Accepted for publication in ***Archives of Virology***, 16 August, 2016

Whole genome characterisation of G12P[6] rotavirus strains possessing two distinct genotype constellations co-circulating in Blantyre, Malawi, 2008

Running head: G12 rotavirus strains in Malawi

T. Nakagomi[1,2*], **L. P. Do**[1**], **C. A. Agbemabiese**[1], **M. Kaneko**[1], **P. Gauchan**[1], **Y. H. Doan**[1***], **K. C. Jere**[2,3], **A. D. Steele**[4,5], **M. Iturriza-Gomara**[2], **O. Nakagomi**[1,2], **and N. A. Cunliffe**[2]

1. Department of Hygiene and Molecular Epidemiology, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan

2. Department of Clinical Infection, Microbiology and Immunology, Institute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom

3. Malawi-Liverpool Wellcome Trust Clinical Research Programme, College of Medicine, University of Malawi, Blantyre, Malawi

4. Vaccines and Immunization, PATH, Seattle, WA 98121, USA

5. MRC Diarrhoeal Pathogens Research Unit, MEDUNSA, University of Limpopo, Pretoria, South Africa

Word counts: Abstract (234 words), Text (3,700 words), Inserts: Tables (3), Figures (2)

* To whom all correspondence should be addressed:

Dr Toyoko Nakagomi

Department of Hygiene and Molecular Epidemiology, Graduate School of Biomedical Sciences, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan

Telephone/Fax number: +81-95-819-7062/+81-95-819-7064

Email address: tnakagom@nagasaki-u.ac.jp

** Present address: Department of Virology, National Institute of Hygiene and Epidemiology, Hanoi Vietnam

*** Present address: Department of Virology 2, National Institute of Infectious Diseases, Tokyo, Japan

Abstract

Rotavirus A strains detected from diarrhoeal children commonly possess any one of the genotypes G1, G2, G3, G4, and G9, with a recent increase in G12 detection globally. G12P[6] strains possessing short RNA (DS-1-like) and long RNA (Wa-like) migration patterns accounted for 27% of the strains circulating in Blantyre, Malawi between 2007 and 2008. To understand how the G12P[6] strains with two distinct genetic backgrounds emerged in Malawi, we conducted whole genome analysis of two long RNA and two short RNA strains. While the former had a typical Wa-like genotype constellation of G12-P[6]-I1-R1-C1-M1-A1-N1-T1-E1-H1, the latter was found to have G12-P[6]-I2-R2-C2-M1-A2-N2-T2-E2-H2: a VP3 gene mono-reassortant on the DS-1-like backbone. Phylogenetic and Bayesian Markov chain Monte Carlo analyses showed that the short RNA G12P[6] strains were generated around 2006 by reassortment between an African Wa-like G12P[6] strain donating three genes (the VP7, VP4, and VP3 genes) and a G2P[4] strain similar to the one circulating in Thailand or the United States of America that donated the remaining eight genes. On the other hand, the long RNA strains were generated as a result of reassortment events within Wa-like G12 and non-G12 strains commonly circulating in Africa; only the VP4 gene was from

a Malawian G8P[6] strain. In conclusion, this study uncovered the evolutionary pathways through which two distinct G12P[6] strains emerged in Malawi.

Keywords: rotavirus; genotype constellation; reassortment; G12; P[6]; phylogenetic analysis; Malawi

Introduction

Rotavirus A (RVA), a major cause of severe acute gastroenteritis in infants and young children worldwide (1), taxonomically belongs to the genus *Rotavirus* within the family *Reoviridae* (2). The genome of RVA consists of 11 segments of double-stranded RNA, coding for six structural viral proteins (VPs) and six non-structural proteins (NSPs); each genome segment codes for a single viral protein, with the exception of genome segment 11 that encodes two proteins (NSP5 and NSP6) (2). Genotypes are assigned according to the pre-defined nucleotide sequence identity for the open reading frame of each gene, and designated using a letter, which refers to a property or function of the protein encoded by a given gene, and a number (3). Thus, VP7 (glycoprotein) and VP4 (protease-sensitive protein) are referred to as G and P types, respectively. There have been 28 G and 39 P genotypes thus far reported for all mammalian and avian RVA strains in the literature (3, 4), but most human RVA strains are known to carry one of five G and P type combinations: G1P[8], G2P[4], G3P[8], G4P[8] and G9P[8] (5, 6). More recently, however, the G12 genotype in combination with P[6] or P[8] , and less frequently with P[4] or P[9], have been considered to constitute major human RVA genotype combinations as the detection of G12 strains increased (5, 7) and became predominant in some countries including Nepal (8-10), India (11), Spain (12),

Cameroon (13), Malawi (14, 15), USA (16), and Italy (17). Calculation of the evolutionary rate of the G12 VP7 gene estimated that the lineage containing more than 90% of G12 VP7 genes (lineage III) disseminated globally in little more than a decade with the most recent common ancestor emerging around 1995 (7).

A better understanding of the global emergence and dissemination of G12 strains requires information about the full genotype constellation of those strains and their relationships to each other as well as to other strains sharing the same genotype. In this regard, the recent expansion in genotype classification proposed and implemented by the Rotavirus Classification Working Group to include all 11 genes of RVA facilitated such endeavours (3, 18, 20). According to this classification system the genome of individual rotavirus strains is given the complete descriptor of Gx-P[x]-Ix-Rx-Cx-Mx-Ax-Nx-Tx-Ex-Hx ("x" representing the genotype number) denoting the VP7-VP4-VP6-VP1-VP2-VP3-NSP1-NSP2-NSP3-NSP4-NSP5 genes, respectively. Thus, the genotype constellations of prototype human strains Wa, DS-1 and AU-1 are described as G1-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 (18).

During a clinical trial of the attenuated human rotavirus vaccine (Rotarix®; GlaxoSmithKline, Biologicals, Belgium) in Blantyre, Malawi (between 2007 and 2008), the most prevalent genotype was G12P[6] which accounted for 27% of the rotavirus strains recovered from the placebo recipients (14). Upon electrophoresis, these G12P[6] strains showed either short RNA pattern (DS-1-like) or long RNA pattern (Wa-like), and RNA-RNA hybridisation revealed that the short and long RNA pattern viruses shared three genome segments apparently corresponding to the VP7, VP4 and VP2/VP3 genes (19). This observation led us to hypothesise that the DS-1-like G12P[6] strains were generated by reassortment in which Wa-like G12P[6] strains, which are common globally, acquired the DS-1-like internal capsid and non-structural protein genes from locally circulating DS-1-like strains. To examine this hypothesis, we determined the complete genotype constellations of four representative G12P[6] strains and compared them with those of global G12P[6] strains as well as G1P[8], G12P[8] and G2P[4] strains co-circulating during the clinical trial in Malawi.

