Interspecies transmission of rotaviruses and its evolutionary implication: a view from Africa

By

CHANTAL AMA AGBEMABIESE

"A dissertation submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy"

Department of Molecular Epidemiology

Graduate School of Biomedical Sciences

Nagasaki University

2014 - 2018

PhD Thesis Supervisor

Osamu Nakagomi, MD, PhD

Professor of Molecular Epidemiology

Graduate School of Biomedical Sciences

Nagasaki University

ACKNOWLEDGEMENTS

First, I would like to thank the leadership of the Program for Nurturing Global Leaders in Tropical and Emerging Communicable Diseases, Nagasaki University for the opportunity granted me to pursue my Doctoral studies and the immense financial support and guidance. I am very appreciative of the permission and support I received from the leadership of the Noguchi Memorial Institute for Medical Research, University of Ghana, Legon, which enabled me to pursue my PhD program.

With a joyful heart, I would like to sincerely thank my PhD supervisor - Professor Osamu Nakagomi (Head, Department of Molecular Epidemiology, Graduate School of Biomedical Sciences, Nagasaki University). Your acceptance, guidance, contributions, suggestions, support and deadlines have been valuable to my training and the successful completion of my PhD program. I would also like to express my profound gratitude to Dr. Toyoko Nakagomi (Associate Professor, retired; Department of Molecular Epidemiology, Graduate School of Biomedical Sciences, Nagasaki University) for the immense contributions, training, kindness and all the love she has shown me. I thank Professor Suzuki Yoshiyuki (Graduate School of natural Science, Nagoya University, Japan) for the useful guidance and fruitful discussions we had whenever Professor Nakagomi invited you to our department in Nagasaki University. I also thank Drs. Yen Hai Doa, Punita Gauchan, Miho Kaneko, Loan Phuong Do, Thi Nguyen Hoa-Tran, Hanae Takatsuki and Ms. Junko Hiroshige for their support.

I am also very grateful to Professor George Armah (Head, West African Regional Rotavirus Reference Laboratory, Noguchi Memorial Institute for Medical Research, University of Ghana, Legon), Dr. Michael Ofori, (Head, Department of Electron Microscopy and Histopathology, Noguchi Memorial Institute for Medical Research, University of Ghana, Legon) and Dr. Jonathan P. Adjimani (Senior lecturer, Department of Biochemistry, Cell and Molecular Biology, University of Ghana, Legon) for nurturing my growing career in the area

of rotavirus research; you have been outstanding mentors I will forever remain grateful to. A special thank you goes to the staff of the West African Regional Rotavirus Reference Laboratory, Dr. Susan Damanka, Dr. Francis Dennis, Ms. Belinda Lartey, Mr. Fred Asamoah and Mrs. Yorm Abedi-Lartey for their support.

Furthermore, my sincere appreciation goes to my overseas training mentors - Professor Jeevan Sherchand (Institute of Medicine, Tribhuvan University, Nepal) and Dr. John Thomas Patton (Department of Biology, Indiana University, USA) and their awesome staff, post-doctoral fellows and graduate students.

I would like to thank my loving family in a special way for their endless support and encouragement throughout this journey. To my friends, I say a big thank you; I appreciate all the support and the listening ears you gave to all my stories. God bless you all!

OUTLINE OF THESIS

Content	Page
Acknowledgements	iii
Outline of Thesis	v
Abbreviations	vi
Abstract	vii
Chapter I Introduction	1
Chapter II Evolution of a G6P[6] rotavirus strain isolated from a child with acute gastroenteritis in Ghana, 2012	11
Chapter III Genomic constellation and evolution of Ghanaian G2P[4] rotavirus strains from a global perspective	35
Chapter IV Whole genomic constellation of the first human G8 rotavirus strain detected in Japan	65
Chapter V Discussion, conclusion and recommendations	89
References	107
Supplementary materials	129

ABBREVIATIONS

A: interferon Antagonist
AICc: corrected Akaike Information Criterion
ARSN: African Rotavirus Surveillance Network
BEAST: Bayesian Evolutionary Analysis Sampling Trees
BIC: Bayesian Information Criterion
BLAST: Basic Alignment Search Tool
C: Core shell
CI: Confidence Interval
E: Enterotoxin
G: Gamma distribution
G. B. D.: Global Burden of Disease
GTR: General Time Reversible model
H: pHosphoprotein
HPD: Highest Posterior Density
I: Intermediate capsid shell
I: Invariant sites
M: RNA-capping Methyltransferase
MAFFT: Multiple Alignment using Fast Fourier Transform
MCMC: Markov chain Monte Carlo
MEGA: Molecular Evolutionary Genetics Analysis
N: octameric NTPase
NCBI: National Center for Biotechnology Information
NSP: non-structural protein
ORF: Open Reading Frame
R: RNA polymerase
RVA: *Rotavirus A*
T: Translation regulation
T92: Tamura 3-parameter nucleotide substitution model
tMRCA: time of most recent common ancestor
TN93: Tamura-Nei nucleotide substitution model
UI: Uncertainty Interval
VP: viral protein
WHO: the World Health Organisation

ABSTRACT

Rotavirus A (RVA) is a leading cause of diarrhoea and severe dehydration in children and many animal species worldwide. Thus, safe and efficacious, live-attenuated vaccines were developed based on the molecular epidemiology of rotavirus strains in developed countries and rolled out gradually in developing regions in the world including sub-Saharan African countries. The rotavirus genome is notoriously diverse and evolves through rapid point mutations, genetic reassortment and interspecies transmission. Rotavirus strains circulating in Africa are considerably different from the ones circulating elsewhere in the world in that, apart from the globally common strains, the prevalence of unusual strains such as G8, G6P[6] and P[6] strains among others is high. These unusual genotypes at a glance, are indicative of animal rotavirus origin. While there is always a vague speculation that frequent rotavirus interspecies transmission events occur in Africa because people and animals live in close proximity, precise studies making use of the tools of molecular epidemiology and molecular phylogeny to decipher the evolutionary history of the novel strains are limited.

To gain insight into how rotaviruses evolve in Africa with special emphasis on the role of interspecies transmission of animal rotaviruses in human rotavirus infection, I carried out three molecular epidemiology studies that are included in this thesis. In the first study (Chapter II), I showed that the G6 VP7 possessed by G6P[6] strains in Africa as well as Europe originated from a single ancestral VP7 from a human G6P[9] strain around the year 1998 and not directly from bovine G6 strains or bovine-like human G6P[14] strains. Also, it was discovered that the G6 VP7 gene after crossing the host species barrier from cattle to human in the distant past, underwent an accelerated evolutionary rate, a phenomenon which could constitute a post-transfer adaptation process in the new host. Genetic reassortment played a major role in the generation of the G6P[6] strain as there was not a single strain that provided the DS-1-like genetic backbone carried by the G6P[6] strains. It is also

hypothesised that the acquisition of a genetic backbone of already adapted regional circulating DS-1-like strains enabled the G6P[6] strains to establish a sustained transmission chain in the population.

In the second study (Chapter III), contrary to the general notion of frequent RVA interspecies transmission events occurring in Africa, it was noted that the genome of G2P[4] strains from Ghana - the potential donor strains of the DS-1-like backbones to many unusual strains in Africa including G6P[6] strains discussed above, evolved by utilizing a step-wise lineage replacement strategy similar to the pattern described for global G2P[4] strains by Doan et al. (2015). Of note was a frequent expansion of the E2 NSP4 gene at the sub-genotype level in African G2P[4] strains leading to African specific lineages such IX and X in the NSP4 gene. However, this diversity was explained by frequent intra-genotype reassortment events involving regional DS-1-like rotavirus strains such as G2P[6], G3P[6] and G6P[6] strains of human host species origin and not direct introduction of genotype 2 rotavirus genes from animal rotaviruses.

Third, the genome of a G8 strain, an epidemiologically important genotype on the African continent detected for the first time in Japan, was analysed to understand how it was generated, how it relates to G8 detected elsewhere in the world, and to determine the host species origin of its genes. Until recently in 2014, this G8 rotavirus of human host species origin was the only one reported in Japan although infection with G8 strains was a common phenomenon in children on the African continent. This strain was concluded to have been generated by genetic reassortment where co-circulating G2P[4] strains in Japan obtained the VP7, VP1 and NSP2 genes from unknown ruminant G8 RVA strains. Although this strain was detected on a different continent from Africa, the genetic composition and the origin of the genes reflect an attempt of an animal strain to establish itself in the human population by acquiring the genetic backbone of DS-1-like strains believed to be already adapted to the human population – a similar mechanism utilised by the G6P[6] strains in Africa in

establishing a human-to-human transmission chain. On account of the lack of subsequent detection of this strain in the human population in Japan, I interpreted that this strain was an example of failures encountered by some animal rotaviruses in establishing a human to human transmission chain in the population after they have crossed the host species barrier.

In Chapter V, I aimed to explore the major observations made in the preceding chapters (Chapters II-IV) to understand the specific features of the circulating rotavirus strains on the African continent and discussed the role played by prevalent P[6] VP4 genes in reference to the abundance of the Lewis-negative phenotype in Africa. Most importantly, however, what appeared to be African specific G8 VP7 lineages were divided at least into two lineages, namely: the Cameroonian and Malawian lineages, and while their origin was of bovine, after crossing the host species barrier, they seemed to have been transmitted only from human to human which was made possible by the acquisition of either the human RVA Wa-like or DS-1-like genetic backbone. Those G8 strains that gained the Wa-like genetic backbone seem to have died out from Africa after prevailing for some time on the continent. Also noted were the ever-diversifying NSP4 lineages within the E2 genotype which were mostly due to the introduction of the NSP4 sequences of animal rotavirus origin; these lineages were however short-lived with limited geographical distribution.

In conclusion, I postulated a hypothesis that while proximity of people and animals in Africa provides abundant opportunities for animal rotaviruses to cross species barriers into humans, many of such events terminate as dead-end infections without establishing a human to human transmission chain and only a few interspecies transmission events do so after gaining human rotavirus backbone genes through genetic reassortment. Even in such successfully established interspecies transmission cases, the lifespan of such novel lineages within human rotavirus is rather short and limited geographically as they might have been out-competed by the co-circulating parental strains. Nevertheless, such interspecies

transmission events coupled with genetic reassortment provide the source of rich genetic diversity, whether transient or permanent, in African rotavirus strains we observe today.

Chapter I

Introduction

1.1 ROTAVIRUS DISEASE BURDEN

Diarrhoeal diseases remain one of the leading causes of morbidity and mortality in children under the age of five years especially in low socio-economically developing countries. *Rotavirus A (RVA)*, a leading cause of acute gastroenteritis in infants, young children, and the young of many animal species, is a species within the genus *Rotavirus* and family *Reoviridae* (Estes and Greenberg, 2013). Despite the availability of safe and effective rotavirus vaccines such as Rotarix by GlaxoSmithKline Biologicals (Ruiz-Palacios et al., 2006) and Rotateq by Merck & Co. Inc. (Vesikari et al., 2006), current global estimates in children under 5 years revealed that of the top three aetiologies to which diarrhoea mortality is attributed, rotavirus is in the lead with an average of 146,000 deaths (95% uncertainty interval (UI) 118,000-183,000; 29.3%, 24.6 – 35.9%), followed by *Cryptosporidium* spp (60,400 deaths, 95% UI 13,709.1 – 134,506; 12.1%, 2.8-26.9%) and *Shigella* spp (54,900 deaths, 95% UI 27,000-94,700; 11.0%, 5.5-18.7%) (Global Burden of Disease Diarrhoeal Disease Collaborators, 2017).

The World Health Organisation in 2009 recommended the routine use of rotavirus vaccines especially in countries which experienced high mortality rates due to rotavirus diarrhoea - most of which were in Africa and South Asia (WHO, 2009). Notably, four countries accounted for approximately half (49%) of rotavirus associated deaths in children under five years in 2013. These included India, Nigeria, Pakistan and Democratic Republic of Congo; two of which were from the African continent (Tate et al., 2016). In addition, RVA accounts for a median of 39.4% (95% CI: 37.1 – 43.1%) of diarrhoeal hospitalisations (Lanata et al., 2013). With such a high rotavirus disease burden, PATH has adopted a comprehensive approach which includes efforts to increase access to as well as develop new rotavirus vaccines. In this regard, as of March 2017, 92 countries have introduced rotavirus vaccines (http://rotacouncil.org/vaccine-introduction/global-introduction-status/,

retrieved on October 31st, 2017) and these include 85 countries with a national coverage, two ongoing phased, and five sub-national introductions.

1.2 ROTAVIRUS STRUCTURE, GENOME ORGANISATION AND CLASSIFICATION

The rotavirus virion has a triple-layered capsid (Fig. 1.1) which encloses a genome (approximately 18.5 kb in size) of 11 segments of double-stranded RNA (Table 1.1). The genome encodes six structural viral proteins (VP1-VP4, VP6, VP7) and six non-structural proteins (NSP1-NSP6) (Fig. 1.1) (Estes and Greenberg, 2013). The structural proteins constitute the rotavirus virion whereas the non-structural proteins are produced to play diverse roles (Table 1.2) during the complex virus replication cycle that is orchestrated by an interplay between the rotavirus structural and non-structural proteins.

Rotaviruses are classified into groups A to G based on the antigenic determinants of the major capsid protein VP6 (Estes and Greenberg, 2013) and recently, Groups H and I were discovered based on VP6 sequence analysis (Kindler et al., 2013; Matthijnssens et al., 2012; Mihalov-Kovacs et al., 2015). Groups A, B, C, and H rotaviruses infect humans, however, of the four groups that infect humans, group A rotaviruses have been established as the single most important cause of severe acute gastroenteritis in infants and young children in both developed and developing countries (Estes and Greenberg, 2013; Santos and Hoshino, 2005). RVA strains are further classified into G and P genotypes based on the nucleotide sequence diversity of the two outermost capsid proteins VP7 and VP4, respectively. In addition to previous reports (Matthijnssens et al., 2008a; Matthijnssens et al., 2008b; Trojnar et al., 2013) and the latest update from the Rotavirus Classification Working Group (https://rega.kuleuven.be/cev/viralmetagenomics/virus-classification/rcwg) (April, 2017), at least 50 P-genotypes and 35 G genotypes have been identified in humans and animals. Of these, five G/P type combinations namely G1P[8], G2P[4], G3P[8], G4P[8] and G9P[8] are globally commonly detected in humans (Banyai et al., 2012; Gentsch et

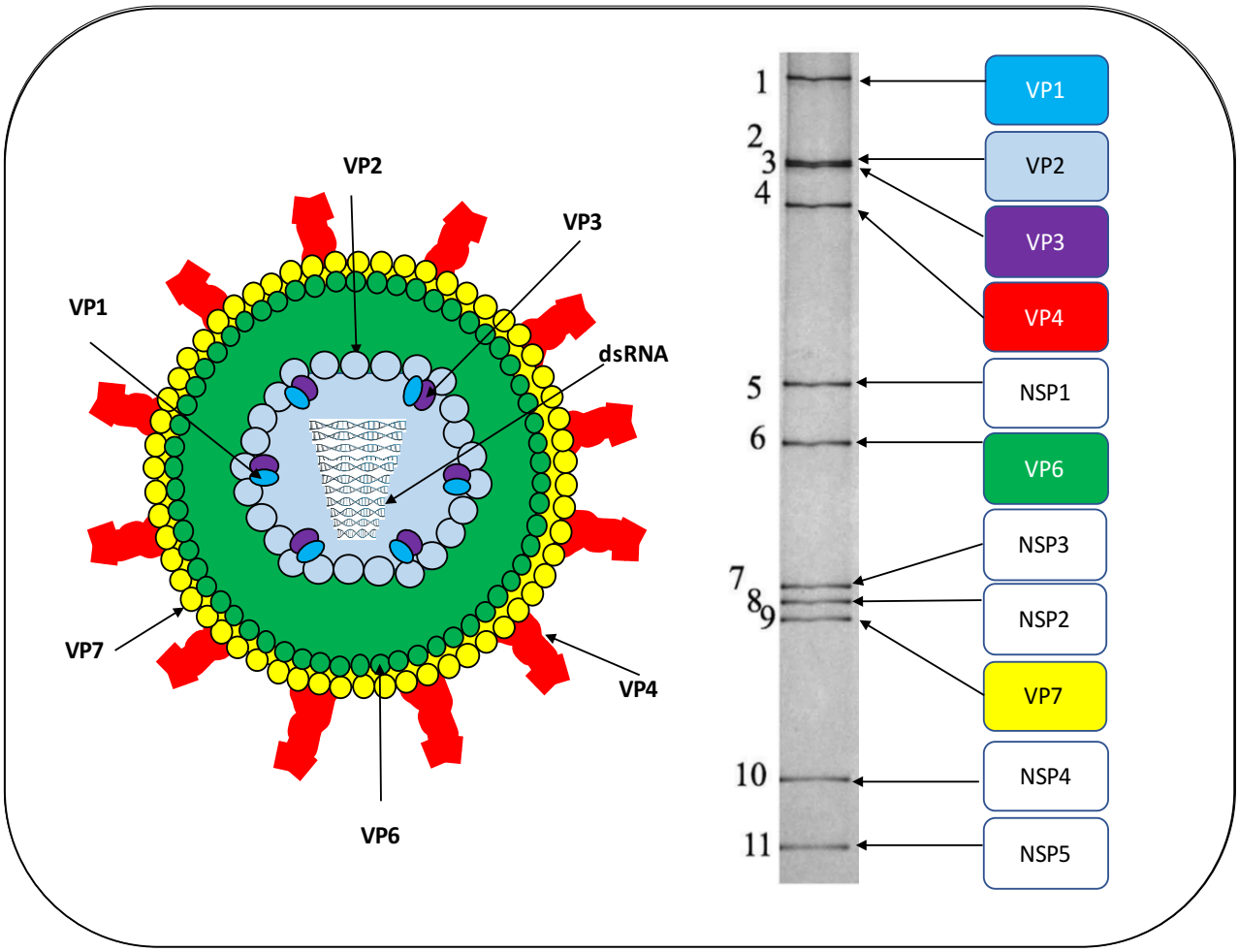


Fig. 1.1: Rotavirus virion structure and genome organization.

A schematic representation of the rotavirus structure. Locations of various structural proteins are shown. Also shown is the electrophoretic migration pattern of the 11 genome segments of double stranded RNA (dsRNA) of the Wa RVA strain on polyacrylamide gel.

Table 1.1: Rotavirus genome segment sizes, proteins encoded and abundance per virion

Genome segment	Size (bp)	Protein encoded	Open Reading Frame (nucleotide number)	Size of protein encoded	Number of molecules per virion
1	3302	VP1	18-3282	1088 aa (125.0 kDa)	12
2	2690	VP2	17-2659	881 aa (102.4 kDa)	120
3	2591	VP3	50-2554	835 aa (98.1 kDa)	12
4	2362	VP4	10-2337	776 aa (86.7 kDa)	120
		VP8* portion		247 aa (1-247) (28 kDa)	
		VP5* portion		529 aa (247-776) (60 kDa)	
5	1611	NSP1	31-1515	495 aa (58.6 kDa)	NA
6	1356	VP6	24-1214	397 aa (48.1 kDa)	780
7	1105	NSP3	26-970	315 aa (34.6 kDa)	NA
8	1059	NSP2	47-997	317 aa (36.7 kDa)	NA
9	1062	VP7	49-1026	326 aa (37.3 kDa)	780
10	751	NSP4	41-569	175 aa (20.3 kDa)	NA
11	667	NSP5	22-615	198 aa (21.7 kDa)	NA
11	667	NSP6	80-355	92 aa (11.0 kDa)	NA

Modified from: http://www.reoviridae.org/dsrna_virus_proteins/Rotavirus.htm

(The RNAs and Proteins of dsRNA Viruses: Edited by Peter. P. C. Mertens and Dennis H. Bamford)

Based on strain RVA/Simian-tc/ZAF/SA11/1958/G3P[2]. *GenBank accession numbers: VP1: X16830; VP2: X16831; VP3: X16062; VP4: X14204; VP6: L15384; NSP1: X14914; NSP3: M87502; NSP2: L04531; VP7: K02028; NSP4: AF087678; NSP5, NSP6: X07831*

al., 2009; Santos and Hoshino, 2005). In recent years, G12 rotaviruses emerged globally as one of the important causes of RVA diarrhoea in children (Castello et al., 2006; Cunliffe et al., 2009; Matthijnssens et al., 2010; Page et al., 2009; Pun et al., 2007; Rahman et al., 2007; Uchida et al., 2006). Also, the G8 genotype which was mostly detected on the African continent than elsewhere in the world (Cunliffe et al., 2000; Dennis et al., 2014; Esona et al., 2009; Heylen et al., 2014; Heylen et al., 2015; Istrate et al., 2015; Nakagomi et al., 2013; Steele et al., 2002; Steele et al., 1999) seems to be gaining grounds in its emergence, persistence and spread in some Asian countries too (Hoa-Tran et al., 2016; Kondo et al., 2017; Tacharoenmuang et al., 2016).

Based on RNA-RNA hybridisation, human rotaviruses were previously classified into three genogroups namely the Wa, DS-1 and AU-1 genogroups (Nakagomi et al., 1989; Nakagomi and Nakagomi, 1989). In line with this, the dual classification system of RVA strains was extended to include the other nine genome segments and based on pre-defined nucleotide and amino acid cut-off values (Table 1.2). The whole genome VP7-VP4-VP6-VP1-VP2-VP3-NSP1-NSP2-NSP3-NSP4-NSP5/6 of RVA strains is therefore respectively denoted by the descriptor Gx-P[x]-Ix-Rx-Cx-Mx-Ax-Nx-Tx-Ex-Hx where x represents the genotype number (Matthijnssens et al., 2008a; Matthijnssens et al., 2011; Matthijnssens et al., 2008b). As such, most human RVA strains can be classified into two major and one minor genotype constellations - the Wa-like, DS-1-like and AU-1-like genotype constellations, which are described as G1/3/4/9-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1, G2-P[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2, and G3-P[9]-I3-R3-C3-M3-A3-N3-T3-E3-H3, respectively (Matthijnssens et al., 2008a; Matthijnssens et al., 2011; Matthijnssens et al., 2008b).

The whole genome classification system further revealed that human Wa-like strains and porcine RVA strains share a common evolutionary origin whereas the human DS-1-

Table 1.2: Pre-defined nucleotide and amino acid identity cut-off values for rotavirus whole genome classification

Gene product	Percentage nucleotide (amino acid) identity cut-off value	Genotype nomenclature (number of genotypes)	Description of gene product and function
VP7	80 (89)	**G** (G1-35)	**G**lycoprotein
VP4	80 (89)	**P** (P[1]-P[50])	**P**rotease sensitive protein
VP6	85	**I** (I1-26)	**I**ntermediate capsid
VP1	83	**R** (R1-21)	**R**NA-dependent RNA polymerase
VP2	84	**C** (C1-19)	**C**ore shell
VP3	81	**M** (M1-19)	**M**ethyltransferase
NSP1	79	**A** (A1-30)	Interferon **A**ntagonist
NSP2	85	**N** (N1-20)	**N**TPase
NSP3	85	**T** (T1-21)	**T**ranslation enhancer
NSP4	85	**E** (E1-26)	**E**nterotoxin
NSP5	91	**H** (H1-21)	p**H**osphoprotein

Adapted from Matthijnssens et al., 2008a; updated with information from the Rotavirus Classification Working Group website: (https://rega.kuleuven.be/cev/viralmetagenomics/virus-classification/rcwg) (April, 2017)

like strains and bovine RVA strains share a common evolutionary origin (Matthijnssens et al., 2008a). Porcine rotaviruses usually possess G3, G4, G9 and G11 in association with P[6] or P[7] whereas G1, G2, G6, G10, G12 and G26 in combination with P[5], P[8], P[11], P[13], P[14], P[19], P[26], P[27] and P[32] are sporadically detected (Papp et al., 2013; Silva et al., 2015; Silva et al., 2016; Theuns et al., 2015).

At the whole genome level, porcine RVA strains typically possess the genotype constellation G3/4/5/9/11-P[6]/[7]/[13]/[19]/[23]-I5-R1-C1-M1-A8-N1-T1/7-E1-H1 (Kim et al., 2012; Martel-Paradis et al., 2013; Matthijnssens et al., 2008a; Monini et al., 2014; Silva et al., 2016; Theuns et al., 2015). A recent comprehensive phylogenetic analysis of the whole genome sequences of genotype 1 genes of RVA strains revealed that, typical modern human Wa-like strains belonged to a separate cluster from that of typical modern porcine RVA strains (Silva et al., 2016). On the other hand, bovine rotaviruses usually possess G6, G8, and G10 in association with P[1], P[5], and P[11] although G1-G3, G5, and G11 in association with P[3], P[6], P[7], and P[14]; G15, G17, G21 and G24 in association with P21], P[29] and P[33] have also been detected in sporadic cases (Matthijnssens et al., 2011; Papp et al., 2013).

The rotavirus genome is incredibly diverse and basically five mechanisms namely: point mutation, genetic reassortment, genetic rearrangement, genetic recombination, and interspecies transmission have been reported to contribute to their genome evolution. Genetic reassortment of rotavirus genes can occur between rotaviruses from the same or different host species as well as between rotaviruses of the same genotype (intra-genotype) or different genotypes (inter-genotype) during co-infection of host cells leading to the generation of novel strains in the human population. In addition, direct interspecies transmission of rotaviruses between multiple host species has also been shown to play a major role in the genome evolution of circulating rotavirus strains.

Given the segmented nature of the rotavirus genome, the natural history of rotavirus infection and the close relationship of cattle and pigs with humans, it is reasonable that porcine and bovine RVA strains serve as a large potential gene pool for the generation of novel human RVA strains. In line with the massive efforts being made to curb the high rotavirus morbidity and mortality in children, molecular epidemiological studies that shed light on the whole genome evolution of rotavirus strains of both human and animal host species origin are vital in understanding the constantly changing landscape of rotavirus strains.

1.3 ROTAVIRUS MOLECULAR EPIDEMIOLOGY IN AFRICA

In Africa, rotavirus detection rates among children with acute diarrhoea ranges between 20% and 63% (Benhafid et al., 2009; Cunliffe et al., 1998; de Villiers et al., 2009; Enweronu-Laryea et al., 2012; Fischer et al., 2010; Mwenda et al., 2010). Despite the general perception that improved sanitation has little effect on the prevalence of rotavirus infection (Parashar et al., 2009; Rodrigues et al., 2007), country specific mortality rates vary more than 10-fold from 32 in China to 300 in Niger, Angola and Afghanistan (Naghipour et al., 2008). Six of the seven countries with the highest mortality from rotavirus (>500 deaths per 100,000 live births) are in sub-Saharan Africa (Parashar et al., 2009).

Prior to the period before rotavirus vaccine was rolled out in many African countries, Sanchez-Padilla et al. (2009) estimated the annual rotavirus mortality rate in children aged below five years in sub-Saharan Africa to be approximately 243/100,000. While this figure translates into 308,579 deaths per year within this age group, the mortality rate varied from country to country, ranging from 6.2 (South Africa) to 301 per 100,000 child-years. These deaths are preventable in principle but current issues regarding vaccine efficacy and coverage in African countries greatly influence whether this goal is achievable or not.

Regarding seasonality, rotavirus infection occurs all year round in sub-Saharan Africa but seasonal epidemics occur in the cool dry seasons whereas infection peaks occur during the winter and early spring in the temperate regions. Thus, it will be interesting to investigate the evolutionary dynamics of rotavirus strains where the rotavirus circulation is year-round to help notice if any, when novel virus variants are introduced into the population.

Knowledge on the age distribution of rotavirus infection in children is important in devising an effective strategy of vaccination schedules. In a rotavirus surveillance study in 11 African countries, majority of rotavirus infections (~90%) occurred in children aged 3-18 months (Mwenda et al., 2010) and this age range did not vary by country. In summary, the age at which children get vaccinated with rotavirus vaccines is critical and needs to be tailored accordingly to suit each population.

Keeping in mind the general background information about rotavirus infection in Africa within a global context as delineated above, I wrote the following three chapters related to the circulating rotavirus strains on the African continent based on my studies that were peer-reviewed and published in academic journals in order to provide the scientific basis on which I developed my thoughts and hypotheses on the evolution of rotavirus genome with special emphasis on the role of interspecies transmission of animal rotaviruses to humans.

Briefly, first, the whole genome of a rare G6P[6] RVA strain detected in a child with diarrhoea in Ghana was examined for its full genome. Second, the whole genome constellation and evolutionary history of the globally common G2P[4] strains detected in Ghana during the 2008-2013 rotavirus seasons were examined. The third study examined the whole genome evolution of a G8 rotavirus strain detected in Japan to understand how it was generated and how it was related to the G8 strains detected elsewhere in the world.

Chapter II

Evolution of a G6P[6] rotavirus strain isolated from a child with acute gastroenteritis in Ghana, 2012

Published in:
Chantal Ama Agbemabiese, Toyoko Nakagomi, Yoshiyuki Suzuki, George Armah, and Osamu Nakagomi. ***Journal of General Virology*** **96**, 2219–2231 **(2015)**

2.1 SUMMARY

Unusual human G6P[6] rotavirus strains were reported sporadically in Europe and Africa, but how they evolved was not fully understood. The whole genome of a Ghanaian G6P[6] strain, designated PML1965, was analysed to understand how it evolved in Africa and to know how its G6 VP7 gene was related to that of rotaviruses of human and artiodactyl origin. The genotype constellation of RVA/Human-wt/GHA/PML1965/2012/G6P[6] was G6-P-[6]-I2-R2-C2-M2-A2-N2-T2-E2-H2. It shared lineages with G6P[6] strains previously detected in Italy and Africa in all genome segments except the VP6 gene of a few Burkinabe and Cameroonian strains and both the VP6 and NSP4 genes of Guinea Bissau strains. The VP7 gene of the G6P[6] strains appeared to have been derived from those of human G6P[9] strains, and they were distantly related to the VP7 genes of artiodactyl G6 or human G6P[14] strains. The time of the most recent common ancestor of the VP7 sequences of G6P[6] strains was estimated to be the year 1998 meaning that the common ancestral sequence from which the current G6 VP7 sequences directly emerged existed in 1998 at the latest. The evolutionary rate of the VP7 genes in bovine and human G6 rotaviruses were 6.93 x 10^{-4} and 3.42 x 10^{-3} nucleotide substitutions/site/year, respectively, suggesting an accelerated adaptive process in the new host. The sequences of the remaining 10 genome segments of PML1965 clustered with those of G2 and G8 human rotaviruses detected in Africa possessing the DS-1-like genetic background. In conclusion, PML1965 evolved by G2 or G8 RVA strains with DS-1-like background acquiring the G6 VP7 gene from a human G6P[9] RVA and not from an artiodactyl G6 RVA strain.

Key words: rotavirus; genotype constellation; G6P[6]; phylogenetic analysis; Bayesian analysis; evolutionary rate

2.2 INTRODUCTION

Over the past few years, the African Rotavirus Surveillance Network (ARSN), co-ordinated by the World Health Organisation intensified its surveillance activities and expanded its sentinel sites from 4 to 34 sentinel sites located in 20 African countries (Mwenda et al., 2014). As more surveillance studies are conducted, the chance of detecting human rotaviruses possessing unusual combinations of G and P genotypes increases. A great variability in circulating RVA strains in children on the African continent has been observed from year to year and from region to region (Ouermi et al., 2017; Sanchez-Padilla et al., 2009; Seheri et al., 2014; Todd et al., 2010).

Human rotaviruses of uncommon G and P type combinations are largely classified into two categories; one comprises strains suggestive of reassortants between the Wa-like and the DS-l-like genotype constellations (Ghosh and Kobayashi, 2011, 2014; Iturriza-Gomara et al., 2001; Matthijnssens and Van Ranst, 2012) such as G1P[6] (Ghosh et al., 2013), G1P[4] (Sasaki et al., 2015), and G3P[4] (Hoa Tran et al., 2013). The other comprises rotavirus strains possessing either G or P genotype suggestive of animal rotavirus origin (Ghosh and Kobayashi, 2011, 2014; Matthijnssens and Van Ranst, 2012; Steyer et al., 2008) such as G3P[9] of probable feline rotavirus origin (Nakagomi and Nakagomi, 1989), G4P[6] of probable porcine rotavirus origin (Martinez et al., 2014), G5P[6] (Ahmed et al., 2007), G6P[1] (Doan et al., 2013), G6P[11] (Steyer et al., 2013), G6P[14] (Cooney et al., 2001) and G8P[1] (Adah et al., 2001) of probable bovine rotavirus origin.

Of the uncommon human rotaviruses, there are 35 G6P[6] strains described in the literature and the GenBank database, and they may outnumber others in the frequency of detection. The G6P[6] strain was for the first time detected in Belgium in a child returning from vacation in Mali (Matthijnssens et al., 2008c), and subsequently in Italy (Ianiro et al., 2013) and Africa (Ndze et al., 2014; Nordgren et al., 2012a; Nordgren et al., 2012b). Two distinct hypotheses were proposed to explain the evolutionary process by which these

G6P[6] strains emerged. Whereas the G6P[6] strain detected in Italy was reported to lack any evidence of zoonotic transmission and linked to interspecies reassortment (Ianiro et al., 2013), G6P[6] strains detected in Belgium and Burkina Faso were linked to interspecies transmission from cattle to humans (Matthijnssens et al., 2008c; Nordgren et al., 2012b). As we had an opportunity to analyse a G6P[6] strain, designated PML1965, detected in Ghana during the 2012 rotavirus surveillance period (Enweronu-Laryea et al., 2014), we carried out a whole genome sequencing analysis of PML1965 in order to gain clues regarding the evolutionary process by which such G6P[6] strains emerged in Africa. To obtain further insight into the adaptation process after jumping into a new host, we carried out a Bayesian evolutionary analysis to determine the evolutionary rates of the G6 VP7 genes possessed by human and bovine rotaviruses.

2.3 MATERIALS AND METHODS

Rotavirus strain

Rotavirus G6P[6] strain PML1965 was detected in an 11 month old male child hospitalised for acute gastroenteritis during the 2012 rotavirus surveillance period in Ghana (Enweronu-Laryea et al., 2014).

Whole Genome Amplification and Sequencing

Viral RNA was extracted from 10% (w/v) stool suspension using the QIAamp Viral RNA Mini Kit (Qiagen Sciences, Germantown, MD, USA) following the manufacturer's protocol. Complementary DNA (cDNA) was generated from the extracted double stranded RNA by reverse transcription using the SuperScript^{TM} III first-strand synthesis system for RT-PCR (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions. Briefly, an initial reaction mixture consisting of viral double stranded RNA and random primers was denatured at 97^{o}C for 5 minutes and quickly chilled on ice for 5 minutes. To this was added

a reverse transcription reaction mixture containing SuperScript™ III reverse transcriptase and dNTPs to make up a final volume of 20 µL, and cDNA was synthesised at 42°C for 1 hour. Each of the 11 genome segments was amplified by PCR from 2 µL of cDNA using gene specific primers (Supplementary Table 2.1) (Doan et al., 2012; Gentsch et al., 1992; Giambiagi et al., 1994; Gouvea et al., 1990; Matthijnssens et al., 2008a) and GoTaq® Green Master Mix System (Promega Corporation, Madison, WI, USA) under the following conditions: 95°C/5 min followed by 35 cycles of PCR at 94°C/30s; 45°C/30s; 72°C/3 min and final extension at 72°C/8 min.

Amplicons of the 11 genome segments were purified using EXOSAP-IT purification system (USB products, Cleveland, OH, USA) following the manufacturer's protocol and sequenced in both forward and reverse directions by the fluorescent dideoxy chain termination chemistry using the Big Dye Terminator Cycle Sequencing Ready Reaction Kit v3.1 (Applied Biosystems). Nucleotide sequences were determined using the ABI-PRISM 3730 Genetic Analyser (Applied Biosystems). For the sequencing of larger genes, the primer-walking method was employed on both strands to cover the complete ORF.

Sequence and Phylogenetic Analysis

Nucleotide sequences for each genome segment were assembled into contigs using the SeqMan program in DNAstar Lasergene core suite software v11 (DNAstar, Inc. Madison, WI, USA) and the genotypes were determined using the RotaC v.2.0 automated online genotyping tool for Group A rotaviruses (Maes et al., 2009). Using the Basic Local Alignment Search Tool on the NCBI website, sequences similar to each of the 11 genome segments of PML1965 were retrieved and included in multiple sequence alignment files constructed using the online version of Multiple Alignment using Fast Fourier Transform (MAFFT version 7) (Katoh and Standley, 2013).

Nucleotide and amino acid similarity matrices were calculated for the multiple aligned sequences for each genome segment using MEGA v6.06. The best-fit nucleotide substitution models were determined for the dataset for each genome segment using MEGA v6.06 based on the corrected Akaike Information Criterion (AICc) values (Tamura et al., 2013). Using the best fit substitution models with the lowest AICc scores and highest log likelihood scores obtained from the model test in MEGA6 for each of the 11 datasets: T92+G+I (VP7, VP4, VP6), GTR+G+I (VP1, VP3), TN93+G+I (VP2), T92+G (NSP2, NSP4, and NSP5), T92+I (NSP1) and TN93+I (NSP3), maximum likelihood phylogenetic trees were constructed using 1000 pseudo-replicate datasets.

Lineages were assigned to closely related collections of sequences with $\geq$70% bootstrap support at the branching point. Where there is further diversification below the lineage level, the term sub-lineage was introduced.

Estimation of the evolutionary rate of G6 VP7 gene and the time of the most recent common ancestor of the G6P[6] VP7 sub-lineage

The divergence times were estimated for the G6 VP7 gene of 50 dated representative G6 rotavirus strains of animal and human host species origin detected from 1971 to 2012 using the Bayesian Markov chain Monte Carlo (MCMC) method implemented in BEAST v1.8.1 (Drummond et al., 2012). Two separate datasets were compiled for the estimation of the evolutionary rates before and after the bovine G6 rotaviruses crossed the host species barrier into humans: (1) 85 dated G6 VP7 genes from bovine rotavirus strains detected from 1971 – 2012 and (2) 53 dated G6 VP7 genes from human rotavirus strains detected from 1987 – 2012 (Supplementary Table 2.2).

The general time reversible (GTR) nucleotide substitution model and Gamma distributed rate variation with invariant sites (G+I), a lognormal relaxed clock (Drummond et

al., 2006) and a coalescent constant size (Drummond et al., 2002) were assumed. Three independent MCMC runs were carried out for 100 million generations and evaluated using Tracer software v1.6 (http://tree.bio.ed.ac.uk/software/tracer/). Maximum clade credibility tree was annotated with the Treeannotator and viewed with FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

Nucleotide sequence accession numbers

Nucleotide sequences were submitted to the International Nucleotide Sequence Database Collaboration under the accession numbers LC026103 to LC026113 (Supplementary Table 2.3).

2.4 RESULTS

Genotype constellation of PML1965

The nucleotide sequence spanning the entire open reading frame of each of the genes except the VP4 gene was determined for PML1965 (Supplementary Table 2.3). The genotype constellation of PML1965 was G6-P[6]-I2-R2-C2-M2-A2-N2-T2-E2-H2, which was identical with that of the prototype G6P[6] strain B1711 detected in Belgium (Matthijnssens et al., 2008c) as well as those of G6P[6] strains detected in Cameroon (Ndze et al., 2014) and Guinea Bissau (Wentworth et al., GenBank data, 2014) (Table 2.1). Furthermore, it appeared identical with the genotype constellation of G6P[6] strains detected in Burkina Faso (Nordgren et al., 2012b) and Italy (Ianiro et al., 2013) although the genotypes of genome segments 1, 2 and 3 were not available (Table 2.1).

Phylogenetic analysis of PML1965

VP7 Gene

In the G6 VP7 phylogenetic tree, PML1965 was located in a distinct sub-lineage composed exclusively of human G6P[6] strains (designated as VIb in Fig. 2.1). This G6P[6] sub-lineage then clustered together with human G6P[9] strains with a 100% bootstrap support, forming a large lineage yet consisting exclusively of human G6P[6] and G6P[9] strains (designated as lineage VI in Fig. 2.1).

Within the lineage VI, the VP7 sequences diverged as follows; first, an Italian G6P[9] strain PA151 detected in 1987 and then an American G6P[9] strain Se584 detected in 1998 branched off, and the remaining strains were divided into two sub-lineages. One sub-lineage was made up of human G6P[9] strains detected in Africa, Asia and Europe (designated as VIa in Fig. 1), and the other comprised only G6P[6] strains detected in Africa and Europe including PML1965 (designated as VIb in Fig. 2.1).

The VP7 sequences within the latter sub-lineage were highly identical with ≥97.0% identity (Table 2.2). The topology that two G6P[9] strains were located outside of the lineage containing all the other G6P[9] and G6P[6] strains indicate that the VP7 sequences of G6P[6] strains originated from those of G6P[9] strains. The branch on which the transition occurred from the VP7 sequences of G6P[9] strains to those of G6P[6] strains is shown by the arrow in Fig. 2.1.

The VP7 sequences in the G6P[9] and G6P[6] sub-lineages VIa and VIb, respectively, were very closely related to each other with average nucleotide and amino acid identities of >95.3% which are almost within the range of intra-lineage identities (Table 2.3). Thus, they clustered into a single lineage VI, and this lineage was distantly related to any other G6 VP7 lineage (designated as lineages I - V in Fig. 2.1) that

Table 2.1: Comparison of the genotype constellation of PML1965 with other G6P[6] strains

Strains	Genome Segment											Reference
	VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5	
RVA/Human-wt/GHA/PML1965/2012/G6P[6]	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	This study
RVA/Human-wt/BEL/B1711/2002/G6P[6]	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Matthijnssens *et al.,* 2008*a*
RVA/Human-wt/GNB/MRC-DPRU5608/XXXX/G6P[6]	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Wentworth *et al.,* 2014, GenBank
RVA/Human-wt/GNB/MRC-DPRU5615/2011/G6[P6]	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Wentworth *et al.,* 2014, GenBank
RVA/Human-wt/GNB/MRC-DPRU5625/2011/G6P[6]	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Wentworth *et al.,* 2014, GenBank
RVA/Human-wt/CMR/MA202/2011/G6P[6]	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze e*t al*., 2014
RVA/Human-wt/CMR/MA228/2011/G6P[6]	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze e*t al*., 2014
RVA/Human-wt/CMR/ES298/2011/G6P[6]	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze e*t al*., 2014
RVA/Human-wt/CMR/BA346/2010/G6P[6]	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze e*t al*., 2014
RVA/Human-wt/CMR/BA369/2010/G6P[6]	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze e*t al*., 2014
RVA/Human-wt/BFA/238-BF/2010/G6P[6]	G6	P[6]	I2	-	-	-	A2	N2	T2	E2	H2	Nordgren *et al.,* 2012*a*
RVA/Human-wt/BFA/263-BF/2010/G6P[6]	G6	P[6]	I2	-	-	-	A2	N2	T2	E2	H2	Nordgren *et al.,* 2012*a*
RVA/Human-wt/BFA/265-BF/2010/G6P[6]	G6	P[6]	I2	-	-	-	A2	N2	T2	E2	H2	Nordgren *et al.,* 2012*a*
RVA/Human-wt/BFA/272-BF/2010/G6P[6]	G6	P[6]	I2	-	-	-	A2	N2	T2	E2	H2	Nordgren *et al.,* 2012*a*
RVA/Human-wt/ITA/CEC06/2011/G6P[6]	G6	P[6]	I2	-	-	-	A2	N2	T2	E2	H2	Ianiro *et al.,* 2013

Dashes indicate no sequence available; shaded strains: sequence data is only available in GenBank
Accession numbers for shaded strains
RVA/Human-wt/GNB/MRC-DPRU5608/XXXX/G6P[6] : KJ751916, KJ751914, KJ751915, KJ751911, KJ751912, KJ751913, KJ751906, KJ751907, KJ751908, KJ751909, KJ751910
RVA/Human-wt/GNB/MRC-DPRU5615/2011/G6[P6]: KJ752355, KJ752353, KJ752354, KJ752350, KJ75235, KJ752352, KJ752345, KJ752346, KJ752347, KJ752348, KJ752349
RVA/Human-wt/GNB/MRC-DPRU5625/2011/G6P[6]: KJ752122, KJ752120, KJ752121, KJ752117, KJ752118, KJ752119, KJ752112, KJ752113, KJ752114, KJ752115, KJ752116

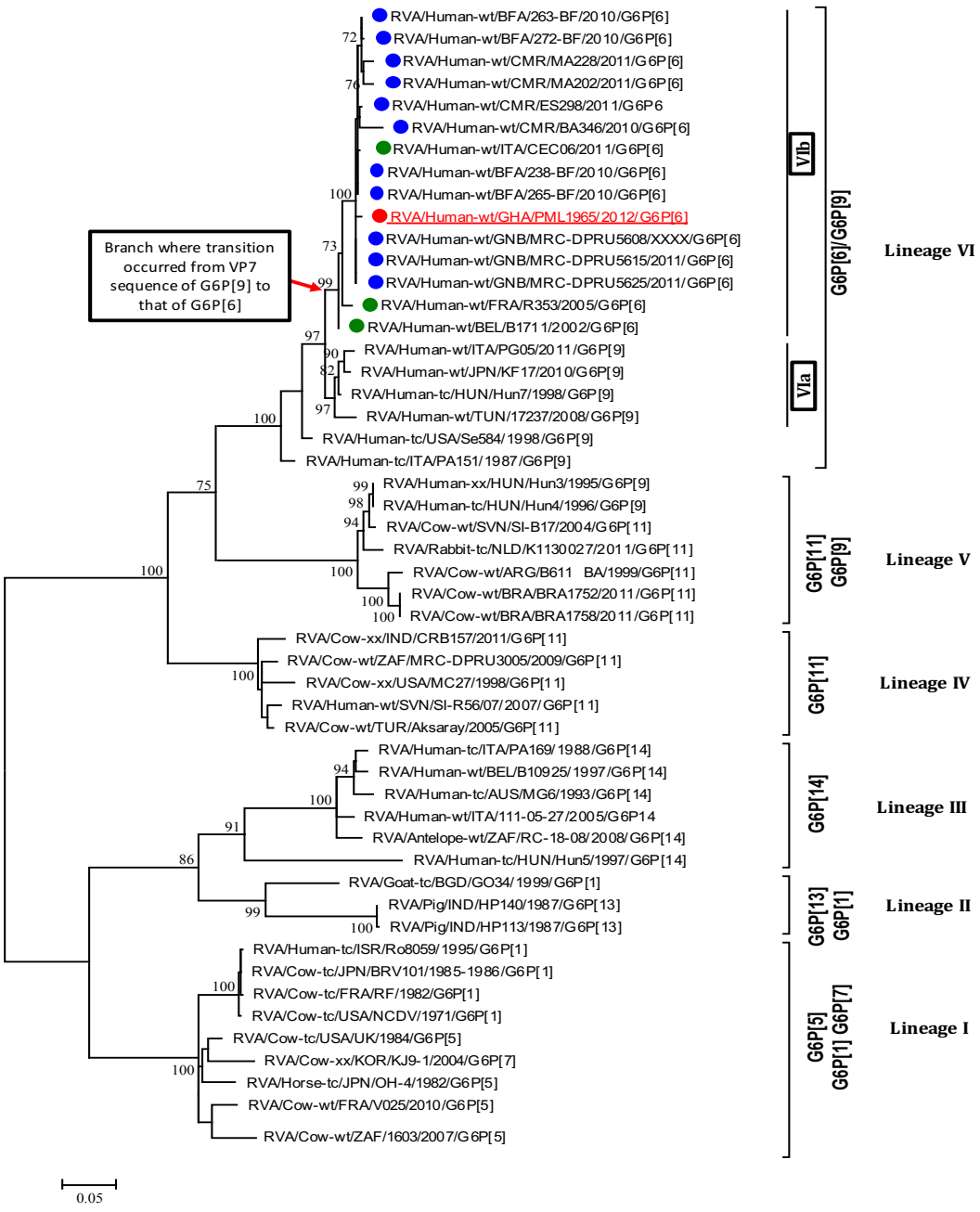


Fig. 2.1: A phylogenetic tree of the VP7 gene of G6 rotavirus strains showing the genetic relationship between PML1965 in this study and other human and animal G6 RVA. The phylogenetic analysis included the VP7 nucleotide sequence of the Ghanaian G6P[6] strain PML1965 in this study (indicated in red font with red dot), other African (indicated by blue dots), European (indicated by green dots) G6P[6] strains and G6 strains of both human and animal origin possessing P-types: [1], [5], [7], [9], [11], [13] or [14]. Maximum likelihood phylogenetic analysis was performed using Tamura 3-parameter substitution model with gamma distributed invariant sites in MEGA6 software package, and the resulting tree presented here is a midpoint-rooted tree. Significant bootstrap values (1000 replicates) of ≥70% are indicated at each node. The scale bar at the bottom of the tree indicates a genetic distance expressed as nucleotide substitutions per site.

