collection of G2P[4] strains and African non-G2P[4] DS-1-like strains. The genotype constellation of the study strains was G2-P[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2. Strains from the same season were highly identical across the whole genome while strains from different seasons were more divergent from each other. The VP7, VP4, VP2, NSP1, and NSP5 genes belonged to lineage IVa; the VP6, VP1, NSP2, and NSP3 genes belonged to lineage V, and all these genes evolved in the same fashion as the global strains. In the NSP4 gene, lineages V (2008) and X (2009) were replaced by VI (2012/2013) whereas in the VP3 gene, lineage V (2008/2009) was replaced by VII (2012/2013) and these replacements coincided with the vaccine introduction period (2012). The evolutionary rate of the NSP4 gene was $1.2 \times 10^{-3}$ substitutions/site/year and was rather comparable to that of the remaining 10 genes. The multiple NSP4 lineages were explained by intra-genotype reassortment with co-circulating African human DS-1-like strains bearing G2[6], G3P[6], G6[6] and G8. There was no explicit evidence of the contribution of animal RVA strains to the genome of the Ghanaian G2P[4] strains. In summary, this study revealed the dynamic evolution of the G2P[4] strains through intra-genotype reassortment events leading to African specific lineages such IX and X in the NSP4 gene. So far, there was no evidence of a recent direct involvement of animal RVA genes in the genome diversity of African G2P[4] strains.



## 1. Introduction

*Rotavirus A* (RVA) within genus *Rotavirus* and family *Reoviridae* causes acute gastroenteritis in children and the young of animals. The virus has a triple-layered capsid with a genome of 11 segments of double-stranded RNA encoding six structural viral proteins (VP1-VP4, VP6, VP7) and six non-structural proteins (NSP1-NSP6) (Estes and Greenberg, 2013). RVA strains are classified into G and P genotypes on the basis of the nucleotide sequence diversity of the two outermost capsids — VP7 and VP4, respectively, and 27 G-types and 37 P-types have been reported so far (Matthijnssens et al., 2008a; Trojnar et al., 2013). A complete genome based classification system classifies human RVA strains into the Wa-like (G1/G3/G4-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) and AU-1-like (G3-P[9]-I3-R3-C3-M3-A3-N3-T3-E3-H3) genotype constellations (Matthijnssens et al., 2008a; Matthijnssens et al., 2008b).

Many countries have introduced either of the two live attenuated rotavirus vaccines pre-qualified by the World Health Organisation (WHO): the pentavalent bovine-human reassortant vaccine RotaTeq™ (Merck & Co. Inc., USA) and the monovalent human rotavirus vaccine Rotarix (GlaxoSmithKline Inc., Belgium) into their national immunisation programmes after the WHO's recommendation (WHO, 2009). The global under-five-mortality due to rotavirus diarrhoea has

*Abbreviations:* RVA, *Rotavirus A*; I, **I**ntermediate capsid shell; R, **R**NA polymerase; C, **C**ore shell; M, RNA-capping **M**ethyltransferase; A, interferon **A**ntagonist; N, octameric **N**TPase; T, **T**ranslation regulation; E, **E**nterotoxin; H, p**H**osphoprotein; VP, viral protein; NSP, non-structural protein; MEGA, Molecular Evolutionary Genetics Analysis; BIC, Bayesian Information Criterion; BEAST, Bayesian Evolutionary Analysis Sampling Trees; tMRCA, time of most recent common ancestor; MCMC, Markov chain Monte Carlo.

* Corresponding author at: Department of Hygiene and Molecular Epidemiology, Graduate School of Biomedical Sciences, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan.
*E-mail address:* onakagom@nagasaki-u.ac.jp (O. Nakagomi).
[1] Present address: Department of Virology 2, National Institute of Infectious Diseases, Tokyo, Japan.
[2] Present address: Department of Virology, National Institute of Hygiene and Epidemiology, Hanoi Vietnam.

http://dx.doi.org/10.1016/j.meegid.2016.08.024
1567-1348/

since declined from 528,000 in 2000 to 215,000 in 2013, and four countries in Africa and Asia account for about half of the deaths (Tate et al., 2016). Ghana, one of the early rotavirus vaccine adopter countries in Africa has also recorded a substantial decline in hospitalisation due to severe diarrhoea (Enweronu-Laryea et al., 2014) after Rotarix introduction in May 2012. As many countries are introducing rotavirus vaccines into their national immunisation programmes (http://sites.path.org/rotavirusvaccine/country-introduction-maps-and-spreadsheet/), it bears key importance to define at the whole genome level the genetic diversity and evolution of such circulating strains as G2P[4] on which the effect of vaccine introduction has been investigated (Donato et al., 2014; Gomez et al., 2014).

At the whole genome level, the global G2P[4] strains detected after 2000 were observed to possess a distinct lineage constellation from those detected before 2000 (Doan et al., 2015; Giammanco et al., 2014). A phylogenetic framework established for global G2P[4] whole genomes suggested that they evolved in a stepwise fashion (Doan et al., 2015) whereas additional emergent lineages noted in the VP3 and NSP4 genes of some G2P[4] strains have had their host species origins debated (Dennis et al., 2014; Doan et al., 2015; Ghosh et al., 2011b; Giammanco et al., 2014).