Materials and Methods

The viruses

A phase III efficacy trial of an attenuated human monovalent rotavirus vaccine, Rotarix® (GlaxoSmithKline Biologicals, Rixensart, Belgium), was conducted in

Blantyre, Malawi between 2007 and 2008 and stool samples were collected from both vaccine and placebo recipients (14). The major G and P genotypes identified in the RVA-positive specimens from the placebo recipients during the clinical trial were G12P[6] (27%), G8P[4] (24%) and G9P[8] (24%), with only 13% of strains being G1P[8] (14). In order to investigate possible reasons for the lower efficacy of the rotavirus vaccine in Malawi (49.4%) (14), molecular characterisation was carried out on 21 strains selected from 88 RVA-positive specimens based on the amount of antigen (higher absorbance in ELISA), RNA migration patterns (absence of extra bands), and the quantity of the specimens (19). Among the 88 samples there were 25 G12P[6] strains; 11 showed identifiable RNA electropherotypes of which three were of long RNA pattern and eight were of short RNA pattern. Out of the 21 strains selected for molecular characterisation, there were four G12P[6] strains (MAL12, MAL39, MAL40, and MAL88). Among these, MAL12 and MAL40 represented long RNA strains whereas MAL39 and MAL88 represented short RNA strains. A single G2P[4] strain (MAL66) present among the 21 strains used in the previous study (19) was included in this whole genome analysis for comparison. Similarly, as representatives of strains with the Wa-like genetic background, one of three G1P[8] strains (MAL38) and the sole G12P[8] strain (MAL65) were selected for whole genome analysis.

RT-PCR and nucleotide sequencing

Genomic RNAs were extracted from the infected culture fluid by using the QIAamp Viral RNA Mini Kit (QIAGEN Sciences, Germantown, MD, USA) according to the manufacturer's instructions. An 8µl portion of genomic RNA was mixed with random primers (Invitrogen, Carlsbad, CA, USA) and dNTPs in a total volume of 9.5 µl and denatured at 97°C for 5 min. Reverse-transcription mixture (Invitrogen Super Script RNAse Reverse Transcriptase) was added to make the final reaction volume 20 µl. The thermal profile included incubation at 25°C for 5 min, 42°C for 60 min for reverse transcription and 70°C for 15 min to terminate the reaction.

From the reverse-transcription products (cDNA), the VP4, VP1, VP2, VP3, NSP1, NSP2, NSP3 and NSP5 genes of MAL12, MAL39, MAL40 and MAL88 as well as the VP7, VP6 and NSP4 genes of MAL39 and MAL40 were amplified by PCR with primers specific to both ends of the relevant genome segment (20, 21) by using the GoTaq® Green Master Mix system (Promega Corporation, Madison, WI, USA) at 95°C for 5 min, and 35 cycles of amplification (at 94°C for 45 sec; at 45°C for 45 sec and at 72°C for 2 to 6 min) followed by final extension at 72°C for 7 min. The amplified products were purified using an ExoSAP-IT purification kit (USB Products, Affymetrix, Cleveland, Ohio USA) according to the manufacturer's instructions. Nucleotide

sequencing reactions were performed by fluorescent dideoxy chain termination chemistry using the BigDye Terminator Cycle Sequencing Ready Reaction Kit, version 3.1 (Applied Biosystems, Foster City, CA, USA), and nucleotide sequences were determined using an ABI Prism 3730 Genetic Analyzer (Applied Biosystems).

Sequence analyses

Nucleotide sequences were aligned using the Megalign program in the Lasergene 11 software package (DNAstar, Inc. Madison, WI, USA). The genotype of each genome segment was classified by using the RotaC2.0 automated genotyping tool for RVA (22). Calculation of nucleotide sequence identity and phylogenetic analysis were performed by using MEGA ver. 6.06 (23). Multiple sequence alignment was carried out using the MUSCLE program, and the genetic distances between sequences were calculated by the p-distance method. Maximum likelihood phylogenetic trees were constructed using the best-fit nucleotide substitution model based on the Bayesian Information Criterion score, i.e. Tamura 3-parameter model with gamma rate distribution (T92+G), for the VP7, VP4 and VP3 genes using MEGA ver.6.06. The statistical significance at each node was calculated with 1000 pseudo-replicate trials.

The Bayesian Markov chain Monte Carlo (MCMC) framework implemented in BEAST ver.1.8.1 (24) was employed to determine the time of the most recent common

ancestor (tMRCA) and the divergence time. The nucleotide sequence dataset for each analysis comprised dated sequences of globally circulating strains similar to the VP7 (n=256), VP4 (n=53), VP6 (n=94) and VP3 (n=108) genes of the Malawian G12P[6] strains. Using the best-fit model - Tamura Nei with a gamma distributed rate variation among sites (TN93+G), an uncorrelated lognormal relaxed clock model (25) and a coalescent constant size tree prior (26), two independent analyses were carried out for 100 million, 20 million, 40 million and 25 million generations for the VP7, VP4, VP6 and VP3 genes, respectively. Log files were examined using the Tracer software ver.1.6 (http://tree.bio.ed.ac.uk/software/tracer/) to ensure that each analysis had reached convergence and effective sampling sizes were >200. Maximum clade credibility trees were annotated using the TreeAnnotator and viewed with FigTree ver.1.42 (http://tree.bio.ed.ac.uk/software/figtree/).

Nucleotide sequence accession numbers

The nucleotide sequences for the VP1, VP2, VP3, VP4, VP6, VP7, NSP1, NSP2, NSP3, NSP4, and NSP5 genes of MAL38, MAL39, MAL40, MAL65 and MAL66, the VP1, VP2, VP3, VP4, NSP1, NSP2, NSP3, and NSP5 genes of MAL12 and MAL88 were deposited in GenBank/DDBJ under accession numbers from AB938240 to AB938310. The VP7, VP6, and NSP4 genes of MAL12 and those of MAL88 were used

from the preceding study (19). The accession numbers of MAL12 are JN59143 (VP7), JN591408 (VP6) and JN591393 (NSP4), and those of MAL88 are JN591407 (VP7), JN591412 (VP6), and JN591397 (NSP4). As only partial VP4 genes of MAL12 and MAL88 were determined in the preceding study (19), the near-full lengths of these genes were newly determined and deposited under new accession numbers AB938245 and AB938248, respectively.

Results

Genotype constellation of Malawian G12P[6] strains

We determined the nucleotide sequences of the VP4, VP1, VP2, VP3, NSP1, NSP2, NSP3 and NSP5 genes of MAL12, MAL39, MAL40 and MAL88 as well as the VP7, VP6 and NSP4 genes of MAL39 and MAL40. When taken together with the sequence information obtained in the previous study (19), the full genotype constellation of MAL12 and MAL40 was G12-P[6]-I1-R1-C1-M1-A1-N1-T1-E1-H1 and that of MAL39 and MAL88 was G12-P[6]-I2-R2-C2-M1-A2-N2-T2-E2-H2 (Table 1). We also sequenced all genome segments of MAL38, MAL65 and MAL66, and determined their genotype constellation as G1-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1, G12-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1, and G2-P[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2, respectively (Table 1).

While the long RNA pattern strains, MAL12 and MAL40, had a Wa-like genotype constellation, the short RNA pattern strains, MAL39 and MAL88, had a DS-1-like genotype constellation except for their VP3 gene (M1); thus, MAL39 and MAL88 were a mono-reassortant in which the VP3 gene was reassorted from a strain possessing the Wa-like genotype constellation. The G12-P[6]-I1-R1-C1-M1-A1-N1-T1-E1-H1 genotype constellation of MAL12 and MAL40 is shared by many global G12P[6] strains such as the ones detected in Bangladesh (Dhaka12-03) (27), the United States of America (US6597) (28), South Korea (CAU195, CAU214) (29), Germany (GER172-08) (30), Cameroon (ES283) (31) and South Africa (3176WC) (32) (Table 1). An exception was Matlab 13-03, another Bangladeshi strain (27) that had T2 as its NSP3 genotype. By contrast, the G12-P[6]-I2-R2-C2-M1-A2-N2-T2-E2-H2 constellation of MAL39 and MAL88 was unique, and the only other G12P[6] strains possessing the DS-1-like backbone reported to date were three Bangladeshi G12P[6] strains each of which possess a unique genotype constellation (Table 1). However, none of them possess an M1VP3 gene.