Table 2.2: Comparison of p-distances between PML1965, representative G6P[6] and closely related DS-1-like strains

		RVA/Human-wt/GHA/PML1965/2012/G6P[6]										
		VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5
	Strains	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2
G6P[6] Strains	RVA/Human-wt/BEL/B1711/2002/G6P[6]	0.019	0.03	0.066	0.092	0.027	0.159	0.034	0.009	0.023	0.098	0.025
	RVA/Human-wt/GNB/MRC-DPRU5608/XXXX/G6P[6]	0.006	0.004	0.07	0.013	0.003	0.003	0.015	0.003	0.008	0.125	0.015
	RVA/Human-wt/GNB/MRC-DPRU5615/2011/G6P[6]	0.006	0.005	0.07	0.013	0.002	0.003	0.014	0.003	0.008	0.125	0.013
	RVA/Human-wt/GNB/MRC-DPRU5625/2011/G6P[6]	0.006	0.004	0.07	0.013	0.002	0.003	0.015	0.001	0.008	0.125	0.014
	RVA/Human-wt/CMR/MA202/2011/G6P6	0.03	0.027	0.073	0.013	0.012	0.013	0.016	0.034	0.005	0.153	0.023
	RVA/Human-wt/CMR/MA228/2011/G6P6	0.016	0.011	0.079	0.019	0.013	0.013	0.014	0.035	0.004	0.009	0.008
	RVA/Human-wt/CMR/ES298/2011/G6P6	0.008	0.01	0.073	0.007	0.013	0.01	0.012	0.006	0.002	0.006	0.023
	RVA/Human-wt/CMR/BA346/2010/G6P6	0.021	0.01	0.074	0.007	0.013	0.01	0.01	0.006	0.003	0.006	0.023
	RVA/Human-wt/BFA/238-BF/2010/G6P[6]	0.005	0.008	0.075	_	_	_	0.015	0.002	0.005	0.002	0.007
	RVA/Human-wt/BFA/265-BF/2010/G6P[6]	0.005	0.008	0.005	_	_	_	0.013	0.006	0.004	0.001	0.005
	RVA/Human-wt/BFA/263-BF/2010/G6P[6]	0.008	0.037	0.077	_	_	_	0.016	0.003	0.004	0.006	0.011
	RVA/Human-wt/BFA/272-BF/2010/G6P[6]	0.011	0.011	0.077	_	_	_	0.018	0.004	0.004	0.007	0.011
	RVA/Human-wt/ITA/CEC06/2011/G6P[6]	0.007	0.012	0.003	_	_	_	0.008	0.003	0.003	0.001	0.006
	RVA/Human-wt/FRA/R353/2005/G6P[6]	0.018	–	–	–	–	–	–	–	–	–	–
Non-G6P[6] DS-1-like Strains	RVA/Human-wt/GMB/MRC-DPRU3180/2010/G2P[6]	DG	0.009	0.07	0.007	0.015	0.009	0.006	0.128	0.002	0.093	0.008
	RVA/Human-wt/GHA/GH018-08/2008/G8P[6]	DG	0.011	0.022	0.088	0.158	0.172	0.02	0.082	0.01	0.11	DG
	RVA/Human-wt/GHA/GH019-08/2008/G8P[6]	DG	0.01	0.021	0.012	0.154	0.172	0.02	0.083	0.012	0.11	DG
	RVA/Human-wt/COD/DRC86/2003/G8P[6]	DG	0.044	0.066	0.027	0.04	0.026	0.037	0.023	0.019	0.02	0.031
	RVA/Human-wt/GHA/MRC-DPRU1818/1999/G2P[6]	DG	0.027	0.071	0.088	0.022	0.032	0.03	0.011	0.019	0.097	0.026
	RVA/Human-wt/ZAF/MRC-DPRU1815/1999/G2P[6]	DG	0.026	0.069	0.088	0.022	0.032	0.03	0.01	0.019	0.095	0.026
	RVA/Human-wt/ZAF/MRC-DPRU1845/1999/G2P[6]	DG	0.026	0.07	0.088	0.022	0.032	0.031	0.01	0.018	0.095	0.026
	RVA/Human-wt/TGO/MRC-DPRU5124/2010/G2P[4]	DG	DG	0.068	0.013	0.005	0.005	0.016	0.035	0.006	0.12	0.018
	RVA/Human-wt/TGO/MRC-DPRU2201/XXXX/G2P[4]	DG	DG	0.068	0.012	0.005	0.004	0.013	0.034	0.007	0.123	0.01

Key: DG: Different Genotype, _: Sequence not available, Boxed values: G6P[6] strains that obtained their NSP4 gene from a different ancestral sequence

contained the sequences possessed by rotavirus strains with common bovine G and P type combinations such as G6P[1], G6P[5], and G6P[11] or those possessed by rare human G6P[14] strains that were reported to be of bovine rotavirus origin (Banyai et al., 2003; Gerna et al., 1992; Matthijnssens et al., 2008a). It is noted that the VP7 gene of G6 strains from diverse host species origin clustered in accordance with the P genotype they possessed, and the nucleotide sequence identity within each lineage was ≥93.6% (Table 2.3).

VP4 Gene

Similarly, in the P[6] VP4 phylogenetic tree, PML1965 belonged to a lineage that contained only human rotavirus strains (designated as lineage VII in Fig. 2.2). Together with Guinea Bissau G6P[6] strains PML1965 formed a sub-lineage with G1, G2, G4 and G8 rotavirus strains detected in Africa (designated as VIIb in Fig. 2.2). Among the non-G6P[6] strains, a Gambian G2P[6] strain MRC-DPRU3180 had the closest VP4 sequence to that of PML1965 (Table 2.2). The prototype G6P[6] strain B1711, however, was 3.0% divergent from PML1965 (Table 2.2), forming a separate sub-lineage (designated as VIIa in Fig. 2.2) and was closely related to two G2P[6] strains from Ghana and South Africa.

VP6 Gene and other internal and non-structural protein genes (VP1, VP2, VP3, NSP1-NSP5)

In the VP6 phylogenetic tree, PML1965 belonged to the same lineage with G6P[6] strains from Burkina Faso (265/BF) and Italy (CEC06) (indicated in a box in Fig. S1(a)), and their minimum nucleotide sequence identity was 99.5% (Table 2.2). In the VP1, VP2 and VP3 phylogenetic trees, PML1965 belonged to the same lineage with Guinea Bissau G6P[6] strains (boxed in Fig. S1(b), (c) and (d)), and their minimum nucleotide sequence identity in the VP1, VP2 and VP3 genes was 98.7%, 99.7%, and 99.7%, respectively (Table 2.2).

In the NSP1, NSP2, NSP3, and NSP5 phylogenetic trees, PML1965 belonged to the same lineage with G6P[6] strains from Guinea Bissau, Burkina Faso and Italy (Fig. S1(e), (f), (g), (i)) their minimum nucleotide sequence identity in the NSP1, NSP2, NSP3, and NSP5 genes was 98.4%, 99.6%, 99.2% and 98.5%, respectively (Table 2.2). The NSP4 gene of PML1965 belonged to the same lineage with G6P[6] strains detected in Burkina Faso and Italy (Fig. S1(h)), and their minimum nucleotide sequence identity was 99.3% (Table 2.2). Whereas the NSP4 genes of B1711 and the Guinea Bissau G6P[6] strains were derived from different ancestral RVA strains (Fig. S1(h), boxed in Table 2.2), the high nucleotide sequence identities between the remaining G6P[6] strains imply that these non-structural protein genes were derived from the same ancestral sequences.

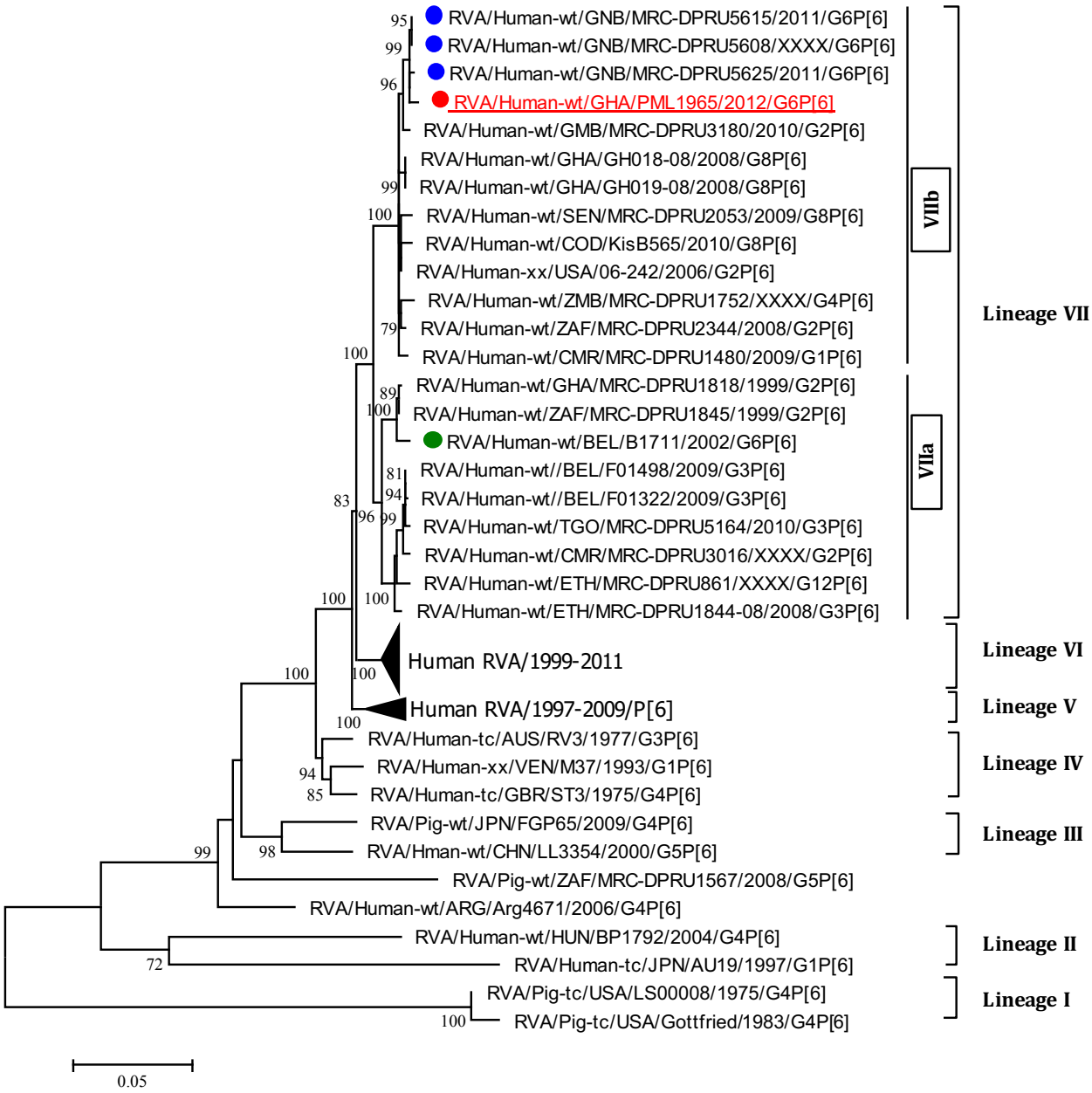


Fig. 2.2: A Phylogenetic tree of the VP4 gene of P[6] rotavirus strains showing the genetic relationship between PML1965 in this study and other human and animal P[6] RVA. The phylogenetic analysis included the VP4 nucleotide sequence of the Ghanaian G6P[6] strain PML1965 in this study (indicated in red font with red dot), the near-full length ORF of VP4 gene of other African (indicated by blue dots), European (indicated by green dot) G6P[6] strains and representative human and animal P[6] strains. Maximum likelihood phylogenetic analysis was performed using Tamura 3-parameter substitution model with Gamma distributed invariant sites in MEGA6 software package, and the resulting tree presented here is a midpoint-rooted tree. Significant bootstrap values (1000 replicates) of ≥70% are indicated at each node. The scale bar at the bottom of the tree indicates a genetic distance expressed as nucleotide substitutions per site.

Table 2.3: Comparison of the average nucleotide and amino acid sequence identities between and within G6 VP7 lineages in the VP7 phylogenetic tree

Average amino acid sequence identities between G6 VP7 lineages

G/P genotype	G6P[6]	G6P[9]	G6P[9], G6P[11]	G6P[11]	G6P[14]	G6P[13], G6P[1]	G6P[5], G6P[1], G6P[7]
Linage/sub-lineage	**VIb**	**VIa**	**V**	**IV**	**III**	**II**	**I**
VIb	**98.6 (98.4)**	97.4	91.3	91.1	91.7	91.9	88.6
VIa	95.3	**96.7 (98.1)**	92.2	91.9	92.8	93.1	89.8
V	86.1	87.0	**96.4 (97.7)**	92.2	91.7	92.1	89.4
IV	85.0	86.1	85.5	**96.7 (97.4)**	92.5	92.1	90.0
III	80.3	81.9	81.0	82.3	**93.6 (99.1)**	96.8	92.8
II	80.8	82.0	81.3	80.6	86.3	**93.8 (99.9)**	92.3
I	80.5	81.8	81.8	81.7	84.2	83.7	**95.0 (97.3)**

Average nucleotide sequence identities between G6 VP7 lineages

Lower left portion and top right portion of the table correspond respectively to average nucleotide and amino acid sequence identities between G6 VP7 lineages. Bold fonts are identities within lineages, values in brackets correspond to average amino acid sequence identities within the G6 VP7 lineages.

Table 2.4: Comparison of parameters of rotavirus G6 VP7 gene to other G genotypes previously obtained from Bayesian phylogenetic reconstruction using MCMC analysis in BEAST.

Parameter	Genotype								
	G6 (VP7) BRV	**G6 (VP7) HRV**	**G1(VP7)**	**G2(VP7)**	**G2(VP7)**	**G3(VP7)**	**G9(VP7)**	**G9(VP7)**	**G12(VP7)**
Number of samples	87	53	85	77	328	80	82	356	140
Sampling Interval	1971-2012	1987-2012	1991-2012	1991-2012	1975-2012	1976-2012	1996-2012	1980-2009	1988-2009
Sampling area	Global	Global	Bangladesh	Bangladesh	Global	Global	Bangladesh	Global	Global
Evolutionary rate (10^{-3}) substitutions/site/year	0.69	3.42	0.93	1.45	1.45	1.47	1.07	1.87	1.66
(95% HPD interval)	(0.45 – 0.95)	(1.53 – 6.11)	(0.68-1.18)	(1.12-1.78)	(1.12-1.63)	(0.75-2.33)	(0.78-1.39)	(1.45-2.27)	(1.30-2.32)
Coefficient of Variation (95% HPD interval)	0.45 (0.26-0.66)	1.31 (0.61-2.21)	NA	NA	0.8 (0.63-1.05)	NA	NA	1.36 (1.026 - 1.681)	0.8 (0.406 - 1.307)
Reference	This study	This study	Afrad *et al.,* 2014	Afrad *et al.,* 2014	Dennis *et al.*, 2014	He *et al*., 2013	Afrad *et al.,* 2014	Matthijnssens *et al.,* 2010	Matthijnssens *et al.,* 2010

Key: NA: Not available; BRV: Bovine rotavirus; HRV: Human rotavirus

Of the non-G6P[6] DS-1-like strains, PML1965 had the closest nucleotide sequence with a Gambian G2P[6] strain MRC-DPRU3180 in four genome segments: i.e., the VP1, NSP1, NSP3 and NSP5 genes (Table 2.2), and they belonged to the same lineage (Fig. S1(b), (e), (g), (i)). PML1965 had the closest nucleotide sequence with two Togolese G2P[4] strains MRC-DPRU5124 and MRC-DPRU2201 in two genome segments: i.e., the VP2 and VP3 genes (Table 2.2), and they belonged to the same lineage (Fig. S1(c) and (d)). PML1965 had the closest nucleotide sequence with three G2P[6] strains detected in Ghana and South Africa in the NSP2 gene (Table 2.2), and they belonged to the same lineage (Fig. S1(f)). PML1965 had the closest nucleotide sequence with two Ghanaian G8P[6] strains GH019-08 and GH018-08 in the VP6 gene (Table 2.2), and they belonged to the same lineage (Fig. S1(a)). No single rotavirus strain provided PML1965 with the DS-1-like genetic backbone since the individual gene segments of PML1965 showed high nucleotide sequence identities ($\geq$97.8%) with genes from no single, but multiple non-G6P[6] DS-1-like strains.

On the other hand, the prototype G6P[6] strain B1711 belonged to the same lineage only in the NSP2 gene (Fig. S1(f)), and their nucleotide sequence identity was 99.1% (Table 2.2). Among the genome segments in which B1711 did not share the lineage with PML1965, the VP3 gene of B1711 showed the lowest nucleotide sequence identity with PML1965 (84.1%) (Table 2.2).

With respect to the VP6 gene and other internal and non-structural protein genes, we aligned the deduced protein sequences of PML1965 together with other G6P[6] strains and non-G6P[6] DS-1-like strains to explore whether there were any amino acid residues unique to the G6P[6] strains. We found that the proteins encoded by the VP1, VP3, NSP1 and NSP4 genes of the G6P[6] strains and strains that shared the same lineage with them contained a few unique amino acid residues: i.e., 289Q in VP1, 87N and 199V in VP3, 190I in NSP1, and 62N, 95M, and 129H in NSP4. Their biological implications were not clear, however.

Evolutionary rate of G6 VP7 gene and the time of most recent common ancestor of the G6P[6] VP7 sub-lineage

The evolutionary rates of the G6 VP7 genes from strains of bovine origin and of human origin were estimated as 6.93 x 10^{-4} substitutions/site/year (Highest Posterior Density interval [HPD]: 4.49 x 10-4 - 9.54 x 10^{-4} substitutions/site/year) and 3.42 x 10^{-3} substitutions/site/year (HPD: 1.53 x 10-3 - 6.11 x 10^{-3}), respectively (Table 2.4). Thus, the evolutionary rate of the G6 VP7 genes of human rotavirus origin was approximately 5 times faster than that estimated for the G6 VP7 genes of bovine rotavirus origin (Fig. 2.3a), suggesting that the evolutionary rate was accelerated after the bovine G6 rotaviruses crossed the host species barrier into humans. The credible interval of the coefficient of variation of 0.26 to 0.66 and 0.61 to 2.21 for both estimates (Table 2.4) clearly excluded zero, and it indicated variation in rates among branches. This result validated the use of the relaxed clock model in estimating the evolutionary rate of the G6 VP7 gene.

To explore whether the frequency of sampling over time could lead to an opposite result, we calculated the evolutionary rates for the human and bovine datasets under two different sampling scenarios; (i) one strain each was selected at random from each isolation year, and (ii) one sample each was selected from the years of detection shared by both human and bovine strains. Under both sampling scenarios we observed consistently accelerated evolutionary rates in the human G6 VP7 dataset over the bovine G6 VP7 dataset (data not shown).

In the maximum clade credibility tree (Fig. 2.3b), the ancestor of the VP7 gene of G6P[6] strains - a G6P[9] strain, diverged from the contemporary G6P[9] strains around 1990. The VP7 gene of the ancestral human G6P[9] strain accumulated point mutations and transitioned into the VP7 gene of G6P[6] strains. The time of the most recent common ancestor of the VP7 gene of G6P[6] sub-lineage was calculated to be around the year 1998

(Fig. 2.3b), and it appears that the VP7 gene of G6P[6] strains detected after the prototype Belgian strain B1711 including PML1965 evolved and spread from this single introduction point of strain B1711 which was detected in a child returning from a vacation in Mali (Matthijnssens et al., 2008c; Rahman et al., 2003).

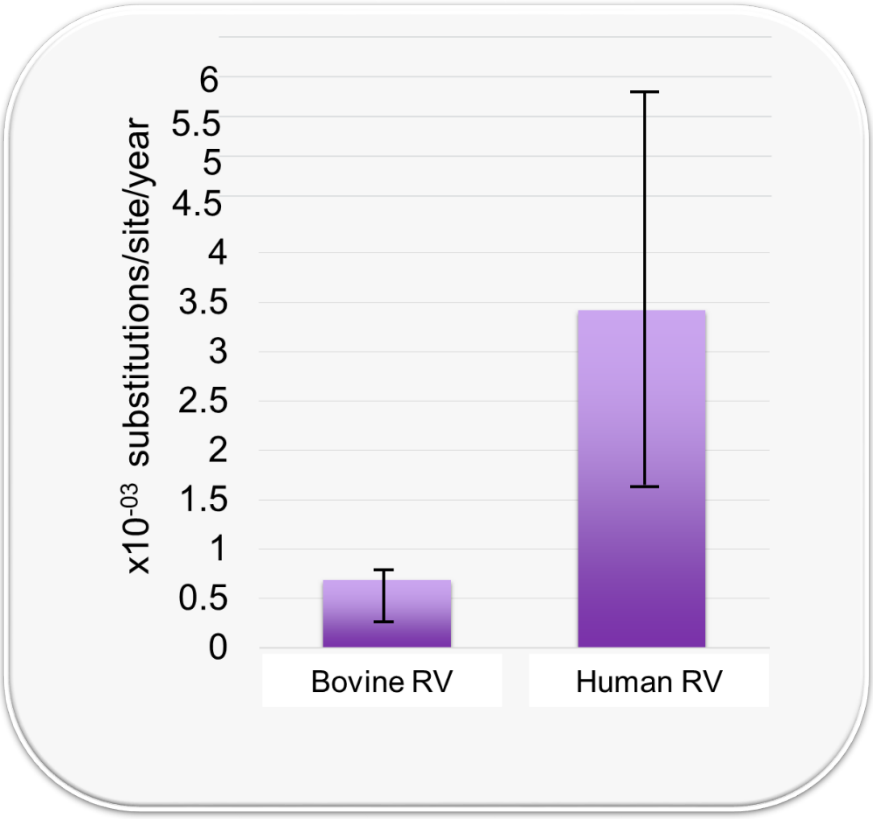


Fig. 2.3a: Increased evolutionary rate was observed in the human G6 VP7 gene upon comparison of the evolutionary rate of bovine G6 VP7 gene to that of G6 VP7 gene of human rotavirus origin. Error bars represent the 95% highest posterior density intervals.

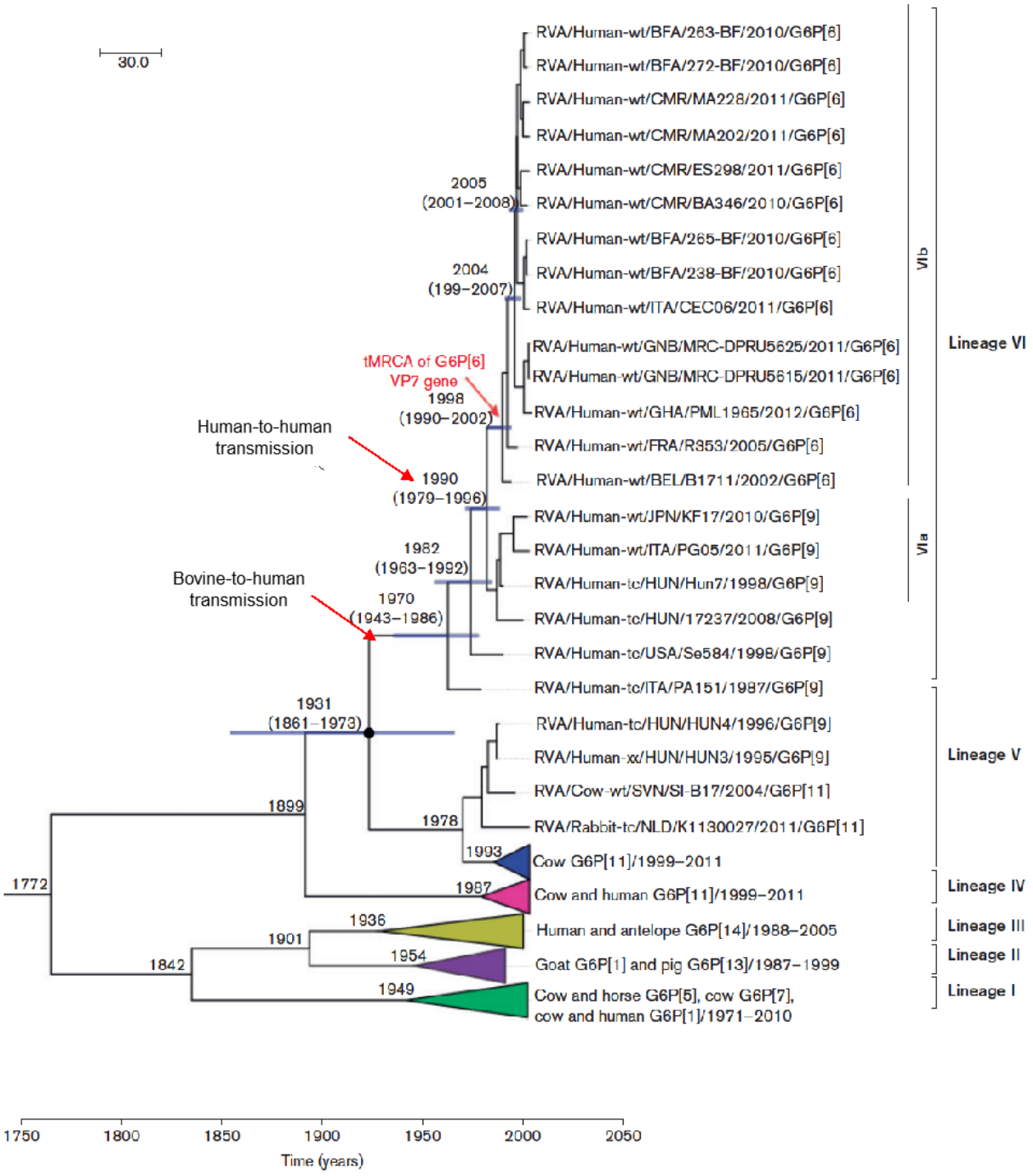


Fig. 2.3b: A simplified maximum clade credibility tree of 50 dated G6 VP7 nucleotide sequences constructed using the Bayesian MCMC framework. The 95% highest posterior density (HPD) interval of each significant node is indicated with bars. The time of most recent common ancestor (tMRCA) is indicated for the G6P[6] sub-lineage VIb and the remaining lineages. Lineages far away from the G6P[6] and G6P[9] sub-lineages VIb and VIa have been collapsed for simplicity. The black dot indicates the node at which the VP7 sequence of PA151-like G6P[9] strain diverged from its ancestral rotavirus VP7 sequence

2.5 DISCUSSION

The VP7 gene of the Ghanaian G6P[6] rotavirus strain PML1965 was shown to be closely related to that of G6P[6] strains detected in Europe and Africa. Their VP7 genes shared a common ancestral VP7 sequence with human G6P[9] strains. In the remaining 10 genome segments, PML1965 shared the same lineage with some G6P[6] strains and DS-1-like G2 or G8 strains circulating in Africa. Thus, we hypothesised that the VP7 gene of a G6P[9] strain was reasssorted into the DS-1 backbone of regional circulating African DS-1-like strains, which resulted in the ancestor of the G6P[6] strains that emerged in Africa.

Previously, Matthijnssens et al. (2008c) suggested that the VP7 and VP3 genes of the prototype G6P[6] strain from Belgium, B1711, were of bovine rotavirus origin. Subsequently, the VP7 genes of G6P[6] strains detected in Burkina Faso were speculated to have been derived from an interspecies transmission event of bovine rotavirus to humans because people live in close proximity with cattle in Burkina Faso, thereby increasing the chance of interspecies transmission (Nordgren et al., 2012b). However, Ianiro et al. (2013) concluded that there was no evidence of zoonosis or interspecies reassortment after analysing eight genome segments of a G6P[6] strain detected in Italy in 2011. We extended their observations and suggested that the VP7 gene of PML1965 together with those of previously reported G6P[6] strains evolved from a single ancestral VP7 sequence originating from a human G6P[9] strain that occurred around 1998 (Fig. 2.3b). In this regard, it should be interesting to know how closely the VP7 genes of human G6P[6] strains are related to those of feline G6P[9] rotaviruses that were recently reported to be the most prevalent genotype among cats in the United Kingdom (German et al., 2015).

An interesting observation in this study was the demonstration, with a statistically significant difference, of a much faster evolutionary rate for the G6 VP7 genes possessed by human rotaviruses (3.42×10^{-3} substitutions/site/year) than that for the G6 VP7 genes possessed by bovine rotaviruses (6.93×10^{-4} substitutions/site/year). As it can be taken for

granted that the original host species of G6 rotaviruses are artiodactyls including cattle (Cashman et al., 2010; Midgley et al., 2012; Monini et al., 2008; Suzuki et al., 1993), it is reasonable to presume that the G6 VP7 genes have already been well adapted to bovine rotaviruses whereas humans are a new host species to the G6 VP7 genes. Thus, the increase in evolutionary rate observed for the G6 VP7 genes after their crossing the host species barrier into humans could constitute a post-transfer adaptation process through which the virus achieved increased replication and transmissibility after the initial transfer to a human host. An increased evolutionary rate after jumping into the new host species was demonstrated for the SARS coronaviruses which appeared to gain some host-adaptive changes during its spread among humans (Parrish et al., 2008; Zhang et al., 2006). In this regard, it is of note that the point estimates for the evolutionary rate for G9 and G12 VP7 genes calculated from a global collection of sequences were also high as these two VP7 genotypes are thought to have emerged recently in humans (Table 2.4) (Afrad et al., 2014; Dennis et al., 2014; He et al., 2013; Matthijnssens et al., 2010).

The maximum clade credibility tree also showed the divergence times of the VP7 genes of human G6 rotavirus strains from their ancestral animal VP7 sequences. Around 1931, an interspecies transmission event of a G6 bovine rotavirus to humans occurred, and this event gave rise to G6P[9] strains (PA151-like strains) that were perceived to possess the ancestral sequence of the VP7 gene carried by G6P[6] and G6P[9] strains (indicated with a black dot in Fig. 2.3b).

We provided further evidence in support of our hypothesis regarding how PML1965 evolved by taking advantage of the current availability of the whole genome sequence data of many DS-1-like rotavirus strains in the GenBank. To determine which DS-1-like strain aside from the G6P[6] strains was highly similar to PML1965 in the remaining 10 genome segments, we carried out maximum likelihood phylogenetic analyses. There was not a single DS-1-like strain that possessed all 10 genome segments highly similar to the corresponding

genome segments of PML1965 in terms of the nucleotide sequence identity (≥97.8% identity) and the phylogenetic relationships (belonging to the same lineage in the phylogenetic tree). However, at least five different regional circulating DS-1-like rotavirus strains detected in humans were identified that possessed at least one genome segment highly similar to the corresponding genome segment of PML1965 (at the bottom panel of Table 2.3). A Gambian G2P[6] strain MRC-DPRU3180 detected in 2010 was the closest to PML1965 in the VP4, VP1, NSP1, NSP3 and NSP5 genes (Fig. 2.2, Fig. S1(b), (e), (g) and (i)). The presence of African DS-1-like strains that possessed highly similar genome segments with those of PML1965 suggested that the backbone genotype constellation of PML1965 and other G6P[6] strains was configured by multiple intra-genotype reassortment events involving regional circulating DS-1-like strains. Because these parental DS-1-like strains are considered to have already been adapted to humans, the acquisition of such genetic backbones allowed PML1965 and the likes to have the ability to spread continuously from human to human.

In summary, the VP7 gene of PML1965, highly identical to those of previously reported G6P[6] strains, was shown to evolve from the VP7 genes of human G6P[9] strains at a higher evolutionary rate than that of bovine G6 VP7 genes. The remaining 10 genome segments were closely related to those of typical African G2P[4], G2P[6] and G8P[6] rotaviruses possessing the DS-1-like genotype constellation. These observations led us to the following hypotheses: 1) PML1965 is a single gene reassortant strain generated when a human P[6] RVA strain possessing the DS-1-like genetic background acquired the G6 VP7 gene from human G6P[9] RVA strains; 2) PML1965 is a double gene reassortant strain in which human P[4] RVA strains possessing the DS-1-like genetic background acquired the G6 VP7 gene from human G6P[9] RVA strains and P[6] VP4 gene from co-circulating strains. These observations led us to the hypothesis that reassortment events in which human P[6] or P[4] RVA strains possessing DS-1-like genetic background acquire the G6

VP7 gene from human G6P[9] RVA strains gave rise to G6P[6] strains, thereby spreading more efficiently from human to human. A follow up monitoring of the G6 strains is necessary since they may further acquire the Wa-like genetic background with P[8] and make a swift and global spread as was observed for G9 and G12 rotaviruses (Matthijnssens et al., 2010).

2.6 ACKNOWLEDGEMENTS

We acknowledge the immense support of the Program for Nurturing Global Leaders in Tropical and Emerging Communicable Diseases, Graduate School of Biomedical Sciences, Nagasaki University. This study was in part supported by grants-in-aid for scientific research from the Ministry of Health, Labour and Welfare of Japan, as well as a grant from Japan Initiative for Global Research Network on Infectious Diseases. We also thank the staff of the Regional Rotavirus Reference Laboratory at the Noguchi Memorial Institute for Medical Research in Ghana for providing samples for the study and the preliminary assays.

2.7 CONFLICT OF INTEREST

The authors declare no conflict of interest.

Chapter III

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

Published in:

Chantal Ama Agbemabiese, Toyoko Nakagomi, Yen Hai Doan, Loan Phuong Do, Susan Damanka, George E. Armah, Osamu Nakagomi. ***Infection, Genetics and Evolution 45,*** 122–131 **(2016)**

3.1 SUMMARY

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 examined G2P[4] strains detected from 2008-2013 in Ghana to understand their evolution within a global context. Representative G2P[4] strains were sequenced for their whole genomes and analysed 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. Unlike previous studies in Australia and Brazil, 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 1.2×10^{-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 G3P[6], G2[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.

Keywords: Ghana; rotavirus; G2P[4]; whole genome evolution; NSP4; reassortment,

3.2 INTRODUCTION

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^{TM} (Merck & Co. Inc.) 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 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 natural course of variation and evolution of epidemiologically relevant strains circulating before and after vaccine introduction. In this regard, the whole genomes of the G2P[4] strains detected before and during vaccine use were compared in populations in Australia (Donato et al., 2014). Strains detected during both periods shared high genetic relatedness with each other and with globally circulating G2P[4] strains. When Gomez et al. (2014) compared the whole genomes of G2P[4] strains detected in vaccinated children with those from non-vaccinated children and global G2P[4] strains, genes of the strains shared up to 99% nucleotide sequence identity indicating that the introduction of the rotavirus vaccine might not influence G2P[4] genetic diversity.

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., 2011c; 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 1997-2006 (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 evolution and variation in the context of global 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 encountered 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.

3.3 MATERIALS AND METHODS

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) (Supplementary Fig. 3.1a) 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.

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 SuperScript^{TM} 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. 3.1a and b). Briefly, genes were amplified by PCR using 2 µL of cDNA with gene specific primers (Supplementary Table 3.1) (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 Analyser (Applied Biosystems).

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 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) with bootstrap supports >70% and designated using the scheme proposed for global G2P[4] whole genomes (Doan et al., 2015). Where further diversification occurred within a lineage, the term sub-lineage was introduced.

Estimation of divergence time of lineages and evolutionary rate of genes using Bayesian Evolutionary Analysis by Sampling Trees (BEAST)

Bayesian evolutionary analysis was done to estimate when divergence occurred for lineages and the time of the most recent common ancestor (tMRCA) of the designated lineages. The NSP4 gene was noted to have more lineages containing G2P[4] strains compared to the other ten genes. We therefore calculated and compared the evolutionary rate for all genome segments to determine whether the frequent NSP4 lineage expansion was associated with higher evolutionary rate. Each gene's dataset comprised at least 177

taxa sampled globally and strain detection years included 1976 to 2013. The Bayesian Markov chain Monte Carlo (MCMC) method implemented in BEAST v1.8.1 (Drummond et al., 2012) was employed. An uncorrelated lognormal relaxed-clock model (Drummond et al., 2006), a coalescent constant size tree prior (Drummond et al., 2002) together with the best-fit nucleotide substitution models (based on Bayesian Information Criterion) were used. The MCMC calculations were carried out for 50 to 100 million generations depending on the size of the gene as well as the nucleotide substitution and site heterogeneity model with 25% burn-in. The BEAGLE library was used to enhance the computational speed of the analysis (Suchard and Rambaut, 2009). The results as examined by the Tracer software v1.6 showed effective sampling sizes of continuous parameters to be $\geq$ 200 (http://tree.bio.ed.ac.uk/software/tracer/). Maximum clade credibility trees were annotated with the TreeAnnotator (v1.8.1) and viewed in FigTree.

Nucleotide sequence accession numbers

Nucleotide sequences were submitted to the GenBank/DDBJ/EMBL under the accession numbers LC105533-LC105598 (Supplementary Table 3.2).

3.4 RESULTS

Genotype constellation of the Ghanaian G2P[4] strains

The full length ORFs of the 11 genome segments were determined for the six representative strains detected before and after vaccine introduction (Supplementary Table 3.2). Their genotype constellation was G2-P[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2 — a typical DS-1-like genetic backbone (Table 3.1).

Sequence and phylogenetic analysis

Genotype constellation of the Ghanaian G2P[4] strains

The full length ORFs of the 11 genome segments were determined for the six representative strains detected before and after vaccine introduction (Supplementary Table 3.2). Their genotype constellation was G2-P[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2 — a typical DS-1-like genetic backbone (Table 3.1).

Sequence and phylogenetic analysis

The NSP4 gene

The nucleotide sequence identity of the NSP4 gene of strains circulating in the same rotavirus season were highly identical with identities ranging from 99.7-100%. By contrast, between different seasons, NSP4 nucleotide sequence identities ranged from as low as 89.0% to 100% (mean: 91.8%) (Table 3.1). The maximum likelihood phylogenetic analysis revealed 10 G2P[4] containing lineages I to X (Fig. 3.1a), a rather higher number of G2P[4] lineages compared to the other genome segments (Table 3.2, Fig. 3.1b, c, d; Supplementary Fig. 3.1). The sequence diversity among the NSP4 gene of the study strains segregated them into three different lineages according to their year of detection; strains detected before vaccine introduction belonged to lineages V (2008) and a new African specific lineage designated X in this study (2009), whereas those detected after vaccine introduction belonged to lineage VI (2012/2013) (Table 3.2). Lineage X was shared with only Ghanaian G2 strains detected in 2009 (Heylen et al., 2015) and a few African DS-1-like RVA strains possessing G2P[6] and G3P[6] (Fig. 3.1a).

Extending our examination to all African G2 strains, it was observed that just around the turn of the century until 2013, a diverse pool of E2 genes belonging to lineages V (1999-2010), VI (2009-2013), VII (2007-2012), IX (2008-2011) and X (2008-2010) were carried by G2 strains on the African continent (Fig. 3.1a, Table 3.2). It is worthy to note that so far, lineage IX and X of the NSP4 gene contained sequences of RVA strains detected only in Africa.

The mean nucleotide sequence divergence within each of the 10 G2P[4] NSP4 lineages excluding lineage III (only one taxa) ranged from 0.5 to 5.7% (Table 3.3). The divergence between lineages I to V sequences ranged from 3.5 to 6.5% whereas divergence between any of the five aforementioned lineages and lineages VI to X sequences ranged from 9.1 to 15.6% (Table 3.3). The large sequence divergence between these aggregated lineages (I-V versus VI-X) suggests that sequences from these groups originated from distinct ancestral NSP4 sequences.

The internal capsid and remaining non-structural protein genes

Comparison of the VP3 gene of the six study strains revealed a high genetic identity between strains detected within the same season with identities ranging from 99.7-100%. Unlike the NSP4 gene, the VP3 gene of the 2008 and 2009 strains were closely related but were together distantly related to the VP3 gene of the 2012-2013 strains with sequence identities as low as 87.3% compared to the other genome segments (Table 3.1). The VP3 gene of the six study strains despite the high nucleotide sequence diversity only changed lineages from V to VII after vaccine introduction (Fig. 3.1b) and was not as diverse as the NSP4 gene at the lineage level.

The nucleotide sequences of the remaining internal capsid and non-structural protein genes of strains detected in the same season were highly identical (99.6-100%) except the NSP2 gene of the 2008 strains which differed by 2.5% (Table 3.1). However, comparison of sequences of the six strains from different seasons revealed lower mean identities ranging from 96.9-98.5 (Table 3.1). These remaining genes belonged to lineages that had been described for global G2P[4] strains; lineage IVa for VP2, NSP1, NSP5 and lineage V for VP6, VP1, NSP2, NSP3 (Table 3.2, Supplementary Fig. 3.1). Within lineage IVa/V, it was

observed that the pre-vaccine introduction strains belonged to the same sub-cluster with African DS-1-like strains whereas those detected after vaccine introduction were mostly found together with Australian post-vaccine introduction strains.

Table 3.1: Genotype constellation and nucleotide sequence identities of the 11 genes among the six study strains

	Genotype Constellation and % nucleotide identity between strains from the same season									Overall % nucleotide sequence identity	
	GHDC514	GHLA104		GHNAV482	GHNAV483		GHPML1989	GHDC1581		Mean	(range)
Gene	(2008)	(2008)		(2009)	(2009)		(2012)	(2013)			
VP7	G2	G2	99.7	G2	G2	99.9	G2	G2	99.7	97.1	(95.1-99.9)
VP4	P[4]	P[4]	99.9	P[4]	P[4]	100.0	P[4]	P[4]	99.9	98.0	(96.6-100)
VP6	I2	I2	99.8	I2	I2	100.0	I2	I2	99.8	98.4	(97.04-100)
VP1	R2	R2	99.9	R2	R2	99.9	R2	R2	99.8	96.9	(94.4-99.9)
VP2	C2	C2	99.6	C2	C2	99.9	C2	C2	99.9	98.0	(96.7-99.9)
VP3	M2	M2	99.9	M2	M2	100.0	M2	M2	99.7	92.9	(87.3-100)
NSP1	A2	A2	99.7	A2	A2	99.8	A2	A2	99.7	98.1	(96.9-99.8)
NSP2	N2	N2	97.5	N2	N2	100.0	N2	N2	99.8	97.7	(96.0-100)
NSP3	T2	T2	99.9	T2	T2	100.0	T2	T2	99.7	98.5	(97.0-100)
NSP4	E2	E2	99.7	E2	E2	100.0	E2	E2	99.8	91.8	(89.0-100)
NSP5	H2	H2	99.8	H2	H2	99.7	H2	H2	99.6	97.8	(95.9-99.8)

Key: identity across genome 

Fig. 3.1a. NSP4:E2

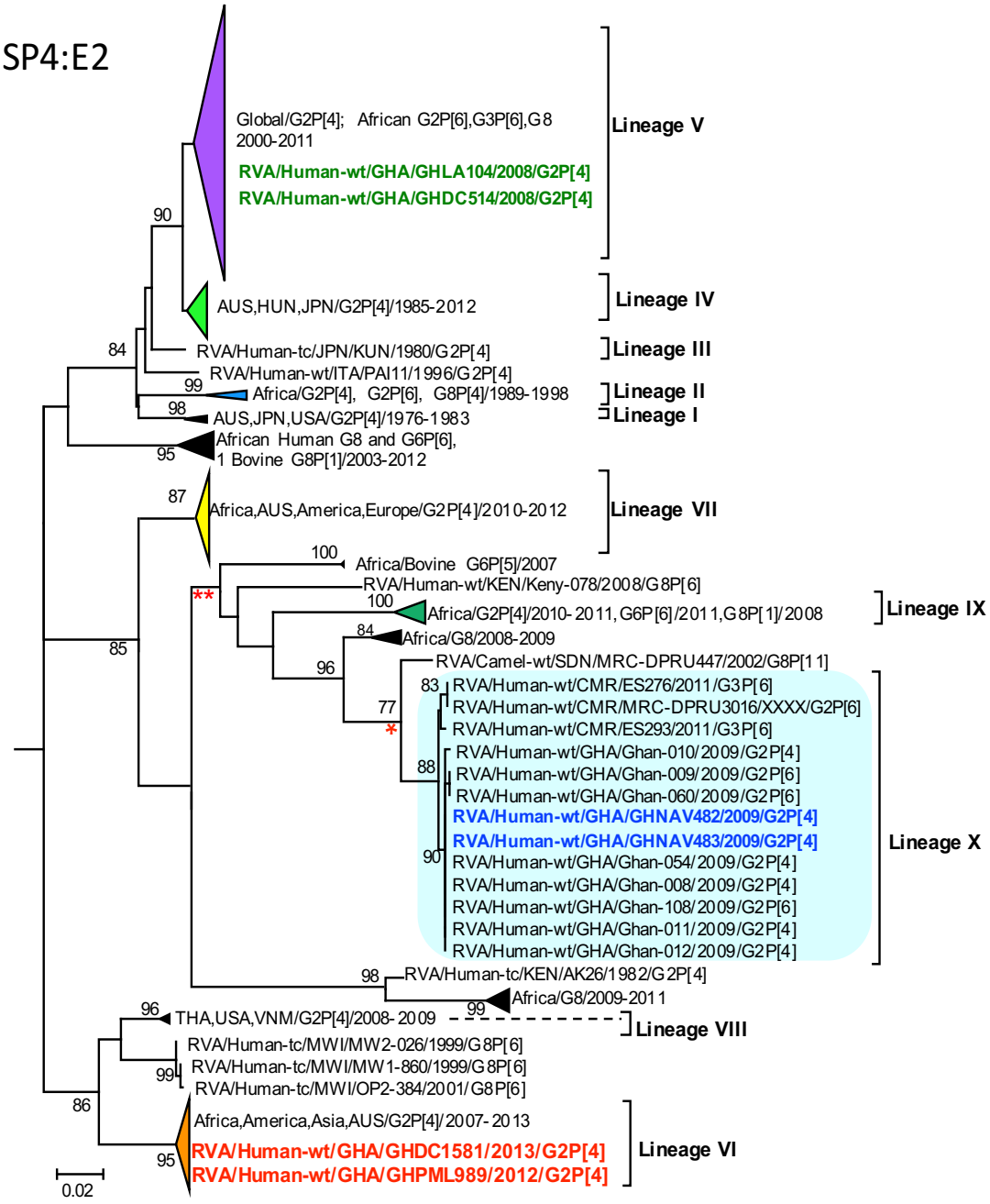
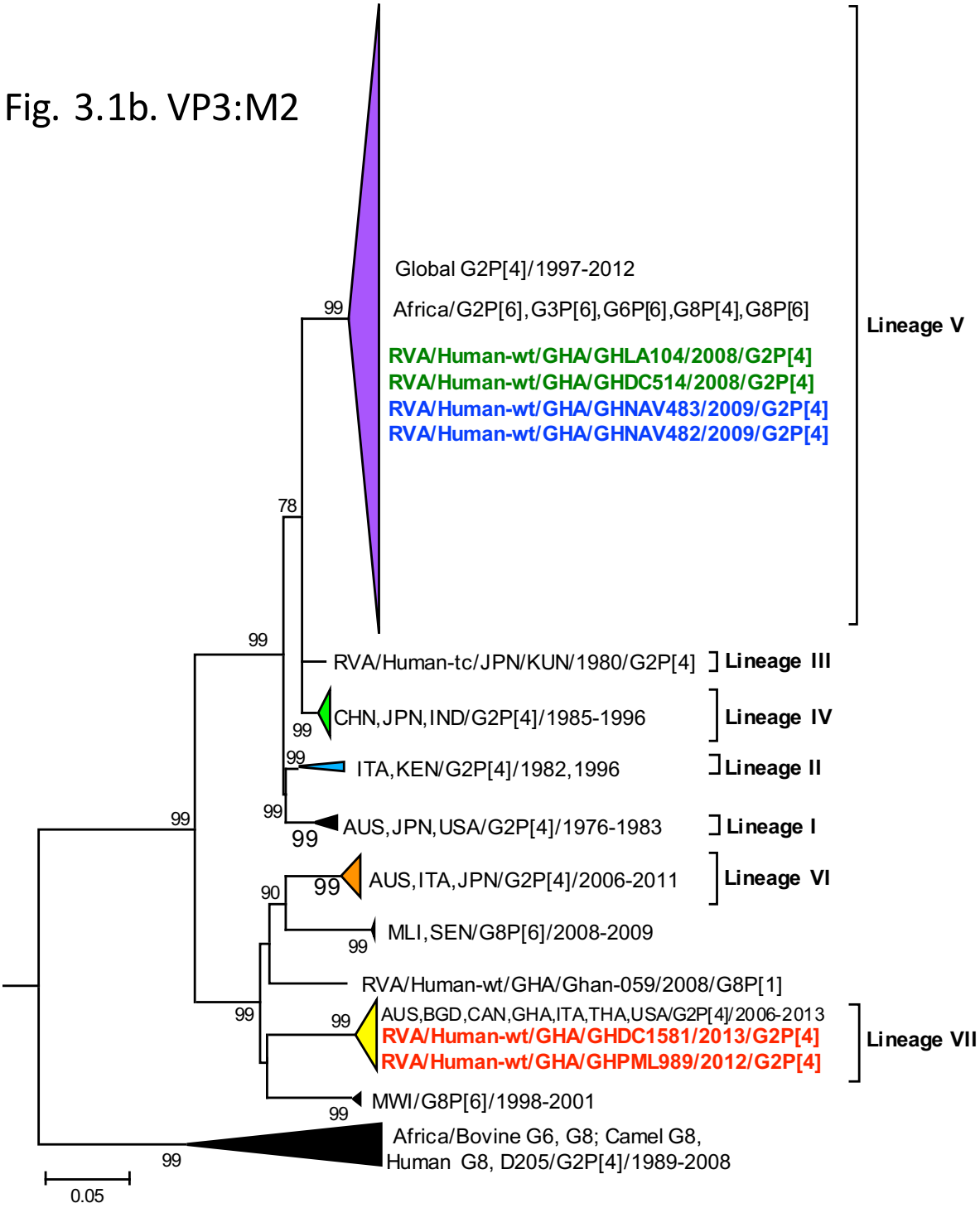


Fig. 3.1: Phylogenetic trees of genes of the: **(a)** NSP4 E2 genotype, **(b)** VP3 M2 genotype, **(c)** VP7 G2 genotype and (**d)** VP4 P[4] genotype of the study strains together with global G2P[4] and some DS-1-like RVA strains showing the evolutionary relationship between Ghanaian pre and post vaccine introduction period G2P[4] strains and other DS-1-like RVA strains. For each phylogenetic tree, analysis included at least 207 taxa consisting of the nucleotide sequences of the six Ghanaian G2P[4] strains in this study (indicated in green, blue (pre vaccine period) and red font (post vaccine period)), globally circulating strains G2P[4] with available whole genome sequence data and some African DS-1-like RVA strains from the GenBank database. Maximum likelihood phylogenetic analyses were performed using Tamura 3-parameter substitution model with gamma distributed rate variation (NSP4), General Time Reversible substitution model with gamma distributed rate variation and invariant sites (VP3, VP4) and Tamura 3-parameter substitution model with gamma distributed rate variation and invariant sites (VP7) in MEGA6 software package. The trees were rooted using the respective genes of the Wa strain. Significant bootstrap values (1000 replicates) are indicated at each node. Lineage designation was based on the scheme by Doan et al. (2015). The scale bar at the bottom of the tree indicates genetic distance expressed as nucleotide substitutions per site.

Fig. 3.1b. VP3:M2



The outer capsid protein genes: VP7 and VP4

A preliminary phylogenetic tree of the VP7 gene of the 38 study samples placed all the VP7 genes in lineage IVa (Supplementary Fig. 3.1a). Within lineage IVa, the 2008-2009 strains belonged to sub-lineage IVa-1 whereas the 2012-2013 strains belonged to sub-lineage IVa-3 (Supplementary Fig. 3.1a). The clustering pattern corresponded to that observed in the phylogenetic tree of the VP8* gene (Supplementary Fig. 3.1b). Unlike the NSP4 and VP3 genes, nucleotide sequence diversity of the outer capsid protein genes of the six representative Ghanaian G2P[4] strains were comparatively less and identities ranged from 95.1 - 100% (Table 3.1).

In addition, in the VP7 and VP4 phylogenetic trees (Fig. 3.1c and d), the study strains belonged to lineage IVa (coloured according to year of detection) together with globally circulating G2P[4] strains detected after 2000. The VP7 gene of the study strains further diverged into sub-lineage IVa-1 which contained the study strains detected before vaccine introduction from 2008 and 2009 and sub-lineage IVa-3 which contained the study strains detected after vaccine introduction from 2012 and 2013 (Fig. 3.1c). Their VP4 gene also clustered into two distinct sub-lineages within lineage IVa (Fig. 3.1d, within the triangle). Extending our description to all African G2 strains detected after 2000, with the exception of two South African G2P[4] strains detected in 2003 and 2012 (Nyaga et al., 2014) whose VP4 genes belonged to lineage II together with Kenyan G2P[4] strains isolated in the 1980s (Ghosh et al., 2011a), both the VP7 and VP4 genes of all other strains belonged to lineage IVa (Fig. 3.1c, d, Table 3.2).