The African continent is rather known for the presence of a diverse pool of rotavirus strains. The five globally common human RVA genotypes G1P[8], G2P[4], G3P[8], G4P[8] and G9P[8] (Banyai et al., 2012; Santos and Hoshino, 2005) accounted for about only 36.5% of circulating strains in Africa during the period from 1997 to 2006 (Todd et al., 2010). Unlike other parts of the world, uncommon genotypes such as G6, G8, and P[6] are frequently detected in Africa (Agbemabiese et al., 2015; Dennis et al., 2014; Heylen et al., 2014; Heylen et al., 2015; Nakagomi et al., 2013; Ndze et al., 2014; Nordgren et al., 2012a; Nordgren et al., 2012b; Nyaga et al., 2014). Currently, the uncommon genotypes together with G2P[6], G3P[6] and G9P[6] strains possess the DS-1-like backbone and account for about 40% of the nearly 500 fully or partially sequenced whole genomes of human RVA strains detected in Africa (http://www.ncbi.nlm.nih.gov/genomes/VirusVariation/Database/nph-select.cgi?cmd=database&taxid=28875) (Brister et al., 2014). Mixed infections with more than one G/P type and genetic background have also been commonly detected in children in Africa (Mwenda et al., 2010; Nyaga et al., 2015; Sanchez-Padilla et al., 2009; Seheri et al., 2014; Todd et al., 2010). In addition, interspecies transmission events are thought to occur frequently in Africa as some of the G6 and G8 strains detected in humans possessed genes some of which were of animal RVA origin (Ben Hadj Fredj et al., 2013; Dennis et al., 2014; Nordgren et al., 2012b).

In this study, we sequenced and examined the outer capsid genes of G2P[4] strains from Ghana detected around rotavirus vaccine introduction period (2008–2013) of which six representative strains were analysed for their whole genomes, intending to understand their genetic diversity and evolution in the context of global collection of G2P[4] strains. Taking into consideration the large numbers of co-circulating DS-1-like strains other than G2P[4] and the frequent mixed infections occurring on the African continent, this study further investigated the contribution of the commonly detected DS-1-like strains in Africa to the genetic diversity of the backbone of the G2P[4] strains.

## 2. Materials and methods

### 2.1. Rotavirus strains

Samples comprised a total of thirty-eight G2P[4] strains detected in children <5 years with acute gastroenteritis who sought medical care in sentinel hospitals in Ghana during the 2008–2009 (n = 16) (Enweronu-Laryea et al., 2013) and 2012–2013 (n = 22) rotavirus seasons (unpublished data). Six representative strains - four from 2008–2009 season and two from the 2012–2013 season from unvaccinated children (age ineligible for vaccination) were selected for further examination by whole genome sequencing and phylogenetic analysis.

### 2.2. Whole genome amplification and sequencing

Viral RNA was extracted from 140 μL of supernatant obtained from 10% stool suspension (w/v) using the QIAamp Viral RNA Mini Kit (Qiagen) according to the manufacturer's protocol. Using the SuperScriptTM III first-strand synthesis system for Reverse Transcription-PCR (Invitrogen), cDNA was generated according to the manufacturer's protocol. The VP7 and VP4 (VP8* portion) genes of all the thirty-eight strains were amplified followed by sequencing and preliminary phylogenetic analysis. Six representative strains were selected for whole genome investigation taking into consideration the year of detection, sampling location and lineage designation of their VP7and VP4 genes (Supplementary Fig. 1a and b). Briefly, genes were amplified by PCR using 2 μL of cDNA with gene specific primers (Supplementary Table 1) (Doan et al., 2012; Gentsch et al., 1992; Gouvea et al., 1990; Matthijnssens et al., 2008a) and GoTaq® Green Master Mix System (Promega). The PrimeSTAR GXL DNA Polymerase (Takara) was used together with primers by Fujii et al. (2012) to amplify portions of larger genes that could not be previously amplified.

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

### 2.3. Sequence and phylogenetic analyses

Sequence contigs of the individual genes were assembled for each strain using the SeqMan program in DNAstar Lasergene core suite software v11 (DNAstar Inc.) and genotyped using the RotaC v.2.0 automated online genotyping tool for Group A rotaviruses (Maes et al., 2009). For phylogenetic analysis, G2P[4] strains for which near full length open reading frames (ORF) of whole genome sequence data were available in the GenBank were included for comparison. Sequences included for each gene's dataset were retrieved from the NCBI website with the Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) using nucleotide sequences of the 11 genes of the prototype G2P[4] strain DS-1 as the query sequence. In addition, the following criteria were employed to justify the inclusion of other sequences for the phylogenetic comparisons. 1. Granted that a substantial proportion of G2 strains in Africa possess the P[6] genotype (Heylen et al., 2015; Mwenda et al., 2010; Seheri et al., 2014; Todd et al., 2010), the G2 VP7 gene of African G2P[6] strains with available whole genome sequence data as well as the oldest available African G2 VP7 sequences were included. 2. Non G/P genes of DS-1-like RVA strains from Africa were included in order to understand their contribution to the genetic backbone of African G2 strains and to address the origin of any unique lineages.

Multiple sequence alignment files were constructed using the online version of Multiple Alignment using Fast Fourier Transform (MAFFT version 7) (Katoh and Standley, 2013). Nucleotide and amino acid sequence identity matrices were calculated for each dataset for all genome segments using the p-distance algorithm in MEGA v6.06 (Tamura et al., 2013). Based on the best fit nucleotide substitution models with the lowest Bayesian Information Criterion scores (Schwarz, 1978), i.e. the Tamura 3-parameter model with gamma distribution and invariant sites (T92 + G + I) for VP7; the General Time Reversible model (GTR) + G + I for VP4, VP1, VP3; the Tamura-Nei model (TN93) + G + I for VP2 and T92 + G for VP6, NSP1, NSP2, NSP3, NSP4 and NSP5, maximum likelihood phylogenetic analysis was carried out with 1000 bootstrap replicates. Lineages in this study were defined as a collection of closely related sequences (with ≤5% mean diversity)