Phylogenetic relationships at the lineage level and Bayesian analysis

Relationship within the G12P[6] strains possessing the same genotype constellation

The two strains possessing the Wa-like genotype constellation (MAL12 and

MAL40) shared high nucleotide sequence identities of 99.7 to 100% across the 11 genes. Similarly, the two strains possessing the DS-1-like genotype constellation (MAL39 and MAL88) shared high nucleotide sequence identities of 99.6 to 100% across the 11 genes. Thus, MAL12 and MAL40 were virtually indistinguishable from each other as were MAL39 and MAL88.

VP7 genes

When a phylogenetic tree was drawn for the G12 VP7 genes, the four G12P[6] strains (MAL39, MAL88, MAL12 and MAL40) and one G12P[8] strain (MAL65) belonged to a well-established lineage III as did many global G12 strains (Fig. 1A). The sequences of the VP7 gene of the Malawian G12 strains were very similar to each other (99.4 to 99.9%) as well as to that of some African G12P[6] strains such as those of Kenyan strains KDH633 and KDH684 (99.7 to 99.8% identical with MAL39/MAL88) (Table 2A). Thus, the G12 VP7 gene of the Malawian and Kenyan G12P[6] strains were considered to share an immediate and direct ancestor. Bayesian MCMC analysis estimated the tMRCA of the cluster containing the VP7 gene of all the five Malawian and the two Kenyan G12 strains together with nearby G12 strains from Asia and South Africa to be 2002 (Fig. 2A).

VP4 genes

When a phylogenetic tree was drawn for the VP4 genes of Malawian G12P[6] strains together with representative P[6] strains, the VP4 genes of MAL12 and MAL40 (Wa-like constellation) were under the same cluster together with those of the G8P[6] strains that predominantly circulated in Malawi between 2003 and 2005 (33) (shaded in grey near the bottom of Fig. 1B) with the OP2-668 strain possessing the highest nucleotide sequence identity (99.2%) to MAL12 (Table 3). On the other hand, the VP4 genes of MAL39 and MAL88 (a VP3 gene mono-reassortant with the DS-1-like backbone) clustered together with the VP4 genes of G1P[6], G9P[6] and G12P[6] strains detected in Africa (Kenya, Zambia, South Africa, etc.) (contained in the triangle highlighted in grey at the top of Fig. 1B), and the cluster was located far from the one containing MAL12 and MAL40 (shaded in grey near the bottom of Fig. 1B). The nucleotide sequence identity between MAL12 and MAL40, and that between MAL39 and MAL88 were high (99.7% and 99.8%, respectively), whereas the nucleotide sequence identities between MAL12/MAL40 and MAL39/MAL88 were much lower (97.9%) (Table 3). When the phylogenetic relationships and the nucleotide sequence identities were taken together, Malawian G12P[6] strains with two distinct genetic backgrounds did not appear to share an immediate ancestor regarding the P[6] VP4 gene. Bayesian MCMC analysis estimated the time of divergence of the cluster containing

MAL12 and MAL40 from the cluster containing MAL39 and MAL88 to be 1996 (indicated by an arrow in Fig. 2B).

VP3 genes

The VP3 genes of the four G12P[6] (MAL12/MAL40 and MAL39/MAL88), one G1P[8] (MAL38) and one G12P[8] (MAL65) strains had the same M1 genotype (Table 1). The nucleotide sequence identity of MAL39 and MAL88 were 100%, and the nucleotide sequence identities between any pair of MAL12, MAL40, MAL38 and MAL65 were 99.2 to 99.7% (Table 3). By contrast, the nucleotide sequence identity of any pair between the former (the MAL39/MAL88 pair) and latter strains (MAL12, MAL40, MAL38 and MAL65) were much lower (98.3 to 98.5%) (Table 3); the cluster containing MAL39 and MAL88 was distinct from the cluster containing MAL12, MAL40, MAL38 and MAL65 (indicated by the clusters shaded in grey near the top and bottom of Fig. 1C). Bayesian MCMC analysis estimated the time of divergence of the cluster containing MAL39 and MAL88 from the one containing MAL12, MAL40, MAL38, and MAL65 to be 1999 (indicated by an arrow in Fig.2C). When a BLAST search was carried out with the MAL39 VP3 sequence as the query sequence, many VP3 genes for African G12P[6] strains had VP3 genes that were 99.2 to 99.7% identical with MAL39/MAL88 (Table 2A), and they clustered together with MAL39/MAL88

(shaded in grey near the top of Fig.1C).

The backbone genes of the DS-1-like G12P[6] strains

When BLAST searches were repeated with the MAL39 sequences as the query sequence, a few G2P[4] strains such as CU209-KK/08 detected in Thailand as well as VU05-06-45 and VU08-09-11 detected in the United States of America were found to possess the cognate genes that were ≥99.2% identical to MAL39 except the NSP3 identity between MAL39 and CU209-KK (98.5%) (Table 2B). A locally co-circulating G2P[4] strain (MAL66), however, showed much lower sequence identities in any of the VP1, VP2, VP6, and non-structural protein genes that shared the same genotypes (92.9 to 98.6%) (Table 2B); hence MAL66 was unlikely to share the backbone genes with the DS-1-like G12P[6] strains. Bayesian MCMC analysis estimated that even the VP6 genes that showed the highest sequence identity (98.6%) among the 11 genes of MAL88 and MAL66 diverged in 1994 (data not shown).

The backbone genes of the Wa-like G12P[6] strains

To understand the origin of the internal capsid (VP1, VP2, VP3, and VP6) and non-structural protein (NSP1, NSP2, NSP3, NSP4, and NSP5) genes of the G12P[6] strains possessing the Wa-like genetic backbone, the BLAST search was performed using each gene of the MAL40 as the query sequence and identified the strains that had

the closest sequences. In all of these genes there were African strains that were ≥99.2% identical to MAL12 (Table 3). No single strain showed the highest sequence identity to MAL12 in all the genes; DPRU309 (G12P[8]) and DPRU1417 (G9P[8]) showed the highest identities in six genes shown in Table 3. When each of such sequences was included in the phylogenetic tree of the respective genome segment, they clustered together with MAL12 and MAL40 (data not shown). Taken together, these observations suggest that the Wa-like G12P[6] strains shared their backbone genes with Wa-like strains commonly circulating in Africa.

When the internal capsid (VP1, VP2, VP3, and VP6) and non-structural protein (NSP1, NSP2, NSP3, NSP4, and NSP5) genes of the MAL12/MAL40 pair were compared with the cognate genes of the contemporary Malawian strains possessing the same genotypes, i.e., MAL38 (G1P[8]) and MAL65 (G12P[8]), the MAL12/MAL40 pair and MAL65 shared high nucleotide sequence identities (99.1 to 99.7%) (Table 3) and clustered together in the phylogenetic trees except the NSP4 and NSP5 genes (data not shown). On the other hand, the MAL12/MAL40 pair and MAL38 shared high nucleotide sequence identities (99.2 to 99.7%) and clustered together only in the VP1, VP2, VP3, and NSP1 genes (data not shown).