Fig. 3.1c. VP7:G2

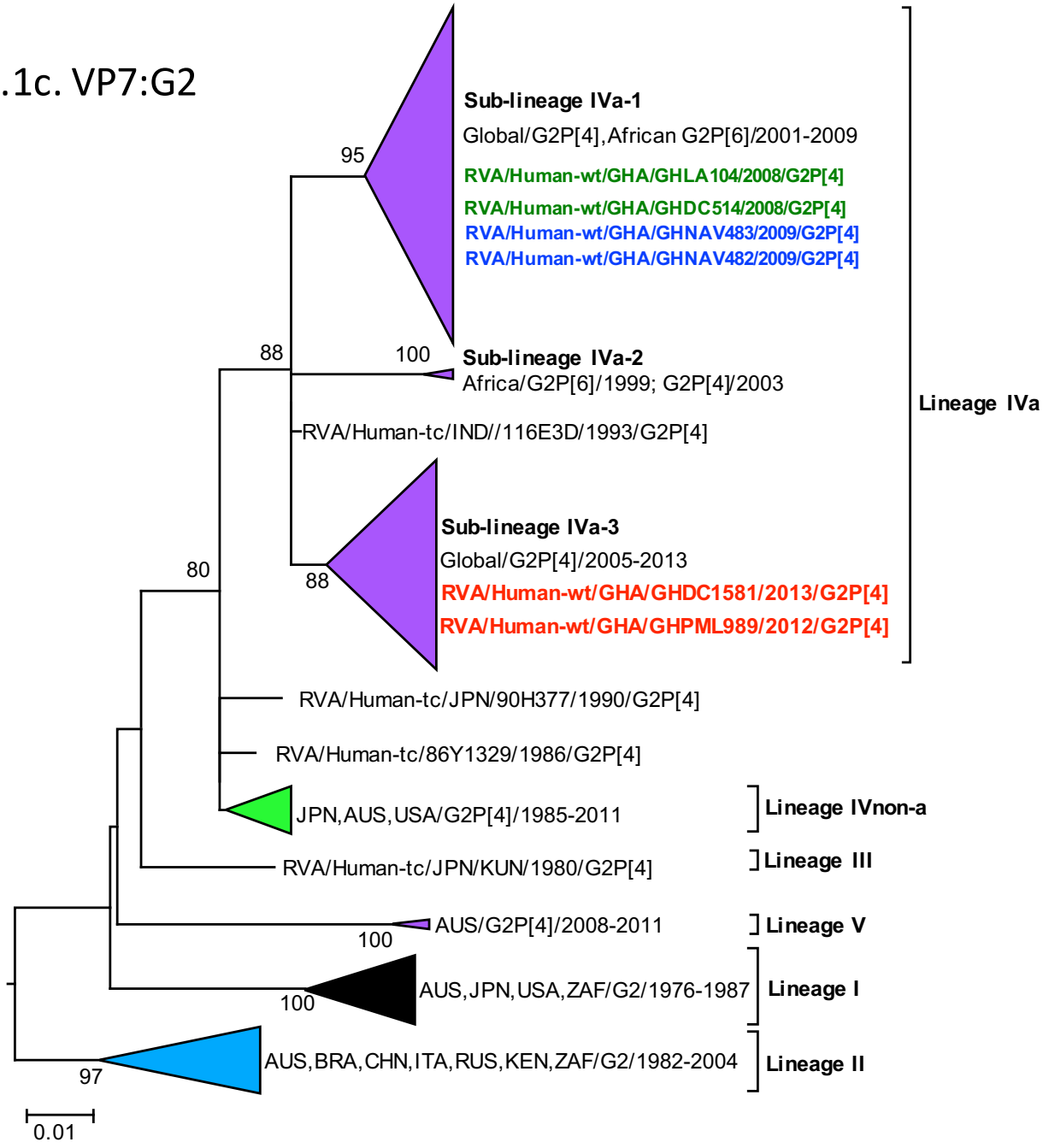


Fig. 3.1d. VP4:P[4]

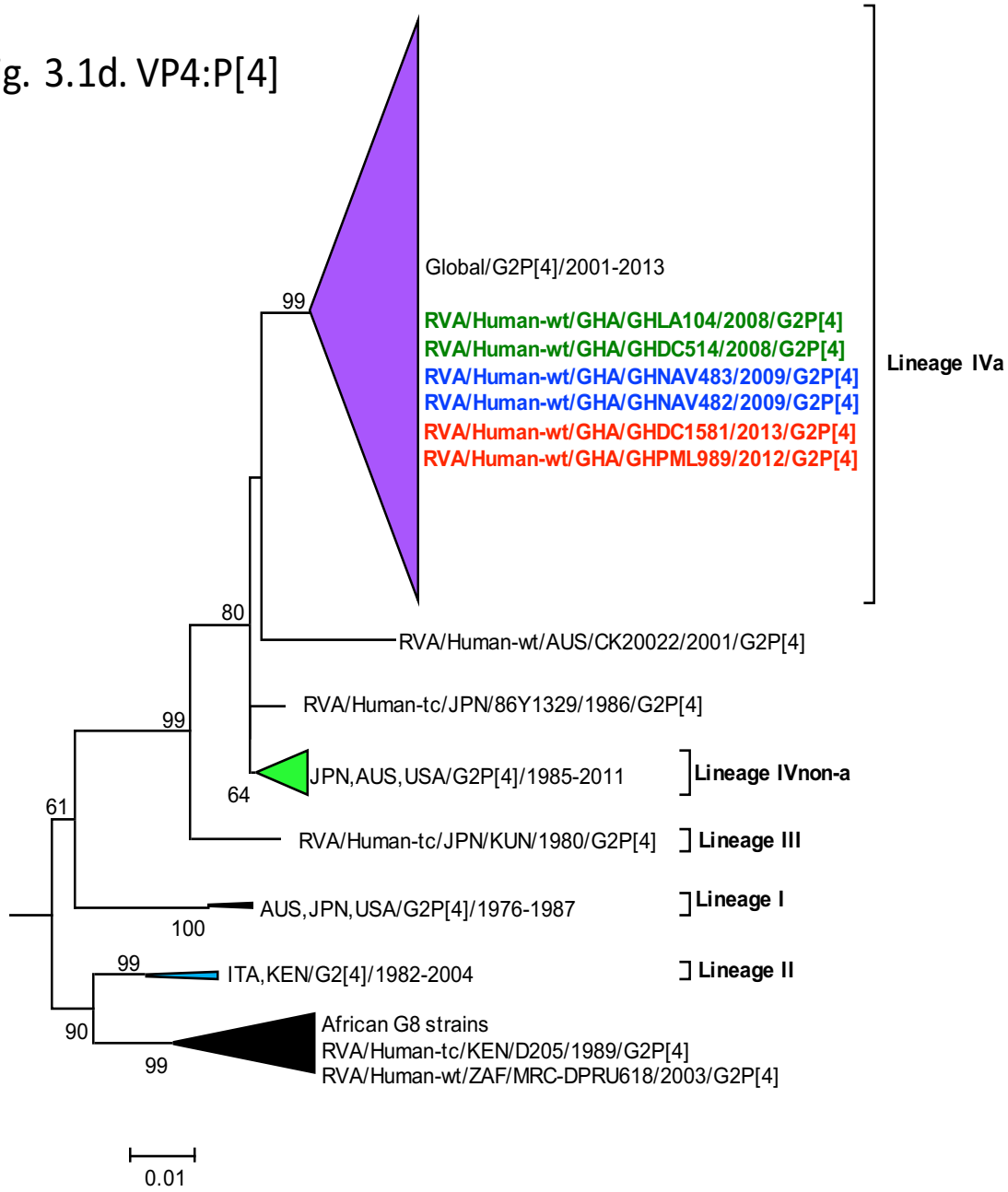


Table 3.2 Lineage constellation at a glance of African G2 RVA strains (1982-2013)

Year	Strain	Vaccine use period	Country	Lineage constellation											Reference
				VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5	
1982	AK26	Before	Kenya	II	II	II	II	II	II	Dist	N1	II	Dist	II	Ghosh et al., 2011
1984	1296GR	Before	South Africa	I	-	-	-	-	-	-	-	-	-	-	Page and Steele, 2004
1985	410GR	Before	South Africa	I	-	-	-	-	-	-	-	-	-	-	Page and Steele, 2004
1986	659GR	Before	South Africa	I	-	-	-	-	-	-	-	-	-	-	Page and Steele, 2004
1987	405GR	Before	South Africa	I	-	-	-	-	-	-	-	-	-	-	Page and Steele, 2004
1987	514GR	Before	South Africa	I	-	-	-	-	-	-	-	-	-	-	Page and Steele, 2004
1989	D205	Before	Kenya	II	II	Dist	II	II	Dist	IVnon-a	V	IV	II	II	Ghosh et al., 2011
1989	7PE	Before	South Africa	II	-	-	-	-	-	-	-	-	-	-	Page and Steele, 2004
1993	1831GR	Before	South Africa	II	-	-	-	-	-	-	-	-	-	-	Page and Steele, 2004
1996	64SB	Before	South Africa	II	-	-	-	-	-	-	-	-	-	-	Page and Steele, 2004
1998	906SB	Before	South Africa	II	-	-	-	-	-	-	-	-	-	-	Page and Steele, 2004
2003	MRC-DPRU618	Before	South Africa	IVa (IVa-2)	II	V	Dist	IVa	V	IVa	V	V	V	IVa	GenBank data, 2014
2007	MRC-DPRU81	Before	South Africa	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	VII	IVa	Nyaga et al., 2014
2007	MRC-DPRU1362	Before	South Africa	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	VII	IVa	Nyaga et al., 2014
2007	Ghan-086	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	V	IVa	Heylen et al., 2015
2007	Ghan-013	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	V	IVa	Heylen et al., 2015
2007	Ghan-148	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	V	IVa	Heylen et al., 2015
2008	Ghan-002	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	V	IVa	Heylen et al., 2015
2008	Ghan-004	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	V	IVa	Heylen et al., 2015
2008	Ghan-085	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	V	IVa	Heylen et al., 2015
2008	GHLA104	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	V	IVa	This study
2008	GHDC514	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	V	IVa	This study
2009	3203WC	Before	South Africa	IVa (IVa-3)	IVa	V	V	IVa	V	IVa	V	V	V	IVa	Jere et al., 2011
2009	MRC-DPRU1673	Before	Zambia	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	VII	IVa	GenBank data, 2015
2009	MRC-DPRU1061	Before	South Africa	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	VII	IVa	Nyaga et al., 2014
2009	Ghan-008	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	X	IVa	Heylen et al., 2015
2009	Ghan-010	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	X	IVa	Heylen et al., 2015
2009	Ghan-011	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	X	IVa	Heylen et al., 2015
2009	Ghan-012	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	X	IVa	Heylen et al., 2015
2009	Ghan-054	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	X	IVa	Heylen et al., 2015
2009	GHNAV-482	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	X	IVa	This study
2009	GHNAV-483	Before	Ghana	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	X	IVa	This study
2009	MRC-DPRU3710	Before	Uganda	IVa (IVa-3)	IVa	V	V	IVa	V	IVa	V	V	VI	IVa	GenBank data, 2015
2010	MRC-DPRU3199	Before	Gambia	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	IX	IVa	GenBank data, 2015
2010	MRC-DPRU2201	Before	Togo	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	IX	IVa	Nyaga et al., 2014
2010	MRC-DPRU5124	Before	Togo	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	IX	IVa	Nyaga et al., 2014
2010	BA366	Before	Cameroon	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	IX	IVa	Ndze et al., 2014
2010	BA368	Before	Cameroon	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	IX	IVa	Ndze et al., 2014
2011	MRC-DPRU5594	After	South Africa	IVa (IVa-1)	IVa	V	V	IVa	V	IVa	V	V	IX	IVa	GenBank data, 2015
2012	MRC-DPRU82	After	South Africa	IVa (IVa-1)	II	V	V	IVa	V	IVa	V	V	VII	IVa	Nyaga et al., 2014
2012	GHPML1989	After	Ghana	IVa (IVa-3)	IVa	V	V	IVa	VII	IVa	V	V	VI	IVa	This study
2013	GHDC1851	After	Ghana	IVa (IVa-3)	IVa	V	V	IVa	VII	IVa	V	V	VI	IVa	This study
1999	MRC-DPRU1815	Before	South Africa	IVa (IVa-2)	P[6]	V	Dist	IVa	V	IVa	V	V	V	IVa	GenBank data, 2014
1999	MRC-DPRU1845	Before	South Africa	IVa (IVa-2)	P[6]	V	Dist	IVa	V	IVa	V	V	V	IVa	GenBank data, 2014
1999	MRC-DPRU1818	Before	Ghana	IVa (IVa-2)	P[6]	V	Dist	IVa	V	IVa	V	V	V	IVa	GenBank data, 2014
2008	Ghan-052	Before	Ghana	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	V	V	V	IVa	Heylen et al., 2015
2008	MRC-DPRU2344	Before	South Africa	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	V	V	VII	IVa	GenBank data, 2015
2008	Mali-028	Before	Mali	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	Heylen et al., 2015
2008	Mali-029	Before	Mali	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	Heylen et al., 2015
2008	Mali-030	Before	Mali	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	Heylen et al., 2015
2008	Mali-035	Before	Mali	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	Heylen et al., 2015
2008	Mali-036	Before	Mali	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	Heylen et al., 2015
2008	Mali-072	Before	Mali	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	Heylen et al., 2015
2008	Mali-104	Before	Mali	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	Heylen et al., 2015
2009	Ghan-009	Before	Ghana	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	X	IVa	Heylen et al., 2015
2009	Ghan-053	Before	Ghana	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	Heylen et al., 2015
2009	Ghan-060	Before	Ghana	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	X	IVa	Heylen et al., 2015
2009	Ghan-108	Before	Ghana	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	X	IVa	Heylen et al., 2015
2009	Mali-045	Before	Mali	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	Heylen et al., 2015
2009	Mali-070	Before	Mali	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	Heylen et al., 2015
2009	Mali-071	Before	Mali	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	Heylen et al., 2015
2009	Mali-074	Before	Mali	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	Heylen et al., 2015
2010	MRC-DPRU3180	Before	Gambia	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	Dist	V	V	IVa	GenBank data, 2015
2011	MA104	Before	Cameroon	IVa (IVa-1)	P[6]	V	V	IVa	V	IVa	V	V	IX	IVa	Ndze et al., 2014

Lineage colouring according to phylogenetic trees; Dist: Distinct; -: Sequence data not available

Table 3.3 Comparison of genetic distances within and between NSP4 gene lineages

Lineage	I	II	III	IV	V	VI	VII	VIII	IX	X
I	1.6									
II	6.1	5.7								
III	4.8	5.2								
IV	5.1	5.8	3.7	1.3						
V	5.9	6.5	4.9	3.5	2.5					
VI	10.7	10.4	10.1	10.8	11.5	1.5				
VII	10.5	11.1	10.3	11.8	11.7	10.8	1.4			
VIII	9.5	10.6	9.1	10.3	10.5	7.5	9.6	0.5		
IX	15.6	14.4	15.6	15.3	15.2	13.7	12.2	12.0	0.6	
X	13.7	13.1	14.8	13.6	13.5	12.6	10.1	11.0	10.5	0.6

Key

Evolutionary rate of genes and years of divergence of designated lineages

All the 11 genome segments evolved at relatively similar rates ranging from 0.85 x 10^{-3} - 1.26 x 10^{-3} with overlapping 95% Highest Posterior Density (HPD) intervals ranging from 0.69 x 10^{-3} - 1.63 x 10^{-3} (Fig. 3.2a). A varying tMRCA for lineage IVa/V to which most genes of our study strains belonged was observed ranging from 1977 (VP1) to 1992 (NSP5) (Fig. 3.2b). Also, the tMRCA of emergent lineages VI, VII (VP3, NSP4) and VIII (NSP4) ranged from 2002 to 2005 whereas tMRCA for the African E2 specific lineages IX and X were more recent and around 2007-2008 (Fig. 3.2b, Fig. 3.3). Lineage X NSP4 sequences diverged from a cluster of African G8 NSP4 sequences speculated in this study to be the progenitors of lineage X sequences around 1991 (HPD interval: 1983-1997) (Fig. 3.3).

Fig. 3.2a Evolutionary rates

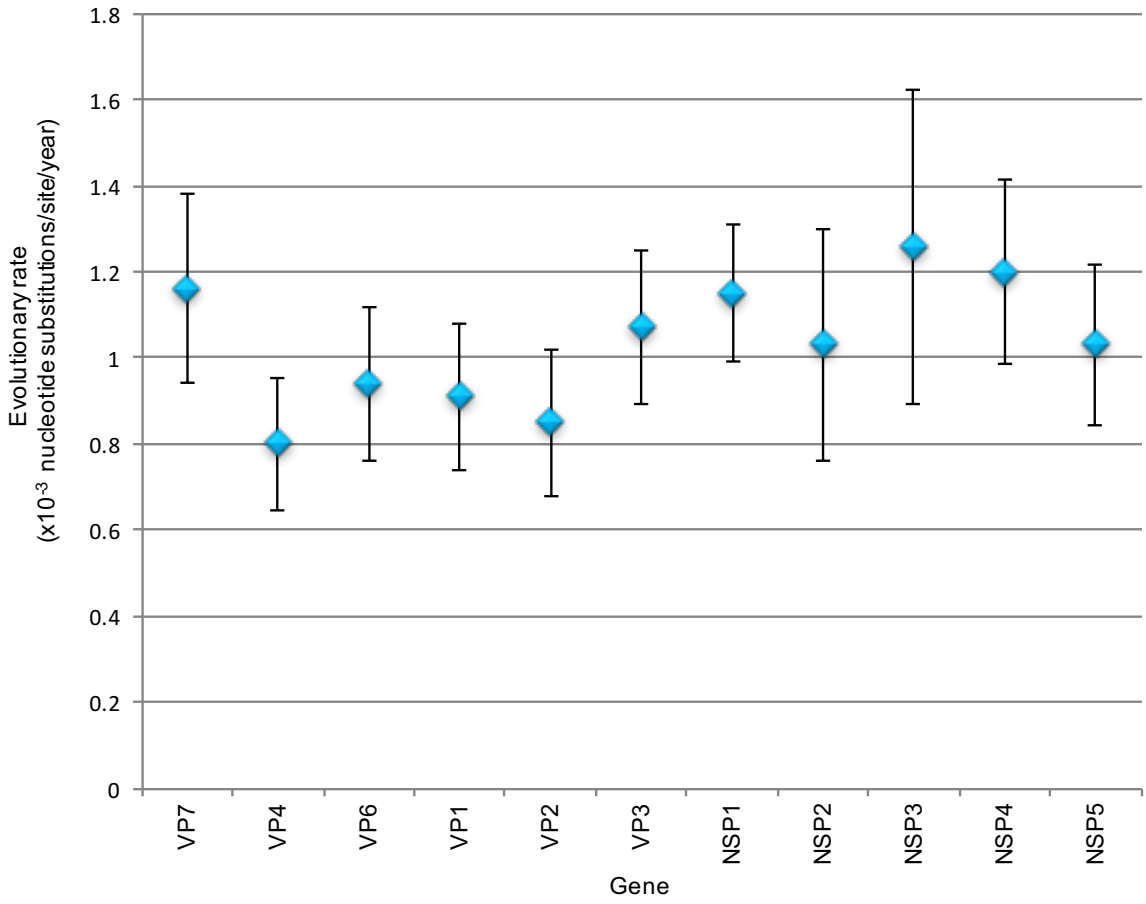


Fig. 3.2: Comparison of evolutionary rates of genome segments **(a)** and tMRCA of lineages **(b)** for the 11 genome segments of genotype 2 genes analyzed in this study. The nucleotide substitution rate per site per year of each gene and tMRCA of lineages were obtained using the Bayesian Markov chain Monte Carlo method implemented in BEAST v1.8.1 under the uncorrelated lognormal relaxed clock model. Each gene's dataset comprised at least 207 taxa sampled globally and strain detection period included 1976 to 2013. The mean substitution rates for each genome segment (light blue diamonds) and tMRCA for lineages (deep blue diamonds) are shown with the 95% highest posterior density intervals indicated as vertical error bars. Emergent lineages are indicated with curly brackets.

Fig. 3.2b The tMRCA of designated lineages

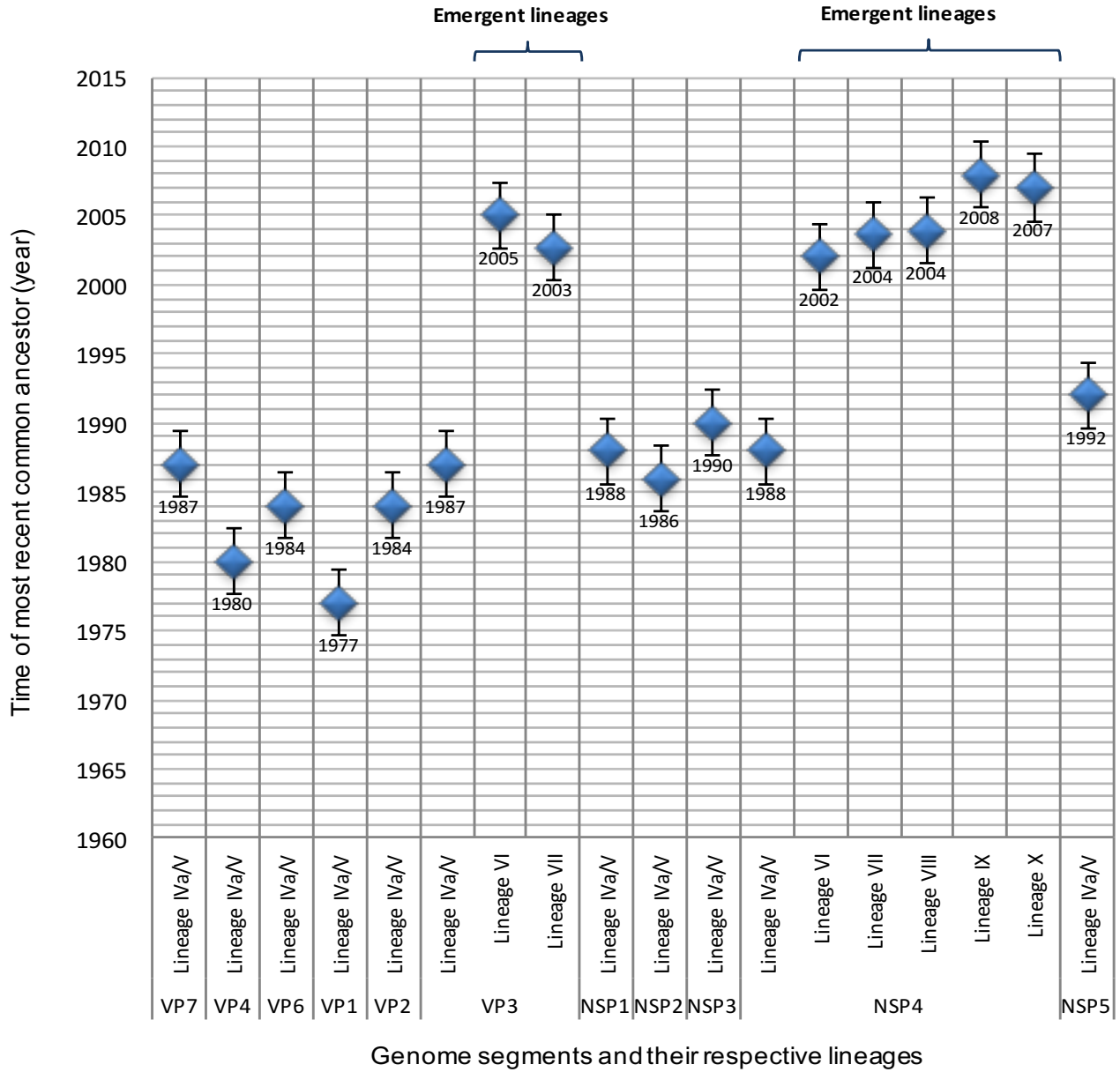


Fig. 3.3 Maximum clade credibility tree of NSP4 gene

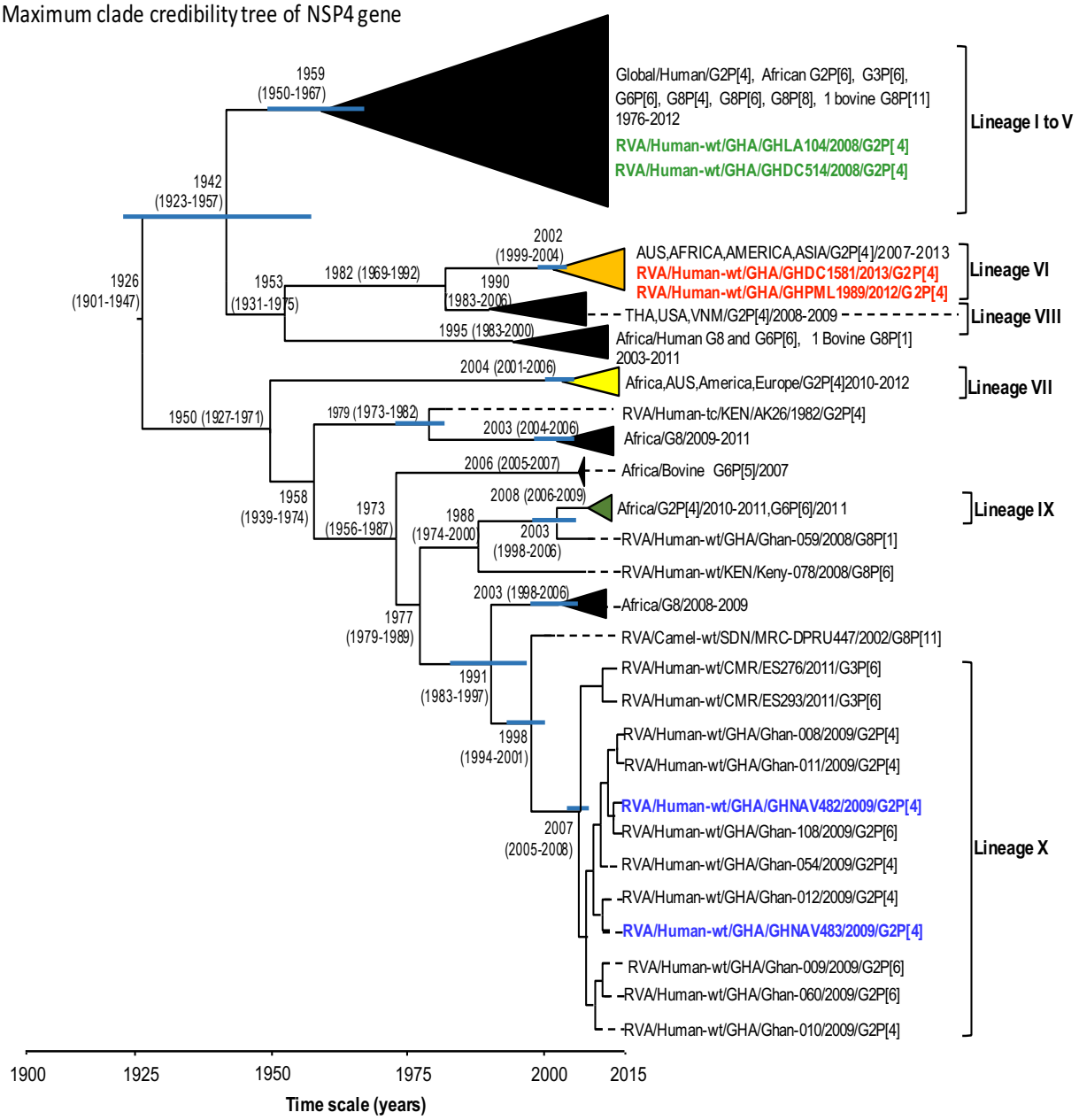


Fig. 3.3: A Simplified maximum clade credibility tree of 300 dated E2 NSP4 nucleotide sequences including those of the six Ghanaian G2P[4] strains described in this study (indicated in green, blue (pre vaccine period) and red font (post vaccine period) constructed using the Bayesian Markov chain Monte Carlo method implemented in BEAST v1.8.1. The years of divergence and the time to the most recent common ancestor of lineages are indicated at the nodes. The 95% highest posterior density interval of each significant node is indicated with blue bars. Lineages apart from lineage X have been collapsed into triangles for simplicity. The horizontal axis at the bottom of the figure indicates time scale in years.

3.5 DISCUSSION

This study comprehensively examined within a global context the whole genome evolution of G2P[4] RVA strains detected after 2000 during rotavirus seasons before and after Rotarix vaccine introduction in an African country — Ghana. Keeping in mind the general notion of frequent RVA interspecies transmission events occurring in Africa, our extensive phylogenetic analyses revealed that the genes of the study strains which were typically DS-1-like, belonged to at least one of the human G2P[4] RVA lineages proposed by Doan et al. (2015) except the NSP4 gene of the 2009 strains. Specifically, there was a change from lineage V to VII in the VP3 gene after vaccine introduction. The NSP4 gene of the study strains was most diverse and strains detected before vaccine introduction belonged to lineages V (2008) and a new lineage X in this study (2009), whereas those detected after vaccine introduction (2012-2013) belonged to lineage VI. We also noted that the NSP4 gene diversity with multiple emergent lineages in African G2P[4] strains could be explained by intra-genotype reassortment events involving co-circulating human DS-1-like strains bearing G2[6], G3P[6], and G6[6]. The remaining genes, VP7, VP4, VP2, NSP1, and NSP5 genes belonged to lineage IVa whereas the VP6, VP1, NSP2, and NSP3 genes belonged to lineage V irrespective of the period of detection.

An important evolutionary mechanism suggested to be used by rotaviruses to adapt to different immunological environments is lineage replacement (Jere et al., 2017; McDonald et al., 2009; Roy et al., 2014; Zeller et al., 2015). As selective pressure on the rotavirus genome due to vaccine use at early stages of vaccine introduction may be subtle, changes may only be apparent after many years of monitoring. Studies examining the impact of rotavirus vaccine use on rotavirus genome diversity and evolution are limited as most vaccination programs were established relatively recently; however, available evidence from the following studies provides insight into what impact the introduction of vaccines might have on rotavirus genome diversity and evolution. In Nicaragua, a country that vaccinates

with Rotateq, Bucardo et al. (2012) provided evidence on the detection of an NSP2 gene identical to that of Rotateq in the genome of G1P[8] strains detected in two children. This finding suggests the possibility that genome segments from vaccine strains are likely to get reassorted into wild-type circulating human rotaviruses resulting in novel gene pools of circulating strains. In a recent study conducted by Zeller et al. (2015) on Australian and Belgian samples collected pre- and post-vaccine introduction (between 1999 and 2011), there was no evidence of changes in rotavirus population sizes after vaccine introduction. Of note were different phylogenetic clusters after vaccine introduction but these changes were ascribed to either natural fluctuations in circulating lineages or subtle signs of vaccine-driven evolution. Also, six amino acid sites in the structural proteins VP2, VP3 and VP7 and the non-structural protein NSP1 were identified to be under positive selection pressure. Furthermore, Jere et al. (2017) upon analysis of whole genomes of G1P[8] strains detected in Malawi before (1998-2012) and after (2013-2014) nationwide vaccine introduction, reported the emergence of atypical DS-1-like G1P[8] strains during the post-vaccine era. Also, three distinct G1P[8] lineages were reported to have circulated chronologically and lineage turnover were driven by point mutations and genetic reassortment events. The emergence of the DS-1-like G1P[8] strains were interpreted to be unrelated to vaccine use but likely due to natural strain evolutionary pressure.

The theoretical basis of the genetic diversity in G2P[4] strains has been explained by point mutations and genetic reassortment with other DS-1-like RVA strains (Doan et al., 2015) but evidence about the acquisition of genes from animal RVA strains by G2P[4] strains has been controversial (Dennis et al., 2014; Doan et al., 2015; Ghosh et al., 2011c; Giammanco et al., 2014). Recently, Do et al. (2015) provided evidence for the sharing of common ancestors by the NSP2 genes of Vietnamese G2P[4] strains and ruminant RVA strains, suggesting the involvement of interspecies transmission events in the evolution of G2P[4] strains. Contrary to the general notion of frequent RVA interspecies transmission

events in Africa, we did not detect any trace of animal RVA sequences among the genome segments of the Ghanaian G2P[4] strains.

Large sequence diversity in the NSP4 genes of African DS-1-like strains compared to the other 10 RVA genome segments was reported by Nyaga et al. (2014). Regarding G2P[4] strains, Do et al. (2015) recently found an additional lineage IX in the NSP4 gene which was thus far African specific. In this study, the lineage X is another evidence supporting the continuous diversification of the NSP4 gene at the lineage level in African G2P[4] strains. We found that the immediate ancestral sequence of the additional NSP4 lineages were not of a recent direct animal RVA origin. Given the abundance of non-G2P[4] strains possessing the DS-1-like genetic backbone on the African continent (Table 3.4) (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) and mixed infections in Africa (Mwenda et al., 2010; Nyaga et al., 2015; Sanchez-Padilla et al., 2009; Seheri et al., 2014; Todd et al., 2010), these strains could serve as the donors of genotype 2 genes to G2P[4] strains in Africa. An unexplainable observation is why this frequent reassortment was not seen in the other 10 genome segments of African G2P[4] strains.

In the NSP4 phylogenetic trees (Fig. 3.1a, Fig. 3.3), further downstream from the root of lineage X, at the node, a divergence event occurred (indicated with an asterisk) and the common ancestral sequence was shared with the NSP4 gene of an African camel G8P[11] strain MRC-DPRU447. However, camel RVA origin of this lineage is disputed because this camel RVA strain was reported to have emerged from multiple reassortment events involving several mammalian rotavirus strains (Jere et al., 2014). On the other hand, the 96% bootstrap support at the downstream node to the ancestor of the camel RVA and lineage X sequences suggested that the lineage X sequences originated from non-G2P[4] strains most probably human G8 RVA strains (Fig. 3.1a, Fig. 3.3). A similar observation was

made about the relationship between the African specific lineage IX NSP4 sequences which also shared a common ancestor with a Ghanaian human G8P[1] strain Ghan-059 around the year 2003 (Fig. 3.3). While one might argue that the origin of lineage IX and X NSP4 sequences was bovine RVA (Fig. 3.1a), the bootstrap value for the cluster was not significant enough (node marked with double asterisk; value below 70% hence not shown) to support this assertion. Additional evidence of intra-genotype reassortment in the generation of emergent NSP4 lineage X and VI, VII, VIII, IX is that the nucleotide sequence divergence of these lineages from lineages I to V were relatively high reaching over 15% (Table 3.3, Fig. 3.2). Given the recent years of detection of strains in the emergent lineages coupled with the evolutionary rate of the NSP4 gene (Fig. 3.2a), it is unlikely that accumulated point mutations accounted for the high genetic diversity.

The tMRCA of lineage IVa/V (1977-1992) revealed that most genome segments evolved within the same lineage with a similar level of accumulation of point mutations resulting in a co-evolving pattern of lineages across the genome whereas lineages VI and VII in the VP3 gene and lineages IX and X in the NSP4 gene were recently introduced by reassortment (Fig. 3.2b). The similar evolutionary rates for all the genome segments (Fig. 3.2a) suggest that the sequence variations over time occurred at similar rates across the genome and the acquisition of genotype 2 genes from non-G2P[4] DS-1-like strains (in the NSP4 and VP3) did not cause a significant change in the rate of evolution compared to the other genes.

Table 3.4: Genotype constellation of non-G2P[4] strains with DS-1-like backbone genes that were detected in Africa

Detection Year	Strain	Country	VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5	Reference
1997	MW1-006	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
1997	MW1-333	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
1997	MW1-023	Malawi	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
1997	MW1-131	Malawi	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
1998	OP2-506	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
1998	MW1-467	Malawi	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
1998	NeO2-007	Malawi	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
1998	NeO2-025	Malawi	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
1999	MRC-DPRU1815	South Africa	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	GenBank data, 2014
1999	MRC-DPRU1845	South Africa	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	GenBank data, 2014
1999	MRC-DPRU1818	Ghana	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	GenBank data, 2014
1999	MW1-860	Malawi	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
1999	MW2-026	Malawi	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2000	MW2-489	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2001	MW2-624	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2001	OP2-384	Malawi	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2002	MW2-924	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2003	DRC86	DRC	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Matthijnssens et al., 2006
2003	DRC88	DRC	G8	P[8]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Matthijnssens et al., 2006
2003	OP2-668	Malawi	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2004	MW2-1114	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2004	MW2-1189	Malawi	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2005	MW2-1238	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2005	MW2-1246	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2005	QOP002	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2005	QEC29	Malawi	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2006	QEC257	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2006	QEC287	Malawi	G8	P[8]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2006	QEC289	Malawi	G8	P[8]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2007	QOP250	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2007	QOP340	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2007	QOP387	Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
2008	GH018-08	Ghana	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H3	Dennis et al., 2014
2008	GH019-08	Ghana	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H3	Dennis et al., 2014
2008	Ghan-052	Ghana	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2008	MRC-DPRU2344	South Africa	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	GenBank data, 2015
2008	Mali-028	Mali	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2008	Mali-029	Mali	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2008	Mali-030	Mali	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2008	Mali-035	Mali	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2008	Mali-036	Mali	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2008	Mali-072	Mali	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015

Detection Year	Strain	Country	VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5	Reference
2008	Mali-104	Mali	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2008	Ghan-113	Ghana	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H3	Heylen et al., 2015
2008	Ghan-149	Ghana	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H3	Heylen et al., 2015
2008	Mali-039	Mali	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2008	Mali-048	Mali	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2008	Mali-119	Mali	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2008	Mali-135	Mali	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2008	Keny-078	Kenya	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2008	MRCDPRU1621	Zambia	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nyaga et al., 2014
2008	MRCDPRU1922	Uganda	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nyaga et al., 2014
2009	Ghan-009	Ghana	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Ghan-053	Ghana	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Ghan-060	Ghana	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Ghan-108	Ghana	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Mali-045	Mali	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Mali-070	Mali	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Mali-071	Mali	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Mali-074	Mali	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Ghan-055	Ghana	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Mali-120	Mali	G1	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Ghan-006	Ghana	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Ghan-007	Ghana	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Ghan-106	Ghana	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Ghan-107	Ghana	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	Ghan-056	Ghana	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2015
2009	MRCDPRU1606	Kenya	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nyaga et al., 2014
2009	MRCDPRU3463	Zambia	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nyaga et al., 2014
2010	KisB554	DRC	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2014
2010	KisB565	DRC	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Heylen et al., 2014
2010	MRC-DPRU3180	Gambia	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	GenBank data, 2015
2010	MRCDPRU3347	Zimbabwe	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nyaga et al., 2014
2010	MRCDPRU4576	Tanzania	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nyaga et al., 2014
2010	238-BF	Burkina Faso	G6	P[6]	I2	[redacted]			A2	N2	T2	E2	H2	Nordgren et al., 2012b
2010	263-BF	Burkina Faso	G6	P[6]	I2				A2	N2	T2	E2	H2	Nordgren et al., 2012b
2010	265-BF	Burkina Faso	G6	P[6]	I2				A2	N2	T2	E2	H2	Nordgren et al., 2012b
2010	272-BF	Burkina Faso	G6	P[6]	I2				A2	N2	T2	E2	H2	Nordgren et al., 2012b

Detection Year	Strain	Country	VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5	Reference
2011	MRC-DPRU5608	Guinea Bissau	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	GenBank data, 2014
2011	MRC-DPRU5615	Guinea Bissau	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	GenBank data, 2014
2011	MRC-DPRU5625	Guinea Bissau	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	GenBank data, 2014
2011	MA114	Cameroon	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	MA155	Cameroon	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	ES276	Cameroon	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	ES293	Cameroon	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	MA202	Cameroon	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	MA228	Cameroon	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	ES298	Cameroon	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	BA346	Cameroon	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	BA369	Cameroon	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	MRCDPRU4568	Tanzania	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nyaga et al., 2014
2011	MA104	Cameroon	G2	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	MA109	Cameroon	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	MA114	Cameroon	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	MA155	Cameroon	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	ES276	Cameroon	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	ES293	Cameroon	G3	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	MA202	Cameroon	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	MA228	Cameroon	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	ES298	Cameroon	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	BA346	Cameroon	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	BA369	Cameroon	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Ndze et al., 2014
2011	MRCDPRU4568	Tanzania	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nyaga et al., 2014
2012	PML1965	Ghana	G6	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Agbemabiese et al., 2015

In summary, this study provided evidence that the currently circulating G2P[4] strains in Ghana and in Africa are similar to those detected globally except the NSP4 gene of the Ghanaian 2009 strains and some of the contemporary African G2P[4] strains which belonged to African specific lineages IX and X. Although changes occurred at the lineage level in the VP3 and NSP4 genes of the study strains after vaccine introduction, it is too early to draw conclusions as these changes were observed not long after the countrywide vaccine introduction. The evolutionary rate of the genes was similar across the whole genome and the African specific lineages emerged through genetic reassortment with co-circulating human DS-1-like strains bearing G2[6], G3P[6], and G6[6]. It however remains unknown why the intra-genotype reassortment event observed in the NSP4 gene was not equally frequent in the other genome segments. It will be worthwhile to understand what characteristics enable certain genes of RVA strains to be reassorted into other RVA strains while the virus still maintains its replicative ability and continues the transmission chain within the population.

3.6 ACKNOWLEDGEMENT

We acknowledge the immense support of the Program for Nurturing Global Leaders in Tropical and Emerging Communicable Diseases, Graduate School of Biomedical Sciences, Nagasaki University. We also thank the staff of the Regional Rotavirus Reference Laboratory at the Noguchi Memorial Institute for Medical Research in Ghana for providing samples for the study and the preliminary assays.

3.7 CONFLICT OF INTEREST

The authors declare that they have no conflict of interest

Chapter IV

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

Published in:

Chantal Ama Agbemabiese, Toyoko Nakagomi, Yen Hai Doan, Osamu Nakagomi.

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

4.1 SUMMARY

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

Key words: rotavirus; G8; genotype constellation; ruminant rotavirus; phylogenetic analysis; reassortment.

4.2 INTRODUCTION

In cattle, RVA genotype G8 accounts for about 3.5% of rotavirus diarrhoeal infections (Papp et al., 2013) and have also been detected in other less explored animal host species such as guanacos, sheep, goats and camels (Ciarlet et al., 2008; Jere et al., 2014; Louge Uriarte et al., 2014; Sieg et al., 2015). In humans, the first G8 RVA strain 69M bearing a DS-1-like genotype constellation was detected in an Indonesian child in 1980 (Matsuno et al., 1985). Since then, a large number of G8 RVA strains bearing either DS-1-like or Wa-like genotype constellations were detected in children on the African continent (Adah et al., 2003; Adah et al., 2001; Dennis et al., 2014; Esona et al., 2009; Ghosh et al., 2011b; Heylen et al., 2014; Heylen et al., 2015; Istrate et al., 2015; Matthijnssens et al., 2006) and recently in large numbers in a few countries in Asia (Hoa-Tran et al., 2016; Kondo et al., 2017; Tacharoenmuang et al., 2016) but only sporadically in Europe and the Americas (Banyai et al., 2010; Delogu et al., 2013; Gautam et al., 2015; Ianiro et al., 2014; Pietsch et al., 2009).

In Japan, the G8 RVA strain is one of the commonly detected bovine rotaviruses (Fukai et al., 2004a; Fukai et al., 2004b). A single G8P[4] RVA strain designated AU109 was for the first time detected during the 1993-1994 rotavirus season in a hospital-based study conducted between 1987-1996 in Akita prefecture (Nakagomi et al., 2009). In 2010, Banyai and colleagues published a Hungarian human G8P[14] (Banyai et al., 2010) RVA strain whose VP7 gene was closely related to that of AU109 and whose genotype constellation was identical to that of a Spanish sheep G8 RVA strain described by (Ciarlet et al., 2008) . In this regard, the whole genome sequence of AU109 was determined to obtain insight into how it relates to the previously reported bovine G8 strains in Japan as well as those reported from humans and other host species elsewhere in the world.

4.3. MATERIALS AND METHODS

Rotavirus strain

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

Whole genome amplification and sequencing

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

Briefly, an initial reaction mixture consisting of viral RNA and 50 mM random primers was denatured at 97°C for 5 minutes and quickly chilled on ice for 5 minutes. After addition of a reverse transcription reaction mix containing 20 µM dNTP and 20 U/µL SuperScript™ III reverse transcriptase to the reaction mix to make up a final volume of 20 µL, cDNA was synthesised under the following condition: 25°C for 5 min; 42°C for 60 min; 70°C for 15 min. With the exception of the VP7 gene whose sequence was already deposited in the GenBank (Nakagomi et al., 2009) under the accession number AB272753, the remaining 10 genome segments were amplified by PCR from 2 µL of cDNA using gene specific primers (Supplementary Table 4.1) (Doan et al., 2012; Gentsch et al., 1992; Gouvea et al., 1990; Matthijnssens et al., 2008a) and GoTaq® Green Master Mix System (Promega) under the following conditions: 95°C/5 min followed by 40 cycles of PCR at 94°C/1 min; 45°C/1 min; 72°C/2-6 min depending on the size of the gene and final extension at 72°C/8 min.

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

Sequence and Phylogenetic Analyses

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

Bayesian evolutionary analysis using BEAST

The time of most recent common ancestor (tMRCA) was determined for the VP7, VP1 and NSP2 genes of AU109. The sequence datasets were the same as those used for

the maximum likelihood phylogenetic analysis except that sequences of strains whose year of detection were not available were omitted. The final datasets for the VP7, VP1 and NSP2 genes consisted of 187, 207 and 214 dated sequence respectively. The Bayesian Markov chain Monte Carlo (MCMC) framework implemented in BEAST v1.8.1 (Drummond et al., 2012) was employed. Briefly, using the following substitution models: VP7 and NSP2 (TN93+G), VP1 (GTR+G), a lognormal relaxed clock (Drummond et al., 2006) and a coalescent constant size (Drummond et al., 2002) model, two independent MCMC runs were carried out for 50 million generations for each genome segment and evaluated using Tracer software v1.6 (http://tree.bio.ed.ac.uk/software/tracer/). The MCMC runs were combined using a LogCombiner v.1.8.1 to achieve effective sampling sizes of $\geq$ 200. Maximum clade credibility trees were annotated with the Treeannotator and viewed with FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

Nucleotide sequence accession numbers

Nucleotide sequences were submitted to the GenBank under accession numbers LC065018-LC065027 (Supplementary Table 4.2).

4.4 RESULTS

Genotype constellation ***of AU109***

The complete length of each of the 11 genes were sequenced for AU109 (Supplementary Table 4.2). The genotype constellation of AU109 as determined by RotaC was G8-P[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2 (Table 4.1), a genetic background commonly associated with a vast majority of human G8 RVA strains in Africa (Nakagomi et al., 2013) .

Sequence and phylogenetic analysis

In a maximum likelihood phylogenetic tree constructed for 246 G8 VP7 genes of both human and animal origin, five lineages designated lineage I to V were identified according to our definition of a lineage as a collection of sequences in a cluster with bootstrap supports >80% (except for lineage III) (Fig. 4.1a). The VP7 gene of AU109 was located in a distinct lineage (Fig. 4.1a) together with only the Hungarian G8P[14] strain BP1062. AU109 and BP1062 VP7 gene differed at the nucleotide and amino acid sequence level by 3.5% and 3.6%, respectively (Table 4.2). This distinct lineage shared an ancestor with one of the five lineages we identified in this study, i.e. lineage IV G8 VP7 sequences which originated from both human and bovine G8 RVA strains, majority of which were detected in India with additional sporadic cases from Egypt, Taiwan and Thailand. The nucleotide sequences of the sequences in Lineage IV were at least 10.3% divergent from AU109 whereas their amino acid sequences were at most 6.4% divergent from AU109.