Discussion

The confirmation and extension of the hypothesis postulated in the preceding study

This study determined the full genotype constellations of four Malawian G12P[6] strains, and identified two pairs each of which had a distinct genotype constellation; one pair possessing long RNA patterns (MAL12 and MAL40) had a complete Wa-like genotype constellation (I1-R1-C1-M1-A1-N1-T1-E1- H1), whereas the other pair possessing short RNA patterns (MAL39 and MAL88) were a single VP3 gene substitution reassortant on the background of the DS-1-like genotype constellation (I2-R2-C2-M1-A2-N2-T2-E2-H2). Thus, the three hybrid bands that were previously shown to appear around the regions of genome segments 2/3, 4, and 7/8/9 across the four G12P[6] strains (19) were determined to be the VP3, the VP4, and the VP7 gene, respectively. However, the hypothesis that the DS-1-like G12P[6] strains (MAL39 and MAL88) were generated by a locally-occurring reassortment event between the Wa-like G12P[6] strains (MAL12 and MAL40) and G2P[4] strains (such as MAL66) required revision. Whole genome sequencing of the four G12P[6] strains together with three locally co-circulating G2P[4], G1P[8] and G12P[8] strains followed by phylogenetic and Bayesian MCMC analyses provided an accurate and more precise understanding of how these G12P[6] strains with two different genetic backgrounds emerged in Malawi.

The evolutionary pathway through which the DS-1-like G12P[6] strains likely

emerged in Malawi.

While the two pairs of the G12P[6] strains shared the same genotype in the VP3, the VP4, and the VP7 gene, only the VP7 genes showed a high degree (99.4 to 99.6%) of nucleotide sequence identity and were considered to share a recent common ancestor. Nevertheless, Bayesian MCMC analysis of the VP7 genes estimated that the tMRCA of the DS-1-like G12P[6] pair (MAL39 and MAL88) and Kenyan G12P[6] strains was more recent (2006) than the tMRCA of the DS-1-like pairs and Wa-like pairs (2002). Furthermore, the fact that all of the VP7, VP4 and VP3 genes of the MAL39/MAL88 pair shared the nearly the same tMRCA (between 2005 and 2006) with African Wa-like G12P[6] strains but not with Malawian Wa-like G12P[6] strains suggests the occurrence of a single reassortment event in which an African Wa-like G12P[6] strain donated the three genes namely the VP7, VP4 and VP3 genes to a strain possessing the DS-1-like genetic backbone. The time point of the reassortment (between 2005 and 2006) was immediately before the year of detection of the DS-1-like G12P[6] strains in Malawi (2007). The observation that four Wa-like strains (MAL12, MAL40, MAL65, and MAL38) co-circulating with MAL39/MAL88 did not share high sequence identities in the VP3 and the VP4 genes lent support to the hypothesis that the ressortment event likely occurred outside Malawi.

Similarly, the source of the DS-1-like strain that provided the remaining eight genes to MAL39/MAL88 was unlikely in Malawi as a co-circulating G2P[4] strain (MAL66) did not show high sequence identities. Instead, the source likely existed among globally circulating G2P[4] strains since high nucleotide sequence identities (99.2 to 99.9%) of the eight genes (VP1, VP2, VP6, NSP1, NSP2, NSP3, NSP4, and NSP5) were observed between the Malawian DS-1-like G12P[6] pair (MAL39/MAL88) and G2P[4] strains detected in either Thailand (CU209-KK) or the United States of America (VU05-06-45). Thus, we hypothesised that the DS-1-like G12P[6] strains were generated around 2006 by reassortment between a Wa-like G12P[6] strain elsewhere in Africa and a globally-circulating G2P[4] strain.

The evolutionary pathway through which the Wa-like G12P[6] strains likely emerged in Malawi.

The Wa-like G12P[6] strains (MAL12 and MAL40) shared high sequence identities (>99%) in eight genes (VP1, VP2, VP3, VP6, VP7, NSP1, NSP2, and NSP3) with a co-circulating Malawian G12P[8] strain (MAL65), and the NSP4 and NSP5 genes with a South African G12P[8] strain (DPRU123). The remaining VP4 gene was 99.2% identical with a Malawian G8P[6] strain (OP2-688). Furthermore, many African strains possessing G1, G9 and G12 showed high sequence identities (>99%) in any one

of the genes carried by the MAL12/MAL40 pair. This suggests that MAL12 and MAL40 were one of those Wa-like G12P[6] strains commonly circulating in Africa. The observation that no single strain showed a high level of sequence identities with MAL12 and MAL40 across the 11 genes was taken as evidence for the frequent occurrence of genetic reassortment after the introduction of G12 strains in Africa.

Interpretation of diverse lineage constellations of African G12P[6] strains

Despite a presumed short history of circulation of G12 strains in Africa, there was a degree of genomic diversity at the level of lineage of the internal capsid and non-structural protein genes of the strains that carried the G12 VP7 gene. Even the two pairs of G12P[6] strains detected in a single year in the same place lacked high sequence identities enough to be considered to derive from a recent and direct ancestor except for the VP7 gene. This suggests frequent occurrence of reassortment events among co-circulating G12 strains in Africa. In Africa, the source of DS-1-like internal capsid and non-structural protein genes is not limited to G2P[4] strains but there is abundance of P[6] strains with the DS-1-like backbone. Thus, there is ample opportunities for G12 strains to reassort and acquire DS-1-like genetic backbone.

Culture-adaptation of RVA strains from clinical specimens

A caveat may be necessary as the RVA strains used in this study as well as in the

preceding study (19) were all adapted in cell culture. While culture-adaption allows virtually unlimited amount of materials for analysis (the amount of the original faecal suspensions was limited), selection of variants may occur from the quasispecies population present in the original host. There are a few studies relevant to this point. After 33 passages in cell culture, one non-synonymous substitution (0.2%) was reported to occur in the NSP4 gene of strain 89-12 (a precursor to the Rotarix vaccine strain) (34). In another study in which a porcine rotavirus strain with partially-duplicated NSP5 gene was passaged by serial plaque-to-plaque transfer, Blackhall, et al (35) observed that plaque-to-plaque transfer resulted in selection of variants of reduced fitness with occasional nucleotide substitutions. Such fitness changes resulting from repeated bottlenecking effects were not observed when viruses were passaged in cell culture by the standard moderate dilutions. They reported overall mutation rate (of the NSP5 gene) was a maximum value of 5×10^{-5} per nucleotide per replication (35). Under rather complicated alternating experimental conditions in vivo and in cell culture, a slightly more frequent (0.4-1.7%) mutations were noted in the murine RVA system (36). While the number of passages (up to 5 times) during culture adaptation in this study was far fewer than the above examples, caution in interpretation is warranted. Caution also needs to be exercised when the host is infected with more than one RVA strain as

reassortment may occur during the cell-culture adaptation (37). However, neither polyacrylamide gel electrophoresis nor RT-PCR genotyping suggested the presence of an extra strain in the faecal samples used in this study.

Conclusions

This study was intended to confirm the prediction by RNA-RNA hybridisation studies of what had appeared to be an uncomplicated reassortment event occurring between locally co-circulating, Wa-like and DS-1-like G12P[6] strains. However, this study uncovered the evolutionary pathways through which two distinct G12P[6] strains emerged in Malawi at the precision level that could be achieved by whole genome sequencing and Bayesian phylogenetics. Co-circulation of Wa-like and DS-1-like G12 strains let alone the predominance of DS-1-like G12 strains over the Wa-like has so far been observed rarely. Further investigation will be required to see whether the DS-1-like G12 strains will continue to spread efficiently among children in Malawi and elsewhere.