Lineage I contained 15 G8 VP7 sequences of human host species origin detected from 2000-2011; majority of which were detected in African countries and the remaining were from Croatia, Germany, Slovenia and the United States. (Delogu et al., 2013; Esona et al., 2009; Ianiro et al., 2014; Pietsch et al., 2009; Weinberg et al., 2012) Lineage II contained mainly bovine and porcine RVA strains, a few guanaco, goat, sheep, rhesus G8 sequences and human G8 RVA VP7 sequences described to be of animal RVA origin (Medici et al., 2008; Mukherjee et al., 2013). Lineage III supported by a rather low bootstrap value — 58% — contained mainly human RVA G8 detected in Indonesia, Australia, Finland and New Zealand together with a UK (678) and a South African (1604)

Table 4.1: Comparison of genotype constellation of AU109 with G8 RVA strains of diverse host species origin and geographical locations

Strain	Host	Country	Genotype constellation											Reference
			VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5	
RVA/Human-tc/JPN/AU109/1994	Human	Japan	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	**This study**
RVA/Human-wt/GER/GER1H/2009		Germany	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Pietsch et al., 2009
RVA/Human-tc/MWI/MW333/1997		Malawi	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
RVA/Human-wt/IND/mcs65/2011		India	G8	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H3	Mukherjee et al., 2013
RVA/Human-wt/IND/mcs72/2011		India	G8	P[4]	I2	R2	C2	M5	A2	N2	T2	E2	H3	Mukherjee et al., 2013
RVA/Human-wt/MWI/MW1-131/1997		Malawi	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
RVA/Human-wt/COD/DRC86/2003		DR Congo	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Matthijnssens et al., 2006
RVA/Human-wt/GHA/GH019-08/2008		Ghana	G8	P[6]	I2	R2	C2	M2	A2	N2	T2	E2	H3	Dennis et al., 2014
RVA/Human-wt/COD/DRC88/2003		DR Congo	G8	P[8]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Matthijnssens et al., 2006
RVA/Human-wt/MWI/QEC287/2006		Malawi	G8	P[8]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Nakagomi et al., 2013
RVA/Human-wt/CIV/6736/2004		Cote d'Ivoire	G8	P[8]	I1	R1	C1	M1	A1	N1	T1	E1	H1	Esona et al., 2009
RVA/Human-wt/CMR/6782/2000		Cameroon	G8	P[8]	I1	R1	C1	M1	A1	N1	T1	E1	H1	Esona et al., 2009
RVA/Human-wt/ETH/6810/2000		Ethiopia	G8	P[8]	I1	R1	C1	M1	A1	N1	T1	E1	H1	Esona et al., 2009
RVA/Human-wt/TUN/6854/2000		Tunisia	G8	P[8]	I1	R1	C1	M1	A1	N1	T1	E1	H1	Esona et al., 2009
RVA/Human-wt/HRV/CR2006/2006		Croatia	G8	P[8]	I1	R1	C1	M1	A1	N1	T1	E1	H1	Delogu et al., 2013
RVA/Human-tc/IDN/69M/1980		Indonesia	G8	P[10]	I2	R2	C2	M2	A2	N2	T2	E2	H2	Matthijnssens et al., 2008
RVA/Human-wt/HUN/BP1062/2004		Hungary	G8	P[14]	I2	R2	C2	M2	A11	N2	T6	E2	H3	Banyai et al., 2010
RVA/Human-wt/GTM/2009726790/2009		Guatemala	G8	P[14]	I2	R2	C2	M2	A3	N2	T6	E2	H3	Gautam et al., 2015
RVA/Human-wt/ITA/PR1300/2004		Italy	G8	P[14]	I2	R2	C2	M2	A3	N2	T6	E2	H3	Medici et al., 2015
RVA/Human-wt/KEN/B12/1987		Kenya	G8	P[1]	I2	R2	C2	M2	A3	N2	T6	E2	H3	Ghosh et al., 2011
RVA/Cow-wt/ZAF/1604/2007	Bovine	South Africa	G8	P[1]	I2	R2	C2	M2	A3	N2	T6	E2	H3	Jere et al., 2012
RVA/Rhesus-tc/USA/PTRV/1990	Rhesus	Unites States	G8	P[1]	I2	R2	C2	M2	A3	N2	T6	E2	H3	Matthijnssens et al., 2010
RVA/Goat-wt/ARG/0040/2011	Goat	Argentina	G8	P[1]	I2	R5	C2	M2	A3	N2	T6	E12	H3	Louge Uriarte et al., 2014
RVA/Guanaco-wt/ARG/Rio_Negro/1998	Guanaco	Argentina	G8	P[1]	I2	R5	C2	M2	A11	N2	T6	E12	H3	Matthijnssens et al., 2009
RVA/Dog-wt/GER/88977/2013	Dog	Germany	G8	P[1]	I2	-	C2	-	A3	N2	T6	E2	H3	Sieg et al., 2015
RVA/Camel-wt/SDN/MRC-DPRU447/2002	Camel	Sudan	G8	P[11]	I2	R2	C2	M2	A18	N2	T6	E2	H1	Jere et al., 2014
RVA/Guanaco-wt/ARG/Chubut/1999	Guanaco	Argentina	G8	P[14]	I2	R5	C2	M2	A11	N2	T6	E12	H3	Matthijnssens et al., 2009
RVA/Sheep-tc/ESP/OVR762/2002	Sheep	Spain	G8	P[14]	I2	R2	C2	M2	A11	N2	T6	E2	H3	Ciarlet et al., 2008

Fig. 4.1a: VP7

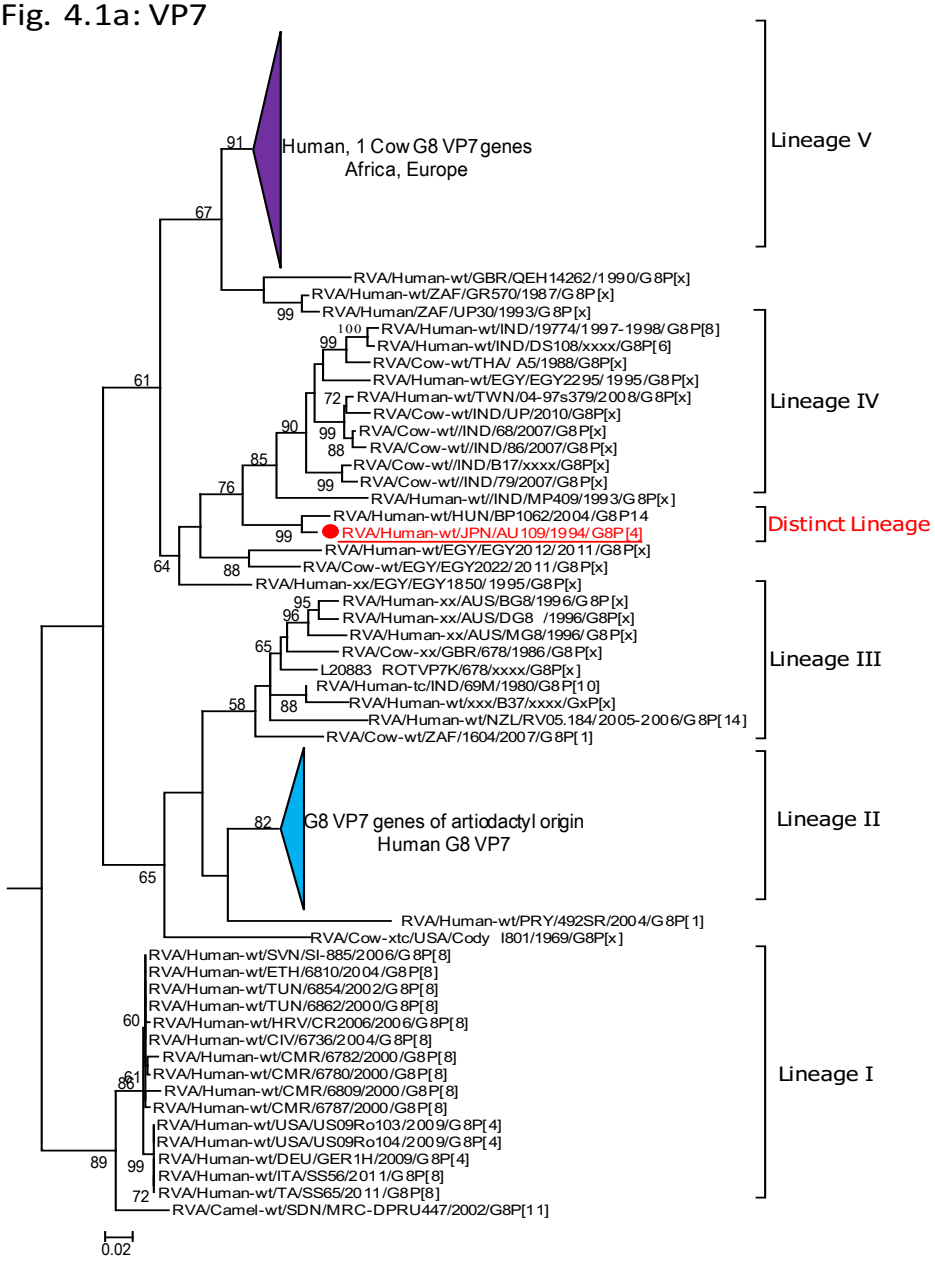
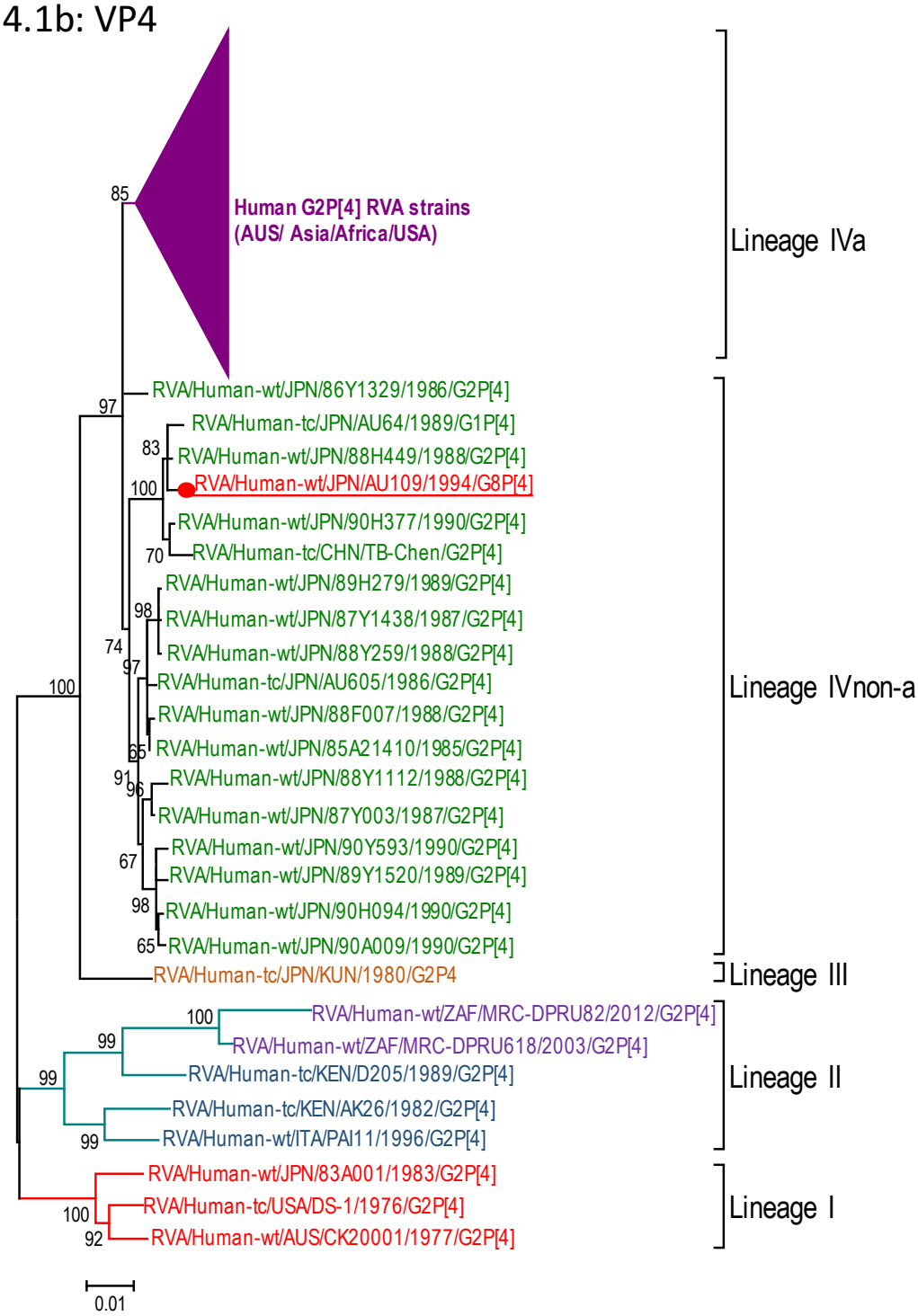


Fig. 4.1: Phylogenetic trees of the (a) VP7 gene of G8 RVA strains and (b) VP4 gene of P[4] RVA strains showing the genetic relationship between AU109 in this study and other animal and/or human RVA strains. For each phylogenetic tree, analyses included the nucleotide sequence of the Japanese human G8P[4] strain AU109 in this study (indicated in red font with a dot) and globally circulating strains bearing the same genotype as AU109 retrieved from the GenBank database. Maximum likelihood phylogenetic analyses were performed using Tamura 3-parameter substitution model with gamma distributed rate variation in MEGA6 software package. The VP7 tree was rooted using AU-1 VP7 gene the VP4 tree is presented here as a midpoint-rooted tree. Significant bootstrap values (1000 replicates) are indicated at each node. The scale bar at the bottom of the tree indicates a genetic distance expressed as nucleotide substitutions per site.

Fig. 4.1b: VP4


Human G2P[4] RVA strains
(AUS/ Asia/Africa/USA)
Lineage IVa
RVA/Human-wt/JPN/86Y1329/1986/G2P[4]
RVA/Human-tc/JPN/AU64/1989/G1P[4]
RVA/Human-wt/JPN/88H449/1988/G2P[4]
RVA/Human-wt/JPN/AU109/1994/G8P[4]
RVA/Human-wt/JPN/90H377/1990/G2P[4]
RVA/Human-tc/CHN/TB-Chen/G2P[4]
RVA/Human-wt/JPN/89H279/1989/G2P[4]
RVA/Human-wt/JPN/87Y1438/1987/G2P[4]
RVA/Human-wt/JPN/88Y259/1988/G2P[4]
RVA/Human-tc/JPN/AU605/1986/G2P[4]
RVA/Human-wt/JPN/88F007/1988/G2P[4]
RVA/Human-wt/JPN/85A21410/1985/G2P[4]
RVA/Human-wt/JPN/88Y1112/1988/G2P[4]
RVA/Human-wt/JPN/87Y003/1987/G2P[4]
RVA/Human-wt/JPN/90Y593/1990/G2P[4]
RVA/Human-wt/JPN/89Y1520/1989/G2P[4]
RVA/Human-wt/JPN/90H094/1990/G2P[4]
RVA/Human-wt/JPN/90A009/1990/G2P[4]
Lineage IVnon-a
RVA/Human-tc/JPN/KUN/1980/G2P4
Lineage III
RVA/Human-wt/ZAF/MRC-DPRU82/2012/G2P[4]
RVA/Human-wt/ZAF/MRC-DPRU618/2003/G2P[4]
RVA/Human-tc/KEN/D205/1989/G2P[4]
RVA/Human-tc/KEN/AK26/1982/G2P[4]
RVA/Human-wt/ITA/PAI11/1996/G2P[4]
Lineage II
RVA/Human-wt/JPN/83A001/1983/G2P[4]
RVA/Human-tc/USA/DS-1/1976/G2P[4]
RVA/Human-wt/AUS/CK20001/1977/G2P[4]
Lineage I
0.01

bovine strains. Lineage V contained mainly G8 RVA VP7 genes of human host species origin from Africa, Iraq, France and Germany. It also contained a single bovine G8 VP7 gene detected in Nigeria. The genetic background of majority of the strains in Lineage V with available information was DS-1-like which seemed to have been well adapted hence were able to spread well from human to human on the African continent.

When the other 10 genome segments of the AU109 were examined, close relationship with Hungarian BP1062 like the VP7 gene occurred only in the VP1 gene (Fig. 4.2a), and these two strains were rather distantly related in different lineages or even belonged to different genotypes in the VP4, NSP1, NSP3 and NSP5 genes (Table 4.1). On the other hand, except the VP1 and NSP2 genes, AU109 showed close relationship with its contemporary G2P[4] strains detected in Japan and China; 88H449, 90H377, and TB-Chen, in particular. Thus, in accordance with the recently-proposed lineage designation for globally circulating G2P[4] strains (Doan et al., 2015) AU109 belonged to the following lineages; lineage IVnon-a for VP4 (Fig. 4.1b), lineage IV for VP6, lineage IVnon-a for VP2, lineage IV for VP3, lineage IVnon-a for NSP1, lineage IV for NSP3, lineage IV for NSP4, and lineage IVnon-a for NSP5 (Supplementary Fig. 4.1a – g).

However, in the VP1 and NSP2 phylogenetic trees, AU109 did not share lineages with its contemporary Japanese G2P[4] strains. Instead, the VP1 gene of AU109 clustered with the Hungarian G8P[14] strain (BP1062), Chinese bovine, goat and sheep strains with a 93% bootstrap support (Fig. 4.2a). Within this lineage, AU109 was the closest to BP1062 with a nucleotide sequence divergence of 1.8% whereas the bovine, goat and sheep strains were 9.1% divergent. Similarly, in the NSP2 tree (Fig. 4.2b), AU109 shared close identity with the same Chinese sheep and goat NSP2 RVA sequences of strains encountered in the VP1 tree. In this case, however, the AU109 NSP2 gene was a mean divergence of 4.3% from the sheep and goat NSP2 genes (Table 4.2).

Fig. 4.2a: VP1

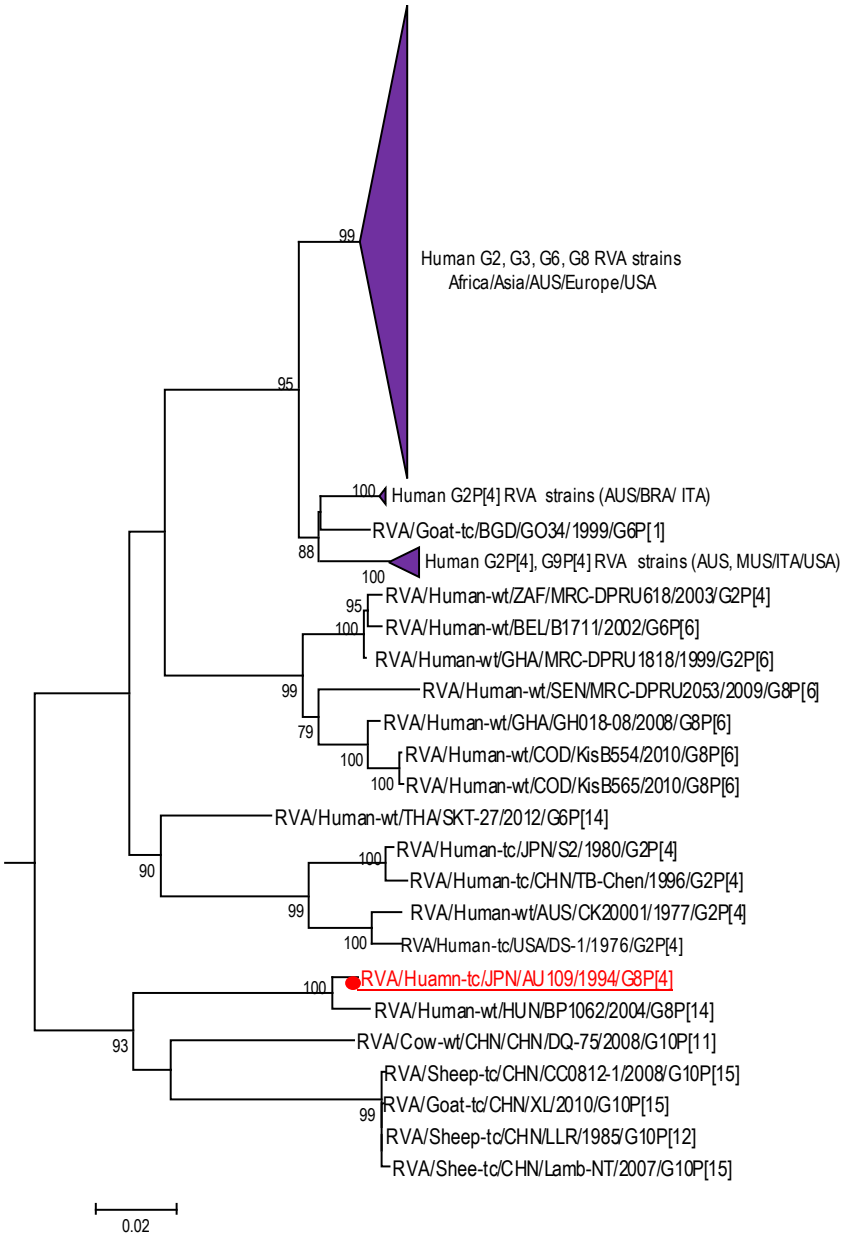


Fig. 4.2: Phylogenetic trees of the (a) VP1 gene of R2 RVA strains and (b) NSP2 gene of N2 RVA strains showing the genetic relationship between AU109 in this study and other animal and/or human RVA strains. For each phylogenetic tree, analyses included the nucleotide sequence of the Japanese human G8P[4] strain AU109 in this study (indicated in red font with a dot) and globally circulating strains bearing the same genotype as AU109 retrieved from the GenBank database. Maximum likelihood phylogenetic analyses were performed using Tamura Nei substitution model with gamma distributed rate variation (VP1); Tamura 3-parameter substitution model with gamma distributed rate variation (NSP2) in MEGA6 software package. The trees were respectively rooted using the VP1 and NSP2 gene of the Wa strain. Significant bootstrap values (1000 replicates) of ≥70% are indicated at each node. The scale bar at the bottom of the tree indicates a genetic distance expressed as nucleotide substitutions per site.

Fig. 4.2b: NSP2

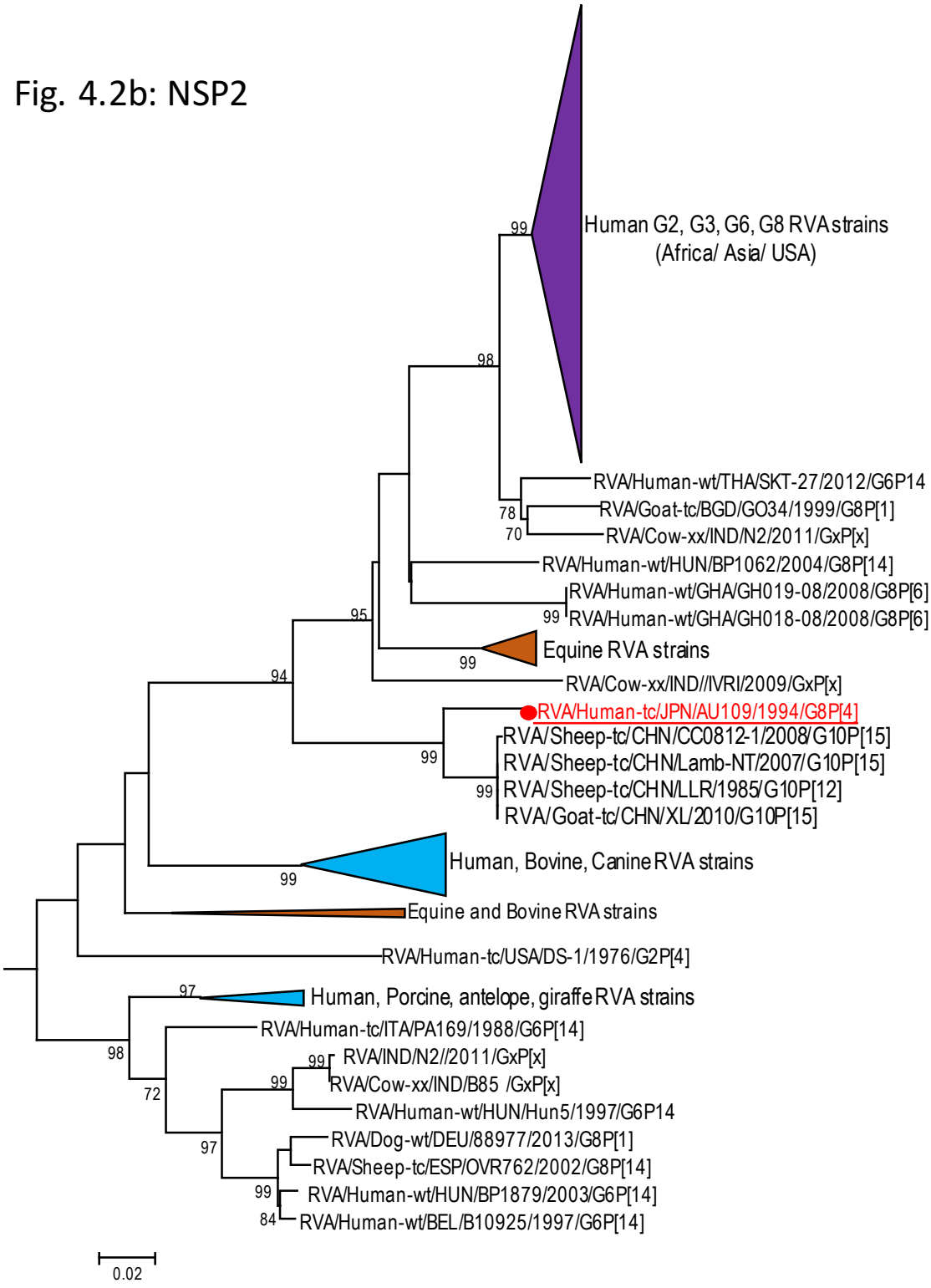

Human G2, G3, G6, G8 RVA strains
(Africa/ Asia/ USA)
RVA/Human-wt/THA/SKT-27/2012/G6P14
RVA/Goat-tc/BGD/GO34/1999/G8P[1]
RVA/Cow-xx/IND/N2/2011/GxP[x]
RVA/Human-wt/HUN/BP1062/2004/G8P[14]
RVA/Human-wt/GHA/GH019-08/2008/G8P[6]
RVA/Human-wt/GHA/GH018-08/2008/G8P[6]
Equine RVA strains
RVA/Cow-xx/IND//IVRI/2009/GxP[x]
RVA/Human-tc/JPN/AU109/1994/G8P[4]
RVA/Sheep-tc/CHN/CC0812-1/2008/G10P[15]
RVA/Sheep-tc/CHN/Lamb-NT/2007/G10P[15]
RVA/Sheep-tc/CHN/LLR/1985/G10P[12]
RVA/Goat-tc/CHN/XL/2010/G10P[15]
Human, Bovine, Canine RVA strains
Equine and Bovine RVA strains
RVA/Human-tc/USA/DS-1/1976/G2P[4]
Human, Porcine, antelope, giraffe RVA strains
RVA/Human-tc/ITA/PA169/1988/G6P[14]
RVA/IND/N2//2011/GxP[x]
RVA/Cow-xx/IND/B85 /GxP[x]
RVA/Human-wt/HUN/Hun5/1997/G6P14
RVA/Dog-wt/DEU/88977/2013/G8P[1]
RVA/Sheep-tc/ESP/OVR762/2002/G8P[14]
RVA/Human-wt/HUN/BP1879/2003/G6P[14]
RVA/Human-wt/BEL/B10925/1997/G6P[14]
0.02

Table 4.2: Nucleotide sequence and amino acid identities between AU109 and its closest RVA strain in each genome segment

Genome segment	Closest related RVA strain to AU109	% identity		Accession number	Predominant host species in phylogenetic lineage*	Country	Reference
		Nucleotide	Amino acid				
VP7	RVA/Human-wt/HUN/BP1062/2004/G8P[14]	96.5	96.4	FN665696	Human	Hungary	Ciarlet et al., 2008
VP4	RVA/Human-wt/JPN/88H449/1989/G2P[4]	99.5	99.6	AB971550	Human	Japan	Sasaki et al., 2015
VP6	RVA/Human-wt/JPN/88H449/1989/G2P[4]	99.8	100	AB971552	Human	Japan	Sasaki et al., 2015
VP1	RVA/Human-wt/HUN/BP1062/2004/G8P[14]	98.2	99.8	FN665688	Human	Hungary	Ciarlet et al., 2008
VP2	RVA/Human-wt/JPN/88H449/1989/G2P[4]	99.3	99.6	AB971556	Human	Japan	Sasaki et al., 2015
VP3	RVA/Human-wt/JPN/88H449/1989/G2P[4]	99.4	99.6	AB971558	Human	Japan	Sasaki et al., 2015
NSP1	RVA/Human-wt/JPN/88H449/1989/G2P[4]	99.3	98.4	AB971560	Human	Japan	Sasaki et al., 2015
NSP2	RVA/Lamb-tc/CHN/Lamb-NT/2007/G10P[15]	95.6	97.6	FJ031020	Sheep	China	Chen et al., 2009
NSP3	RVA/Human-wt/JPN/90H377/1990/G2P[4]	99.7	100	LC002027	Human	Japan	Doan et al., 2015
	RVA/Human-wt/JPN/88H449/1989/G2P[4]	99.7	99.7	AB971563			Sasaki et al., 2015
NSP4	RVA/Human-wt/JPN/88H449/1989/G2P[4]	99.6	98.8	AB971564	Human	Japan	Sasaki et al., 2015
NSP5	RVA/Human-wt/JPN/90H377/1990/G2P[4]	100	100	LC002065	Human	Japan	Doan et al., 2015
	RVA/Human-wt/JPN/88H449/1989/G2P[4]	99.6	100	AB971566			Sasaki et al., 2015

Predominant host species in phylogenetic lineage*: the dominant host species origin of sequences in the lineages to which AU109 sequences belong in the phylogenetic trees

We also attempted to determine the time when the VP7, VP1 and NSP2 genes of AU109 diverged from the shared ancestors of these animal rotavirus VP7, VP1 and NSP2 sequences by conducting BEAST analysis. The topology of the MCC trees generated by the Bayesian evolutionary analyses were similar to the maximum likelihood trees even though a few minor branching differences occurred due to removal of strains without dates from the dataset. In the VP7 MCC tree (Fig. 4.3a), the tMRCA of the distinct lineage to which AU109 belonged was 1988 (95% HPD interval: 1976-1993) (Table 4.3). AU109 and its closest VP7 sequence BP1062 diverged from the ancestor they shared with lineage IV sequences in 1960. In the VP1 MCC tree (Fig. 4.3b), the tMRCA of the distinct lineage AU109 shared with BP1062 was 1982 (95% HPD interval: 1976-1993) (Table 4.3). Towards the root of the tree, the AU109 and BP1062 lineage diverged from a common ancestor shared by a Chinese bovine strain DQ-75 in 1927. In the NSP2 MCC tree (Fig. 4.3c), AU109 NSP2 gene shared a distant past ancestor with the Chinese goat and sheep RVA NSP2 genes in 1937 (Table 4.3).

Table 4.3: Time of most recent common ancestor of AU109 lineage and evolutionary rates of the VP7, VP1 and NSP2 genes analysed in this study

	Genes		
Parameter	**VP7**	**VP1**	**NSP2**
Number of sequences	187	207	214
Sampling period	1965-2013	1976-2013	1975-2013
Geographical coverage	Global	Global	Global
Evolutionary rate (10^{-3}/site/year) (95% HPD)	1.44 (1.10-1.79)	0.56 (0.46-0.67)	0.44 (0.35-0.54)
tMRCA of AU109 lineage (95% HPD)	1988 (1976-1993)	1982 (1968-1992)	1937 (1902-1966)
Year of divergence from animal or animal-derived sequences	1960	1927	1937

Fig. 4.3a: VP7 G8 Maximum clade credibility tree

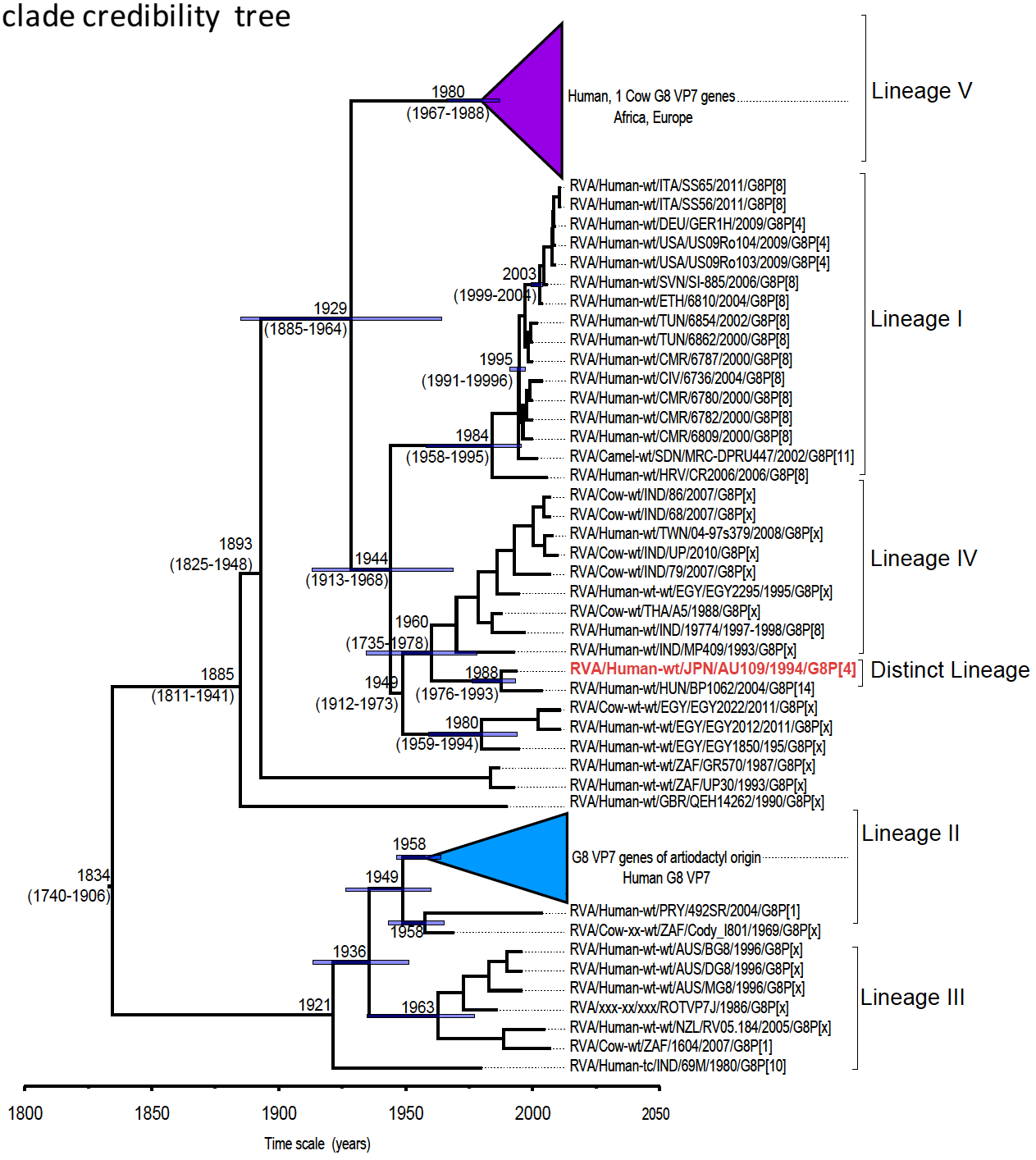


Fig 4.3: Simplified maximum clade credibility trees of dated (a) VP7, (b) VP1 and (c) NSP2 nucleotide sequences constructed using the Bayesian MCMC framework. The 95% highest posterior density (HPD) interval of each significant node is indicated with bars. The time of most recent common ancestor (tMRCA) is indicated for each lineage (cluster). Lineages far away from the AU109 sequences have been collapsed for simplicity. The scale below each tree represents the time scale in years.

Fig. 4.3b: VP1 R2 Maximum clade credibility tree

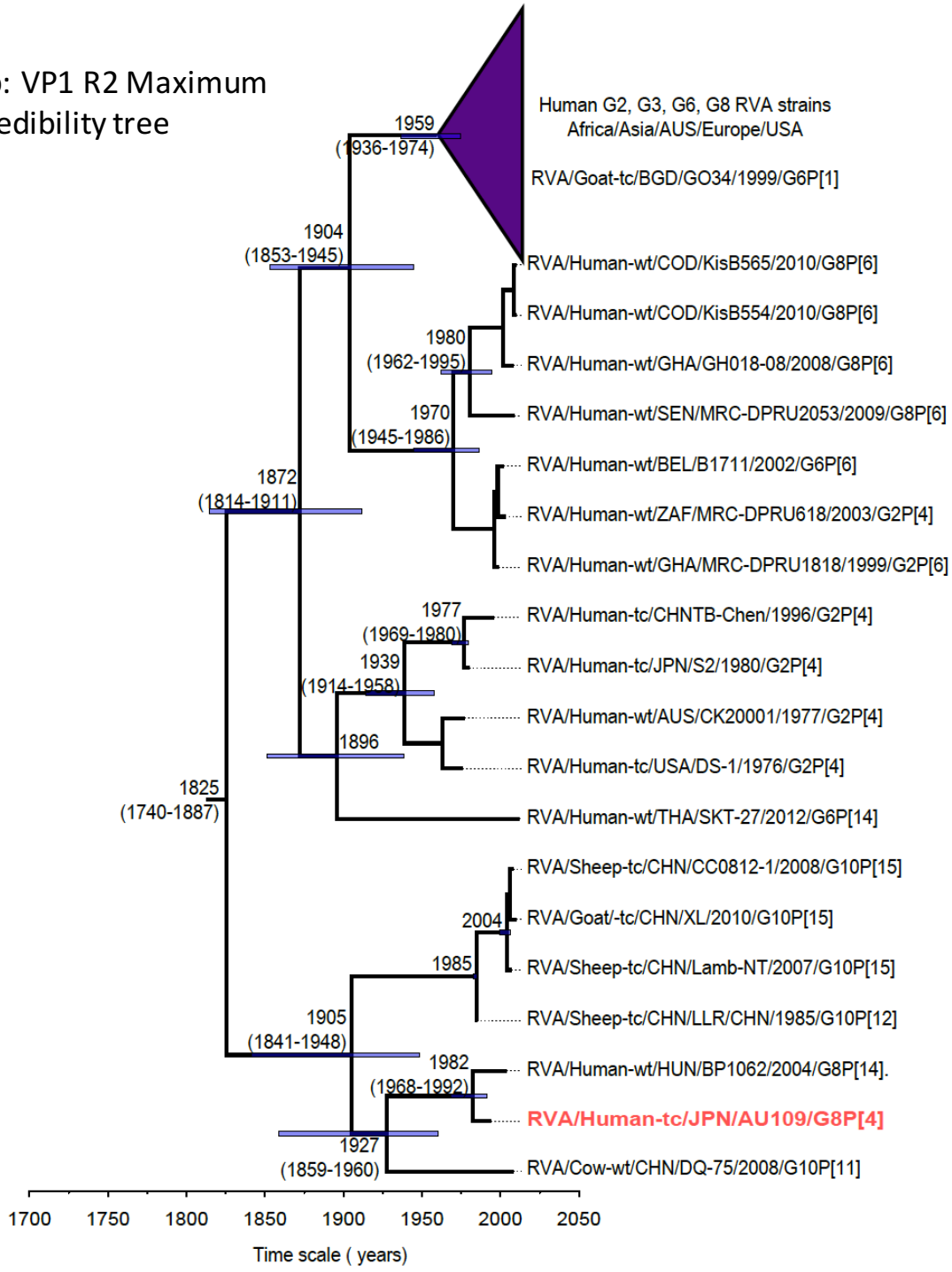
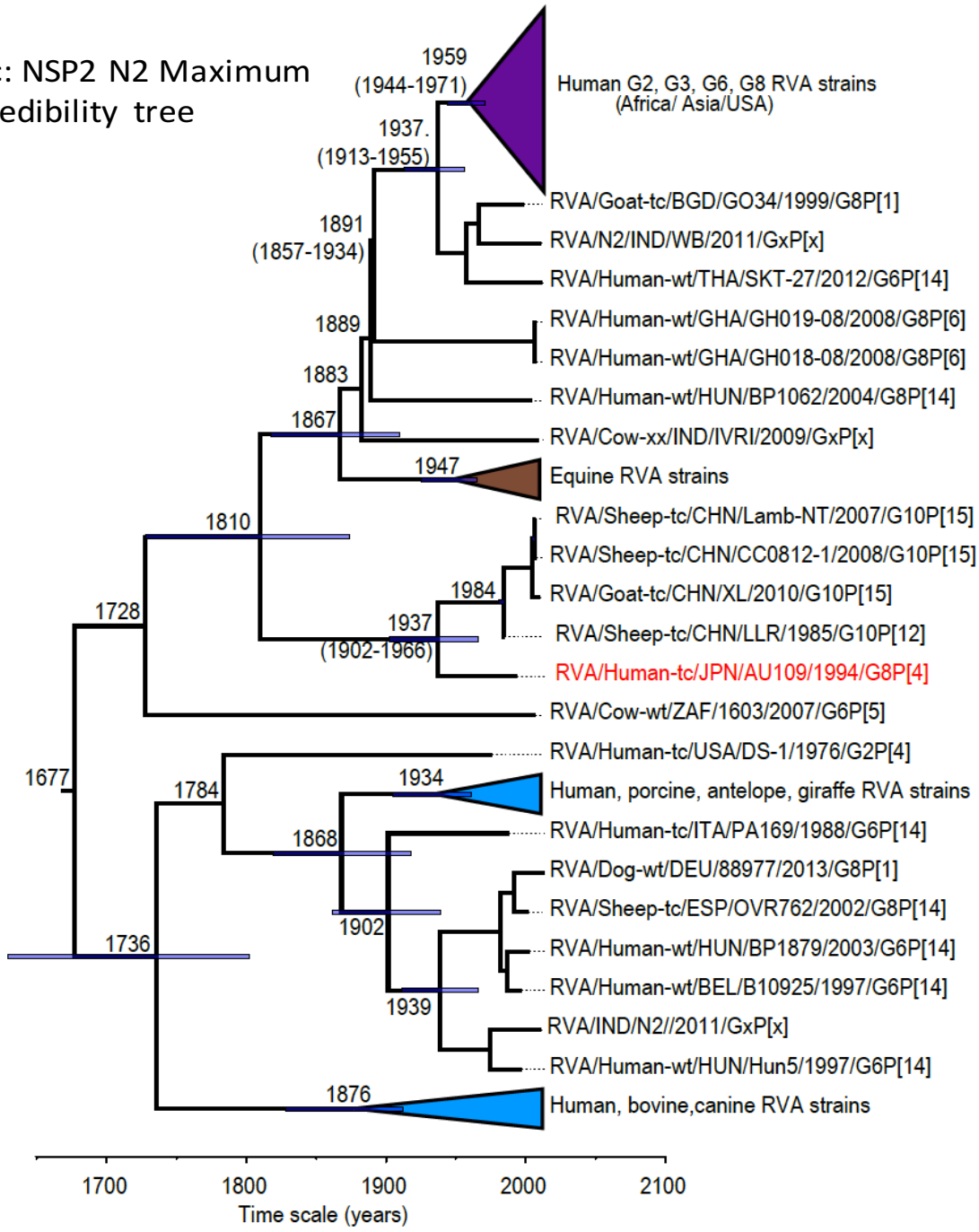


Fig. 4.3c: NSP2 N2 Maximum clade credibility tree



4.5 DISCUSSION

In this study, the first human G8 RVA strain in Japan was characterized for its complete genome to understand how the strain was generated under natural conditions. Evidence gathered upon sequence and phylogenetic analyses of the 11 genome segments revealed that the genome of AU109 was generated by genetic reassortment events in which Japanese DS-1-like G2P[4] strains circulating between 1985 and 1990 obtained three genes namely the VP7, VP1 and NSP2 from unknown ruminant G8 RVA strains.

Whereas we were able to provide a robust evidence of human host species origin of eight of the eleven genome segments of AU109, those of the VP7, VP1 and NSP2 genes were unclear as there was lack of sufficient whole genome sequence data on rotaviruses from less explored host species such as goats and sheep in the GenBank. We however recognized the relatedness of the VP7, VP1 and NSP2 genes of AU109 to those of unknown ruminant RVA strains. Regarding the VP7 gene, even though it is tempting to hypothesize that it is of bovine RVA origin, there were a few observations that clearly disputed this assumption.

First, while the host species origin of the VP7 gene of the majority of human G8 rotavirus strains such as those detected in Malawi (Nakagomi et al., 2013) as well as those from India (Mukherjee et al., 2013) could be easily elucidated based on the predominant host species in their lineage, nucleotide and amino acid sequence identity matrices, the only G8 VP7 gene which shared such properties with AU109 was the Hungarian G8P[14] strain BP1062, both of which were of human host species origin (Table 4.2, Fig. 4.1a). Second, the monophyletic cluster with a high bootstrap support to which AU109 and BP1062 belonged could not be grouped together with adjacent lineages in the VP7 phylogenetic tree (Fig. 4.1a) as the genetic distance of AU109 and BP1062 from the other G8 VP7 genes was more than 10%. Third, since the detection of AU109 in 1994 and then BP1062 after 10 years, it is reasonable to hypothesize that there was an intermediate unknown host species

which served as the reservoir of this type of G8 VP7 gene as we observed that the VP7 gene of AU109 and BP1062 did not spread readily from human to human but were more likely dead-end infections.

Similarly, the VP1 phylogenetic tree strongly suggested that the VP1 of AU109 was obtained from the same ancestor that provided BP1062 with its VP1 gene (Fig. 4.2a). The nucleotide sequence divergence of 1.8% (Table 4.2) between them could be due to the accumulation of point mutations over the years when BP1062 was later detected. It is however interesting that of all the many R2 sequences deposited in the GenBank, AU109 and BP1062 were the only sequences that shared such a close identity. On the other hand, the presence of one bovine, one goat and three sheep RVA VP1 sequences all of Chinese origin nearby AU109 and BP1062 (Fig. 4.2a) provided evidence that AU109 VP1 gene shared ancestors with RVA strains of ruminant host species origin.

The time when AU109 and BP1062 diverged from the ancestor shared with the closest ruminant RVA strain —DQ-75 of bovine origin was 1927 (Fig. 4.3b). It is worthy to note that between 1927 and the time of most recent common ancestor of AU109 and BP1062 VP1 genes (1982), there was no record of any VP1 gene of similar identity that was characterized during this period as the closest sequence (DQ-75) was more than 9.1% distant from them. This led us to hypothesize that, perhaps, the VP1 gene of AU109 as well as that of BP1062 were obtained from unknown ruminant RVA strains that existed between 1927 and 1982.

In the case of the NSP2 gene of AU109, the nucleotide sequence divergence of 4.3% (Table 4.2) from the Chinese lamb and goat RVA strains and the high bootstrap support of 100% (Fig. 4.2b) showed that these sequences were of a common ancestral origin. With an evolutionary rate of 4.41×10^{-4} nucleotide substitutions/per/year (Table 4.3) until their time of detection, AU109 and the Chinese ruminant NSP2 genes accumulated point mutations after diverging from their common ancestor which resulted in the 4.3% nucleotide sequence

divergence (Supplementary Fig. 4.2c).

On the other hand, the following eight genes: VP2-VP4, VP6, NSP1, NSP3-NSP5 of AU109 distinctly clustered together with Japanese DS-1-like G2P[4] strains at high bootstrap supports. As rotaviruses utilize reassortment as one of the evolutionary processes to generate diverse genomes, the genome of AU109 was no exception to this phenomenon. One evidence suggesting the occurrence of genetic reassortment events during the time of circulation these Japanese G2P[4] strains is worth mentioning. As AU64, a Japanese G1P[4] RVA strain detected in 1989 was found to be a reassortant from a sibling of 88H449 (G2P[4]) which donated 10 genes and a G1 VP7 from another strain (Sasaki et al., 2015) , it is not surprising that AU109 as revealed by our analysis carried eight genome segments which were closely related to Japanese G2P4] strains circulating at the time of detection of AU109 (Table 4.2).

In 2010, Banyai and colleagues who could not predict the origin of the VP7 gene of their G8P[14] strain BP1062, could only refer to AU109 whose VP7 sequence was available in the GenBank. In addition, BP1062 was hypothesized to be of direct zoonotic origin due to the unrelatedness of its nine non-G/P genome segments to that of canonical human RVA strains (Banyai et al., 2010). It was reasonable to predict that the availability of the whole genome sequence of AU109 could help confirm the host species origin of the genome of BP1062. On the contrary, AU109 even though shared the same lineage with BP1062 in the VP7 and VP1 phylogenetic trees, the remaining nine genome segments were either located in different lineages or were of different genotypes.

Another interesting observation that confirmed the diversity achieved by rotaviruses through genetic reassortment events was the fact that even though AU109 possessed the same genotype constellation as the first G8P[4] strains reported from India (Mukherjee et al., 2013) — the first to be reported from Asia as well as G8P[4] strains reported in Africa (Nakagomi et al., 2013), AU109 genome was unique at the lineage level with the exception

of its NSP5 gene which shared the same lineage with Malawian G8P[4] RVA sequences (data not shown).

It is worthy to note that in recent years, G8 rotaviruses possessing the P[8] VP4 genotype emerged in Japan, Thailand and Vietnam (Hoa-Tran et al., 2016; Kondo et al., 2017; Tacharoenmuang et al., 2016). Although both kinds of G8 strains were reassortants and their genetic backbones consist of genotype 2 genes from both RVAs of human and animal origin, there are two important differences to take note of. First, while molecular phylogeny established the bovine rotavirus origin of the G8 VP7 gene possessed by the G8P[8] strains (Hoa-Tran et al., 2016; Kondo et al., 2017; Tacharoenmuang et al., 2016), the G8 VP7 gene of AU109 is interpreted to be from an unknown ruminant source. Second, although both the G8P[8] strains and AU109 possess the DS-1-like genetic backbone, AU109 resulted in a dead-end infection whereas the G8P[8] strains established a human-to-human transmission chain in the population. It will be interesting to investigate what genetic differences enabled one strain and not the other to spread successfully within the same population.

In summary, AU109 was generated via genetic reassortment events in which Japanese G2P[4] strains circulating during the mid-eighties to the nineties obtained the VP7, VP1 and NSP2 genes from unknown ruminant RVA strains. There is however the need to examine the whole genome of RVA strains from less explored host species especially those animals that are regularly in close proximity with the human population as these data are currently lacking in the DNA databases. These will help gain better insight into the origins of such unique RVA sequences.

4.6 ACKNOWLEDGEMENTS

We acknowledge the immense support of the Program for Nurturing Global Leaders in Tropical and Emerging Communicable Diseases, Graduate School of Biomedical Sciences, Nagasaki University. This study was in part supported by grants-in-aid for scientific research from the Ministry of Health, Labour and Welfare of Japan, as well as a grant from Japan Initiative for Global Research Network on Infectious Diseases.

4. 7 CONFLICT OF INTEREST

The authors declare no conflict of interest.

Chapter V

Discussion, Conclusion and Recommendations

5.1 INTRODUCTION

This chapter aims to exploit the major observations obtained from the published papers of mine (Chapters II-IV) to feature and understand the circulating rotavirus strains on the African continent and to gain insight into the evolution of rotavirus genome with special emphasis on the role of interspecies transmission of animal rotaviruses to humans.

Approximately 90% of human rotavirus infections in the world is caused by one of five types of strains: G1P[8], G2P[4], G3P[8], G4P[8], and G9P[8] (Banyai et al., 2012; Gentsch et al., 2005; Santos and Hoshino, 2005); the most prevalent worldwide are the G1P[8] strains. While this is true for industrialised countries, these globally common strains accounted for only 36.5%, 41.0%, and 42.7% of fully genotyped strains circulating on the African continent during the 1997 - 2007, 2007-2011, and 2006 - 2016 rotavirus season, respectively (Ouermi et al., 2017; Seheri et al., 2014; Todd et al., 2010). The remaining portion of strains in Africa as evidenced by previous research and review articles is diverse and is made up of globally uncommon G/P genotype combinations which were generated either by genetic reassortment or interspecies transmission events (Mwenda et al., 2010; Ouermi et al., 2017; Sanchez-Padilla et al., 2009; Santos and Hoshino, 2005; Seheri et al., 2014; Steele and Ivanoff, 2003; Todd et al., 2010).

Specifically, apart from the five globally common strains, previous studies in Africa have recorded unusual strains such as G1P[6], G2P[6], G3P[6], G6P[6], G8P[6]/P[8]/P[4], G9P[6] (Adah et al., 2001; Agbemabiese et al., 2015b; Armah et al., 2001; Asmah et al., 2001; Binka et al., 2011; Dennis et al., 2014b; Enweronu-Laryea et al., 2013; Esteves et al., 2016; Heylen et al., 2016; Matthijnssens et al., 2006; Nakagomi et al., 2013; Nordgren et al., 2012a; Nordgren et al., 2012b), to mention just a few. It is noteworthy that most of the unusual strains reported on the African continent possess the P[6] VP4 genotype, pointing to the tendency of the P[6] genotype to associate with a wide variety of G-genotypes in Africa. Granted that the African P[6] genotype is part of the DS-1-like genotype constellation

which is not the case elsewhere in the world, one does not have to hypothesise complicated reassortment events to explain, for example, the generation of a G8P[6] strain from a rotavirus with a DS-1-like backbone in Africa - a single reassortment event in which the G8P[6] strain acquired the G8 VP7 gene from a bovine RVA and the rest from a G2P[6] strain is adequate.

5.2 UNUSUAL P[6] STRAINS IN AFRICA ASSOCIATE WITH DS-1-LIKE GENES

The P-genotype P[6] is generally perceived to be of porcine rotavirus origin and it is the second most detected genotype (15.9%) after the P[7] genotype (47.4%) in pigs (Papp et al., 2013). Phylogenetic inference of the P[6] genotype supports this assertion that P[6] originates from porcine rotavirus (Fig. 5.1) as P[6] detected from pigs were closest to the root of the tree whereas more than 90% of the P[6] detected in humans are located at the periphery of the tree. One feature of African RVA strains is the abundance of the P[6] genotype. In this regard, one would expect that P[6] bearing rotavirus strains in Africa would rather be associated with Wa-like backbone genes which are often associated with porcine and porcine-like human rotavirus strains. However, that is not the case.

Here, a caveat may be worthy. The African P[6] genotype is virtually unrelated to porcine RVA strains, and a half of G2 strains possess P[6] (the other half possess P[4]). This means that the African P[6] genotype is part of the DS-1-like genotype constellation which is not the case elsewhere in the world. Thus, P[6] rotaviruses detected on the African continent most often possess the DS-1-like genetic backbone and are more likely to possess traces of bovine or bovine-like human rotavirus genes rather than a porcine rotavirus gene. This was evident from full genome analysis of African rotavirus strains such as G8P[6], G6P[6], G2P[6] (Nakagomi et al., 2013, Dennis et al., 2014, Heylen et al., 2015, Nordgren et al., 2012, Agbemabiese et al., 2015; Ndze et al., 2015). In these studies, one or more genome segments of these P[6] strains shared common ancestors with segments of these

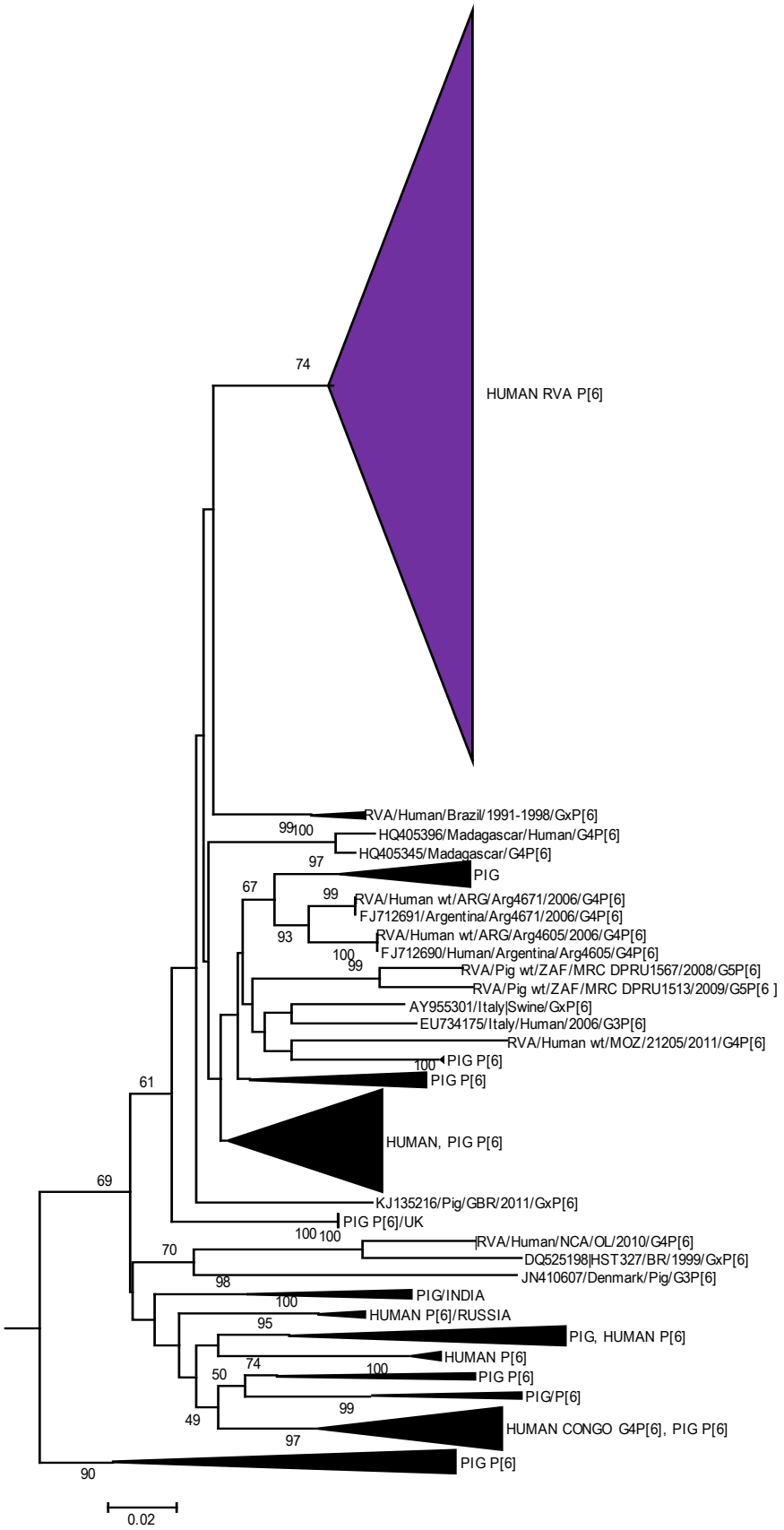


Fig. 5.1: Neighbour joining phylogenetic tree of global P[6] VP4 sequences. Typical human P[6] VP4 sequences are located in the purple triangle. Black triangles contain porcine and porcine-like human P[6] RVA

P[6] strains shared common ancestors with adjacent bovine or bovine-like human genotype 2 genes.

Apart from the many studies that reported P[6] strains with the DS-1-like genetic backbone in Africa, it is important to also mention that G1P[8] strains with the DS-1-like genetic backbone emerged recently in Malawi during the post-vaccine introduction period - 2013-2014 (Jere et al., 2017). Such DS-1-like G1P[8] strains were previously reported in Japan, Thailand and Vietnam (Fujii et al., 2014; Komoto et al., 2015, 2016; Kuzuya et al., 2014; Nakagomi et al., 2017; Yamamoto et al., 2012). In this regard, it is of note that Jere et al. (2017) stated based on phylogenetic analysis that these strains were not imported from the Asian continent. Rather, they were generated through reassortment events between co-circulating Wa-like and DS-1-like strains in Malawi.

In Chapter V, I shall discuss and summarise the facts and interpretation of the evolutionary history of some of the unusual rotavirus genotypes that have thus far been mostly detected on the African continent and the impact of the occurrence of these novel strains on the evolution of human rotavirus populations in Africa. Since G6 and G8 genotypes have been described as pathogens from various animal species especially cattle (Alkan et al., 2012; Browning et al., 1992; Papp et al., 2013; Papp et al., 2014; Snodgrass et al., 1990), their emergence and spread in the African population is worth examining. Thus, in the following sections, I shall focus on the genome of G6P[6] and G8 strains to provide insight into the evolutionary history of such unusual rotavirus strains on the African continent.

5.3 MOLECULAR EPIDEMIOLOGY AND GENOME EVOLUTION OF THE G6P[6] RVA

The G6 VP7 genotype is a major genotype found in bovine rotaviruses (Papp et al., 2013). They have also been reported in other animals including cats (German et al., 2015;

Kaneko et al., 2016), pigs (Ghosh et al., 2007), goat (Ghosh et al., 2010), horse (Ghosh et al., 2013a), and antelope (Matthijnssens et al., 2009).

Rotavirus G6 was for the first time reported in human in 1987 (PA151/G6P[9]), 1988 (PA169/G6P[14]) from two hospitalised children suffering from diarrhoea in Italy (Gerna et al., 1992). Subsequently in the 1990s, G6P[14] strains MG6/1993 and MG6.01/1996-1997, AG6.01/1996-1997 (Cooney et al., 2001; Palombo and Bishop, 1995) were detected in children with acute gastroenteritis in Australia, while a G6P[9] strain Se584/1998 was reported in the United States (Griffin et al., 2002). Also in Hungary, G6 with either P[9] or P[14] specificity were reported during the 1995-1998 rotavirus season in infants and young children admitted with acute gastroenteritis (Banyai et al., 2003a; Banyai et al., 2003b).

G6 from human and artiodactyl origin have been noted to differ by possession of different VP4 genotypes (Gerna et al., 1992). Specifically, bovine G6 strains usually possess P[1], P[5], P[7] and P[11]; porcine G6 strains possess P[13], feline G6 strains possess P[9], and human G6 strains possess P[6], P[9], P[14] (Fig. 2.1).

In recent years, the G6 strains with P[6] VP4 genotype emerged in a few of countries in Europe (Belgium and Italy) and Africa (Burkina Faso, Ghana and Cameroon) (Agbemabiese et al., 2015b; Ianiro et al., 2013; Matthijnssens et al., 2008c; Ndze et al., 2014; Nordgren et al., 2012a; Nordgren et al., 2012b; Rahman et al., 2003). While they were sporadically detected in Europe, they were found in large numbers in children in Burkina Faso and accounted for 23% and 13% of rotavirus positive cases in an urban and a rural setting, respectively (Nordgren et al., 2012a; Nordgren et al., 2012b).

Evidence available indicates that the G6 VP7 found in G6P[14] and G6P[9] strains are as a result of interspecies transmission (Cooney et al., 2001; Gerna et al., 1992; Palombo and Bishop, 1995) whereas those of G6P[6] strains most likely originated from a PA151-like G6P[9] rotavirus strain either of human or feline rotavirus origin (Agbemabiese et al., 2015b).

In the G6 VP7 phylogenetic tree (Fig. 2.1), even though the interpretation of the observed topology far supports the hypothesis that the G6 carried by the G6P[6] strains originated from human G6P[9] strains and not directly from bovine G6 rotavirus strains, the true host donor of this G6 VP7 gene possessed by the G6P[6] strains remains debatable for the following reasons. First, the human G6P[9] strains e.g. ITA/PA151/1987, TUN//17237/2008, JPN/KF17/2010, that share their most recent common ancestor with the G6P[6] VP7 sequence cluster have traces of genes related to feline rotavirus genes (Ben Hadj Fredj et al., 2013; Gerna et al., 1992; Yamamoto et al., 2011). Second, recently in the United Kingdom, G6P[9] strains were the most predominantly detected genotype reported in cats (German et al., 2015). Taken together these observations, domestic cats are speculated to be the source of the G6 VP7 gene carried by human G6P[9] strains which later acquired mutations and transitioned into the G6 VP7 sequence possessed by G6P[6] strains.

Phylogenetic inference of the G6 VP7 gene showed a clear segregation of sequences according to species origin and P genotype (Agbemabiese et al., 2015b; Cooney et al., 2001). With the limited number of G6 sequences available, Cooney et al. (2001) observed that the species-specific phylogenetic clustering did not correlate with the geographical origin of the strains. Also, human G6 rotaviruses were thought to display a restricted geographical distribution because they were detected only in two locations, i.e. Italy and Australia (Palombo et al., 1995; Gerna et al., 1992). With the recent high detection rate of G6P[6] strains in the African population, the above observations need to be revised to reflect the phylogenetic relationship of these G6P[6] strains to other G6 strains. With an update in the phylogenetic inference of the G6 VP7 gene, the following facts and interpretations are worth summarising. First, it remains a fact that G6 VP7 sequences segregate at the lineage level according to the VP4 genotype. Second, the G6P[6] strains characterised for their full genomes thus far, possess pure DS-1-like backbone genes which were provided by co-

circulating DS-1-like strains on the African continent with traces of genes with animal rotavirus origin (Matthijnssens et al., 2008; Agbemabiese et al., Nordgren et al., 2012) as observed for the genome of G8 strains detected in humans in Africa (Nakagomi et al., 2013, Dennis et al., 2014).

Recent epidemiological studies have reported associations between secretor status and Lewis status of human populations and rotavirus infection (Hu et al., 2012; Imbert-Marcille et al., 2014; Jiang et al., 2017; Nordgren et al., 2014; Van Trang et al., 2014). One of such studies by Nordgren et al. (2014) showed that P[6] infections were seen exclusively in Burkinabe children (an African population) with the Lewis-negative phenotype irrespective of their secretor status whereas in Nicaragua (a Central American population), infection with P[8] rotaviruses was seen exclusively in children with the Lewis - and secretor-positive phenotype. As the Lewis-positive phenotype is present in approximately 90% of the world's population but much lower in the African population (Jiang et al., 2017), it is not surprising that the P[6] genotype is more prevalent in Africa. It is therefore logical to speculate that the human to human transmission event of G6P[6] strains in the African population was made possible by both the acquisition of already adapted DS-1-like rotavirus genes and the possession of the P[6] genotype.

5.4 MOLECULAR EPIDEMIOLOGY AND EVOLUTION OF G8 RVA STRAINS IN AFRICA

Until recently, in humans, the G8 genotype occurred more frequently on the African continent mainly in combination with the P[6] genotype (Dennis et al., 2014; Esteves et al., 2016; Heylen et al., 2014; Heylen et al., 2015; Istrate et al., 2015; Nakagomi et al., 2013; Nielsen et al., 2005) than in other populations elsewhere in the world. On the Asian continent, there has been recent reports of the abrupt emergence as well as the persistence and spread of G8P[8] strains (Hoa-Tran et al., 2016; Kondo et al., 2017; Tacharoenmuang et al., 2016). In this section, first, I shall examine and describe the phylogenetic relationship

of the African G8 VP7 gene within a global context in reference to their possible host species origin, and second, I shall provide insight into the possible genome characteristics that enabled the survival or extinction of the different G8 VP7 sequences detected in the African population.

African human G8 rotaviruses can be categorised (but not limited) into two major groups based on evolutionary history gathered from their phylogenetic groupings (Fig. 5.2). In addition to these two major categories, there are a few sporadically detected

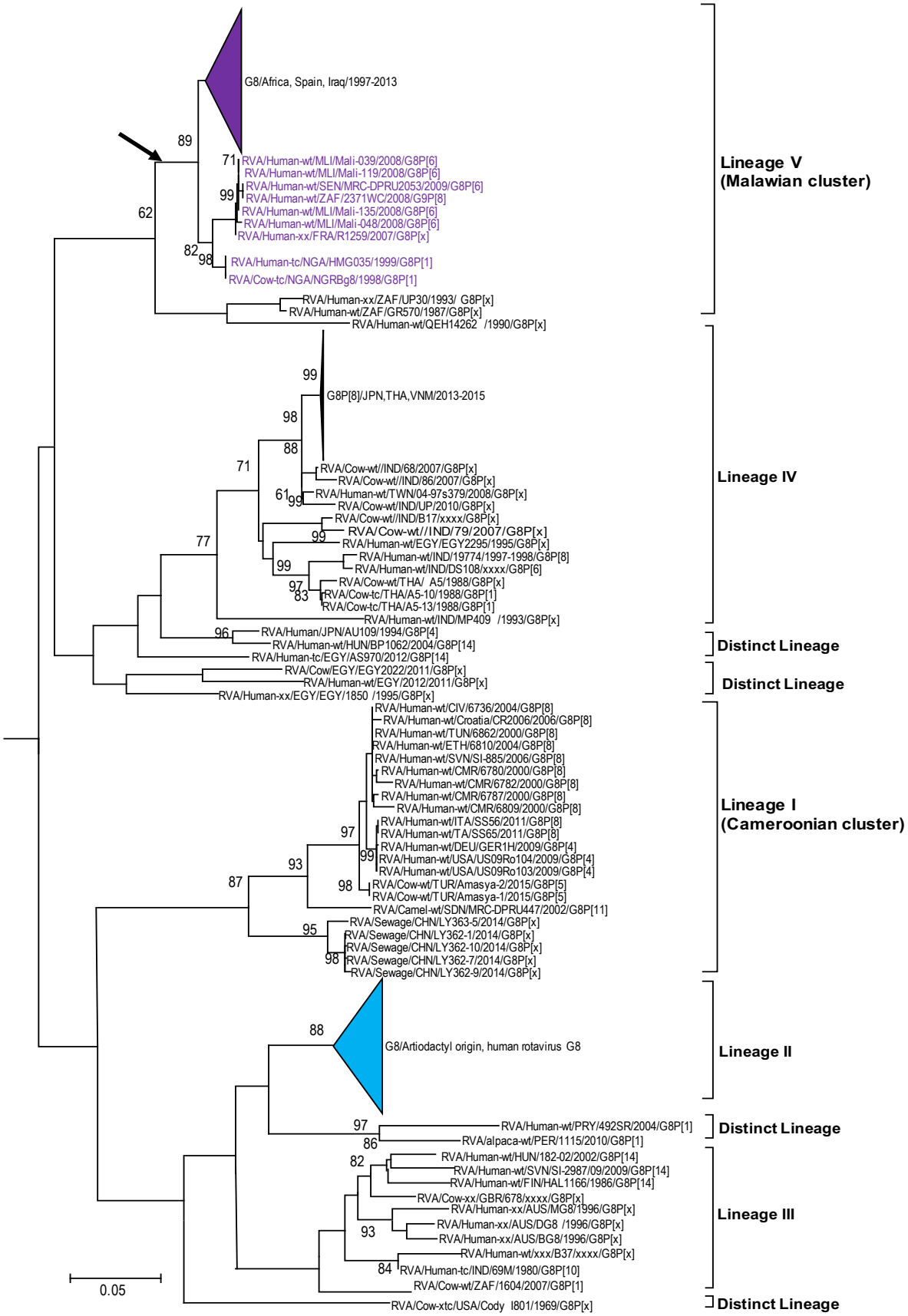


Fig. 5.2: Maximum likelihood phylogenetic tree of global G8 VP7 sequences

human rotavirus G8 sequences from Egypt which share a common ancestor with a cluster of animal and human G8 rotaviruses mainly from Asia (Fig. 5.2, Lineage IV).

The first major category, designated 'the Cameroonian cluster' (Fig. 5.2, Lineage I) consists of human G8 strains from Cameroon, Cote D'Ivoire, Ethiopia, and Tunisia detected in the early 2000s under the African Rotavirus Surveillance Network (Esona et al., 2009). These viruses possessed typical human Wa-like backbone genes with the exception of their NSP2 and NSP5 genes which were probably derived from animal rotaviruses through genetic reassortment events. Even though they were detected in countries across Africa and within the past decade in Europe (Aladin et al., 2010; Delogu et al., 2013; Pietsch et al., 2009; Steyer et al., 2007), the Americas (Weinberg et al., 2012), and China (in sewage) (Zhou et al., 2016), they seem to have gone extinct from the African continent as contemporary G8 strains detected on the continent do not cluster in this lineage anymore. Tracing back their phylogenetic history, this cluster of human G8 sequences diverged from their most recent common ancestor which was shared with a sub-cluster of two recently deposited Turkish bovine G8P[5]/2015 sequences (Karayel et al., GenBank data, February, 2017; accession numbers: KX212865 – KX212866). The upstream ancestral root of both sub-clusters was shared with a Sudanese camel G8P[11]/2002 strain (GenBank accession number: KC257096) described previously to have been generated via complex genetic reassortment events by Jere et al., (2014).

The second major category of African G8 strains, designated 'the Malawian cluster' (Fig. 5.2, Lineage V), consists of the most prevalent G8 strains on the African continent and have been detected in all the regions in Africa (Dennis et al., 2014; Heylen et al., 2014; Heylen et al., 2015; Matthijnssens et al., 2006; Nakagomi et al., 2013; Nyaga et al., 2014). These strains mainly possess in the following order: P[6], P[4] and P[8] VP4 genotypes and for those whose full genome sequence information are available, they possess the genotype 2 genetic backbone genes which are closely related to human DS-1-like rotavirus sequences

(Dennis et al., 2014; Heylen et al., 2014; Matthijnssens et al., 2006; Nakagomi et al., 2013; Nyaga et al., 2014). They can be described as “typical human G8 VP7 genes” since for the past decade, they seem to be well adapted to the human population in Africa and so far, only person-to-person transmission chain has been observed in this cluster. Interestingly, these widely distributed typical human G8 strains in Africa have not spread widely across the world apart from their occurrence in Brazil in 2002 (Gomez et al., 2010), Iraq in 2008 (Ahmed et al., 2013) and Spain in 2008-2009 (Fernandez-Jimenez et al., GenBank data, July 2016, accession numbers: HQ638096 - HQ638098; HQ638100 - HQ638102; HQ638105 - HQ638108).

Phylogenetic inference of the G8 VP7 gene tree revealed that the lineage V was formed following an interspecies transmission event similar to, but independent of the interspecies transmission event involving a Nigerian bovine G8 strain NGRBg8/1998 (Adah et al., 2003; Komoto et al., 2016) (Fig. 5.2, indicated with an arrow). In other words, NGRBg8/1998 is an example that failed to establish a viable transmission chain in the human population.

5.5 THE ROLE OF ROTAVIRUS GENOME COMPOSITION IN THE SPREAD OF G8 STRAINS

The G8 genotype has been persistent on the African continent despite the considerable fluctuation in frequency of detection in the population over the years. It is important to note that whereas a greater proportion of the African G8 rotaviruses characterised in humans possess the P[6] VP4 genotype, those detected recently in Asian countries such as Japan, Thailand and Vietnam strictly possess the P[8] genotype.

Against the background that the Lewis-positive phenotype is present in approximately 90% of the world’s population but much lower in the African population (Jiang et al., 2017),

it is not surprising that the P[6] genotype is more prevalent in Africa and G8 strains are likely to be associated with P[6] — a predominant VP4 genotype on the continent.

Animal G8 rotavirus strains or confirmed interspecies transmitted animal G8 strains to humans possess the following internal capsid and non-structural protein gene constellation: I2-R2-C2-M2-A3/11-N2-T6-E2-H3 whereas G8 strains of typical human rotavirus origin either possess (1) pure DS-1-like constellation I2-R2-C2-M2-A2-N2-T2-E2-H2 (Agbemabiese et al., 2015a; Heylen et al., 2014; Hoa-Tran et al., 2016; Kondo et al., 2017; Nakagomi et al., 2013; Nyaga et al., 2014; Tacharoenmuang et al., 2016) with a few reassortants detected in Ghana possessing the constellation I2-R2-C2-M2-A2-N2-T2-E2-H3 indicative of traces of bovine rotavirus genotypes (Dennis et al., 2014; Heylen et al., 2015); or (2) pure Wa-like backbone genes I1-R1-C1-M1-A1-N1-T1-E1-H1 (Esona et al., 2009).

While Palombo et al. (2000) noted that rotavirus G8 from humans and cattle differ in their VP4 and NSP1 genes and these genes are determinants of host-range restriction (Burke and Desselberger, 1996), Cooney et al. (2001) hypothesised that perhaps, transfer of the G8 genotype to humans from cattle may have occurred first through an intermediate host most likely domestic animals in which reassortment event (s) occurred.

Apart from the recognition of the broad categories of human G8 VP7 sequences in Africa, the phylogenetic grouping of the African G8 VP7 genes, the VP4 and the type of genome backbone seem to determine the survival and the extent of spread or extinction of G8 rotaviruses in Africa and elsewhere in the world. For instance, the G8 VP7 genotype which is often detected in the African population (the Malawian cluster) seems to possess the DS-1-like backbone genes mostly of human host species origin. On the other hand, the G8 strains in the Cameroonian cluster possessed genotype 1 genes some of which were related to porcine rotavirus genes (Esona et al., 2009) and they were able to spread to parts of Europe.

5.6 FATE OF ANIMAL ROTAVIRUS GENES IN HUMAN G8 ROTAVIRUS GENOME

As the G8 genotype is thought to be of animal rotavirus origin, many of the G8 strains detected in Africa irrespective of the type of backbone genes they possess seem to possess traces of animal rotavirus genes although a few such as MW2-1114, and MW2-1189 (Nakagomi et al., 2013) have completely lost traces of animal rotavirus genes. Evidence available indicates that some G8 strains in the Malawian cluster most of which possess the DS-1-like backbone genes, have some genes closely related to genotype 2 sequences of animal host species origin upon phylogenetic analysis (Heylen et al., 2014, 2016; Dennis et al., 2014; Nakagomi et al., 2013). Notable are G8 strains detected in children from Congo (KisB565, KisB554), Ghana (GH018-08, GH019-08, Ghan-113, Ghan-149), Kenya (Keny-078), Mali (Mali-039, 048, 119, 135), and many others from Malawi (Heylen et al., 2014; 2016; Dennis et al., 2014; Nakagomi et al., 2013).

The bovine-like human G8 VP7 strains detected in Asia in Lineage IV also possess traces of animal rotavirus genes (Hoa-Tran et al., 2016; Kondo et al., 2017; Tacharoenmuang et al., 2016). In these Asian bovine-like human G8 strains, the internal capsid and non-structural protein genes that share close relationships with rotaviruses of artiodactyl origin include the VP6 (2013 Thai G8P[8] strains), VP1 (Japanese, Thai and Vietnamese G8P[8] strains), NSP2 (Thai and Vietnamese G8P[8] strains) and NSP4 (Japanese G8P[8] strains). Both G8 rotavirus strains in the two separate lineages V and IV and different geographical regions are presumed to have adapted or still undergoing the adaptation process to the human population.

As to the traces of animal rotavirus genes in the backbone of the genome of the human G8 strains, the key question is whether further human-to-human spread of the G8P[8] strains in Asia and G8P[6]/P[4]/P[8] strains in Africa will lead to replacement of those backbone genes from artiodactyl rotavirus strains by co-circulating typical human rotavirus genes.

Although not recognised outright, it is possible that animal RVA genes are lost after circulating in humans over time. A continuous monitoring of such strains over a substantial study period in a restricted geographical location similar to the Malawian G8 study by Nakagomi et al. (2013) will be required to understand the fate of these animal-like rotavirus genes in the genome of the G8 strains in the human population.

5.7 AN APPARENTLY EVER-DIVERSIFYING E2 NSP4 GENOTYPE OBSERVED IN AFRICAN DS-1-LIKE STRAINS

While studying the genome evolution of the G2P[4] and the unusual strains in Africa, the E2 NSP4 genotype of these African DS-1-like strains was noted to be ever diversifying at the lineage level compared to the other ten genome segments (Nyaga et al., 2012; Do et al., 2015, Agbemabiese et al., 2016). To investigate this observation further, I extended the phylogenetic analysis and observed the following.

The human rotavirus E2 genotype appeared to have a characteristic large number of lineages compared to the other genotype 2 genes. Despite the diversity and the numerous lineages, many of the lineages were short-lived and existed in a particular period with a limited geographical distribution (Fig. 5.3). A typical example is the case of Malawian G8 strains which were studied at the full genome level over a period of ten years (Nakagomi et al., 2013) (Fig. 5.3). Their NSP4 gene formed distinct lineages which I propose to be designated lineages XII, XIII, XIV and XV as an update of the lineage designation for DS-1-like genes (Doan et al., 2015; Do et al., 2015; Agbemabiese et al., 2016). Again, these lineages were short lived and limited to the African continent. Also, the currently circulating number of lineages are not hugely different from the other genotype 2 genes as majority of the contemporary strains belonged to lineage IVa or V as noted for G2P[4] strains by Doan et al. (2015).

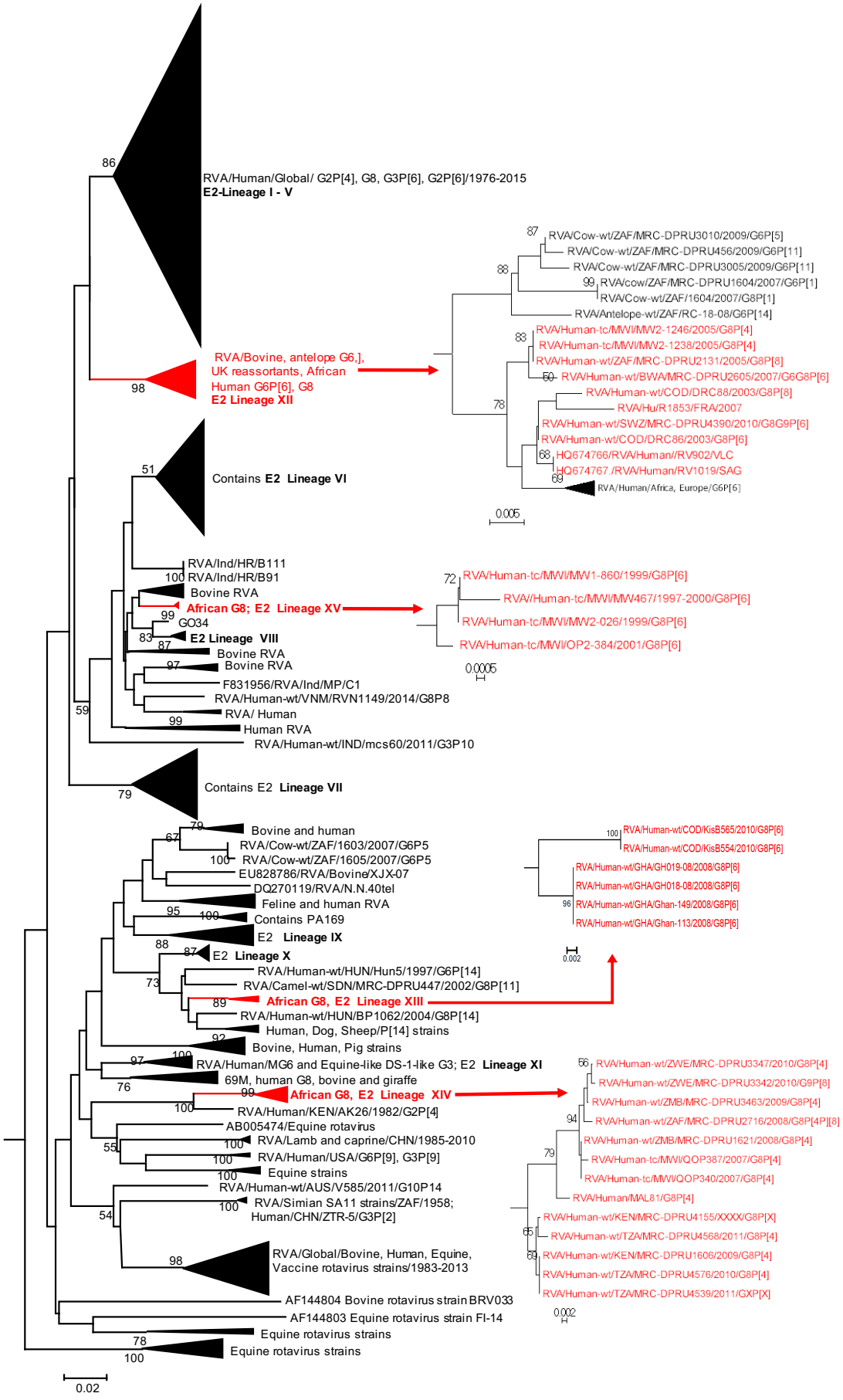


Fig. 5.3: Ever-diversifying E2 lineages with short-lived lineages of limited geographical distribution shown in red font.

5.8 RECOMMENDATIONS

It is commendable that there has been a substantial improvement in rotavirus surveillance activities in Africa; however, there is the need for more quantitative approaches to understand how frequent interspecies transmission events occur in nature in Africa. Also, keeping in mind recent discoveries on the correlation between susceptibility to rotaviruses bearing certain P-genotypes and histo-blood group antigens, studies can be tailored to elucidate beyond just numbers what evolutionary features of these P[6] RVAs allow them to preferentially infect children in Africa as such results might shed light on why the numerous unusual strains are abundant on the continent. It is also recommended that where possible, previously non-typeable strains should be sequenced to find out if some of the unusual strains we see today had already been in circulation and over time, evolved and spread in large numbers. We encourage comprehensive phylogenetic analysis of rotavirus sequences as this will help clarify origins of such unusual genotypes and rectify the belief that animal rotaviruses constantly infect humans and cause disease in Africa. Based on the observation that most of the unusual P[6] strains detected in Africa most often associate with the DS-1-like genetic backbone, it will be interesting to investigate the role of the DS-1-like genetic backbone in the emergence and spread of such unusual strains on the continent.

REFERENCES

Adah, M.I., Nagashima, S., Wakuda, M., Taniguchi, K., 2003. Close relationship between G8-serotype bovine and human rotaviruses isolated in Nigeria. J. Clin. Microbiol. 41, 3945-3950.

Adah, M.I., Wade, A., Taniguchi, K., 2001. Molecular epidemiology of rotaviruses in Nigeria: detection of unusual strains with G2P[6] and G8P[1] specificities. J. Clin. Microbiol. 39, 3969-3975.

Afrad, M.H., Matthijnssens, J., Afroz, S.F., Rudra, P., Nahar, L., Rahman, R., Hossain, M.E., Rahman, S.R., Azim, T., Rahman, M., 2014. Differences in lineage replacement dynamics of G1 and G2 rotavirus strains versus G9 strain over a period of 22years in Bangladesh. Infect. Genet. Evol. 28, 214-222.

Agbemabiese, C.A., Nakagomi, T., Doan, Y.H., Nakagomi, O., 2015a. Whole genomic constellation of the first human G8 rotavirus strain detected in Japan. Infect. Genet. Evol. 35, 184-193.

Agbemabiese, C.A., Nakagomi, T., Suzuki, Y., Armah, G., Nakagomi, O., 2015b. Evolution of a G6P[6] rotavirus strain isolated from a child with acute gastroenteritis in Ghana, 2012. J. Gen. Virol. 96, 2219-2231.

Ahmed, K., Anh, D.D., Nakagomi, O., 2007. Rotavirus G5P[6] in child with diarrhea, Vietnam. Emerg. Infect. Dis. 13, 1232-1235.

Ahmed, S., Klena, J., Albana, A., Alhamdani, F., Oskoff, J., Soliman, M., Heylen, E., Teleb, N., Husain, T., Matthijnssens, J., 2013. Characterization of human rotaviruses circulating in Iraq in 2008: atypical G8 and high prevalence of P[6] strains. Infect. Genet. Evol. 16, 212-217.

Aladin, F., Nawaz, S., Iturriza-Gomara, M., Gray, J., 2010. Identification of G8 rotavirus strains determined as G12 by rotavirus genotyping PCR: updating the current genotyping methods. J. Clin. Virol. 47, 340-344.

Alkan, F., Gulyaz, V., Ozkan Timurkan, M., Iyisan, S., Ozdemir, S., Turan, N., Buonavoglia, C., Martella, V., 2012. A large outbreak of enteritis in goat flocks in Marmara, Turkey, by G8P[1] group A rotaviruses. Arch. Virol. 157, 1183-1187.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403-410.

Armah, G.E., Pager, C.T., Asmah, R.H., Anto, F.R., Oduro, A.R., Binka, F., Steele, D., 2001. Prevalence of unusual human rotavirus strains in Ghanaian children. J. Med. Virol. 63, 67-71.

Asmah, R.H., Green, J., Armah, G.E., Gallimore, C.I., Gray, J.J., Iturriza-Gomara, M., Anto, F., Oduro, A., Binka, F.N., Brown, D.W., Cutts, F., 2001. Rotavirus G and P genotypes in rural Ghana. J. Clin. Microbiol. 39, 1981-1984.

Banyai, K., Gentsch, J.R., Glass, R.I., Szucs, G., 2003a. Detection of human rotavirus serotype G6 in Hungary. Epidemiol. Infect. 130, 107-112.

Banyai, K., Gentsch, J.R., Griffin, D.D., Holmes, J.L., Glass, R.I., Szucs, G., 2003b. Genetic variability among serotype G6 human rotaviruses: identification of a novel lineage isolated in Hungary. J. Med. Virol. 71, 124-134.

Banyai, K., Laszlo, B., Duque, J., Steele, A.D., Nelson, E.A., Gentsch, J.R., Parashar, U.D., 2012. Systematic review of regional and temporal trends in global rotavirus strain diversity in the pre rotavirus vaccine era: insights for understanding the impact of rotavirus vaccination programs. Vaccine 30 Suppl 1, A122-130.

Banyai, K., Papp, H., Dandar, E., Molnar, P., Mihaly, I., Van Ranst, M., Martella, V., Matthijnssens, J., 2010. Whole genome sequencing and phylogenetic analysis of a zoonotic human G8P[14] rotavirus strain. Infect. Genet. Evol. 10, 1140-1144.

Ben Hadj Fredj, M., Heylen, E., Zeller, M., Fodha, I., Benhamida-Rebai, M., Van Ranst, M., Matthijnssens, J., Trabelsi, A., 2013. Feline origin of rotavirus strain, Tunisia, 2008. Emerg. Infect. Dis. 19, 630-634.

Benhafid, M., Youbi, M., Klena, J.D., Gentsch, J.R., Teleb, N., Widdowson, M.A., Elaouad, R., 2009. Epidemiology of rotavirus gastroenteritis among children <5 years of age in

Morocco during 1 year of sentinel hospital surveillance, June 2006-May 2007. J. Infect. Dis. 200 Suppl 1, S70-75.

Binka, E., Vermund, S.H., Armah, G.E., 2011. Rotavirus diarrhea among children less than 5 years of age in urban Ghana. Pediatr. Infect. Dis. J. 30, 716-718.

Brister, J.R., Bao, Y., Zhdanov, S.A., Ostapchuck, Y., Chetvernin, V., Kiryutin, B., Zaslavsky, L., Kimelman, M., Tatusova, T.A., 2014. Virus Variation Resource - recent updates and future directions. Nucleic Acids Res. 42, D660-665.

Browning, G.F., Snodgrass, D.R., Nakagomi, O., Kaga, E., Sarasini, A., Gerna, G., 1992. Human and bovine serotype G8 rotaviruses may be derived by reassortment. Arch. Virol. 125, 121-128.

Bucardo, F., Rippinger, C.M., Svensson, L., Patton, J.T., 2012. Vaccine-derived NSP2 segment in rotaviruses from vaccinated children with gastroenteritis in Nicaragua. Infect. Genet. Evol. 12, 1282-1294.

Burke, B., Desselberger, U., 1996. Rotavirus pathogenicity. Virology 218, 299-305.

Cashman, O., Lennon, G., Sleator, R.D., Power, E., Fanning, S., O'Shea, H., 2010. Changing profile of the bovine rotavirus G6 population in the south of Ireland from 2002 to 2009. Vet. Microbiol. 146, 238-244.

Castello, A.A., Arguelles, M.H., Rota, R.P., Olthoff, A., Jiang, B., Glass, R.I., Gentsch, J.R., Glikmann, G., 2006. Molecular epidemiology of group A rotavirus diarrhea among children in Buenos Aires, Argentina, from 1999 to 2003 and emergence of the infrequent genotype G12. J. Clin. Microbiol. 44, 2046-2050.

Ciarlet, M., Hoffmann, C., Lorusso, E., Baselga, R., Cafiero, M.A., Banyai, K., Matthijnssens, J., Parreno, V., de Grazia, S., Buonavoglia, C., Martella, V., 2008. Genomic characterization of a novel group A lamb rotavirus isolated in Zaragoza, Spain. Virus Genes 37, 250-265.

Cooney, M.A., Gorrell, R.J., Palombo, E.A., 2001. Characterisation and phylogenetic analysis of the VP7 proteins of serotype G6 and G8 human rotaviruses. J. Med. Virol. 50, 462-467.

Cunliffe, N.A., Gentsch, J.R., Kirkwood, C.D., Gondwe, J.S., Dove, W., Nakagomi, O., Nakagomi, T., Hoshino, Y., Bresee, J.S., Glass, R.I., Molyneux, M.E., Hart, C.A., 2000. Molecular and serologic characterization of novel serotype G8 human rotavirus strains detected in Blantyre, Malawi. Virology 274, 309-320.

Cunliffe, N.A., Kilgore, P.E., Bresee, J.S., Steele, A.D., Luo, N., Hart, C.A., Glass, R.I., 1998. Epidemiology of rotavirus diarrhoea in Africa: a review to assess the need for rotavirus immunization. Bull. World Health Organ. 76, 525-537.

Cunliffe, N.A., Ngwira, B.M., Dove, W., Nakagomi, O., Nakagomi, T., Perez, A., Hart, C.A., Kazembe, P.N., Mwansambo, C.C., 2009. Serotype g12 rotaviruses, Lilongwe, Malawi. Emerg. Infect. Dis. 15, 87-90.

de Villiers, F.P., Sawyerr, T.N., de Villiers, G.K., 2009. The incidence and clinical presentation of infantile rotavirus diarrhoea in Sierra Leone. S. Afr. Med. J. 99, 249-252.

Delogu, R., Lo Presti, A., Ruggeri, F.M., Cella, E., Giovanetti, M., Ciccozzi, M., Ljubin-Sternak, S., Bukovski-Simonoski, S., Lukic-Grlic, A., Ianiro, G., Fiore, L., 2013. Full-genome characterization of a G8P[8] rotavirus that emerged among children with diarrhea in Croatia in 2006. J. Clin. Microbiol. 51, 1583-1588.

Dennis, A.F., McDonald, S.M., Payne, D.C., Mijatovic-Rustempasic, S., Esona, M.D., Edwards, K.M., Chappell, J.D., Patton, J.T., 2014. Molecular epidemiology of contemporary G2P[4] human rotaviruses cocirculating in a single U.S. community: footprints of a globally transitioning genotype. J. Virol. 88, 3789-3801.

Dennis, F.E., Fujii, Y., Haga, K., Damanka, S., Lartey, B., Agbemabiese, C.A., Ohta, N., Armah, G.E., Katayama, K., 2014. Identification of novel Ghanaian G8P[6] human-bovine reassortant rotavirus strain by next generation sequencing. PLoS One 9, e100699.

Do, L.P., Doan, Y.H., Nakagomi, T., Gauchan, P., Kaneko, M., Agbemabiese, C., Dang, A.D., Nakagomi, O., 2015. Whole genome analysis of Vietnamese G2P[4] rotavirus strains possessing the NSP2 gene sharing an ancestral sequence with Chinese sheep and goat rotavirus strains. Microbiol. Immunol. 59, 605-613.

Doan, Y.H., Nakagomi, T., Aboudy, Y., Silberstein, I., Behar-Novat, E., Nakagomi, O., Shulman, L.M., 2013. Identification by full-genome analysis of a bovine rotavirus transmitted directly to and causing diarrhea in a human child. J. Clin. Microbiol. 51, 182-189.

Doan, Y.H., Nakagomi, T., Agbemabiese, C.A., Nakagomi, O., 2015. Changes in the distribution of lineage constellations of G2P[4] Rotavirus A strains detected in Japan over 32 years (1980-2011). Infect. Genet. Evol. 34, 423-433.

Doan, Y.H., Nakagomi, T., Nakagomi, O., 2012. Repeated circulation over 6 years of intergenogroup mono-reassortant G2P[4] rotavirus strains with genotype N1 of the NSP2 gene. Infect. Genet. Evol. 12, 1202-1212.

Donato, C.M., Zhang, Z.A., Donker, N.C., Kirkwood, C.D., 2014. Characterization of G2P[4] rotavirus strains associated with increased detection in Australian states using the RotaTeq(R) vaccine during the 2010-2011 surveillance period. Infect. Genet. Evol. 28, 398-412.

Drummond, A.J., Ho, S.Y., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.

Drummond, A.J., Nicholls, G.K., Rodrigo, A.G., Solomon, W., 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics 161, 1307-1320.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969-1973.

Enweronu-Laryea, C.C., Boamah, I., Sifah, E., Diamenu, S.K., Armah, G., 2014. Decline in severe diarrhea hospitalizations after the introduction of rotavirus vaccination in Ghana: a prevalence study. BMC Infect. Dis. 14, 431.

Enweronu-Laryea, C.C., Sagoe, K.W., Damanka, S., Lartey, B., Armah, G.E., 2013. Rotavirus genotypes associated with childhood severe acute diarrhoea in southern Ghana: a cross-sectional study. Virol. J. 10, 287.

Enweronu-Laryea, C.C., Sagoe, K.W., Glover-Addy, H., Asmah, R.H., Mingle, J.A., Armah, G.E., 2012. Prevalence of severe acute rotavirus gastroenteritis and intussusceptions in Ghanaian children under 5 years of age. J. Infect. Dev. Ctries. 6, 148-155.

Esona, M.D., Geyer, A., Page, N., Trabelsi, A., Fodha, I., Aminu, M., Agbaya, V.A., Tsion, B., Kerin, T.K., Armah, G.E., Steele, A.D., Glass, R.I., Gentsch, J.R., 2009. Genomic characterization of human rotavirus G8 strains from the African rotavirus network: relationship to animal rotaviruses. J. Med. Virol. 81, 937-951.

Estes, M.K., Greenberg, H.B., 2013. Rotaviruses. In: Knipe, D. M., Howley, P.M (Eds.), Fields Virology. Wolters Kluwer Health/Lippincott, Williams and Wilkins, Philadelphia. pp. 1347–1401.

Esteves, A., Nordgren, J., Pereira, J., Fortes, F., Dimbu, R., Saraiva, N., Mendes, C., Istrate, C., 2016. Molecular epidemiology of rotavirus in four provinces of Angola before vaccine introduction. J. Med. Virol. 88, 1511-1520.

Fischer, T.K., Aaby, P., Molbak, K., Rodrigues, A., 2010. Rotavirus disease in Guinea-Bissau, West Africa: a review of longitudinal community and hospital studies. J. Infect. Dis. 202 Suppl, S239-242.

Fujii, Y., Shimoike, T., Takagi, H., Murakami, K., Todaka-Takai, R., Park, Y., Katayama, K., 2012. Amplification of all 11 RNA segments of group A rotaviruses based on reverse transcription polymerase chain reaction. Microbiol. Immunol. 56, 630-638.

Fujii Y., Nakagomi T., Nishimura N., Noguchi A., Miura S., Ito H., Doan Y.H., Takahashi T., Ozaki T., Katayama K., Nakagomi O. 2014. Spread and predominance in Japan of novel G1P[8] double-reassortant rotavirus strains possessing a DS-1-like genotype constellation typical of G2P[4] strains. Infect. Genet. Evol. 28, 426-433.

Fukai, K., Onoda, H., Itou, T., Sato, M., Miura, Y., Sakai, T., 2004a. Genetic and serological characterization of novel serotype G8 bovine group A rotavirus strains isolated in Japan. J. Vet. Med. Sci. 66, 1413-1416.

Fukai, K., Saito, T., Inoue, K., Sato, M., 2004b. Molecular characterization of novel P[14],G8 bovine group A rotavirus, Sun9, isolated in Japan. Virus Res. 105, 101-106.

Gautam, R., Mijatovic-Rustempasic, S., Roy, S., Esona, M.D., Lopez, B., Mencos, Y., Rey-Benito, G., Bowen, M.D., 2015. Full genomic characterization and phylogenetic analysis of a zoonotic human G8P[14] rotavirus strain detected in a sample from Guatemala. Infect. Genet. Evol. 33, 206-211.

Gentsch, J.R., Glass, R.I., Woods, P., Gouvea, V., Gorziglia, M., Flores, J., Das, B.K., Bhan, M.K., 1992. Identification of group A rotavirus gene 4 types by polymerase chain reaction. J. Clin. Microbiol. 30, 1365-1373.

Gentsch, J.R., Hull, J.J., Teel, E.N., Kerin, T.K., Freeman, M.M., Esona, M.D., Griffin, D.D., Bielfelt-Krall, B.P., Banyai, K., Jiang, B., Cortese, M.M., Glass, R.I., Parashar, U.D., collaborating laboratories of the National Rotavirus Strain Surveillance, S., 2009. G and P types of circulating rotavirus strains in the United States during 1996-2005: nine years of prevaccine data. J. Infect. Dis. 200 Suppl 1, S99-S105.

Gentsch, J.R., Laird, A.R., Bielfelt, B., Griffin, D.D., Banyai, K., Ramachandran, M., Jain, V., Cunliffe, N.A., Nakagomi, O., Kirkwood, C.D., Fischer, T.K., Parashar, U.D., Bresee, J.S., Jiang, B., Glass, R.I., 2005. Serotype diversity and reassortment between human and animal rotavirus strains: implications for rotavirus vaccine programs. J. Infect. Dis. 192 Suppl 1, S146-159.

German, A.C., Iturriza-Gomara, M., Dove, W., Sandrasegaram, M., Nakagomi, T., Nakagomi, O., Cunliffe, N., Radford, A.D., Morgan, K.L., 2015. Molecular epidemiology of rotavirus in cats in the United Kingdom. J. Clin. Microbiol. 53, 455-464.

Gerna, G., Sarasini, A., Parea, M., Arista, S., Miranda, P., Brussow, H., Hoshino, Y., Flores, J., 1992. Isolation and characterization of two distinct human rotavirus strains with G6 specificity. J. Clin. Microbiol. 30, 9-16.

Ghosh, S., Adachi, N., Gatheru, Z., Nyangao, J., Yamamoto, D., Ishino, M., Urushibara, N., Kobayashi, N., 2011a. Whole-genome analysis reveals the complex evolutionary dynamics of Kenyan G2P[4] human rotavirus strains. J. Gen. Virol. 92, 2201-2208.

Ghosh, S., Alam, M.M., Ahmed, M.U., Talukdar, R.I., Paul, S.K., Kobayashi, N., 2010. Complete genome constellation of a caprine group A rotavirus strain reveals common evolution with ruminant and human rotavirus strains. J. Gen. Virol. 91, 2367-2373.

Ghosh, S., Gatheru, Z., Nyangao, J., Adachi, N., Urushibara, N., Kobayashi, N., 2011b. Full genomic analysis of a G8P[1] rotavirus strain isolated from an asymptomatic infant in Kenya provides evidence for an artiodactyl-to-human interspecies transmission event. J. Med. Virol. 83, 367-376.

Ghosh, S., Kobayashi, N., 2011. Whole-genomic analysis of rotavirus strains: current status and future prospects. Future Microbiol. 6, 1049-1065.

Ghosh, S., Kobayashi, N., 2014. Genetic diversity and evolution of human rotaviruses based on whole genome. Br J Virol 1, 1-8.

Ghosh, S., Paul, S.K., Hossain, M.A., Alam, M.M., Ahmed, M.U., Kobayashi, N., 2011c. Full genomic analyses of two human G2P[4] rotavirus strains detected in 2005: identification of a caprine-like VP3 gene. J. Gen. Virol. 92, 1222-1227.

Ghosh, S., Taniguchi, K., Aida, S., Ganesh, B., Kobayashi, N., 2013a. Whole genomic analyses of equine group A rotaviruses from Japan: evidence for bovine-to-equine interspecies transmission and reassortment events. Vet. Microbiol. 166, 474-485.