Compliance with Ethical Standards

Acknowledgements

The clinical trial of Rotarix® in Malawi, through which the strains characterised in this study were identified, was supported by PATH's Rotavirus Vaccine Program (funded by GAVI) and by GlaxoSmithKline Biologicals.

This study was performed under the Agreement on Academic Partnership between The University of Liverpool and Nagasaki University.

Funding

This study was supported by a grant-in-aid for overseas scientific research from Japan Society for the Promotion of Science awarded to Toyoko Nakagomi (No. 22406014). This study was also supported by the Global Centre of Excellence Program on Integrated Global Control Strategy for Tropical and Emerging Infectious Diseases from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Conflict of interest

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

Ethical approval and Informed consent

This study was performed in accordance with the ethical standards of the Review Boards of all relevant institutions. Informed consent was obtained from the parents/guardians of all individual participants included in the original study. The use of

specimens for this study was also approved by both PATH and GlaxoSmithKline Biologicals.

References

1. Tate JE, Burton AH, Boschi-Pinto C, Parashar UD (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

2. Estes MK, Greenberg HB (2013) Rotaviruses. in: Knipe DM, Howley PM, et al (eds.), Fields Virology, (sixth ed.). Wolters Kluwer/Lippincott Williams & Wilkins, Philadelphia, pp. 1347‑1401

3. Matthijnssens J, Ciarlet M, McDonald SM, Attoui H, Bányai K, Brister JR, Buesa J, Esona MD, Estes MK, Gentsch JR, Iturriza-Gómara M, Johne R, Kirkwood CD, Martella V, Mertens PP, Nakagomi O, Parreño V, Rahman M, Ruggeri FM, Saif LJ, Santos N, Steyer A, Taniguchi K, Patton JT, Desselberger U, van Ranst M (2011) Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). Arch Virol 156:1397-1413

4. Trojnar E, Sachsenröder J, Twardziok S, Reetz J, Otto PH, 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

5. Banyai K, Laszlo B, Duque J, Steele AD, Nelson EA, Gentsch JR, Parashar UD (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

6. Gentsch JR, Laird AR, Bielfelt B, Griffin DD, Banyai K, Ramachandran M, Jain V, Cunliffe NA, Nakagomi O, Kirkwood CD, Fischer TK, Parashar UD, Bresee JS, Jiang B, Glass RI (2005) Serotype diversity and reassortment between human and animal rotavirus strains: implications for rotavirus vaccine programs. J Infect Dis 192 (Suppl 1):S146-159

7. 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 Biol Evol 27:2431-2436

8. Uchida R, Pandey BD, Sherchand JB, Ahmed K, Yokoo M, Nakagomi T, Cuevas LE, Cunliffe NA, Hart CA, 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

9. Pun SB, Nakagomi T, Sherchand JB, Pandey BD, Cuevas LE, Cunliffe NA, Hart CA, Nakagomi O (2007) Detection of G12 human rotaviruses in Nepal. Emerg Infect Dis 13:482-484

10. Sherchand JB, Nakagomi O, Dove W, Nakagomi T, Yokoo M, Pandey BD, Cuevas LE, Hart CA, Cunliffe NA (2009) Molecular epidemiology of rotavirus diarrhea among children aged <5 years in Nepal: predominance of emergent G12 strains during 2 years. J Infect Dis 200 (Suppl 1):S182-187

11. Kang G, Desai R, Arora R, Chitamabar S, Naik TN, Krishnan T, Deshpande J, Gupte MD, Venkatasubramaniam S, Gentsch JR, Parashar UD; Indian Rotavirus Strain Surveillance Network, Mathew A, Anita Sr, Ramani S, Sowmynarayanan TV, Moses PD, Agarwal I, Simon A, Bose A, Arora R, Chhabra P, Fadnis P, Bhatt J, Shetty SJ, Saxena VK, Mathur M, Jadhav A, Roy S, Mukherjee A, Singh NB (2013) Diversity of circulating rotavirus strains in children hospitalized with diarrhea in India, 2005-2009. Vaccine 31:2879-2883

12. Cilla G, Montes M, Gomariz M, Alkorta M, Iturzaeta A, Perez-Yarza EG, Perez-Trallero E (2013) Rotavirus genotypes in children in the Basque Country (North of Spain): rapid and intense emergence of the G12[P8] genotype. Epidemiol Infect 141:868-874

13. Ndze VN, Papp H, Achidi EA, Gonsu KH, László B, Farkas S, Kisfali P, Melegh B, Esona MD, Bowen MD, Bányai K, Gentsch JR, Odama AM (2013) One year survey

of human rotavirus strains suggests the emergence of genotype G12 in Cameroon. J Med Virol 85:1485-1490

14. Madhi SA, Cunliffe NA, Steele D, Witte D, Kirsten M, Louw C, Ngwira B, Victor JC, Gillard PH, Cheuvart BB, Han HH, Neuzil KM (2010) Effect of human rotavirus vaccine on severe diarrhea in African infants. N Engl J Med 362:289-298
15. Turner A, Ngwira B, Witte D, Mwapasa M, Dove W, Cunliffe N (2013) Surveillance of rotavirus gastro-enteritis in children in Blantyre, Malawi. Paediatr Int Child Health 33:42-45
16. Payne DC, Szilagyi PG, Staat MA, Edwards KM, Gentsch JR, Weinberg GA, Hall CB, Curns AT, Clayton H, Griffin MR, Fairbrother G, Parashar UD (2009) Secular variation in United States rotavirus disease rates and serotypes: implications for assessing the rotavirus vaccination program. Pediatr Infect Dis J 28:948-953
17. Delogu R, Ianiro G, Camilloni B, Fiore L, Ruggeri FM (2015) Unexpected spreading of G12P[8] rotavirus strains among young children in a small area of central Italy. J Med Virol 87:1292-1302
18. Matthijnssens J, Ciarlet M, Rahman M, Attoui H, Bányai K, Estes MK, Gentsch JR, Iturriza-Gómara M, Kirkwood CD, Martella V, Mertens PP, Nakagomi O, Patton JT, Ruggeri FM, Saif LJ, Santos N, Steyer A, Taniguchi K, Desselberger U, van Ranst

M (2008) Recommendations for the classification of group A rotaviruses using all 11 genomic RNA segments. Arch Virol 153:1621-1629

19. Nakagomi T, Nakagomi O, Dove W, Doan YH, Witte D, Ngwira B, Todd S, Steele DA, Neuzil KM, Cunliffe NA (2012) Molecular characterization of rotavirus strains detected during a clinical trial of a human rotavirus vaccine in Blantyre, Malawi. Vaccine 30 (Suppl 1):A140-151
20. Matthijnssens J, Ciarlet M, Heiman E, Arijs I, Delbeke T, McDonald SM, Palombo EA, Iturriza-Gómara M, Maes P, Patton JT, Rahman M, van Ranst M (2008) Full genome-based classification of rotaviruses reveals a common origin between human Wa-Like and porcine rotavirus strains and human DS-1-like and bovine rotavirus strains. J Virol 82:3204-3219
21. 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
22. 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

23. Tamura K, Stecher G, Peterson D, Filipski A, and Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30: 2725-2729

24. Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29:1969-1973

25. Drummond AJ, Ho SY, Phillips MJ and Rambaut A (2006) Relaxed phylogenetics and dating with confidence. PLoS Biol 4, e88

26. Drummond AJ, Rambaut A, Shapiro B and Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol 22, 1185-1192

27. Rahman M, Sultana R, Ahmed G, Nahar S, Hassan ZM, Saiada F, Podder G, Faruque FAS, Siddique AK, Sack DA, Matthijnssens J, van Ranst M, Azim T (2007) Prevalence of G2P[4] and G12P[6] rotavirus, Bangladesh. Emerg Infect Dis 13:18-24

28. Freeman MM, Kerin T, Hull J, Teel E, Esona M, Parashar U, Glass RI, Gentsch JR (2009) Phylogenetic analysis of novel G12 rotaviruses in the United States: A molecular search for the origin of a new strain. J Med Virol 81:736-746

29. Than VT, Le VP, Lim I, Kim W (2011) Complete genomic characterization of cell culture adapted human G12P[6] rotaviruses isolated from South Korea. Virus Genes 42:317-322

30. Pietsch C, Liebert UG (2009) Human infection with G12 rotaviruses, Germany. Emerg Infect Dis 15:1512-1515

31. Ndze VN, Esona MD, Achidi EA, Gonsu KH, Dóró R, Marton S, Farkas S, Ngeng MB, Ngu AF, Obama-Abena MT, Bányai 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

32. Jere KC, Mlera L, O'Neill HG, Potgieter AC, Page NA, Seheri ML, van Dijk AA (2011) Whole genome analyses of African G2, G8, G9, and G12 rotavirus strains using sequence-independent amplification and 454® pyrosequencing. J Med Virol 83:2018-2042

33. Nakagomi T, Doan YH, Dove W, Ngwira B, Iturriza-Gómara M, Nakagomi O, Cunliffe NA (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

34. Ward RL, Mason BB, Bernstein DI, Sander DS, Smith VE, Zandle GA, Rappaport RS. (1997) Attenuation of a human rotavirus vaccine candidate did not correlate with mutations in the NSP4 protein gene. J Virol 71:6267-6270.

35. Blackhall J, Fuentes A, Magnusson G. 1996. Genetic stability of a porcine rotavirus RNA segment during repeated plaque isolation. Virology 225:181-190

36. Tsugawa T, Tatsumi M, Tsutsumi H. 2014. Virulence-associated genome mutations of murine rotavirus identified by alternating serial passages in mice and cell cultures. J Virol 88:5543-5558

37. Ward RL, Nakagomi O, Knowlton DR, McNeal MM, Nakagomi T, Huda N, Clemens JD,Sack DA. (1991) Formation and selection of intergenogroup reassortants during cell culture adaptation of rotaviruses from dually infected subjects. J Virol. 65:2699-2701

Figure Legends

Fig. 1 - Phylogenetic trees for the G12 VP7 (A), P[6] VP4 (B) and M1 VP3 (C) genes of Malawian G12P[6] strains sequenced in this study and in the preceding study (Nakagomi et al., 2012) and genes bearing the respective genotypes of global RVA strains for which the full genome sequence information is available in the GenBank database. The trees were constructed using the Maximum Likelihood method using MEGA 6.06 software package. Significant bootstrap values >70 obtained after 1000 bootstrap replicates are shown at each node. The G12P[6] strains sequenced in this study which possess the DS-1-like genetic backbone are indicated with filled red circles whereas those with the Wa-like genetic backbone are indicated with filled green circles. For simplicity, large clusters are collapsed into triangles. The scale bar at the bottom of each tree indicates the genetic distance expressed as nucleotide substitutions per site

Fig. 2 - Simplified maximum clade credibility trees of dated G12 VP7, P[6] VP4 and M1 VP3 genes of Malawian G12P[6] strains sequenced in this study and in the preceding study (Nakagomi et al., 2012) and genes bearing the respective genotypes

of global RVA strains for which the full genome sequence information is available in the GenBank database. Trees were constructed using the Bayesian Markov chain Monte Carlo method implemented in BEAST v.1.8.1 software package. The years of divergence and the time of the most recent common ancestor together with the 95% Highest Posterior Density intervals of clusters or lineages of interest are shown at the nodes. The G12P[6] strains sequenced in this study which possess the DS-1-like genetic backbone are indicated with filled red circles whereas those with the Wa-like genetic backbone are indicated with filled green circles. For simplicity, large clusters are collapsed into triangles. The in-lay Table shows the parameters used and obtained in the MCC analysis. The scale bar at the bottom of each MCC tree indicates time scale in years

Table 1. Whole genotype constellations of G12P[6] strains for which the full genome sequence information is available and three Malawian co-circulating strains for which full genome sequences were determined in this study.

Strain name	Year of detection	Country	VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5
RV161-00	2000	Bangladesh	G12	P[6]	I2	R2	C2	M2	A2	N2	T2	E1	H2
RV176-00	2000	Bangladesh	G12	P[6]	I2	R2	C2	M2	A2	N2	T2	E6	H2
N26-02	2002	Bangladesh	G12	P[6]	I2	R2	C2	M2	A2	N1	T2	E6	H2
Dhaka12-03	2003	Bangladesh	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
Matlab13-03	2003	Bangladesh	G12	P[6]	I1	R1	C1	M1	A1	N1	T2	E1	H1
MRC-DPRU2130-05	2005	South Africa	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
US6597	2005-2006	USA	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
CAU195	2006	South Korea	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
CAU214	2006	South Korea	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
MAL39	2007	Malawi	G12	P[6]	I2	R2	C2	M1	A2	N2	T2	E2	H2
MAL88	2007	Malawi	G12	P[6]	I2	R2	C2	M1	A2	N2	T2	E2	H2
MAL12	2007	Malawi	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
MAL40	2007	Malawi	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
GER172-08	2008	Germany	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
CU331	2008	Thailand	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
KisB521	2008	Congo	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
3176WC	2009	South Africa	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
KDH633	2010	Kenya	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
KDH684	2010	Kenya	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
ES283	2010	Cameroon	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
MRC-DPRU4090	2011	South Africa	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
MRC-DPRU4616	2011	Uganda	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
A23	2011	Myanmar	G12	P[6]	I1	R1	C1	M1	A1	N1	T1	E1	H1
MAL65	2008	Malawi	G12	P[8]	I1	R1	C1	M1	A1	N1	T1	E1	H1
MAL38	2007	Malawi	G1	P[8]	I1	R1	C1	M1	A1	N1	T1	E1	H1
MAL66	2007	Malawi	G2	P[4]	I2	R2	C2	M2	A2	N2	T2	E2	H2

The names of strains are in bold for those whose genotype constellations were determined in this study.