Ghosh, S., Urushibara, N., Chawla-Sarkar, M., Krishnan, T., Kobayashi, N., 2013b. Whole genomic analyses of asymptomatic human G1P[6], G2P[6] and G3P[6] rotavirus strains reveal intergenogroup reassortment events and genome segments of artiodactyl origin. Infect. Genet. Evol. 16, 165-173.

Ghosh, S., Varghese, V., Samajdar, S., Bhattacharya, S.K., Kobayashi, N., Naik, T.N., 2007. Evidence for independent segregation of the VP6- and NSP4- encoding genes in porcine group A rotavirus G6P[13] strains. Arch. Virol. 152, 423-429.

Giambiagi, S., Gonzalez Rodriguez, I., Gomez, J., Burrone, O., 1994. A rearranged genomic segment 11 is common to different human rotaviruses. Arch. Virol. 136, 415-421.

Giammanco, G.M., Bonura, F., Zeller, M., Heylen, E., Van Ranst, M., Martella, V., Banyai, K., Matthijnssens, J., De Grazia, S., 2014. Evolution of DS-1-like human G2P[4] rotaviruses assessed by complete genome analyses. J. Gen. Virol. 95, 91-109.

Global Burden of Disease Diarrhoeal Diseases Collaborators, 2017. Estimates of global, regional, and national morbidity, mortality, and aetiologies of diarrhoeal diseases: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Infect. Dis. 17, 909-948.

Gomez, M.M., Carvalho-Costa, F.A., Volotao Ede, M., Rose, T.L., da Silva, M.F., Fialho, A.M., Assis, R.M., de Andrade Jda, S., Sa, A.C., Zeller, M., Heylen, E., Matthijnssens, J., Leite, J.P., 2014. Prevalence and genomic characterization of G2P[4] group A rotavirus strains during monovalent vaccine introduction in Brazil. Infect. Genet. Evol. 28, 486-494.

Gomez, M.M., Volotao, E.M., de Mendonca, M.C., Tort, L.F., da Silva, M.F., Leite, J.P., 2010. Detection of uncommon rotavirus A strains P[8]G8 and P[4]G8 in the city of Rio de Janeiro, 2002. J. Med. Virol. 82, 1272-1276.

Gouvea, V., Glass, R.I., Woods, P., Taniguchi, K., Clark, H.F., Forrester, B., Fang, Z.Y., 1990. Polymerase chain reaction amplification and typing of rotavirus nucleic acid from stool specimens. J. Clin. Microbiol. 28, 276-282.

Griffin, D.D., Nakagomi, T., Hoshino, Y., Nakagomi, O., Kirkwood, C.D., Parashar, U.D., Glass, R.I., Gentsch, J.R., National Rotavirus Surveillance, S., 2002. Characterization of nontypeable rotavirus strains from the United States: identification of a new rotavirus reassortant (P2A[6],G12) and rare P3[9] strains related to bovine rotaviruses. Virology 294, 256-269.

He, B., Yang, F., Yang, W., Zhang, Y., Feng, Y., Zhou, J., Xie, J., Feng, Y., Bao, X., Guo, H., Li, Y., Xia, L., Li, N., Matthijnssens, J., Zhang, H., Tu, C., 2013. Characterization of a novel G3P[3] rotavirus isolated from a lesser horseshoe bat: a distant relative of feline/canine rotaviruses. J. Virol. 87, 12357-12366.

Heylen, E., Batoko Likele, B., Zeller, M., Stevens, S., De Coster, S., Conceicao-Neto, N., Van Geet, C., Jacobs, J., Ngbonda, D., Van Ranst, M., Matthijnssens, J., 2014.

Rotavirus surveillance in Kisangani, the Democratic Republic of the Congo, reveals a high number of unusual genotypes and gene segments of animal origin in non-vaccinated symptomatic children. PLoS One 9, e100953.

Heylen, E., Zeller, M., Ciarlet, M., Lawrence, J., Steele, D., Van Ranst, M., Matthijnssens, J., 2015. Comparative analysis of pentavalent rotavirus vaccine strains and G8 rotaviruses identified during vaccine trial in Africa. Sci. Rep. 5, 14658.

Heylen, E., Zeller, M., Ciarlet, M., Lawrence, J., Steele, D., Van Ranst, M., Matthijnssens, J., 2016. Human P[6] Rotaviruses From Sub-Saharan Africa and Southeast Asia Are Closely Related to Those of Human P[4] and P[8] Rotaviruses Circulating Worldwide. J. Infect. Dis. 214, 1039-1049.

Hoa Tran, T.N., Nakagomi, T., Nakagomi, O., 2013. Evidence for genetic reassortment between human rotaviruses by full genome sequencing of G3P[4] and G2P[4] strains co-circulating in India. Trop. Med. Health 41, 13-20.

Hoa-Tran, T.N., Nakagomi, T., Vu, H.M., Do, L.P., Gauchan, P., Agbemabiese, C.A., Nguyen, T.T., Nakagomi, O., Thanh, N.T., 2016. Abrupt emergence and predominance in Vietnam of rotavirus A strains possessing a bovine-like G8 on a DS-1-like background. Arch. Virol. 161, 479-482.

Hu, L., Crawford, S.E., Czako, R., Cortes-Penfield, N.W., Smith, D.F., Le Pendu, J., Estes, M.K., Prasad, B.V., 2012. Cell attachment protein VP8* of a human rotavirus specifically interacts with A-type histo-blood group antigen. Nature 485, 256-259.

Ianiro, G., Delogu, R., Bonomo, P., Castiglia, P., Ruggeri, F.M., Fiore, L., 2014. Molecular characterization of human G8P[4] rotavirus strains in Italy: proposal of a more complete subclassification of the G8 genotype in three major lineages. Infect. Genet. Evol. 21, 129-133.

Ianiro, G., Delogu, R., Camilloni, B., Lorini, C., Ruggeri, F.M., Fiore, L., 2013. Detection of unusual G6 rotavirus strains in Italian children with diarrhoea during the 2011 surveillance season. J. Med. Virol. 85, 1860-1869.

Imbert-Marcille, B.M., Barbe, L., Dupe, M., Le Moullac-Vaidye, B., Besse, B., Peltier, C., Ruvoen-Clouet, N., Le Pendu, J., 2014. A FUT2 gene common polymorphism determines resistance to rotavirus A of the P[8] genotype. J. Infect. Dis. 209, 1227-1230.

Istrate, C., Sharma, S., Nordgren, J., Videira, E.C.S., Lopes, A., Piedade, J., Zaky, A., Lima, A., Neves, E., Veiga, J., Esteves, A., 2015. High rate of detection of G8P[6] rotavirus in children with acute gastroenteritis in Sao Tome and Principe. Arch. Virol. 160, 423-428.

Iturriza-Gomara, M., Isherwood, B., Desselberger, U., Gray, J., 2001. Reassortment in vivo: driving force for diversity of human rotavirus strains isolated in the United Kingdom between 1995 and 1999. J. Virol. 75, 3696-3705.

Jere, K.C., Chaguza, C., Bar-Zeev, N., Lowe, J., Peno, C., Kumwenda, B., Nakagomi, O., Tate, J.E., Parashar, U.D., Heyderman, R.S., French, N., Cunliffe, N.A., Miren, I.G., Consortium, V., 2017. Emergence of double- and triple-gene reassortant G1P[8] rotaviruses possessing a DS-1-like backbone post rotavirus vaccine introduction in Malawi. J. Virol. doi: 10.1128/JVI.01246-17.

Jere, K.C., Esona, M.D., Ali, Y.H., Peenze, I., Roy, S., Bowen, M.D., Saeed, I.K., Khalafalla, A.I., Nyaga, M.M., Mphahlele, J., Steele, D., Seheri, M.L., 2014. Novel NSP1 genotype characterised in an African camel G8P[11] rotavirus strain. Infect. Genet. Evol. 21, 58-66.

Jiang, X., Liu, Y., Tan, M., 2017. Histo-blood group antigens as receptors for rotavirus, new understanding on rotavirus epidemiology and vaccine strategy. Emerg Microbes Infect 6, e22.

Kaneko, M., Mochizuki, M., Nakagomi, O., Nakagomi, T., 2016. Whole genome characterization of a G6P[5] rotavirus A strain isolated from a stray cat in Japan. Vet. Microbiol. 188, 25-33.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Bio. Evol. 30, 772-780.

Kim, H.H., Matthijnssens, J., Kim, H.J., Kwon, H.J., Park, J.G., Son, K.Y., Ryu, E.H., Kim, D.S., Lee, W.S., Kang, M.I., Yang, D.K., Hyun, B.H., Park, S.I., Park, S.J., Cho, K.O., 2012. Full-length genomic analysis of porcine G9P[23] and G9P[7] rotavirus strains isolated from pigs with diarrhea in South Korea. Infect. Genet. Evol. 12, 1427-1435.

Kindler, E., Trojnar, E., Heckel, G., Otto, P.H., Johne, R., 2013. Analysis of rotavirus species diversity and evolution including the newly determined full-length genome sequences of rotavirus F and G. Infect. Genet. Evol. 14, 58-67.

Komoto, S., Adah, M.I., Ide, T., Yoshikawa, T., Taniguchi, K., 2016. Whole genomic analysis of human and bovine G8P[1] rotavirus strains isolated in Nigeria provides evidence for direct bovine-to-human interspecies transmission. Infect. Genet. Evol. 43, 424-433.

Komoto S., Tacharoenmuang R., Guntapong R., Ide T., Haga K., Katayama K., Kato T., Ouchi Y., Kurahashi H., Tsuji T., Sangkitporn S., Taniguchi K., 2015. Emergence and characterization of unusual DS-1-Like G1P[8] rotavirus strains in children with diarrhea in Thailand. PLoS One 10, e0141739.

Komoto S., Tacharoenmuang R., Guntapong R., Ide T., Tsuji T., Yoshikawa T., Tharmaphornpilas P., Sangkitporn S., Taniguchi K., 2016. Reassortment of human and animal rotavirus gene segments in emerging DS-1-Like G1P[8] rotavirus strains. PLoS One 11, e0148416. 675.

Kondo, K., Tsugawa, T., Ono, M., Ohara, T., Fujibayashi, S., Tahara, Y., Kubo, N., Nakata, S., Higashidate, Y., Fujii, Y., Katayama, K., Yoto, Y., Tsutsumi, H., 2017. Clinical and molecular characteristics of Human Rotavirus G8P[8] Outbreak Strain, Japan, 2014. Emerg. Infect. Dis. 23, 968-972.

Kuzuya M., Fujii R., Hamano M., Kida K., Mizoguchi Y., Kanadani T., Nishimura K., Kishimoto T., 2014. Prevalence and molecular characterization of G1P[8] human rotaviruses possessing DS-1-like VP6, NSP4, and NSP5/6 in Japan. J . Med. Virol. 86, 1056-1064.

Lanata, C.F., Fischer-Walker, C.L., Olascoaga, A.C., Torres, C.X., Aryee, M.J., Black, R.E., Child Health Epidemiology Reference Group of the World Health, O., Unicef, 2013.

Global causes of diarrheal disease mortality in children <5 years of age: a systematic review. PLoS One 8, e72788.

Louge Uriarte, E.L., Badaracco, A., Matthijnssens, J., Zeller, M., Heylen, E., Manazza, J., Mino, S., Van Ranst, M., Odeon, A., Parreno, V., 2014. The first caprine rotavirus detected in Argentina displays genomic features resembling virus strains infecting members of the Bovidae and Camelidae. Vet. Microbiol. 171, 189-197.

Maes, P., Matthijnssens, J., Rahman, M., Van Ranst, M., 2009. RotaC: a web-based tool for the complete genome classification of group A rotaviruses. BMC Microbiol. 9, 238.

Martel-Paradis, O., Laurin, M.A., Martella, V., Sohal, J.S., L'Homme, Y., 2013. Full-length genome analysis of G2, G9 and G11 porcine group A rotaviruses. Vet. Microbiol. 162, 94-102.

Martinez, M., Galeano, M.E., Akopov, A., Palacios, R., Russomando, G., Kirkness, E.F., Parra, G.I., 2014. Whole-genome analyses reveals the animal origin of a rotavirus G4P[6] detected in a child with severe diarrhea. Infect. Genet. Evol. 27, 156-162.

Matsuno, S., Hasegawa, A., Mukoyama, A., Inouye, S., 1985. A candidate for a new serotype of human rotavirus. J. Virol. 54, 623-624.

Matthijnssens, J., Ciarlet, M., Heiman, E., Arijs, I., Delbeke, T., McDonald, S.M., Palombo, E.A., Iturriza-Gomara, M., Maes, P., Patton, J.T., Rahman, M., Van Ranst, M., 2008a. 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. J. Virol. 82, 3204-3219.

Matthijnssens, J., Ciarlet, M., McDonald, S.M., Attoui, H., Banyai, K., Brister, J.R., Buesa, J., Esona, M.D., Estes, M.K., Gentsch, J.R., Iturriza-Gomara, M., Johne, R., Kirkwood, C.D., Martella, V., Mertens, P.P., Nakagomi, O., Parreno, V., Rahman, M., Ruggeri, F.M., Saif, L.J., Santos, N., Steyer, A., Taniguchi, K., Patton, J.T., Desselberger, U., Van Ranst, M., 2011. Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). Arch. Virol. 156, 1397-1413.

Matthijnssens, J., Ciarlet, M., Rahman, M., Attoui, H., Banyai, K., Estes, M.K., Gentsch, J.R., Iturriza-Gomara, M., Kirkwood, C.D., Martella, V., Mertens, P.P., Nakagomi, O., Patton, J.T., Ruggeri, F.M., Saif, L.J., Santos, N., Steyer, A., Taniguchi, K., Desselberger, U., Van Ranst, M., 2008b. Recommendations for the classification of group A rotaviruses using all 11 genomic RNA segments. Arch. Virol. 153, 1621-1629.

Matthijnssens, J., Heylen, E., Zeller, M., Rahman, M., Lemey, P., Van Ranst, M., 2010. Phylodynamic analyses of rotavirus genotypes G9 and G12 underscore their potential for swift global spread. Mol. Bio. Evol. 27, 2431-2436.

Matthijnssens, J., Otto, P.H., Ciarlet, M., Desselberger, U., Van Ranst, M., Johne, R., 2012. VP6-sequence-based cutoff values as a criterion for rotavirus species demarcation. Arch. Virol. 157, 1177-1182.

Matthijnssens, J., Potgieter, C.A., Ciarlet, M., Parreno, V., Martella, V., Banyai, K., Garaicoechea, L., Palombo, E.A., Novo, L., Zeller, M., Arista, S., Gerna, G., Rahman, M., Van Ranst, M., 2009. Are human P[14] rotavirus strains the result of interspecies transmissions from sheep or other ungulates that belong to the mammalian order Artiodactyla? J. Virol. 83, 2917-2929.

Matthijnssens, J., Rahman, M., Van Ranst, M., 2008c. Two out of the 11 genes of an unusual human G6P[6] rotavirus isolate are of bovine origin. J. Gen. Virol. 89, 2630-2635.

Matthijnssens, J., Rahman, M., Yang, X., Delbeke, T., Arijs, I., Kabue, J.P., Muyembe, J.J., Van Ranst, M., 2006. G8 rotavirus strains isolated in the Democratic Republic of Congo belong to the DS-1-like genogroup. J. Clin. Microbiol. 44, 1801-1809.

Matthijnssens, J., Van Ranst, M., 2012. Genotype constellation and evolution of group A rotaviruses infecting humans. Curr. Opin. Virol. 2, 426-433.

McDonald, S.M., Matthijnssens, J., McAllen, J.K., Hine, E., Overton, L., Wang, S., Lemey, P., Zeller, M., Van Ranst, M., Spiro, D.J., Patton, J.T., 2009. Evolutionary dynamics of human rotaviruses: balancing reassortment with preferred genome constellations. PLoS Pathog. 5, e1000634.

Medici, M.C., Abelli, L.A., Martinelli, M., Dettori, G., Chezzi, C., 2008. Molecular characterization of VP4, VP6 and VP7 genes of a rare G8P[14] rotavirus strain detected in an infant with gastroenteritis in Italy. Virus Res. 137, 163-167.

Midgley, S.E., Banyai, K., Buesa, J., Halaihel, N., Hjulsager, C.K., Jakab, F., Kaplon, J., Larsen, L.E., Monini, M., Poljsak-Prijatelj, M., Pothier, P., Ruggeri, F.M., Steyer, A., Koopmans, M., Bottiger, B., 2012. Diversity and zoonotic potential of rotaviruses in swine and cattle across Europe. Vet. Microbiol. 156, 238-245.

Mihalov-Kovacs, E., Gellert, A., Marton, S., Farkas, S.L., Feher, E., Oldal, M., Jakab, F., Martella, V., Banyai, K., 2015. Candidate new rotavirus species in sheltered dogs, Hungary. Emerg. Infect. Dis. 21, 660-663.

Monini, M., Cappuccini, F., Battista, P., Falcone, E., Lavazza, A., Ruggeri, F.M., 2008. Molecular characterization of bovine rotavirus strains circulating in northern Italy, 2003-2005. Vet. Microbiol. 129, 384-389.

Monini, M., Zaccaria, G., Ianiro, G., Lavazza, A., Vaccari, G., Ruggeri, F.M., 2014. Full-length genomic analysis of porcine rotavirus strains isolated from pigs with diarrhea in Northern Italy. Infect. Genet. Evol. 25, 4-13.

Mukherjee, A., Mullick, S., Deb, A.K., Panda, S., Chawla-Sarkar, M., 2013. First report of human rotavirus G8P[4] gastroenteritis in India: evidence of ruminants-to-human zoonotic transmission. J. Med. Virol. 85, 537-545.

Mwenda, J.M., Ntoto, K.M., Abebe, A., Enweronu-Laryea, C., Amina, I., McHomvu, J., Kisakye, A., Mpabalwani, E.M., Pazvakavambwa, I., Armah, G.E., Seheri, L.M., Kiulia, N.M., Page, N., Widdowson, M.A., Steele, A.D., 2010. Burden and epidemiology of rotavirus diarrhea in selected African countries: preliminary results from the African Rotavirus Surveillance Network. J. Infect. Dis. 202 Suppl, S5-S11.

Mwenda, J.M., Tate, J.E., Parashar, U.D., Mihigo, R., Agocs, M., Serhan, F., Nshimirimana, D., 2014. African rotavirus surveillance network: a brief overview. Pediatr. Infect. Dis. J. 33 Suppl 1, S6-8.

Naghipour, M., Nakagomi, T., Nakagomi, O., 2008. Issues with reducing the rotavirus-associated mortality by vaccination in developing countries. Vaccine 26, 3236-3241.

Nakagomi, O., Nakagomi, T., Akatani, K., Ikegami, N., 1989. Identification of rotavirus genogroups by RNA-RNA hybridization. Mol. Cell. Probes 3, 251-261.

Nakagomi, T., Chang, B.R., Nakagomi, O., 2009. Rotavirus hospitalization and molecular epidemiology in northern Japan, 1987-1996. Vaccine 27 Suppl 5, F93-96.

Nakagomi, T., Doan, Y.H., Dove, W., Ngwira, B., Iturriza-Gomara, M., Nakagomi, O., Cunliffe, N.A., 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. J. Gen. Virol. 94, 1273-1295.

Nakagomi, T., Nakagomi, O., 1989. RNA-RNA hybridization identifies a human rotavirus that is genetically related to feline rotavirus. J. Virol. 63, 1431-1434.

Nakagomi T., Nguyen M.Q., Gauchan P., Agbemabiese C.A., Kaneko M., Do L.P., Vu T.D., Nakagomi O., 2017. Evolution of DS-1-like G1P[8] double-gene reassortant rotavirus A strains causing gastroenteritis in children in Vietnam in 2012/2013. Arch. Virol. 162, 739-748.

Ndze, V.N., Esona, M.D., Achidi, E.A., Gonsu, K.H., Doro, R., Marton, S., Farkas, S., Ngeng, M.B., Ngu, A.F., Obama-Abena, M.T., Banyai, K., 2014. Full genome characterization of human Rotavirus A strains isolated in Cameroon, 2010-2011: diverse combinations of the G and P genes and lack of reassortment of the backbone genes. Infect. Genet. Evol. 28, 537-560.

Nielsen, N.M., Eugen-Olsen, J., Aaby, P., Molbak, K., Rodrigues, A., Fischer, T.K., 2005. Characterisation of rotavirus strains among hospitalised and non-hospitalised children in Guinea-Bissau, 2002 A high frequency of mixed infections with serotype G8. J. Clin. Virol. 34, 13-21.

Nordgren, J., Bonkoungou, I.J., Nitiema, L.W., Sharma, S., Ouermi, D., Simpore, J., Barro, N., Svensson, L., 2012a. Rotavirus in diarrheal children in rural Burkina Faso: high prevalence of genotype G6P[6]. Infect. Genet. Evol. 12, 1892-1898.

Nordgren, J., Nitiema, L.W., Sharma, S., Ouermi, D., Traore, A.S., Simpore, J., Svensson, L., 2012b. Emergence of unusual G6P[6] rotaviruses in children, Burkina Faso, 2009-2010. Emerg. Infect. Dis. 18, 589-597.

Nordgren, J., Sharma, S., Bucardo, F., Nasir, W., Gunaydin, G., Ouermi, D., Nitiema, L.W., Becker-Dreps, S., Simpore, J., Hammarstrom, L., Larson, G., Svensson, L., 2014. Both Lewis and secretor status mediate susceptibility to rotavirus infections in a rotavirus genotype-dependent manner. Clin. Infect. Dis. 59, 1567-1573.

Nyaga, M.M., Jere, K.C., Esona, M.D., Seheri, M.L., Stucker, K.M., Halpin, R.A., Akopov, A., Stockwell, T.B., Peenze, I., Diop, A., Ndiaye, K., Boula, A., Maphalala, G., Berejena, C., Mwenda, J.M., Steele, A.D., Wentworth, D.E., Mphahlele, M.J., 2015. Whole genome detection of rotavirus mixed infections in human, porcine and bovine samples co-infected with various rotavirus strains collected from sub-Saharan Africa. Infect. Genet. Evol. 31, 321-334.

Nyaga, M.M., Stucker, K.M., Esona, M.D., Jere, K.C., Mwinyi, B., Shonhai, A., Tsolenyanu, E., Mulindwa, A., Chibumbya, J.N., Adolfine, H., Halpin, R.A., Roy, S., Stockwell, T.B., Berejena, C., Seheri, M.L., Mwenda, J.M., Steele, A.D., Wentworth, D.E., Mphahlele, M.J., 2014. Whole-genome analyses of DS-1-like human G2P[4] and G8P[4] rotavirus strains from Eastern, Western and Southern Africa. Virus Genes 49, 196-207.

Ouermi, D., Soubeiga, D., Nadembega, W.M.C., Sawadogo, P.M., Zohoncon, T.M., Obiri-Yeboah, D., Djigma, F.W., Nordgren, J., Simpore, J., 2017. Molecular epidemiology of rotavirus in children under five in Africa (2006-2016): a systematic review. Pak. J. Biol. Sci. 20, 59-69.

Page, N.A., de Beer, M.C., Seheri, L.M., Dewar, J.B., Steele, A.D., 2009. The detection and molecular characterization of human G12 genotypes in South Africa. J. Med. Virol. 81, 106-113.

Palombo, E.A., Bishop, R.F., 1995. Genetic and antigenic characterization of a serotype G6 human rotavirus isolated in Melbourne, Australia. J. Med. Virol. 47, 348-354.

Palombo, E.A., Clark, R., Bishop, R.F., 2000. Characterisation of a "European-like" serotype G8 human rotavirus isolated in Australia. J. Med. Virol. 60, 56-62.

Papp, H., Laszlo, B., Jakab, F., Ganesh, B., De Grazia, S., Matthijnssens, J., Ciarlet, M., Martella, V., Banyai, K., 2013. Review of group A rotavirus strains reported in swine and cattle. Vet. Microbiol. 165, 190-199.

Papp, H., Malik, Y.S., Farkas, S.L., Jakab, F., Martella, V., Banyai, K., 2014. Rotavirus strains in neglected animal species including lambs, goats and camelids. Virus Dis. 25, 215-222.

Parashar, U.D., Burton, A., Lanata, C., Boschi-Pinto, C., Shibuya, K., Steele, D., Birmingham, M., Glass, R.I., 2009. Global mortality associated with rotavirus disease among children in 2004. J. Infect. Dis. 200 Suppl 1, S9-S15.

Parrish, C.R., Holmes, E.C., Morens, D.M., Park, E.C., Burke, D.S., Calisher, C.H., Laughlin, C.A., Saif, L.J., Daszak, P., 2008. Cross-species virus transmission and the emergence of new epidemic diseases. Microbiol. Mol. Biol. Rev. 72, 457-470.

Pietsch, C., Petersen, L., Patzer, L., Liebert, U.G., 2009. Molecular characteristics of German G8P[4] rotavirus strain GER1H-09 suggest that a genotyping and subclassification update is required for G8. J. Clin. Microbiol. 47, 3569-3576.

Pun, S.B., Nakagomi, T., Sherchand, J.B., Pandey, B.D., Cuevas, L.E., Cunliffe, N.A., Hart, C.A., Nakagomi, O., 2007. Detection of G12 human rotaviruses in Nepal. Emerg. Infect. Dis. 13, 482-484.

Rahman, M., De Leener, K., Goegebuer, T., Wollants, E., Van der Donck, I., Van Hoovels, L., Van Ranst, M., 2003. Genetic characterization of a novel, naturally occurring recombinant human G6P[6] rotavirus. J. Clin. Microbiol. 41, 2088-2095.

Rahman, M., Matthijnssens, J., Yang, X., Delbeke, T., Arijs, I., Taniguchi, K., Iturriza-Gomara, M., Iftekharuddin, N., Azim, T., Van Ranst, M., 2007. Evolutionary history and global spread of the emerging g12 human rotaviruses. J. Virol. 81, 2382-2390.

Rodrigues, A., de Carvalho, M., Monteiro, S., Mikkelsen, C.S., Aaby, P., Molbak, K., Fischer, T.K., 2007. Hospital surveillance of rotavirus infection and nosocomial transmission

of rotavirus disease among children in Guinea-Bissau. Pediatr. Infect. Dis. J. 26, 233-237.

Roy, S., Esona, M.D., Kirkness, E.F., Akopov, A., McAllen, J.K., Wikswo, M.E., Cortese, M.M., Payne, D.C., Parashar, U.D., Gentsch, J.R., Bowen, M.D., National Rotavirus Strain Surveillance, S., New Vaccine Surveillance, N., 2014. Comparative genomic analysis of genogroup 1 (Wa-like) rotaviruses circulating in the USA, 2006-2009. Infect. Genet. Evol. 28, 513-523.

Ruiz-Palacios, G.M., Perez-Schael, I., Velazquez, F.R., Abate, H., Breuer, T., Clemens, S.C., Cheuvart, B., Espinoza, F., Gillard, P., Innis, B.L., Cervantes, Y., Linhares, A.C., Lopez, P., Macias-Parra, M., Ortega-Barria, E., Richardson, V., Rivera-Medina, D.M., Rivera, L., Salinas, B., Pavia-Ruz, N., Salmeron, J., Ruttimann, R., Tinoco, J.C., Rubio, P., Nunez, E., Guerrero, M.L., Yarzabal, J.P., Damaso, S., Tornieporth, N., Saez-Llorens, X., Vergara, R.F., Vesikari, T., Bouckenooghe, A., Clemens, R., De Vos, B., O'Ryan, M., Human Rotavirus Vaccine Study, G., 2006. Safety and efficacy of an attenuated vaccine against severe rotavirus gastroenteritis. N. Engl. J. Med. 354, 11-22.

Sanchez-Padilla, E., Grais, R.F., Guerin, P.J., Steele, A.D., Burny, M.E., Luquero, F.J., 2009. Burden of disease and circulating serotypes of rotavirus infection in sub-Saharan Africa: systematic review and meta-analysis. Lancet Infect. Dis. 9, 567-576.

Santos, N., Hoshino, Y., 2005. Global distribution of rotavirus serotypes/genotypes and its implication for the development and implementation of an effective rotavirus vaccine. Rev. Med. Virol. 15, 29-56.

Sasaki, E., Nakagomi, T., Doan, Y.H., Gauchan, P., Kaneko, M., Nakagomi, O., 2015. Molecular identification of a G2 rotavirus that provided a G1P[4] mono-reassortant with a DS-1-like genotype constellation. J. Med. Virol. 87, 694-701.

Schwarz, G., 1978. Estimating the dimension of a model. Ann Stat 6, 461-464.

Seheri, M., Nemarude, L., Peenze, I., Netshifhefhe, L., Nyaga, M.M., Ngobeni, H.G., Maphalala, G., Maake, L.L., Steele, A.D., Mwenda, J.M., Mphahlele, J.M., 2014.

Update of rotavirus strains circulating in Africa from 2007 through 2011. Pediatr. Infect. Dis. J. 33 Suppl 1, S76-84.

Sieg, M., Ruckner, A., Kohler, C., Burgener, I., Vahlenkamp, T.W., 2015. A bovine G8P[1] group A rotavirus isolated from an asymptomatically infected dog. J. Gen. Virol. 96, 106-114.

Silva, F.D., Espinoza, L.R., Tonietti, P.O., Barbosa, B.R., Gregori, F., 2015. Whole-genomic analysis of 12 porcine group A rotaviruses isolated from symptomatic piglets in Brazil during the years of 2012-2013. Infect. Genet. Evol. 32, 239-254.

Silva, F.D., Gregori, F., McDonald, S.M., 2016. Distinguishing the genotype 1 genes and proteins of human Wa-like rotaviruses vs. porcine rotaviruses. Infect. Genet. Evol. 43, 6-14.

Snodgrass, D.R., Fitzgerald, T., Campbell, I., Scott, F.M., Browning, G.F., Miller, D.L., Herring, A.J., Greenberg, H.B., 1990. Rotavirus serotypes 6 and 10 predominate in cattle. J. Clin. Microbiol. 28, 504-507.

Steele, A.D., Ivanoff, B., 2003. Rotavirus strains circulating in Africa during 1996-1999: emergence of G9 strains and P[6] strains. Vaccine 21, 361-367.

Steele, A.D., Nimzing, L., Peenze, I., De Beer, M.C., Geyer, A., Angyo, I., Gomwalk, N.E., 2002. Circulation of the novel G9 and G8 rotavirus strains in Nigeria in 1998/1999. J. Med. Virol. 67, 608-612.

Steele, A.D., Parker, S.P., Peenze, I., Pager, C.T., Taylor, M.B., Cubitt, W.D., 1999. Comparative studies of human rotavirus serotype G8 strains recovered in South Africa and the United Kingdom. J. Gen. Virol. 80 (Pt 11), 3029-3034.

Steyer, A., Poljsak-Prijatelj, M., Barlic-Maganja, D., Marin, J., 2008. Human, porcine and bovine rotaviruses in Slovenia: evidence of interspecies transmission and genome reassortment. J. Gen. Virol. 89, 1690-1698.

Steyer, A., Poljsak-Prijatelj, M., Bufon, T.L., Marcun-Varda, N., Marin, J., 2007. Rotavirus genotypes in Slovenia: unexpected detection of G8P[8] and G12P[8] genotypes. J. Med. Virol. 79, 626-632.

Steyer, A., Sagadin, M., Kolenc, M., Poljsak-Prijatelj, M., 2013. Whole genome sequence analysis of bovine G6P[11] rotavirus strain found in a child with gastroenteritis. Infect. Genet. Evol. 13, 89-95.

Suchard, M.A., Rambaut, A., 2009. Many-core algorithms for statistical phylogenetics. Bioinformatics 25, 1370-1376.

Suzuki, Y., Sanekata, T., Sato, M., Tajima, K., Matsuda, Y., Nakagomi, O., 1993. Relative frequencies of G (VP7) and P (VP4) serotypes determined by polymerase chain reaction assays among Japanese bovine rotaviruses isolated in cell culture. J. Clin. Microbiol. 31, 3046-3049.

Tacharoenmuang, R., Komoto, S., Guntapong, R., Ide, T., Sinchai, P., Upachai, S., Yoshikawa, T., Tharmaphornpilas, P., Sangkitporn, S., Taniguchi, K., 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.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30, 2725-2729.

Tate, J.E., Burton, A.H., Boschi-Pinto, C., Parashar, U.D., World Health Organization-Coordinated Global Rotavirus Surveillance, N., 2016. Global, Regional, and National Estimates of Rotavirus Mortality in Children <5 Years of Age, 2000-2013. Clin. Infect. Dis. 62 Suppl 2, S96-S105.

Theuns, S., Heylen, E., Zeller, M., Roukaerts, I.D., Desmarets, L.M., Van Ranst, M., Nauwynck, H.J., Matthijnssens, J., 2015. Complete genome characterization of recent and ancient Belgian pig group A rotaviruses and assessment of their evolutionary relationship with human rotaviruses. J. Virol. 89, 1043-1057.

Todd, S., Page, N.A., Duncan Steele, A., Peenze, I., Cunliffe, N.A., 2010. Rotavirus strain types circulating in Africa: Review of studies published during 1997-2006. J. Infect. Dis. 202 Suppl, S34-42.

Trojnar, E., Sachsenroder, J., Twardziok, S., Reetz, J., Otto, P.H., Johne, R., 2013. Identification of an avian group A rotavirus containing a novel VP4 gene with a close relationship to those of mammalian rotaviruses. J. Gen. Virol. 94, 136-142.

Uchida, R., Pandey, B.D., Sherchand, J.B., Ahmed, K., Yokoo, M., Nakagomi, T., Cuevas, L.E., Cunliffe, N.A., Hart, C.A., Nakagomi, O., 2006. Molecular epidemiology of rotavirus diarrhea among children and adults in Nepal: detection of G12 strains with P[6] or P[8] and a G11P[25] strain. J. Clin. Microbiol. 44, 3499-3505.

Van Trang, N., Vu, H.T., Le, N.T., Huang, P., Jiang, X., Anh, D.D., 2014. Association between norovirus and rotavirus infection and histo-blood group antigen types in Vietnamese children. J. Clin. Microbiol. 52, 1366-1374.

Vesikari, T., Matson, D.O., Dennehy, P., Van Damme, P., Santosham, M., Rodriguez, Z., Dallas, M.J., Heyse, J.F., Goveia, M.G., Black, S.B., Shinefield, H.R., Christie, C.D., Ylitalo, S., Itzler, R.F., Coia, M.L., Onorato, M.T., Adeyi, B.A., Marshall, G.S., Gothefors, L., Campens, D., Karvonen, A., Watt, J.P., O'Brien, K.L., DiNubile, M.J., Clark, H.F., Boslego, J.W., Offit, P.A., Heaton, P.M., Rotavirus, E., Safety Trial Study, T., 2006. Safety and efficacy of a pentavalent human-bovine (WC3) reassortant rotavirus vaccine. N. Engl. J. Med. 354, 23-33.

Weinberg, G.A., Payne, D.C., Teel, E.N., Mijatovic-Rustempasic, S., Bowen, M.D., Wikswo, M., Gentsch, J.R., Parashar, U.D., 2012. First reports of human rotavirus G8P[4] gastroenteritis in the United States. J. Clin. Microbiol. 50, 1118-1121.

WHO, 2009. Rotavirus vaccines:an update. Wkly. Epidemiol. Rec. 84, 533-540.

Yamamoto, D., Kawaguchiya, M., Ghosh, S., Ichikawa, M., Numazaki, K., Kobayashi, N., 2011. Detection and full genomic analysis of G6P[9] human rotavirus in Japan. Virus Genes 43, 215-223.

Yamamoto S.P., Kaida A., Kubo H., Iritani N., 2014. Gastroenteritis outbreaks caused by a DS-1-like G1P[8] rotavirus strain, Japan, 2012-2013. Emerg. Infect. Dis. 20, 1030-1033.

Zhang, C.Y., Wei, J.F., He, S.H., 2006. Adaptive evolution of the spike gene of SARS coronavirus: changes in positively selected sites in different epidemic groups. BMC Microbiol. 6, 88.

Zeller, M., Donato, C., Trovao, N.S., Cowley, D., Heylen, E., Donker, N.C., McAllen, J.K., Akopov, A., Kirkness, E.F., Lemey, P., Van Ranst, M., Matthijnssens, J., Kirkwood, C.D., 2015. Genome-Wide Evolutionary Analyses of G1P[8] Strains Isolated Before and After Rotavirus Vaccine Introduction. Genome Biol. Evol. 7, 2473-2483.

Zhou, N., Lv, D., Wang, S., Lin, X., Bi, Z., Wang, H., Wang, P., Zhang, H., Tao, Z., Hou, P., Song, Y., Xu, A., 2016. Continuous detection and genetic diversity of human rotavirus A in sewage in eastern China, 2013-2014. Virol J. 13, 153.

Supplementary Materials

CHAPTER II

Supplementary Table 2.1: Primers used to amplify and sequence the full genome of PML1965

Gene	Name of Primer	Sequence (5'-3')	Reference
VP7	Beg9	GGCTTTAAAAGAGAGAATTTCCGTCTGG	Gouvea *et al.,* 1990
	End9	GGTCACATCATACAATTCTAATCTAAG	Gouvea *et al.,* 1990
VP4	VP4-1F	GGCTATAAAATGGCTTCGCTC	Doan *et al.,* 2012
	Con2	ATTTCGGACCATTTATAACC	Gentsch *et al.*, 1992
	SK88-VP4-454F	GCCGAATTTCAACATAAGCG	This study
	SK88-VP4-1553R	TGCGTCATTGCTATTTCTTG	This study
	SK12-1296F	TAACCGTTGAAGAGCCACC	This study
	P6-VP4-2356R	GGTCACATCCTCTATAGAGCTCTC	This study
VP6	GEN_VP6F	GGCTTTWAAACGAAGTCTTC	Matthijnssens *et al.*, 2008a
	GEN_VP6R	GGTCACATCCTCTCACT	Matthijnssens *et al.*, 2008a
VP1	VP1-1F	GGCTATTAAAGCTATACAATGGG	Doan *et al.,* 2012
	VP1-745F	CTGGTGTCATCTCCAATGTC	Doan *et al.,* 2012
	VP1-921R	AACCAATCCAGCTTTCCTC	Doan *et al.,* 2012
	VP1-1375F	GACGTACCAGGAAGAAGAAC	Doan *et al.,* 2012
	VP1-1604R	TGCTGAGATGAATCCCATTG	Doan *et al.,* 2012
	VP1-2343F	TGGAACAACTGATGATGAAGTG	Doan *et al.,* 2012
	VP1-2567R	CTCTTCTCGAGCGATATTGG	Doan *et al.,* 2012
	VP1-3302-R	GGTCACATCTAAGCGCTCTAATC	Doan *et al.,* 2012
VP2	GEN_VP2Fc	GGCTATTAAAGGYTCAATGGCGTACAG	Matthijnssens *et al.*, 2008a
	GEN_VP2_Rbc	GTCATATCTCCACARTGGGGTTGG	Matthijnssens *et al.*, 2008a
	VP2-1-438/457F	TGCGTAATAGATGGTATTGG	Doan *et al.,* 2012
	VP2-F-888/906	TACCATCAACTGCCAGATA	Doan *et al.,* 2012
	C2V2-1783F	GCTTATTGGAAATGCGACTG	This study
	VP2-1329F	CATTTGGAATGCAACGAATG	Doan *et al.,* 2012
	VP2-1791R	ACAGTCGCGTTTCCAATAAG	Doan *et al.,* 2012
	VP2-R-2355/2373	ACTATTTGAGCGAACACTG	Doan *et al.,* 2012
VP3	MAX-3aF	GGCTATTAAAGCAGTACCAG	Matthijnssens *et al.*, 2008a
	MAX-3bR	GGTCACATCATGACTAGTGTG	Matthijnssens *et al.*, 2008a

	VP3-598F	AGCAAGACTTTCAAATCGCG	Doan *et al*., 2012
	VP3-792R	CCAAGCCATCTCTCTTGTTTG	Doan *et al*., 2012
	VP3-1102F	GGAATGGCGAAAAATGGTAG	Doan *et al*., 2012
	VP3-1592R	AATCCGTTGGCAAAAATGTC	Doan *et al*., 2012
	VP3-1781F	TTTTCGGAGTCAGCCACTTC	Doan *et al*., 2012
	VP3-2054R	GCAGTTTCGCGAATTCTCTC	Doan *et al*., 2012
NSP1	GEN_NSP1F	GGCTTTTTTTTATGAAAAGTCTTG	Matthijnssens *et al*., 2008a
	GEN_NSP1R	GGTCACATTTTATGCTGCC	Matthijnssens *et al*., 2008a
	NSP1-R-1114/1095	ATGTCCCACGTAAAGATTTG	Doan *et al.,* 2012
NSP2	GEN_NSP2F	GGCTTTTAAAGCGTCTCAG	Matthijnssens *et al*., 2008a
	GEN_NSP2R	GGTCACATAAGCGCTTTC	Matthijnssens *et al*., 2008a
NSP3	GEN_NSP3F	GGCTTTTAATGCTTTTCAGTG	Matthijnssens *et al*., 2008a
	GEN_NSP3R	ACATAACGCCCCTATAGC	Matthijnssens *et al*., 2008a
NSP4	GEN_NSP4F	GGCTTTTAAAAGTTCTGTTCC	Matthijnssens *et al*., 2008a
	GEN_NSP4R	GGWYACRYTAAGACCRTTCC	Matthijnssens *et al*., 2008a
NSP5	Beg11	GGCTTTTAAAGCGCTACAGTGATG	Giambiagi *et al*., 1994
	End11	GGTCACAAAACGGGAGTGGG	Watanabe et al., 2001

Supplementary Table 2.2: Dated G6 RVA strains included in the dataset for VP7 evolutionary rate estimation

	Bovine strains	Nucleotide accession number		Human strains	Nucleotide accession number
1	RVA/Cow-tc/USA/NCDV/1971/G6P[1]	JF693034	1	RVA/Human-wt/GHA/GHPML1965/2012/G6P[6]	LC026103
2	RVA/Cow-tc/JPN/BRV101/1986/G6P[1]	AB747358	2	RVA/Human-wt/GNB/MRC-DPRU5615/2011/G6P[6]	KJ752355
3	RVA/Cow-tc/JPN/BRV106/1983/G6P[1]	AB747360	3	RVA/Human-wt/GNB/MRC-DPRU5625/2011/G6P[6]	KJ752122
4	RVA/Cow-tc/JPN/BRV105/1983/G6P[1]	AB747359	4	RVA/Human-wt/BFA/8-BF/2010/G6P[6]	JX154487
5	RVA/Cow-wt/ARG/B517_B_BA/1998/G6P[5]	KC895771	5	RVA/Human-wt/BFA/3-BF/2010/G6P[6]	JX154484
6	RVA/Cow-wt/ARG/B3553_D_BA/2008/G6P[5]	KC895770	6	RVA/Human-wt/BFA/265-BF/2010/G6P[6]	JN116531
7	RVA/Cow-wt/ARG/B3552_D_BA/2008/G6P[11]	KC895756	7	RVA/Human-wt/BFA/48-BF/2010/G6P[6]	JX154523
8	RVA/Cow-wt/ARG/B1730_B_BA/2002/G6P[5]	KC895765	8	RVA/Human-wt/BFA/240-BF/2010/G6P[6]	JN116528
9	RVA/Cow-wt/ARG/B1230_RN/2000/G6P[5]	KC895763	9	RVA/Human-wt/BFA/238-BF/2010/G6P[6]	JN116527
10	RVA/Cow-wt/ARG/B684_B_BA/1999/G6P[5]	KC895758	10	RVA/Human-wt/BFA/2-BF/2010/G6P[6]	JX154483
11	RVA/Cow-wt/ARG/B1225_B_BA/2000/G6P[5]	KC895762	11	RVA/Human-wt/BFA/52-BF/2010/G6P[6]	JX154527
12	RVA/Cow-wt/ARG/B794_BA/1999/G6P[5]	KC895759	12	RVA/Human-wt/BFA/51-BF/2010/G6P[6]	JX154526
13	RVA/Cow-wt/ARG/B2184_B_SF/2004/G6P[5]	KC895766	13	RVA/Human-wt/BFA/19-BF/2010/G6P[6]	JX154497
14	RVA/Cow-wt/ARG/B2188_B_SF/2004/G6P[11]	KC895753	14	RVA/Human-wt/BFA/18-BF/2010/G6P[6]	JX154496
15	RVA/Cow-wt/ARG/B2818_B_BA/2005/G6P[5]	KC895768	15	RVA/Human-wt/BFA/38-BF/2010/G6P[6]	JX154513
16	RVA/Cow-wt/ARG/B2769_B_SF/2005/G6P[5]	KC895767	16	RVA/Human-wt/BFA/21-BF/2010/G6P[6]	JX154499
17	RVA/Cow-wt/ARG/B1186_B_ER/2000/G6P[5]	KC895761	17	RVA/Human-wt/BFA/14-BF/2010/G6P[6]	JX15449
18	RVA/Cow-wt/ARG/B3515_BA/2008/G6P[11]	KC895782	18	RVA/Human-wt/BFA/17-BF/2010/G6P[6]	JX154495
19	RVA/Cow-wt/ARG/B1115_B_Co/2000/G6P[5]	KC895760	19	RVA/Human-wt/BFA/285-BF/2010/G6P[6]	JN116534
20	RVA/Cow-wt/ARG/B2955_BA/2006/G6P[5]	KC895776	20	RVA/Human-wt/BFA/263-BF/2010/G6P[6]	JN116530
21	RVA/Cow-wt/ARG/B2590_B_BA/2004/G6P[5]	KC895773	21	RVA/Human-wt/BFA/249-BF/2010/G6P[6]	JN116529
22	RVA/Cow-wt/ARG/B1501_B_BA/2001/G6P[5]	KC895764	22	RVA/Human-wt/BFA/277-BF/2010/G6P[6]	JN116533
23	RVA/Cow-wt/ARG/B3486/2007/G6P[x]	KC895754	23	RVA/Human-wt/BFA/234-BF/2010/G6P[6]	JN116526
24	RVA/Cow-wt/ARG/B1965_B_BA/2002/G6P[5]	KC895774	24	RVA/Human-wt/BFA/307-BF/2010/G6P[6]	JN116536
25	RVA/Cow-wt/ARG/B195_B_BA/1997/G6P[5]	KC895757	25	RVA/Human-wt/BFA/272-BF/2010/G6P[6]	JN116532
26	RVA/Cow-wt/ARG/B3035_B_BA/2007/G6P[5]	KC895772	26	RVA/Human-wt/BFA/50-BF/2010/G6P[6]	JX154525

27	RVA/Cow-wt/ARG/B3207_BA/2006/G6P[5]	KC895769	27	RVA/Human-wt/ITA/CEC06/2011/G6P[6]	KC152909
28	RVA/Cow-wt/ARG/B2592_B_Co/2004/G6P[11]	KC895755	28	RVA/Human-wt/CMR/ES298/2011/G6P[6]	KM660403
29	RVA/Cow-tc/USA/UK_WT_BRV4A/1986/G6P[5]	JF693067	29	RVA/Human-wt/COG/12-G0868/2012/G6P[6]	KC510183
30	RVA/Cow-tc/USA/UK/1984/G6P[5]	JF693056	30	RVA/Human-wt/BFA/288-BF/2010/G6P[6]	JN116535
31	RVA/Cow/CAN/FMV1089933/2009/G6P[x]	JX470523	31	RVA/Human-wt/FRA/R353/2005/G6P[6]	DQ122400
32	RVA/Cow/CAN/FMV1081508/2009/G6P[x]	JX470519	32	RVA/Human-wt/BEL/B1711/2002/G6P[6]	EF554087
33	RVA/Cow-wt/FRA/V057/2010/G6P[5]	HE646655	33	RVA/Human-wt/CMR/MA228/2011/G6P[6]	KM660402
34	RVA/Cow-wt/FRA/V026/2010/G6P[5]	HE646647	34	RVA/Human-wt/CMR/MA202/2011/G6P[6]	KM660401
35	RVA/Cow-wt/FRA/V056/2010/G6P[5]	HE646654	35	RVA/Human-wt/CMR/BA346/2010/G6P[6]	KM660404
36	RVA/Cow-wt/FRA/V019/2010/G6P[5]	HE646643	36	RVA/Human-wt/ITA/PG05/2011/G6P[9]	KC152917
37	RVA/Cow-wt/FRA/V025/2010/G6P[5]	HE646646	37	RVA/Human-wt/TUN/17237/2008/G6P[9]	JX271006
38	RVA/Cow-wt/ZAF/1603/2007/G6P[5]	JN831214	38	RVA/Human-xx/ITA/PA43/2003/G6P[9]	JF793944
39	RVA/Cow-wt/ZAF/MRC-DPRU1604/2007/G6P[1]	KF636261	39	RVA/Human-xx/ITA/PA17/2003/G6P[9]	JF793943
40	RVA/Cow-wt/BRA/BRA1532/2009/G6P[5]	JQ943578	40	RVA/Human-wt/JPN/KF17/2010/G6P[9]	JF421980
41	RVA/Cow-wt/BRA/BRA1527/2009/G6P[5]	JQ943566	41	RVA/Human-tc/HUN/Hun7/1998/G6P[9]	AJ488134
42	RVA/Cow-wt/BRA/BRA1536/2009/G6P[5]	JQ943567	42	RVA/Human-tc/USA/Se584/1998/G6P[9]	EF672609
43	RVA/Cow-wt/BRA/BRA1526/2009/G6P[5]	JQ943565	43	RVA/Human-tc/ITA/PA151/1987/G6P[9]	AF532202
44	RVA/Cow-wt/BRA/BRA1513/2009/G6P[5]	JQ943561	44	RVA/Human-wt/ITA/PA27-GV1/1993/G6P[9]	JF793942
45	RVA/Cow-wt/BRA/BRA1520/2009/G6P[5]	JQ94356	45	RVA/Human-xx/HUN/Hun3/1995/G6P[9]	AJ487831
46	RVA/Cow-wt/BRA/BRA1510/2009/G6P[5]	JQ943560	46	RVA/Human-tc/HUN/Hun4/1996/G6P[9]	AJ487833
47	RVA/Cow-wt/BRA/BRA1506/2009/G6P[5]	JQ943559	47	RVA/Human-wt/ITA/111-05-27/2005/G6P[14]	EF554142
48	RVA/Cow-wt/BRA/BRA1503/2009/G6P[5]	JQ943557	48	RVA/Human-wt/ITA/PA77/2002/G6P[14]	JF793946
49	RVA/Cow-wt/BRA/BRA1522/2009/G6P[5]	JQ943563	49	RVA/Human-wt/BEL/B10925/1997/G6P[14]	EF554120
50	RVA/Cow-wt/BRA/BRA1505/2009/G6P[5]	JQ943558	50	RVA/Human-tc/ITA/PA169/1988/G6P[14]	EF554131
51	RVA/Cow-wt/ZAF/MRC-DPRU3010/2009/G6P[5]	KJ752067	51	RVA/Human-tc/AUS/MG6/1993/G6P[14]	EF554098
52	RVA/Cow-xx/IND/AM-B71/2012/G6[x]	KF170899	52	RVA/Human-wt/HUN/BP1879/2003/G6P[14]	FN665685
53	RVA/Cow-xx/IND/UKD/PTN/P-43/2009/ G6[x]	HM591496	53	RVA/Human-tc/HUN/Hun5/1997/G6P[14]	EF554109
54	RVA/Cow-xx/IND/UKD-PTN/2010/PBC/G6 [x]	JX442784			
55	RVA/Cow-xx/IND/HR/2011/B111/G6 [x]	JX442777			
56	RVA/Cow-xx/IND/MP-JBP/B82/2008/G6[x]	JX442766			