Table 2A G12P[6] strains possessing high nucleotide sequence identities (%) to the VP7, VP4, VP3 genes of the MAL39/MAL88 pair

	VP7	VP4	VP3
UGA/MRC-DPRU3713/2010/G12P[6]	99.8	99.5	99.7
ZMB/MRC-DPRU3491/2009/ G12P[6]	99.8	99.7	99.5
ZMB/MRC-DPRU1680/2008/ G12P[6]	99.7	99.7	99.5
KEN/KDH633/2010/ G12P[6]	99.7	99.5	99.3
KEN/KDH684/2010/ G12P[6]	99.8	99.4	99.3
TGO/MRC-DPRU1689/2008/ G12P[6]	99.6	99.7	99.6
ZMB/MRC-DPRU3488/2009/ G12P[6]	99.5	99.7	99.5
ZAF/MRC-DPRU1255/2005/G9 G12P[6]	99.2	99.5	99.5
ZAF/MRC-DPRU1911/2007/ G12P[6]	97.5	99.6	99.5
ZMB/MRC-DPRU3507/2009/ G12P[6]	99.8	99.7	99.5
ZAF/MRC-DPRU2130-05/2005/ G12P[6]	99.7	99.5	99.4
UGA/MRC-DPRU4616/2011/ G12P[6]	99.0	99.2	99.3
COD/KisB521/2008/ G12P[6]	99.5	99.6	99.2

Table 3 Strains possessing high nucleotide sequence identities (%) to the cognate genes of the MAL12/MAL40 pair

Strain name	VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5
MAL65 G12P[8]	99.6	P[8]	99.2	99.3	99.7	99.4	99.1	99.0	99.2	98.4	98.2
MAL38 G1P[8]	G1	P[8]	98.6	99.2	99.7	99.4	98.7	98.7	97.4	98.1	97.8
MAL39 G12P[6]	99.5	97.9	I2	R2	C2	98.5	A2	N2	T2	E2	H2
MAL88 G12P[6]	99.4	97.9	I2	R2	C2	98.5	A2	N2	T2	E2	H2
RVA/Human-wt/ZAF/MRC-DPRU2130-05/2005/G12P[6]	99.4	98.2	97.2	97.4	98.0	98.5	95.8	97.4	95.2	93.4	98.6
RVA/Human-tc/MWI/OP2-668/2003/G8P[6]	G8	99.2	I2	R2	C2	M2	A2	N2	T2	E2	H2
RVA/Human-wt/ZAF/MRC-DPRU309/2010/G12P[8]	99.2	P[8]	99.2	99.0	99.4	99.5	99.5	99.0	99.5	99.0	97.6
RVA/Human-wt/CMR/MRC-DPRU1417/2009/G9P[8]	G9	P[8]	96.5	99.4	99.6	99.1	99.0	99.4	99.1	98.3	98.3
RVA/Human-xx/BGD/Matlab13/2003/G12P[6]	98.7	97.9	99.1	98.4	95.8	93.9	98.9	97.4	T2	99.3	98.4
RVA/Human-wt/ZAF/MRC-DPRU123/2009/G12P[8]	99.1	P[8]	96.3	99.2	98.7	99.0	99.0	97.9	97.0	99.6	99.5
RVA/Human-wt/ZAF/MRC-DPRU1277/2004/G1P[8]	G1	P[8]	96.9	94.1	95.5	90.2	96.1	89.6	95.3	93.1	99.5

Only genotypes are shown where the strains to be compared have the genotypes different from those of MAL12/MAL40 (G12-P[6]-I1-R1-C1-M1-A1-N1-T1-E1-H1).