57	RVA/Cow-xx/IND/WB/2011/G6P[x]	JX442787
58	RVA/Cow/C91/IVRI/India/2011/G6P[x]	JN638724
59	RVA/Cow-wt/SVN/SI-B17/2004/G6P[11]	JX094039
60	RVA/Cow-wt/ARG/B1190_B_ER/2000/G6P[11]	KC895784
61	RVA/Cow-wt/BRA/BRA1758/2011/G6P[11]	KC67869
62	RVA/Cow-wt/BRA/BRA1752/2011/G6P[11]	KC678695
63	RVA/Cow-wt/BRA/BRA1744/2011/G6P[11]	KC678694
64	RVA/Cow-wt/BRA/BRA1743/2011/G6P[11]	KC678693
65	RVA/Cow-wt/ARG/B1541/2001/G6P[11]	KC895794
66	RVA/Cow-wt/ARG_B175_D_BA/1997/G6P[11]	KC895783
67	RVA/Cow-wt/ARG/B611_BA_/1999/G6P[11]	KC895781
68	RVA/Cow-wt/ARG/B609_BA_/1999/G6P[11]	KC895780
69	RVA/Cow-wt/ARG/B3206_BA/2006/G6P[5]	KC895775
70	RVA/Cow-wt/ARG/B3355_D_BA/2007/G6P[11]	KC895789
71	RVA/Cow-wt/ARG/B3099_D_/2006/G6P[11]	KC895787
72	RVA/Cow-wt/ARG/B3708_D_BA/2008/G6P[11]	KC895779
73	RVA/Cow-wt/ARG/B3700_D_BA/2008/G6P[11]	KC895778
74	RVA/Cow-wt/ARG/B3702_D_BA/2008/G6P[11]	KC895777
75	RVA/Cow-wt/ARG/B3100_BA_/2006/G6P[11]	KC895788
76	RVA/Cow-wt/ARG/B2932/2006/G6P[5]	KC895786
77	RVA/Cow-wt/ARG/B3538_D_BA/2008/G6P[11]	KC895793
78	RVA/Cow-wt/ARG/B3516_D_BA/2008/G6P[11]	KC895791
79	RVA/Cow-wt/ARG/B3530_D_BA/2008/G6P[11]	KC895792
80	RVA/Cow-wt/ARG/B3893_D_BA/2008/G6P[11]	KC895790
81	RVA/Cow-wt/ARG/B761_D_BA/1999/G6P[11]	KC895785
82	RVA/Cow-wt/ZAF/MRC-DPRU3005/2009/G6P[11]	KJ751927
83	RVA/Cow-wt/TUR/Aksaray/2005/G6P[11]	JX131338
84	RVA/Cow/IND/CRB200/IVRI/2011/G6P[11]	KC416962
85	RVA/Cow-xx/IND/CRB157/IVRI/2011/G6P[11]	KC416961

Supplementary Table 2.3: Regions sequenced for the 11 genes of PML1965 and GenBank accession numbers

Genes	Nucleotide Accession number	Sequence minus primer sequence (bp)
VP7	LC026103	29-1035
VP4	LC026104	22-2335
VP6	LC026105	21-1339
VP1	LC026106	24-3279
VP2	LC026107	28-2659
VP3	LC026108	21-2570
NSP1	LC026109	25-1548
NSP2	LC026110	20-1041
NSP3	LC026111	22-1044
NSP4	LC026112	22-731
NSP5	LC026113	25-796

Supplementary Figures: Fig. S1(a) - Fig. S1(i)

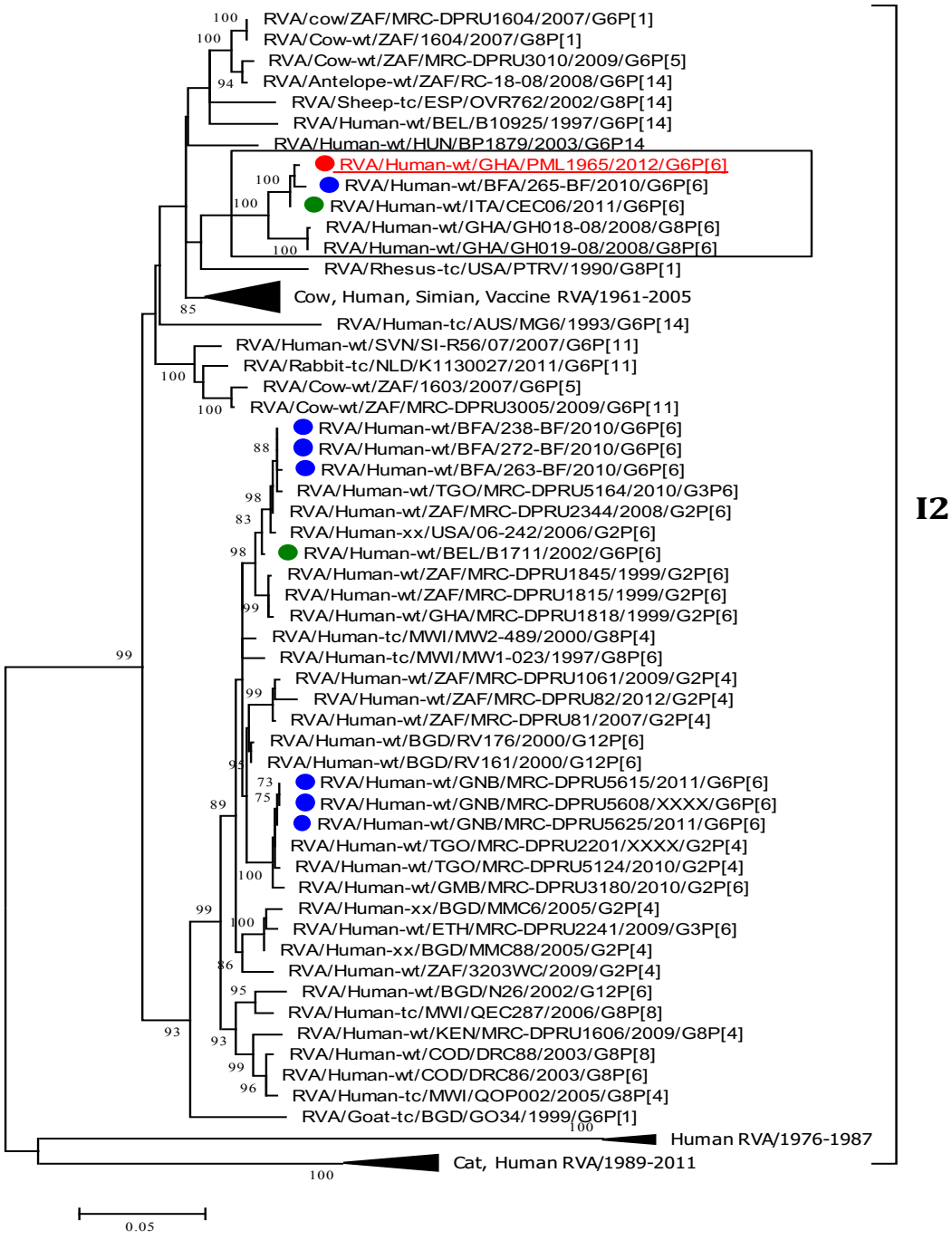


Fig. S1 (a)- (i): Phylogenetic trees based on near-full length ORF of VP6, the internal and non-structural protein genes (a) - (d) VP6, VP1, VP2, VP3; (e) - (i) NSP1-NSP5 respectively of the Ghanaian G6P[6] strain PML1965 in this study (indicated in red font with red dot), other African (indicated by blue dots), European (indicated by green dots) G6P[6] strains and representative human and animal DS-1-like strains. Maximum likelihood phylogenetic trees were constructed in MEGA6 software package, and the resulting tree presented here is a midpoint-rooted tree. Significant bootstrap values (1000 replicates) of ≥70% are indicated at each node. The scale bar at the bottom of each tree indicates a genetic distance expressed as nucleotide substitutions per site.

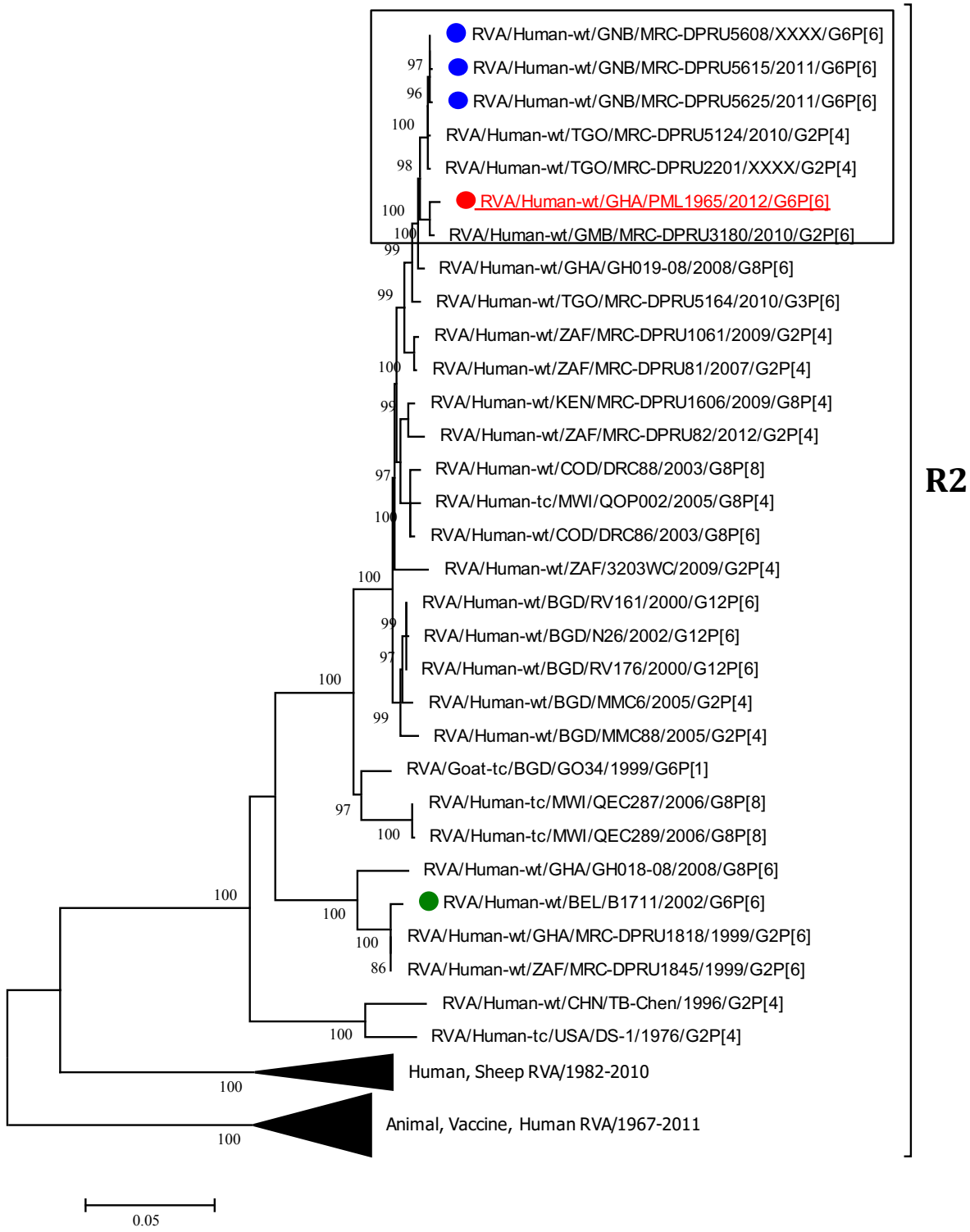


Fig. S1b : VP1 tree

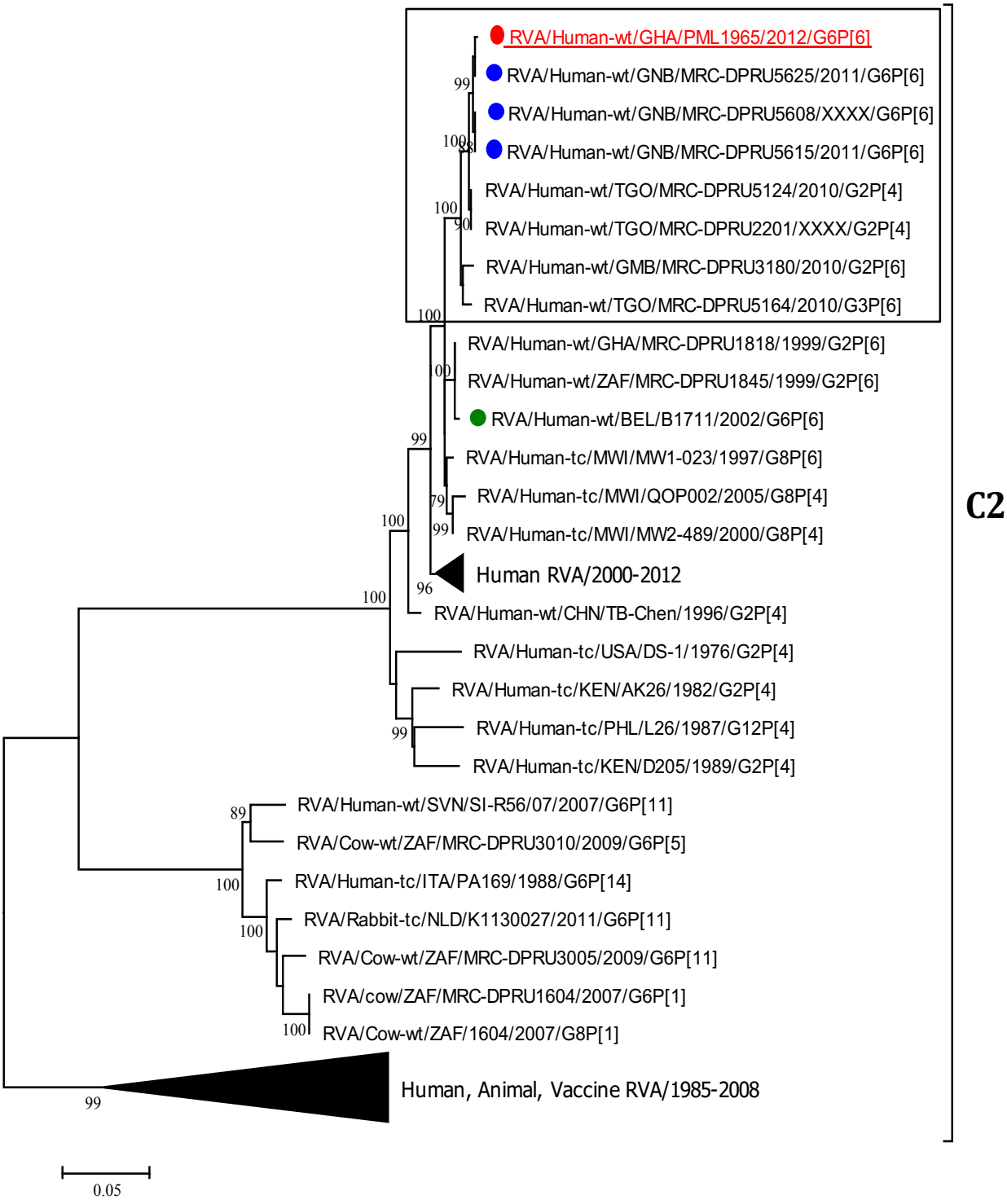


Fig. S1c. VP2 tree

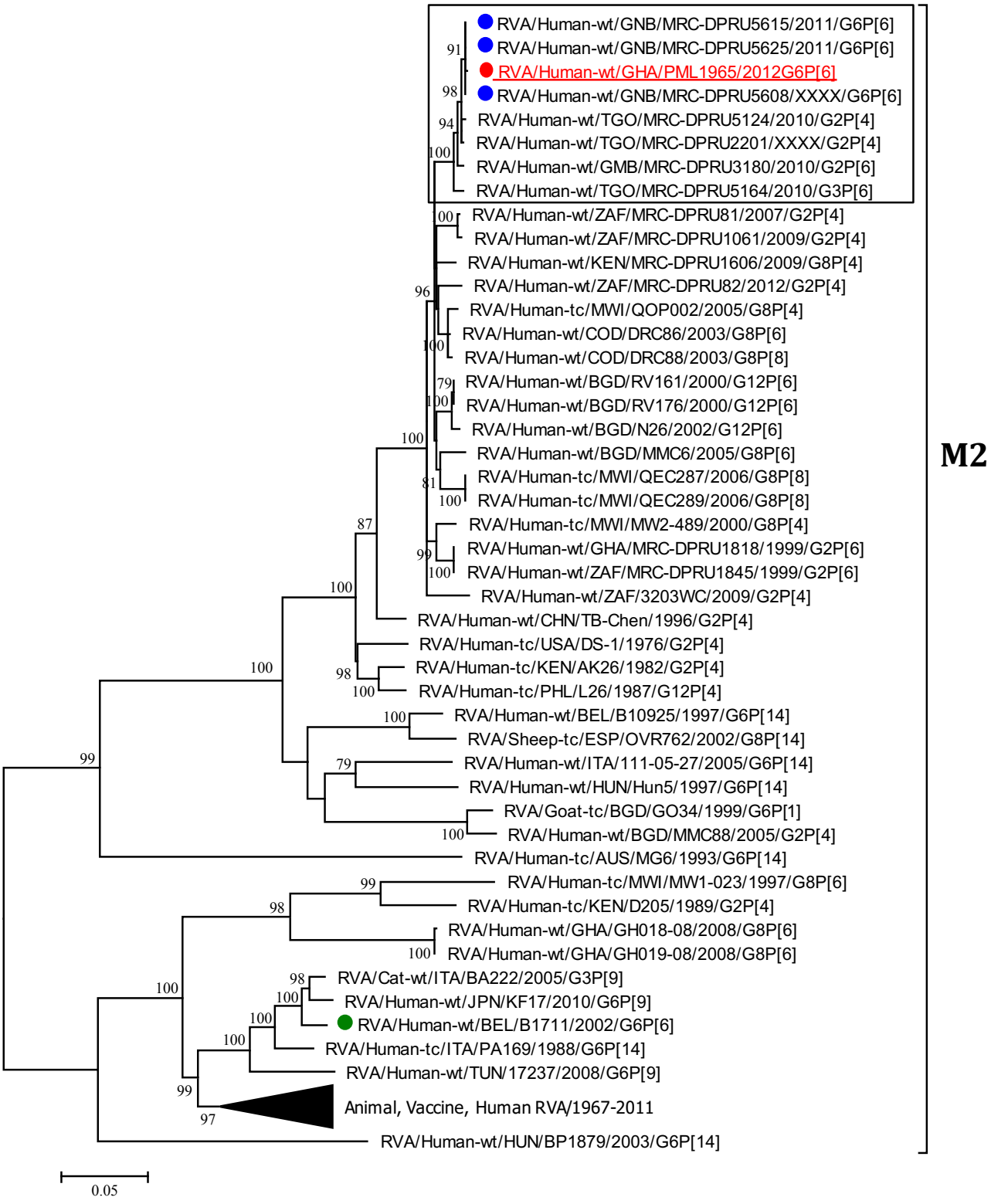


Fig. S1d: VP3 tree

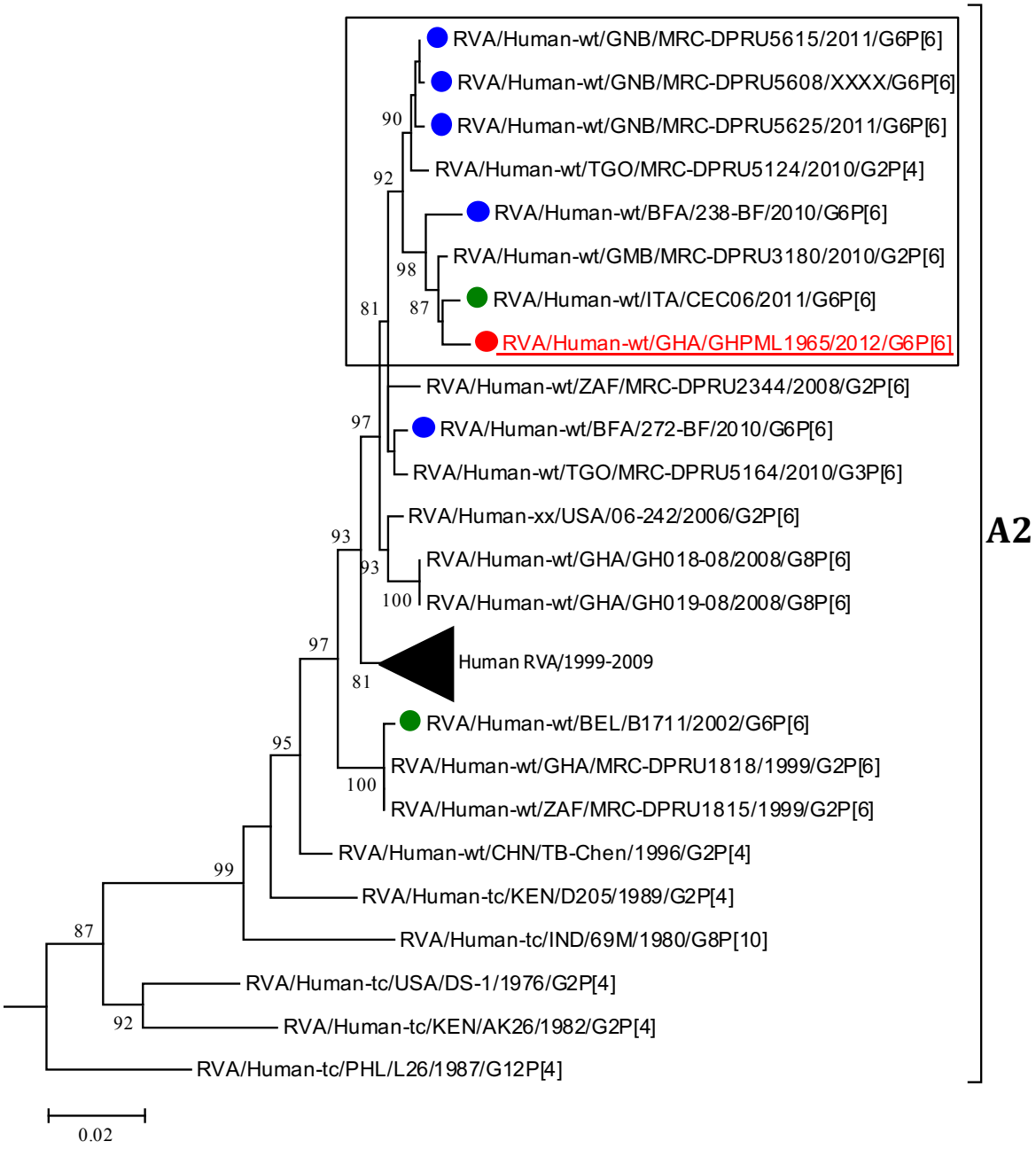


Fig. S1e NSP1 tree

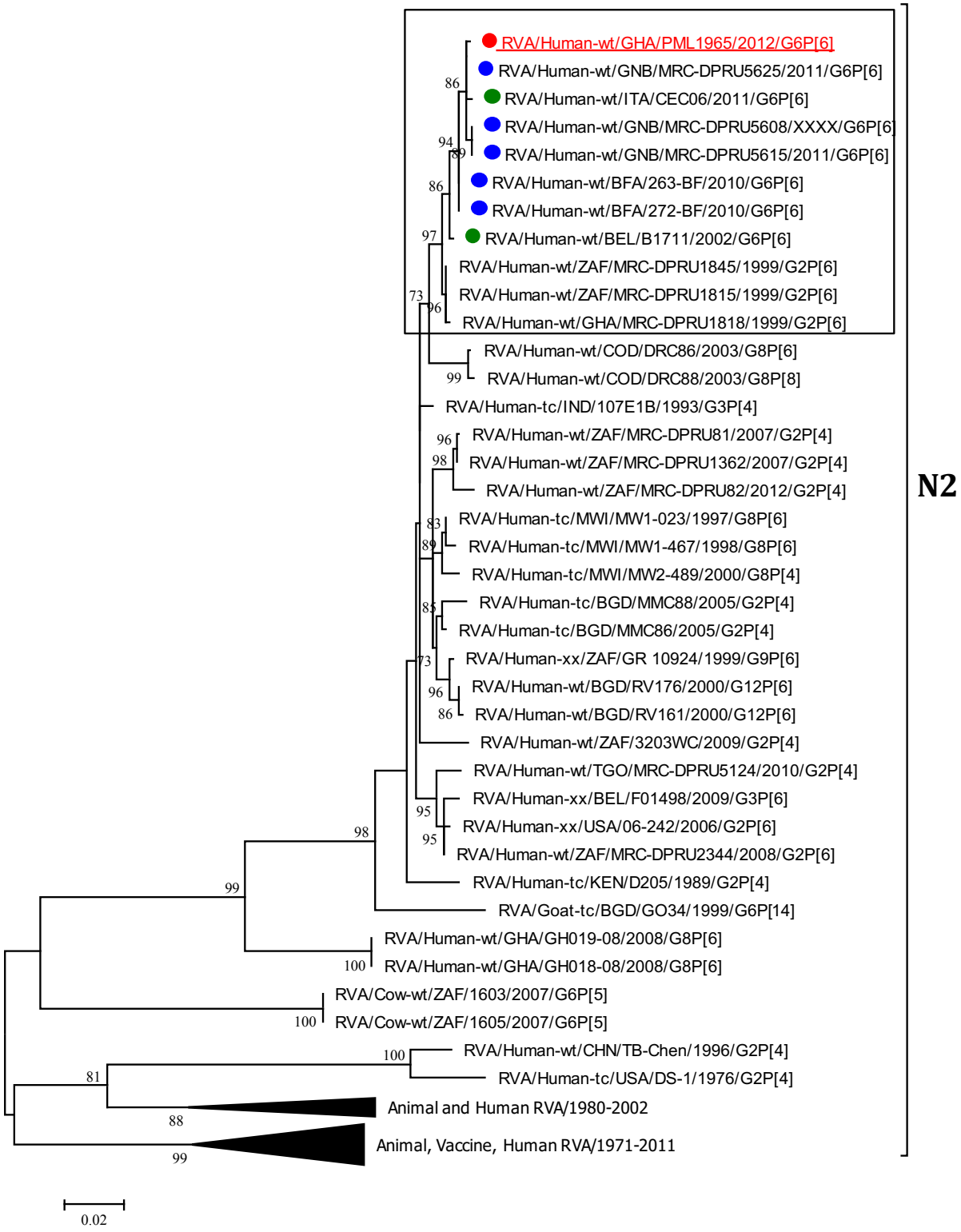


Fig. S1f: NSP2 tree

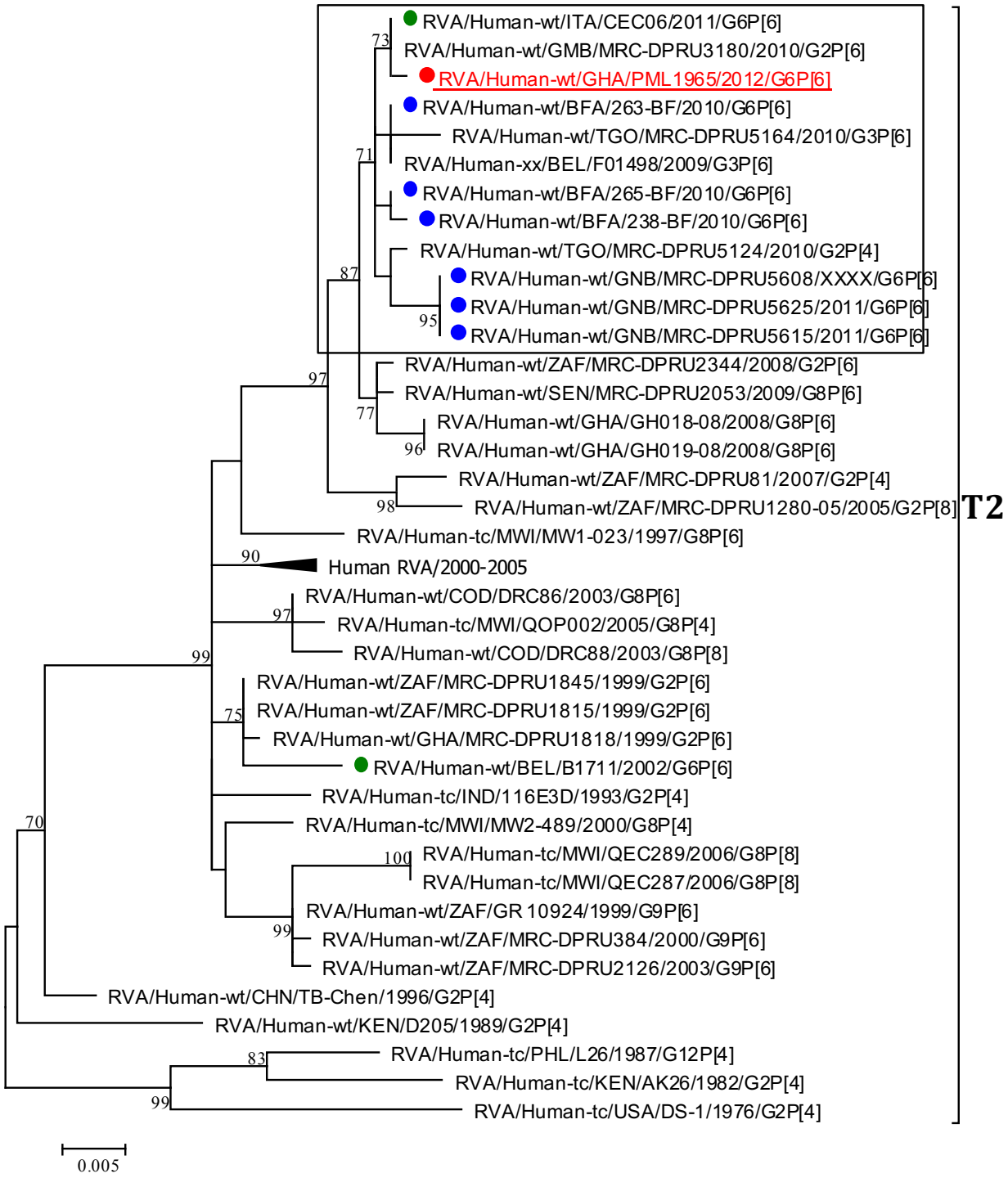


Fig. S1g : NSP3 tree

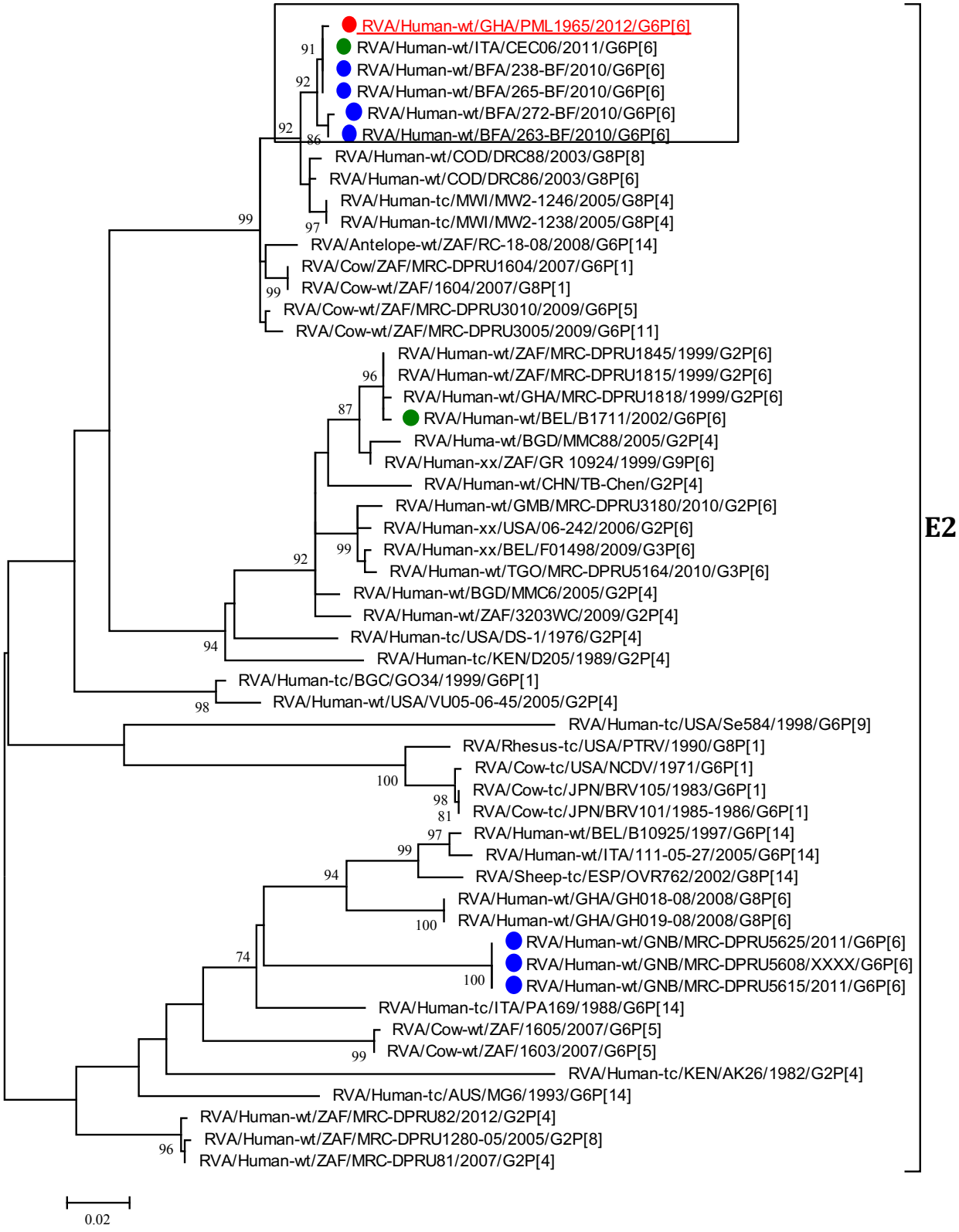


Fig. S1h: NSP4 tree

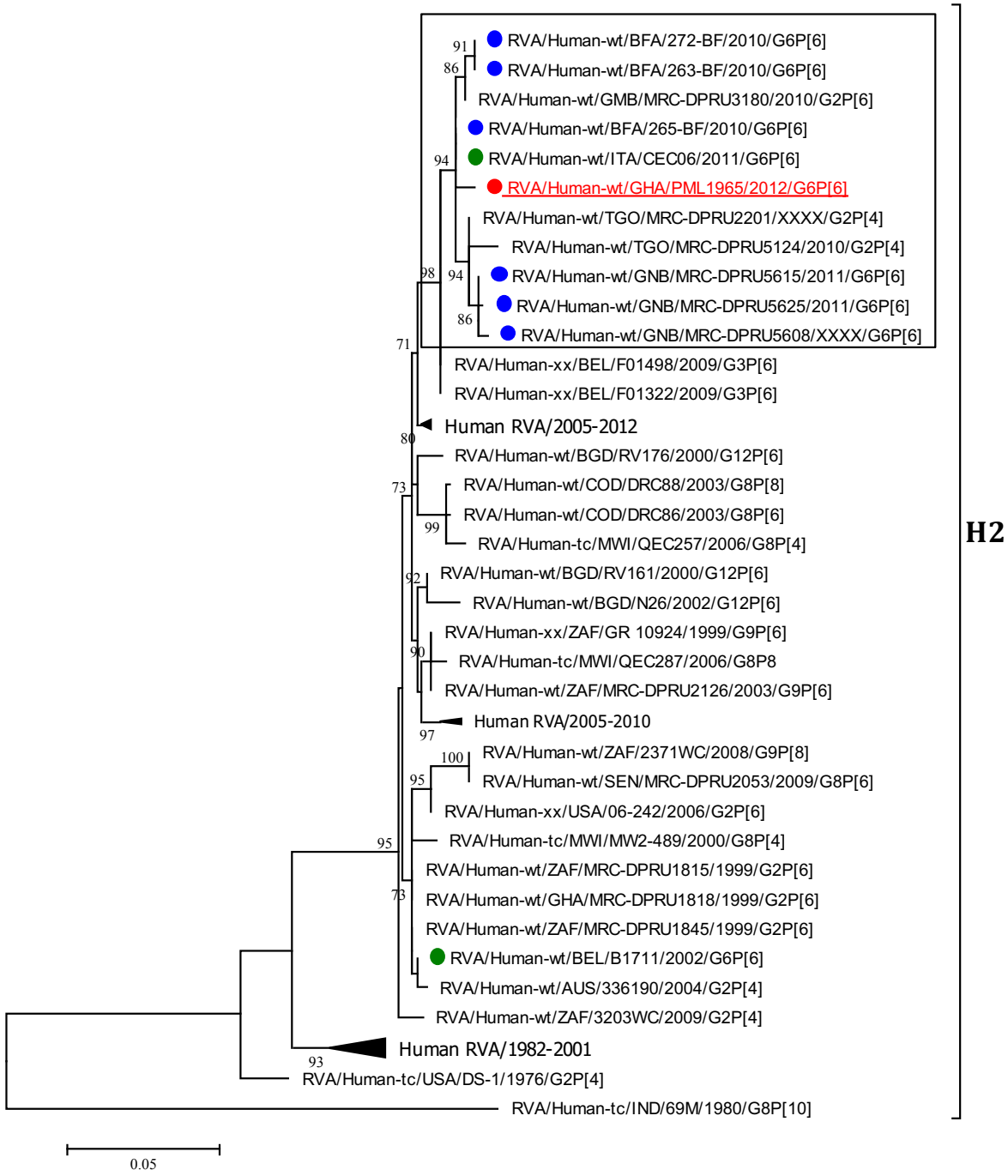


Fig. S1i: NSP5 tree

CHAPTER III

Supplementary Table 3.1: Primers used to amplify the full genome of the Ghanaian G2P[4] strains

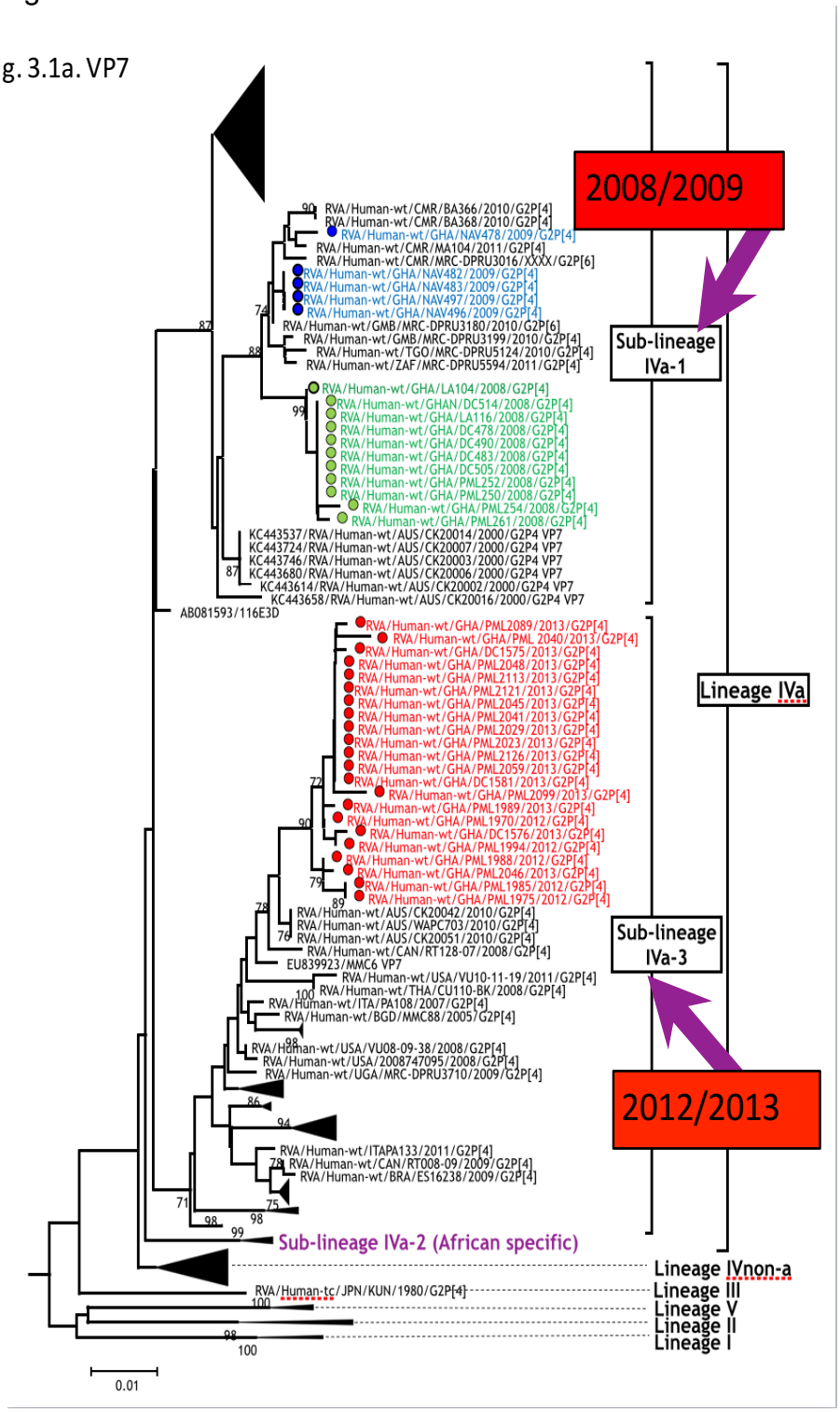
Gene	Fragment name	Name of Primer	Sequence (5'-3')	Reference
VP7	VP7-full	Beg9	GGCTTTAAAAGAGAGAATTTCCGTCTGG	Gouvea et al., 1990
		End9	GGTCACATCATACAATTCTAATCTAAG	Gouvea et al., 1990
VP4	VP4-1	VP4-1F	GGCTATAAAATGGCTTCGCTC	Doan et al., 2012
		Con2	TGGCTTCGCCATTTTATAGACA	Gentsch et al., 1992
	VP4-2	VP4-1523/542F	AGAGTATGGACGTTTCATGG	Doan et al., 2012
		VP4-1506R	ATCATTAAGCTGGCGTTCTA	Doan et al., 2012
	VP4-3	VP4-1180F	GGTGAATGGCCTATTATGAA	Doan et al., 2012
		VP4-2359R	GGTCACATCCTCGATGACATT	Doan et al., 2012
VP6	VP6-full	GEN_VP6F	GGCTTTWAAACGAAGTCTTC	Matthijnssens et al., 2008a
		GEN_VP6R	GGTCACATCCTCTCACT	Matthijnssens et al., 2008a
VP1	VP1-1	VP1-1F	GGCTATTAAAGCTATACAATGGG	Doan et al., 2012
		VP1-921R	AACCAATCCAGCTTTCCTC	Doan et al., 2012
	VP1-2	VP1-745F	CTGGTGTCATCTCCAATGTC	Doan et al., 2012
		VP1-1604R	TGCTGAGATGAATCCCATTG	Doan et al., 2012
	VP1-3	VP1-1375F	GACGTACCAGGAAGAAGAAC	Doan et al., 2012
		VP1-2567R	CTCTTCTCGAGCGATATTGG	Doan et al., 2012
	VP1-4	VP1-2343F	TGGAACAACTGATGATGAAGTG	Doan et al., 2012
		VP1-3302-R	GGTCACATCTAAGCGCTCTAATC	Doan et al., 2012
VP2	VP2-full	GEN_VP2Fc	GGCTATTAAAGGYTCAATGGCGTACAG	Matthijnssens et al., 2008a
		GEN_VP2_Rbc	GTCATATCTCCACARTGGGGTTGG	Matthijnssens et al., 2008a
VP3	VP3-1	MAX-3aF	GGCTATTAAAGCAGTACCAG	Matthijnssens et al., 2008a
		VP3-792R	CCAAGCCATCTCTCTTGTTTG	Doan et al., 2012
	VP3-2	VP3-598F	AGCAAGACTTTCAAATCGCG	Doan et al., 2012
		VP3-1592R	AATCCGTTGGCAAAAATGTC	Doan et al., 2012
	VP3-3	VP3-1102F	GGAATGGCGAAAAATGGTAG	Doan et al., 2012
		VP3-2054R	GCAGTTTCGCGAATTCTCTC	Doan et al., 2012
	VP3-4	VP3-1781F	TTTTCGGAGTCAGCCACTTC	Doan et al., 2012
		MAX-3bR	GGTCACATCATGACTAGTGTG	Matthijnssens et al., 2008a
	VP3-full	VP3F primer	TGCGTTTTACCTCTGATGGTG	Fujii et al., 2012
		VP3R primer	TCACATCATGACYAGTGTGTTAAG	Fujii et al., 2012
NSP1	NSP1	GEN_NSP1F	GGCTTTTTTTTATGAAAAGTCTTG	Matthijnssens et al., 2008a
		GEN_NSP1R	GGTCACATTTTATGCTGCC	Matthijnssens et al., 2008a
NSP2	NSP2	GEN_NSP2F	GGCTTTTAAAGCGTCTCAG	Matthijnssens et al., 2008a
		GEN_NSP2R	GGTCACATAAGCGCTTTC	Matthijnssens et al., 2008a
NSP3	NSP3	GEN_NSP3F	GGCTTTTAATGCTTTTCAGTG	Matthijnssens et al., 2008a
		GEN_NSP3R	ACATAACGCCCCTATAGC	Matthijnssens et al., 2008a
NSP4	NSP4	GEN_NSP4F	GGCTTTTAAAAGTTCTGTTCC	Matthijnssens et al., 2008a
		GEN_NSP4R	GGWYACRYTAAGACCRTTCC	Matthijnssens et al., 2008a
NSP5	NSP5	GEN_NSP5F	GGCTTTTAAAGCGCTACAG	Matthijnssens et al., 2008a
		GEN_NSP5R	GGTCACAAAACGGGAGT	Matthijnssens et al., 2008a

Supplementary Table 3.2 Lengths of genes sequenced and nucleotide sequence accession numbers

Gene	Length of genes without primers (bp)	GenBank nucleotide sequence accession numbers					
		GHDC514	GHLA104	GHNAV482	GHNAV483	GHPML1989	GHDC1581
VP7	1007	LC105533	LC105533	LC105533	LC105533	LC105533	LC105533
VP4	2317	LC105534	LC105534	LC105534	LC105534	LC105534	LC105534
VP6	1319	LC105535	LC105535	LC105535	LC105535	LC105535	LC105535
VP1	3256	LC105536	LC105536	LC105536	LC105536	LC105536	LC105536
VP2	2632	LC105537	LC105537	LC105537	LC105537	LC105537	LC105537
VP3	2550	LC105538	LC105538	LC105538	LC105538	LC105538	LC105538
NSP1	1524	LC105539	LC105539	LC105539	LC105539	LC105539	LC105539
NSP2	1022	LC105540	LC105540	LC105540	LC105540	LC105540	LC105540
NSP3	1023	LC105541	LC105541	LC105541	LC105541	LC105541	LC105541
NSP4	710	LC105542	LC105542	LC105542	LC105542	LC105542	LC105542
NSP5	772	LC105543	LC105543	LC105543	LC105543	LC105543	LC105543

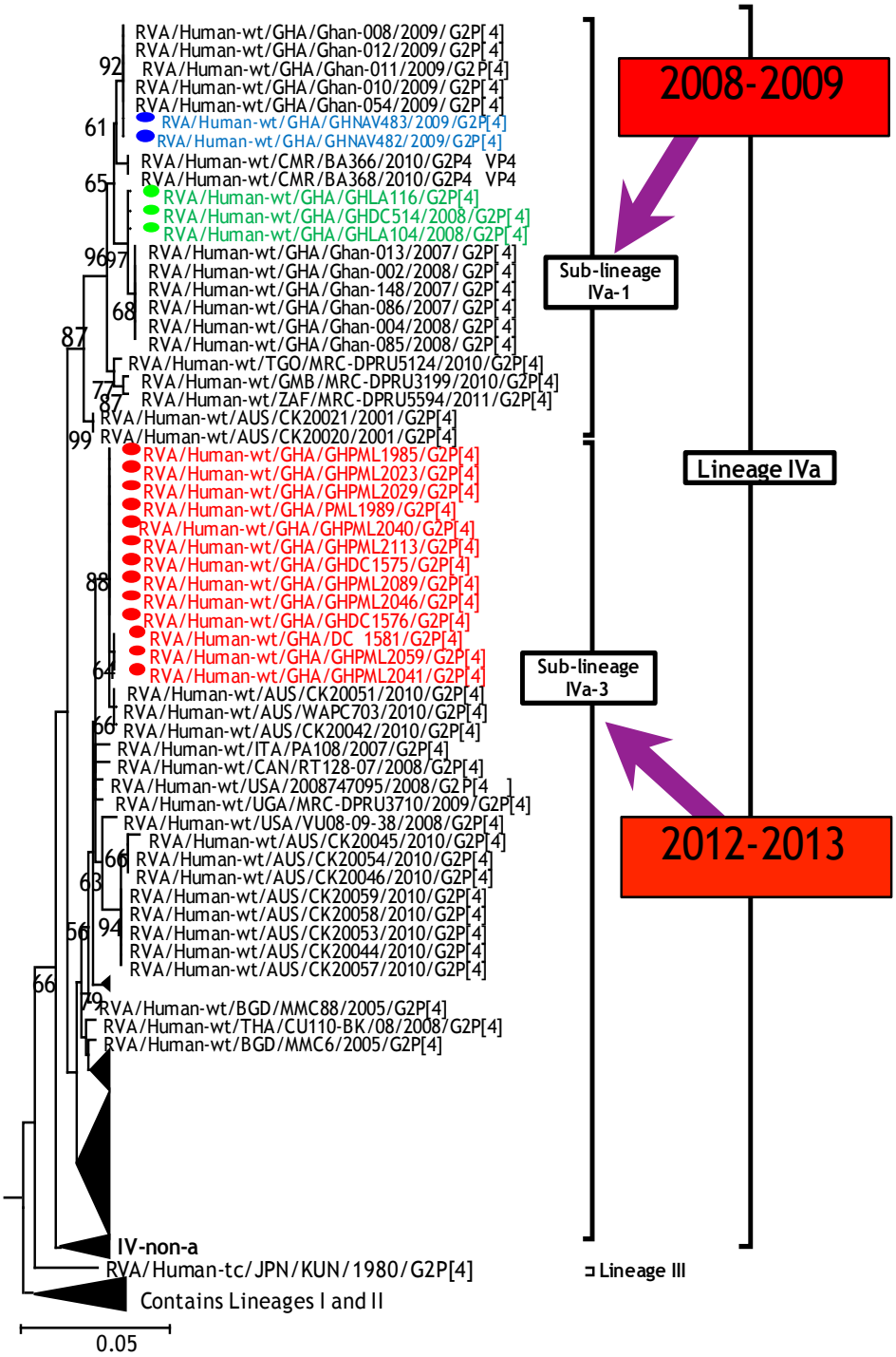
Supplementary Figure 3.1a-3.1b

Supplementary Fig. 3.1a. VP7

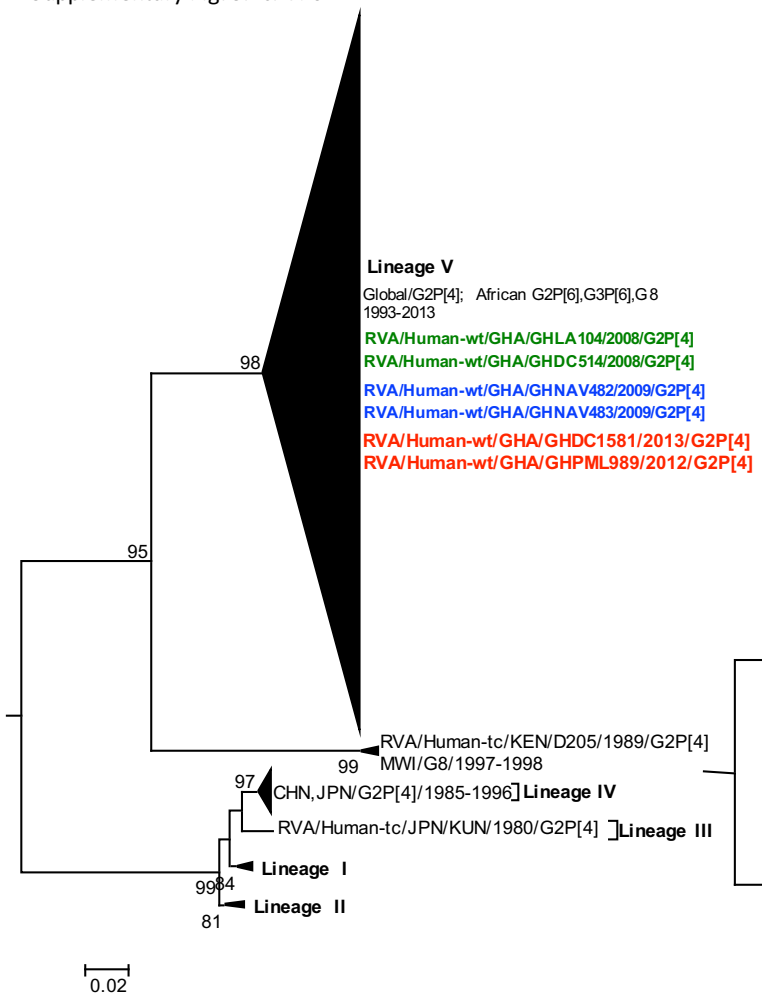


Supplementary Fig. 3.1a-i: Phylogenetic trees of the **(a)** VP7, **(b)** VP8*, **(c)** VP6, **(d)** VP1, **(e)** VP2, **(f)** NSP1, **(g)** NSP2, **(h)** NSP3 and **(i)** NSP5 genes of the study strains together with global G2P[4] and some DS-1-like RVA strains showing the evolutionary relationship between Ghanaian pre and post vaccine introduction period G2P[4] strains and other DS-1-like RVA strains. For each phylogenetic tree, analysis included at least 177 taxa consisting of the nucleotide sequences of the Ghanaian G2P[4] strains in this study (indicated in green, blue (pre vaccine period) and red font (post vaccine period), globally circulating G2P[4] strains with available whole genome sequence data and some African DS-1-like RVA strains from the GenBank database. Maximum likelihood phylogenetic analyses were performed using the best fit nucleotide substitution models selected based on the lowest Bayesian Information Criterion in MEGA6 software package. The trees were rooted using the respective genes of the Wa strain. Significant bootstrap values (1000 replicates) are indicated at each node. Lineage designation was based on the scheme by Doan et al. (2015). The scale bar at the bottom of the tree indicates genetic distance expressed as nucleotide substitutions per site.

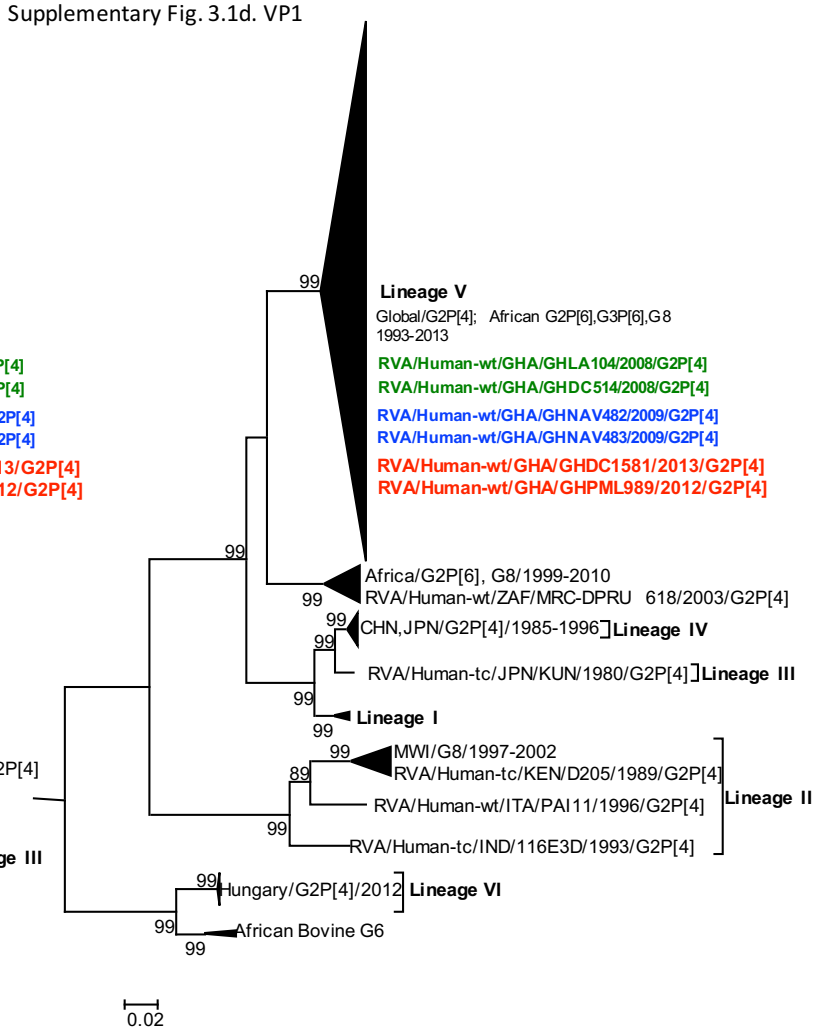
Supplementary Fig. 3.1b. VP4



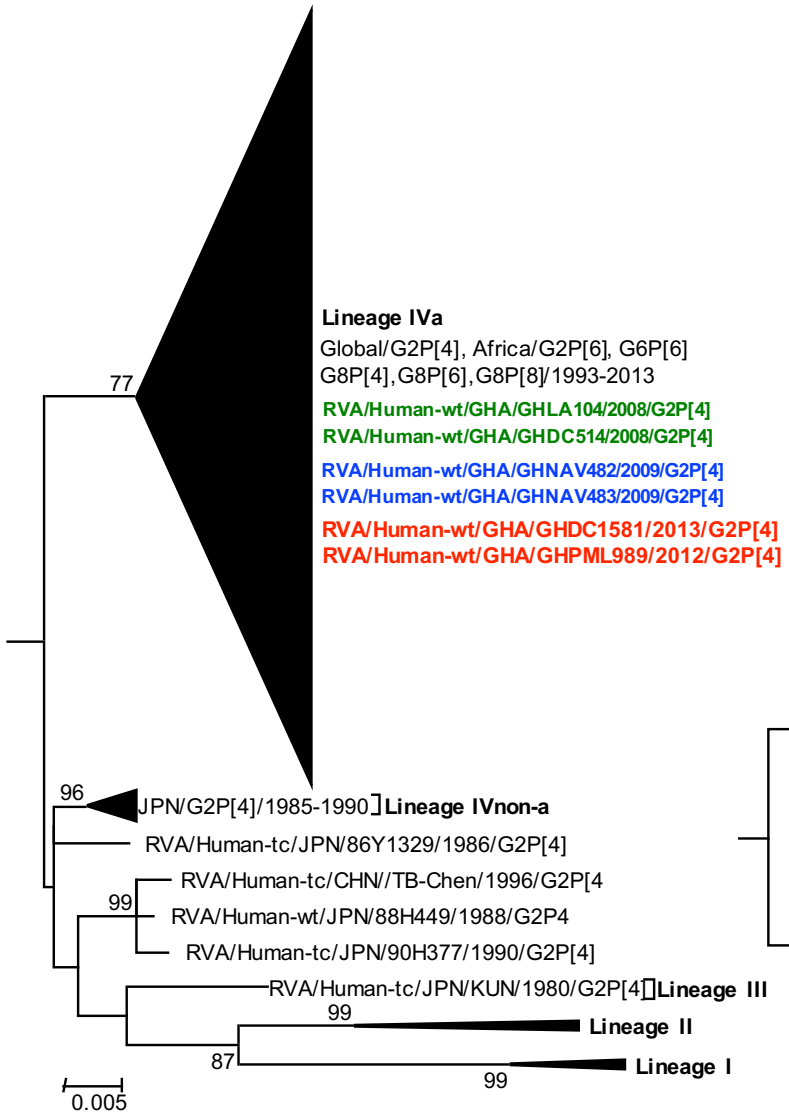
Supplementary Fig. 3.1c. VP6



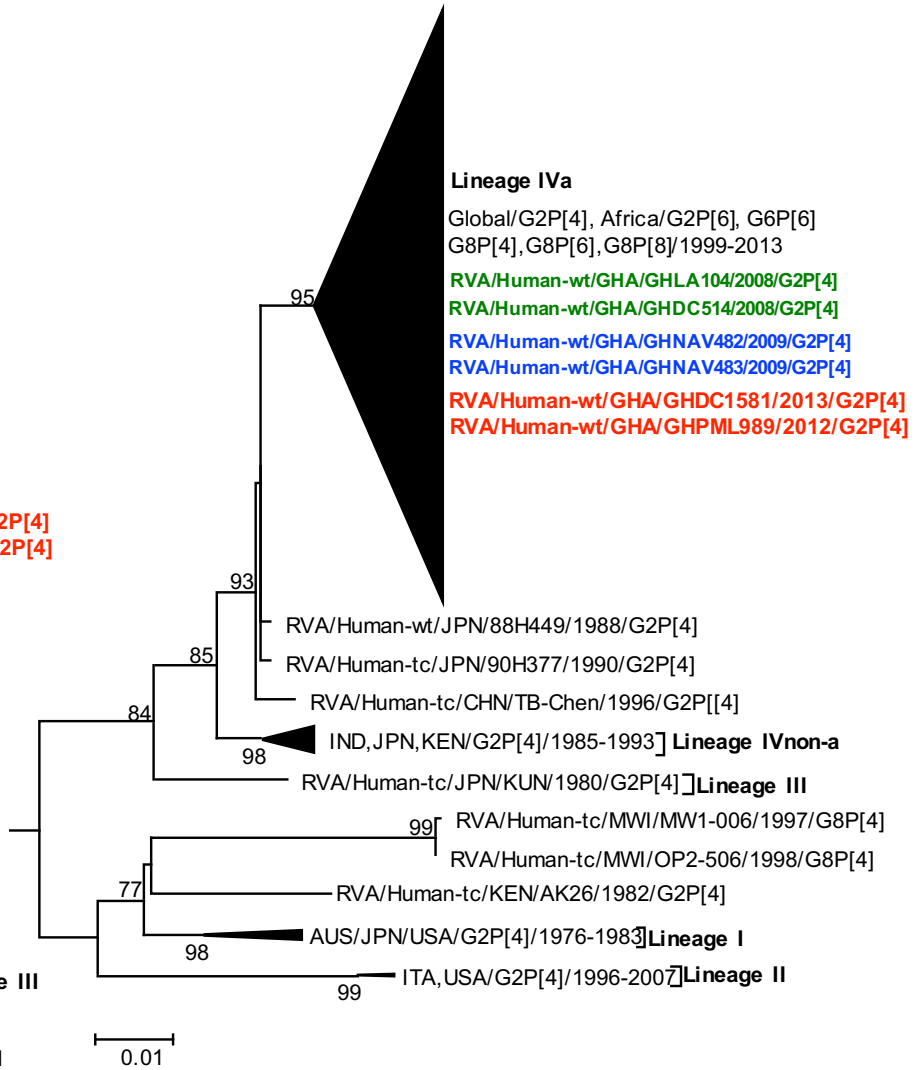
Supplementary Fig. 3.1d. VP1



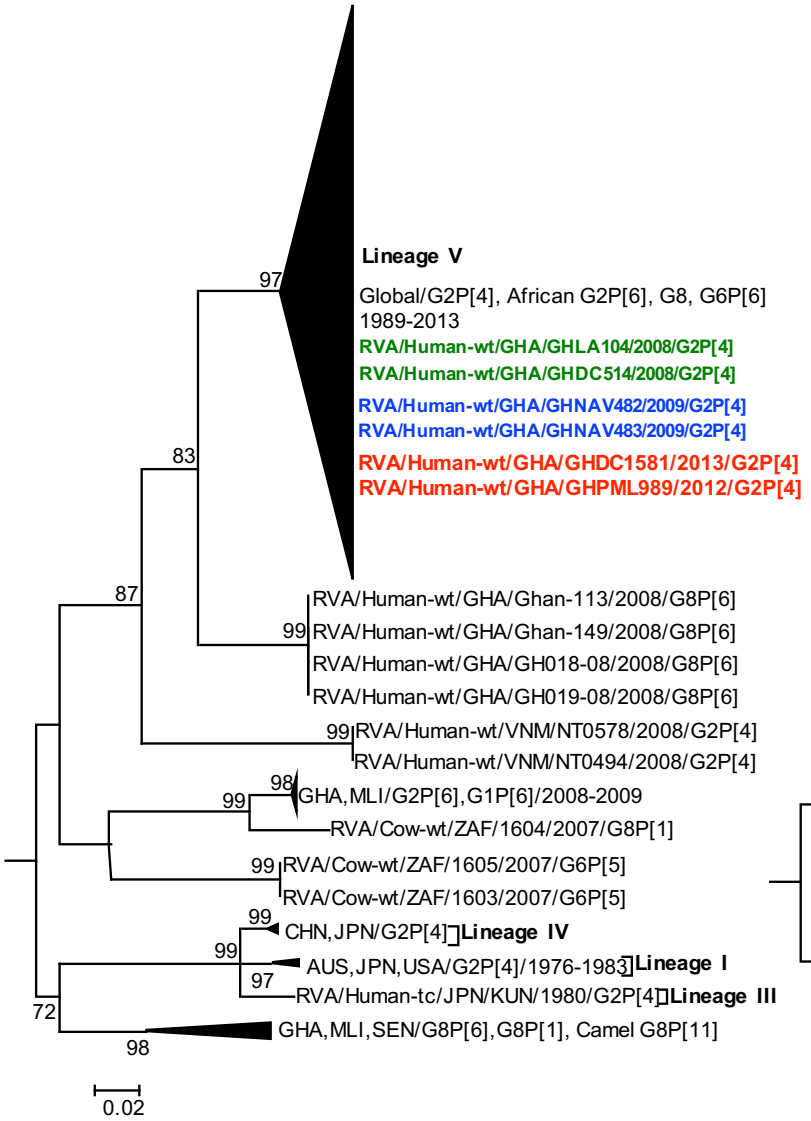
Supplementary Fig. 3.1e. VP2



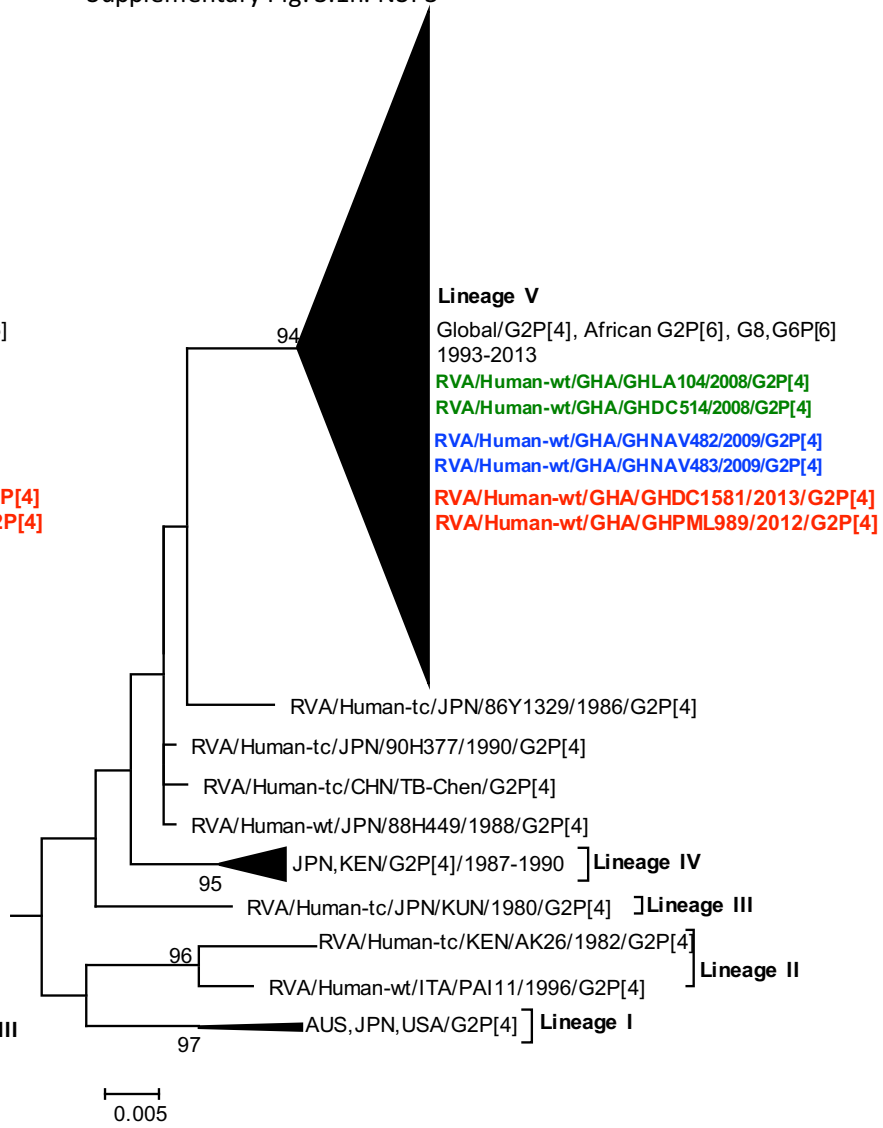
Supplementary Fig. 3.1f. NSP1



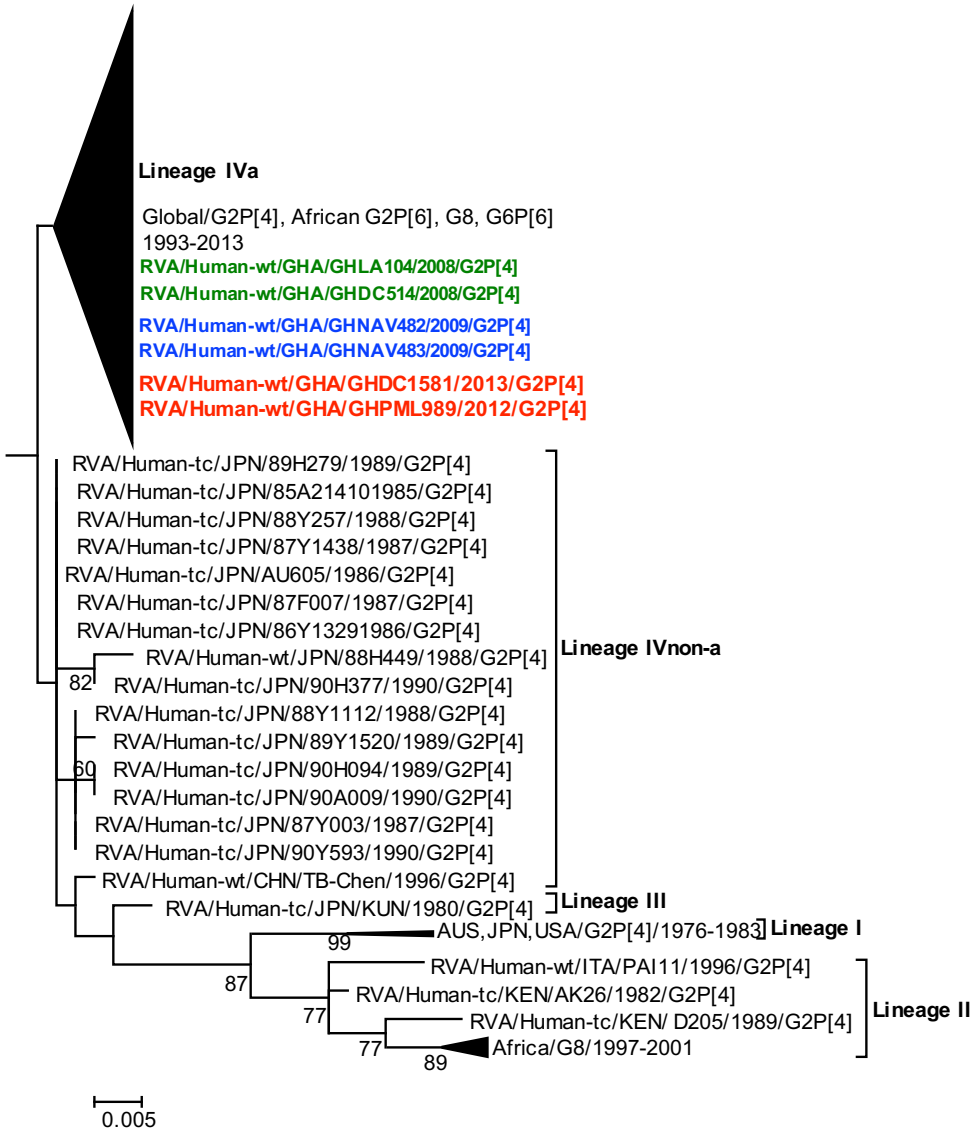
Supplementary Fig. 3.1g. NSP2



Supplementary Fig. 3.1h. NSP3



Supplementary Fig. 3.1i. NSP5



CHAPTER IV

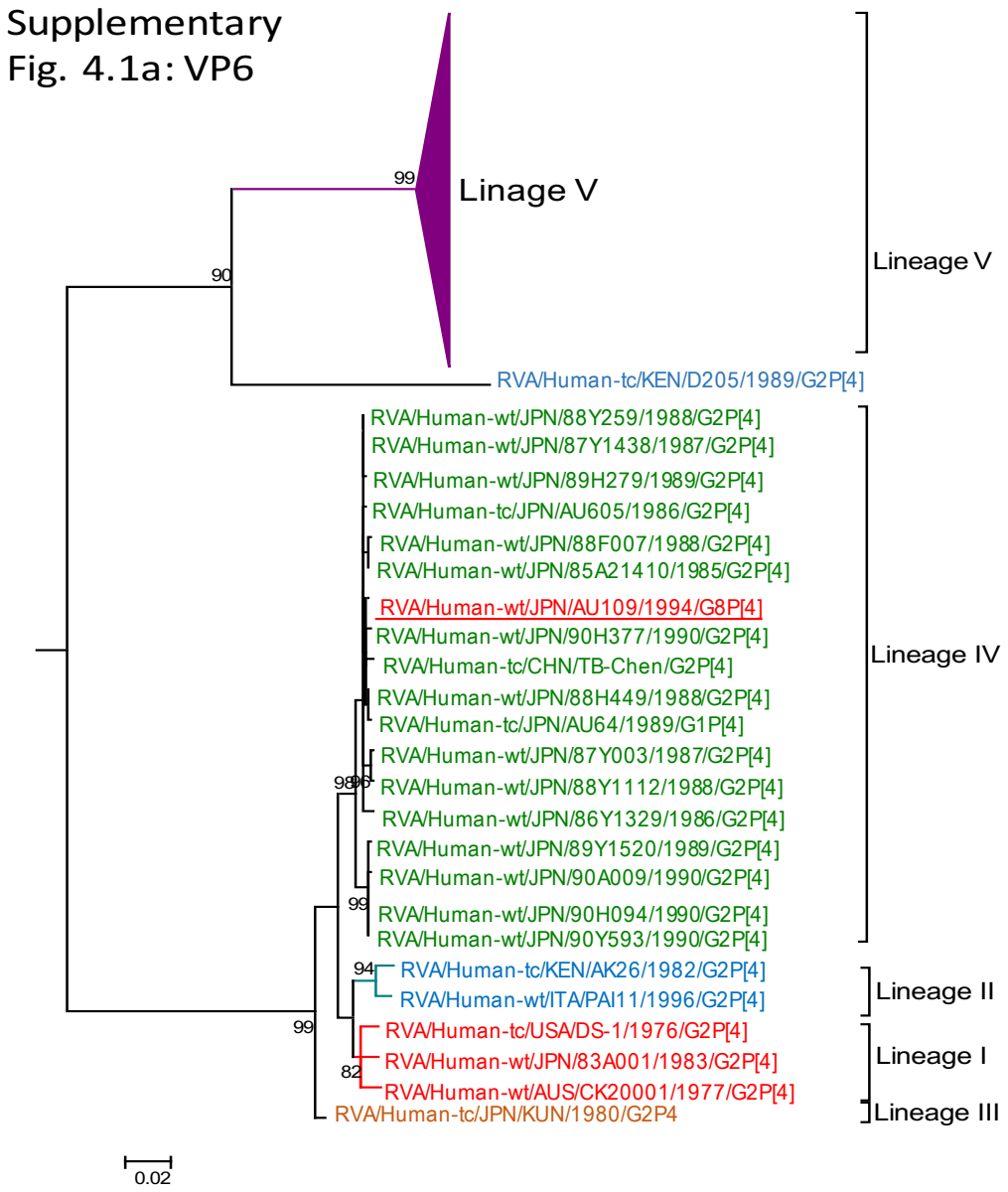
Supplementary Table 4.1: Primers used to amplify the full genome of AU109

Gene	Fragment name	Name of Primer	Sequence (5'-3')	Reference
VP7		Beg9	GGCTTTAAAAGAGAGAATTTCCGTCTGG	Gouvea *et al.,* 1990
		End9	GGTCACATCATACAATTCTAATCTAAG	Gouvea *et al.,* 1990
VP4	VP4-1	VP4-1F	GGCTATAAAATGGCTTCGCTC	Doan *et al.,* 2012
		Con2	TGGCTTCGCCATTTTATAGACA	Gentsch et al., 1992
	VP4-2	VP4-1523/542F	AGAGTATGGACGTTTCATGG	Doan *et al.,* 2012
		VP4-1506R	ATCATTAAGCTGGCGTTCTA	Doan *et al.,* 2012
	VP4-3	VP4-1180F	GGTGAATGGCCTATTATGAA	Doan *et al.,* 2012
		VP4-2359R	GGTCACATCCTCGATGACATT	Doan *et al.,* 2012
VP6	VP6	GEN_VP6F	GGCTTTWAAACGAAGTCTTC	Matthijnssens *et al.*, 2008a
		GEN_VP6R	GGTCACATCCTCTCACT	Matthijnssens *et al.*, 2008
VP1	VP1-1	VP1-1F	GGCTATTAAAGCTATACAATGGG	Doan *et al.,* 2012
		VP1-921R	AACCAATCCAGCTTTCCTC	Doan *et al.,* 2012
	VP1-2	VP1-745F	CTGGTGTCATCTCCAATGTC	Doan *et al.,* 2012
		VP1-1604R	TGCTGAGATGAATCCCATTG	Doan *et al.,* 2012
	VP1-3	VP1-1375F	GACGTACCAGGAAGAAGAAC	Doan *et al.,* 2012
		VP1-2567R	CTCTTCTCGAGCGATATTGG	Doan *et al.,* 2012
	VP1-4	VP1-2343F	TGGAACAACTGATGATGAAGTG	Doan *et al.,* 2012
		VP1-3302-R	GGTCACATCTAAGCGCTCTAATC	Doan *et al.,* 2012
VP2	VP2-full	GEN_VP2Fc	GGCTATTAAAGGYTCAATGGCGTACAG	Matthijnssens *et al.*, 2008a
		GEN_VP2_Rbc	GTCATATCTCCACARTGGGGTTGG	Matthijnssens *et al.*, 2008a
VP3	VP3-1	MAX-3aF	GGCTATTAAAGCAGTACCAG	Matthijnssens *et al.*, 2008a
		VP3-792R	CCAAGCCATCTCTCTTGTTTG	Doan *et al*., 2012
	VP3-2	VP3-598F	AGCAAGACTTTCAAATCGCG	Doan *et al*., 2012
		VP3-1592R	AATCCGTTGGCAAAAATGTC	Doan *et al*., 2012
	VP3-3	VP3-1102F	GGAATGGCGAAAAATGGTAG	Doan *et al*., 2012
		VP3-2054R	GCAGTTTCGCGAATTCTCTC	Doan *et al*., 2012
	VP3-4	VP3-1781F	TTTTCGGAGTCAGCCACTTC	Doan *et al*., 2012
		MAX-3bR	GGTCACATCATGACTAGTGTG	Matthijnssens *et al.*, 2008a
NSP1	NSP1	GEN_NSP1F	GGCTTTTTTTTATGAAAAGTCTTG	Matthijnssens *et al.*, 2008a
		GEN_NSP1R	GGTCACATTTTATGCTGCC	Matthijnssens *et al.*, 2008a
NSP2	NSP2	GEN_NSP2F	GGCTTTTAAAGCGTCTCAG	Matthijnssens *et al.*, 2008a

Supplementary Table 4.2: Regions sequenced for the 11 genes of AU109 and GenBank accession numbers

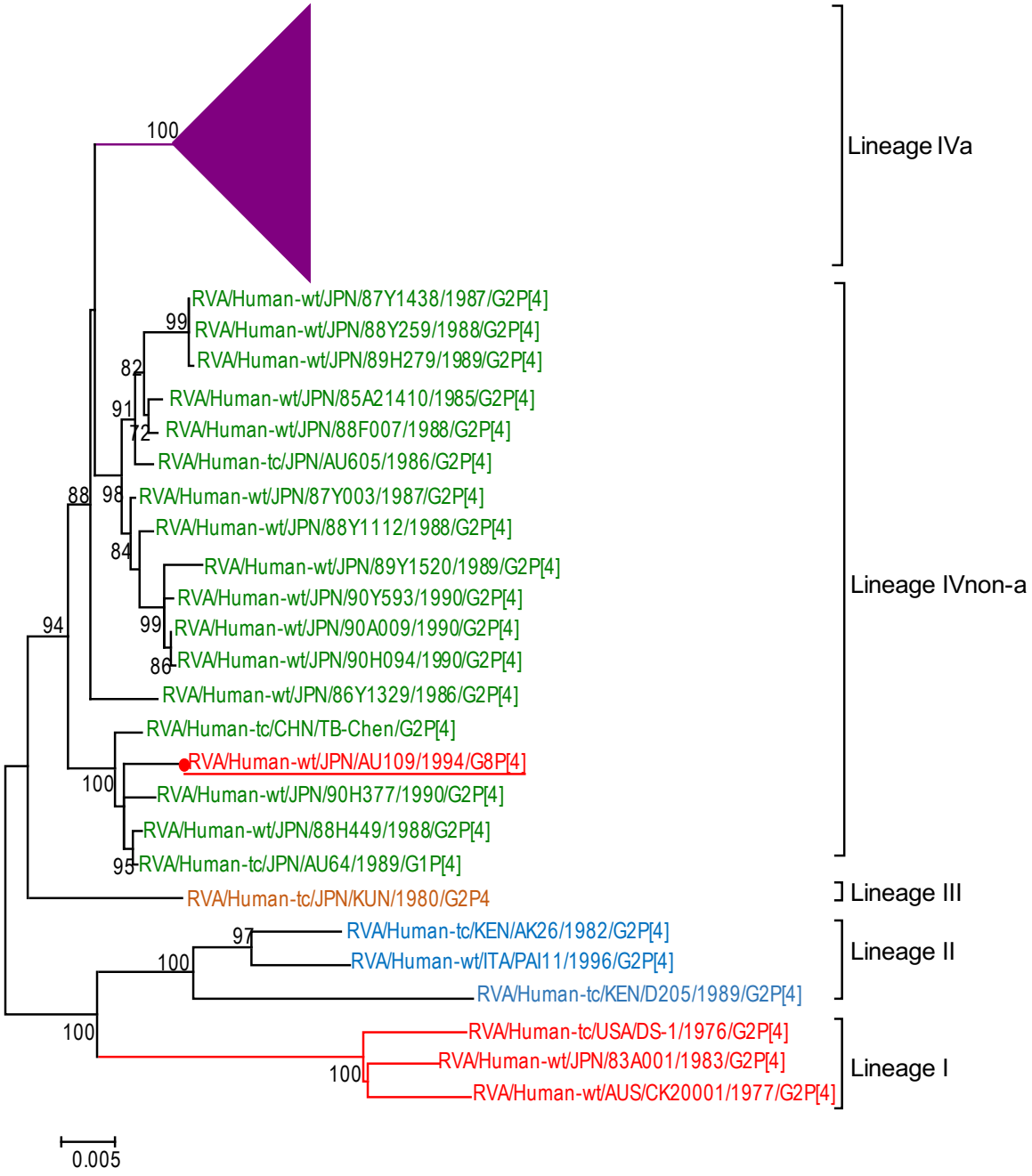
Genes	Nucleotide Accession number	Sequence minus primer sequence (bp)
VP7	AB271753	29-1035
VP4	LC065018	22-2335
VP6	LC065019	21-1339
VP1	LC065020	24-3279
VP2	LC065021	28-2659
VP3	LC065022	21-2570
NSP1	LC065023	25-1548
NSP2	LC065024	20-1041
NSP3	LC065025	22-1044
NSP4	LC065026	22-731
NSP5	LC065027	25-796

Supplementary Fig. 4.1a – g

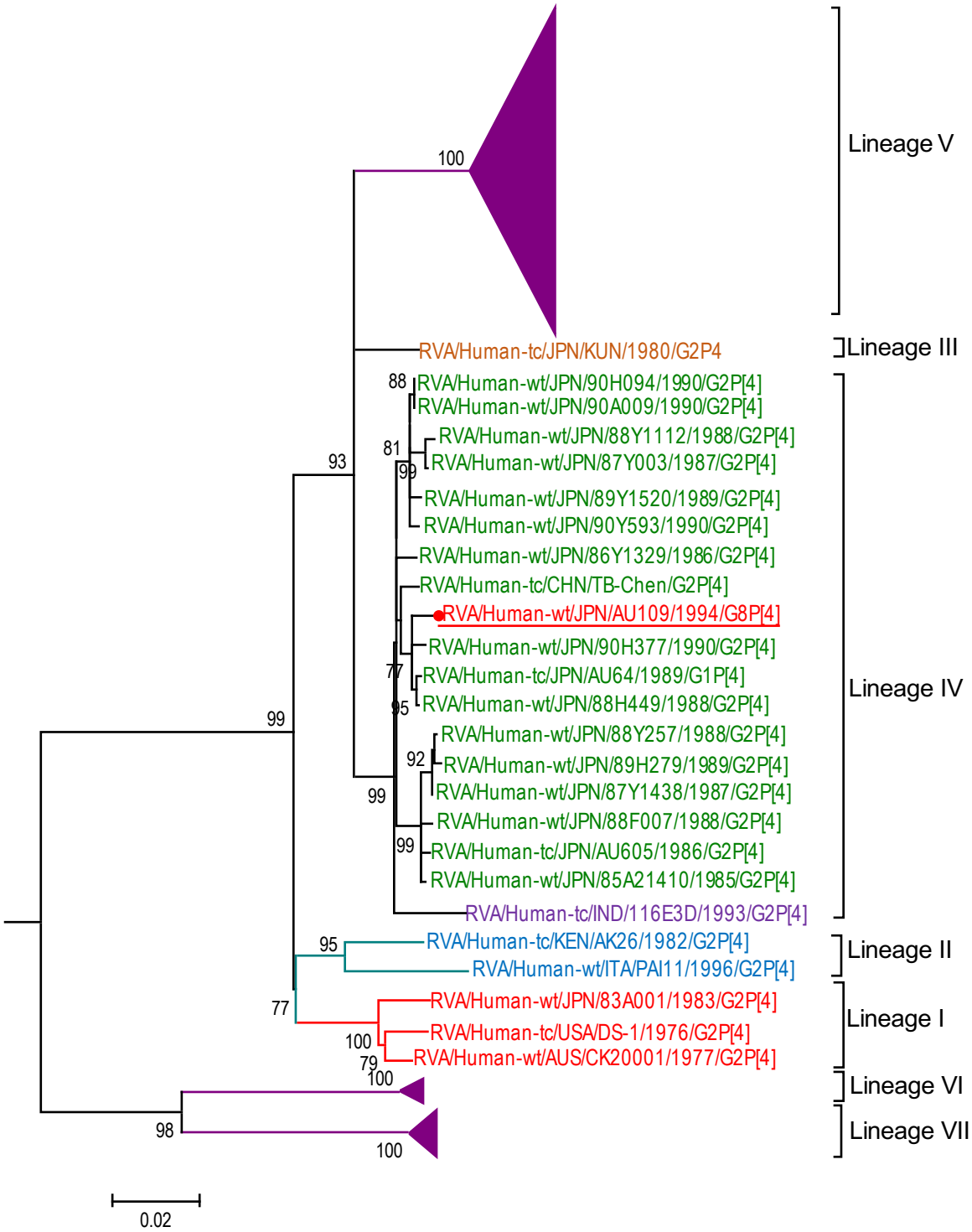


Supplementary Fig.4.1a-g: Phylogenetic trees of the (a) VP6, (b) VP2, (c) VP3, (d) NSP1, (e) NSP3, (f) NSP4, (g) NSP5 genes of DS-1-like RVA strains showing the genetic relationship between AU109 in this study and other human RVA strains. For each phylogenetic tree, analyses included the nucleotide sequence of the Japanese human G8P[4] strain AU109 in this study (indicated in red font) and globally circulating strains bearing the same genotype as AU109 retrieved from the GenBank database. Maximum likelihood phylogenetic analyses were performed using the following models: T92+G (VP6, VP2, NSP1, and NSP3-NSP5) and GTR+G+I (VP3) in MEGA6 software package. Significant bootstrap values (1000 replicates) of ≥70% are indicated at each node. The scale bar at the bottom of the tree indicates a genetic distance expressed as nucleotide substitutions per site.

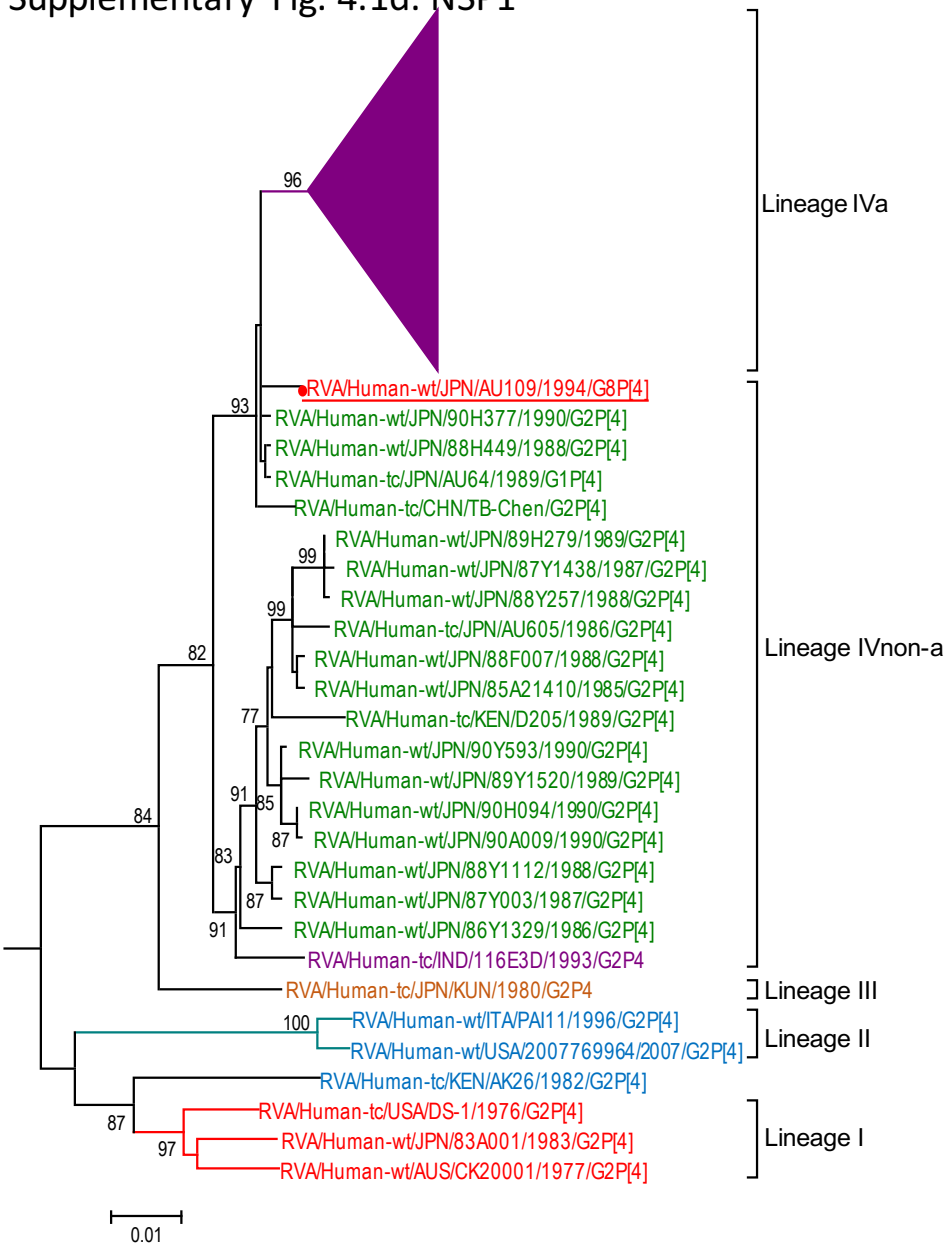
Supplementary Fig. 4.1b: VP2



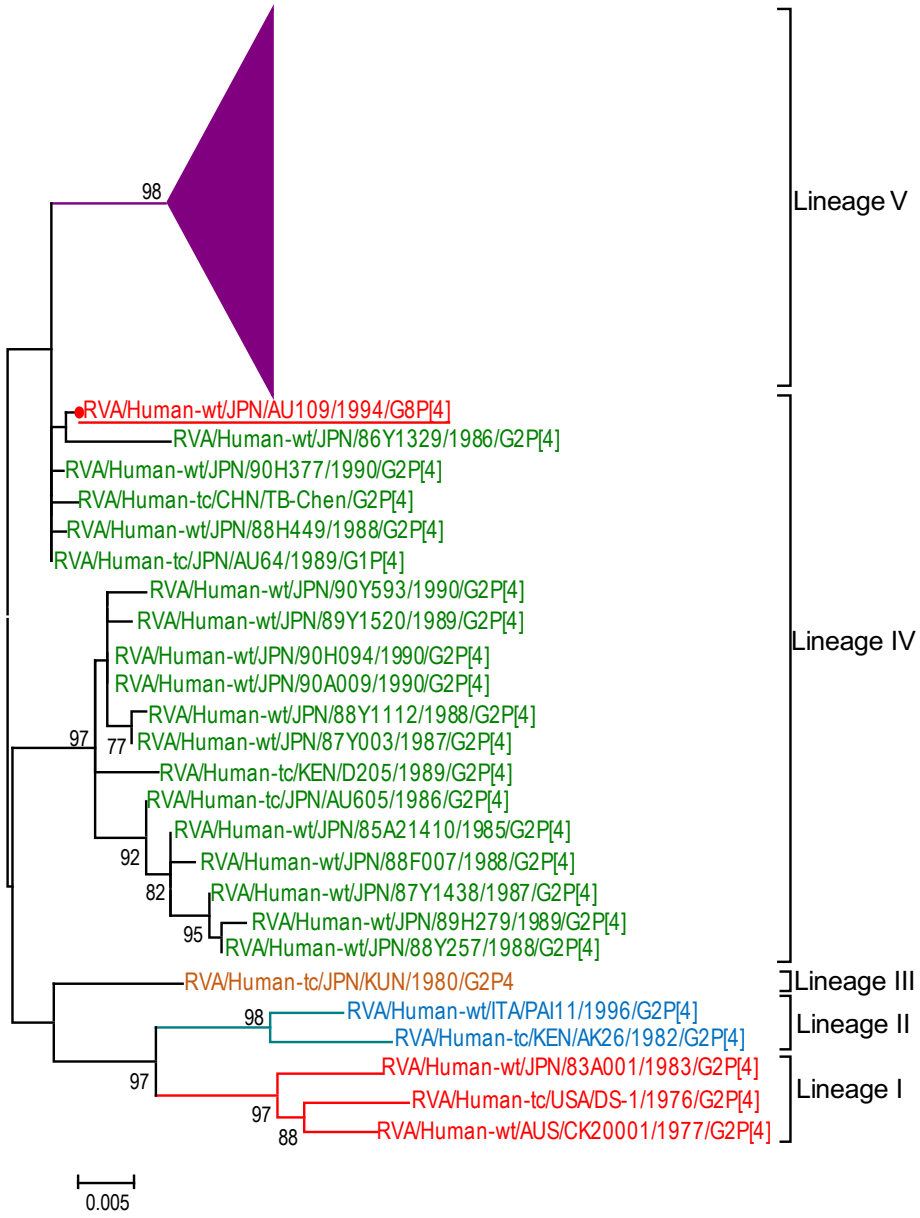
Supplementary Fig. 4.1c: VP3



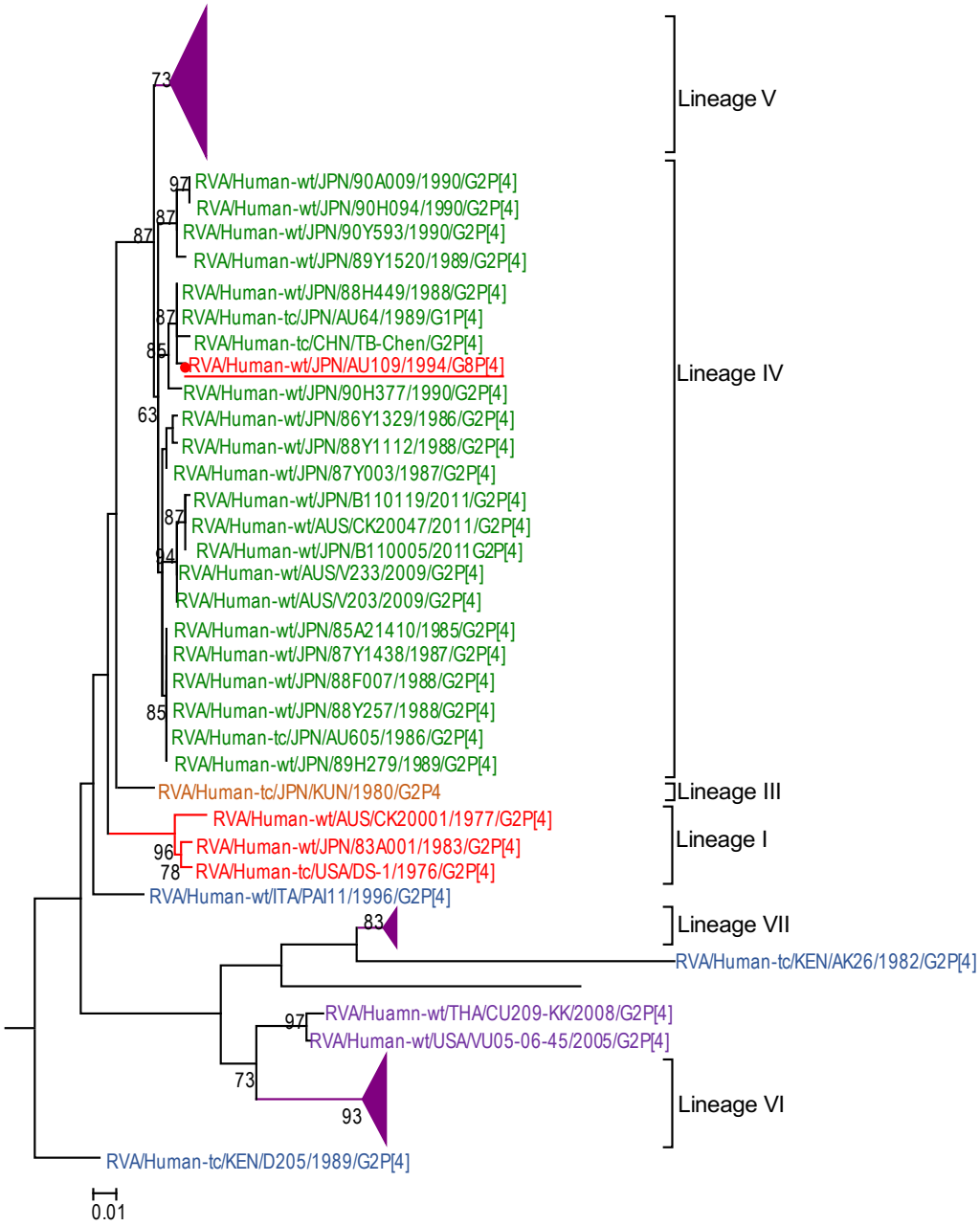
Supplementary Fig. 4.1d: NSP1



Supplementary Fig. 4.1e: NSP3



Supplementary Fig. 4.1f: NSP4



Supplementary Fig. 4.1g: NSP5