Fig. 1A: VP7 ML tree

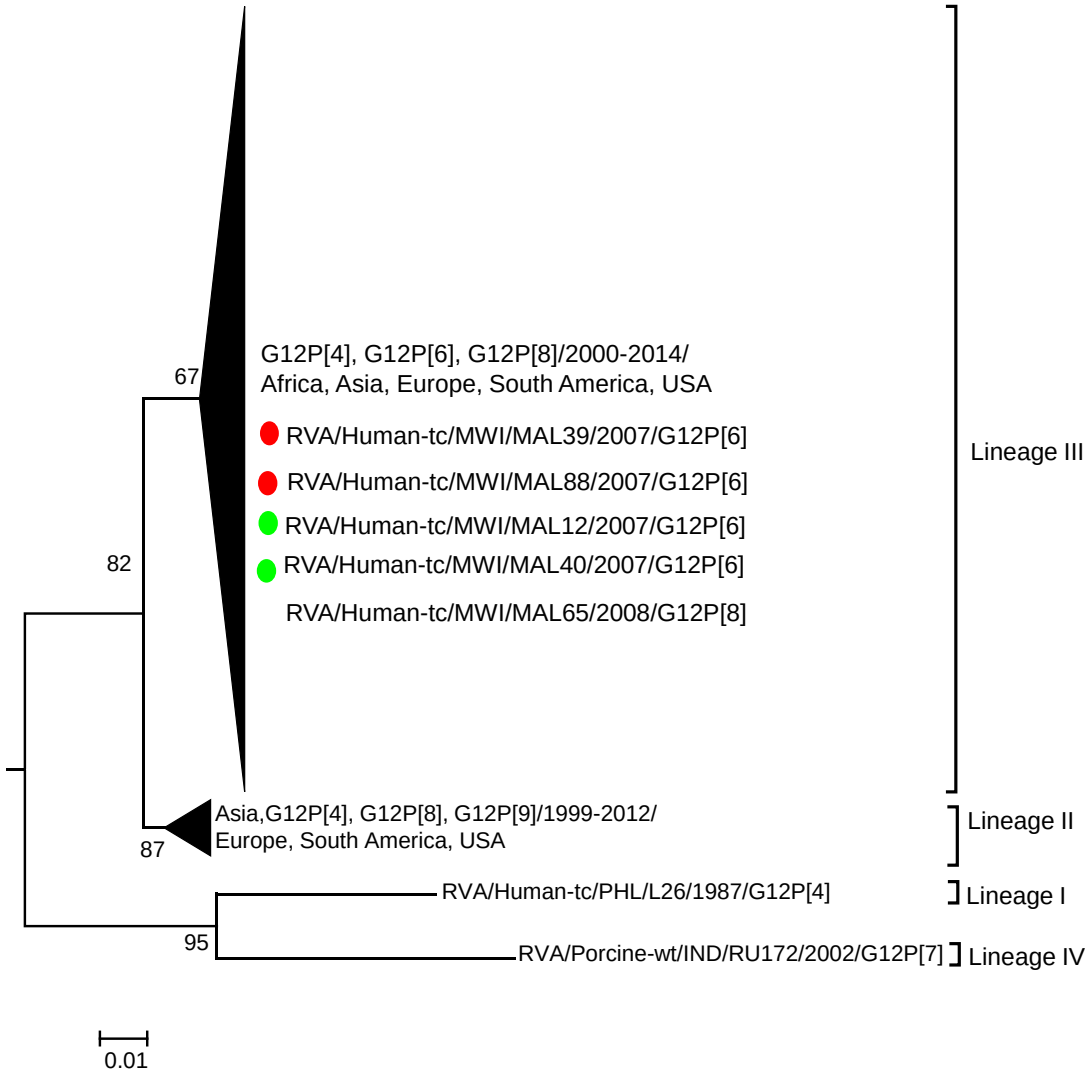

Fig. 1B: VP4 ML tree

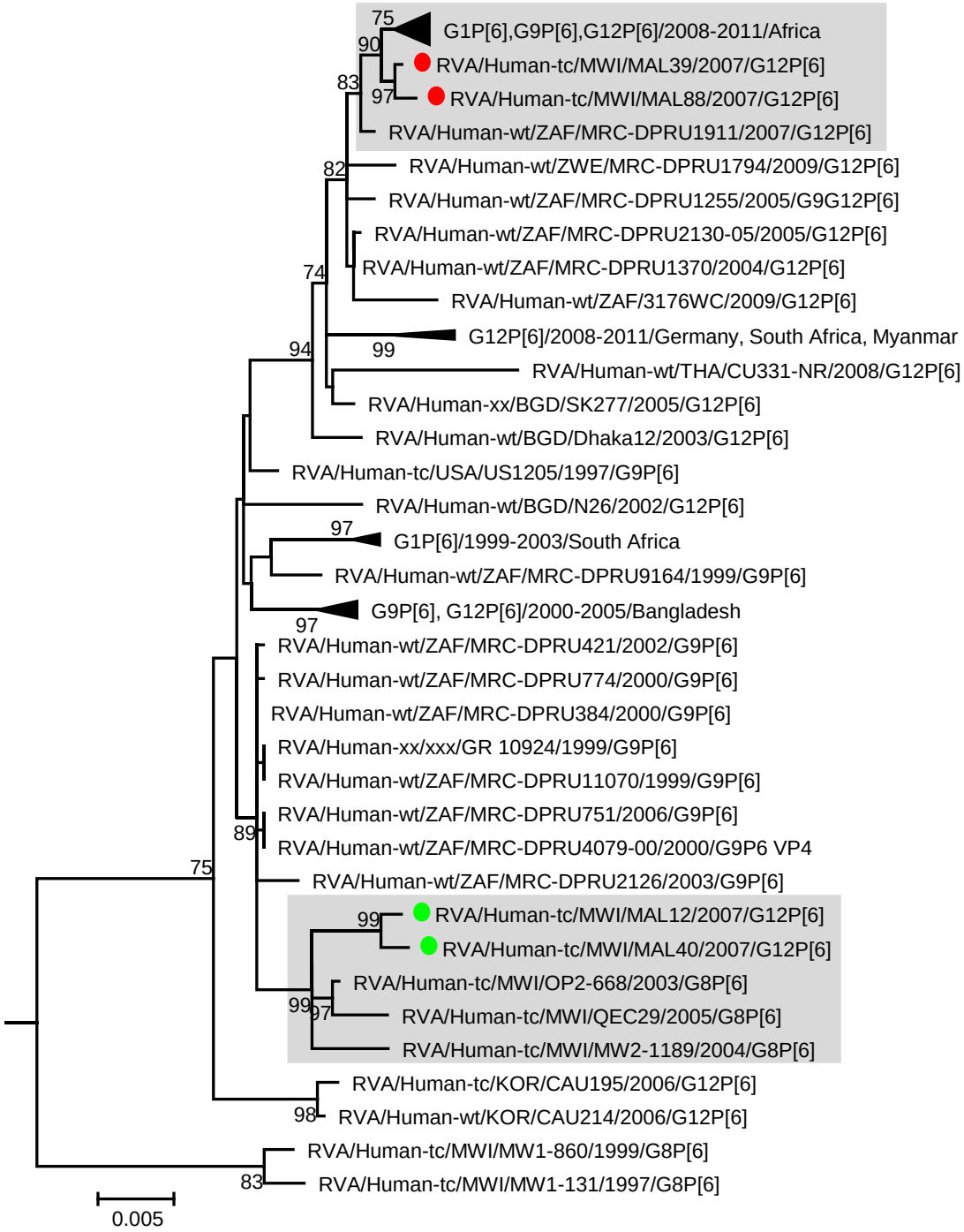

Fig. 1C: VP3 ML tree

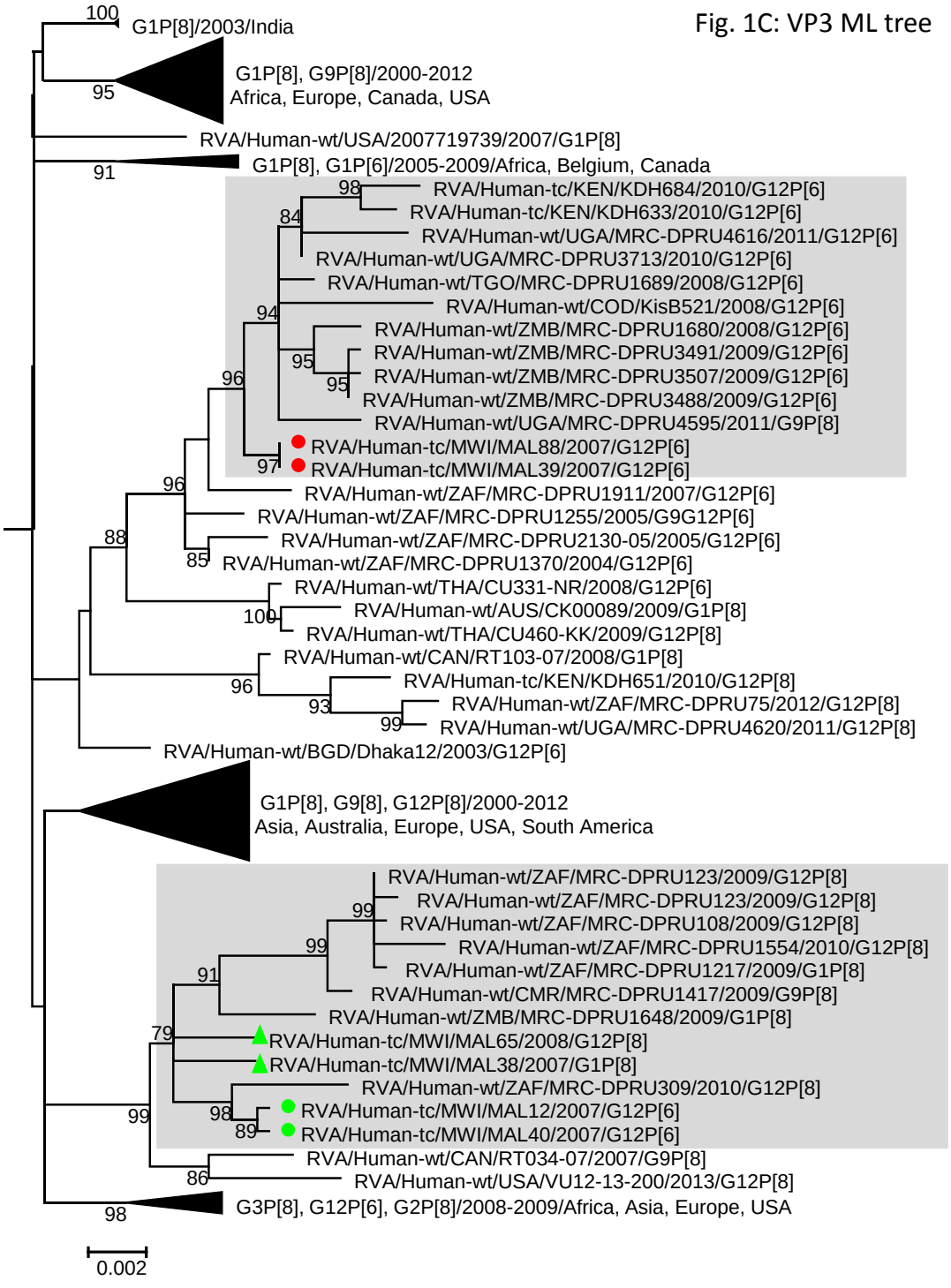

Fig. 2A VP7

Parameter	Value
Number of taxa	256
Sampling area	global
Sampling period	1987-2013
Evolutionary rate (x10E-03) (95% HPD interval)	1.55 (1.20-1.91)

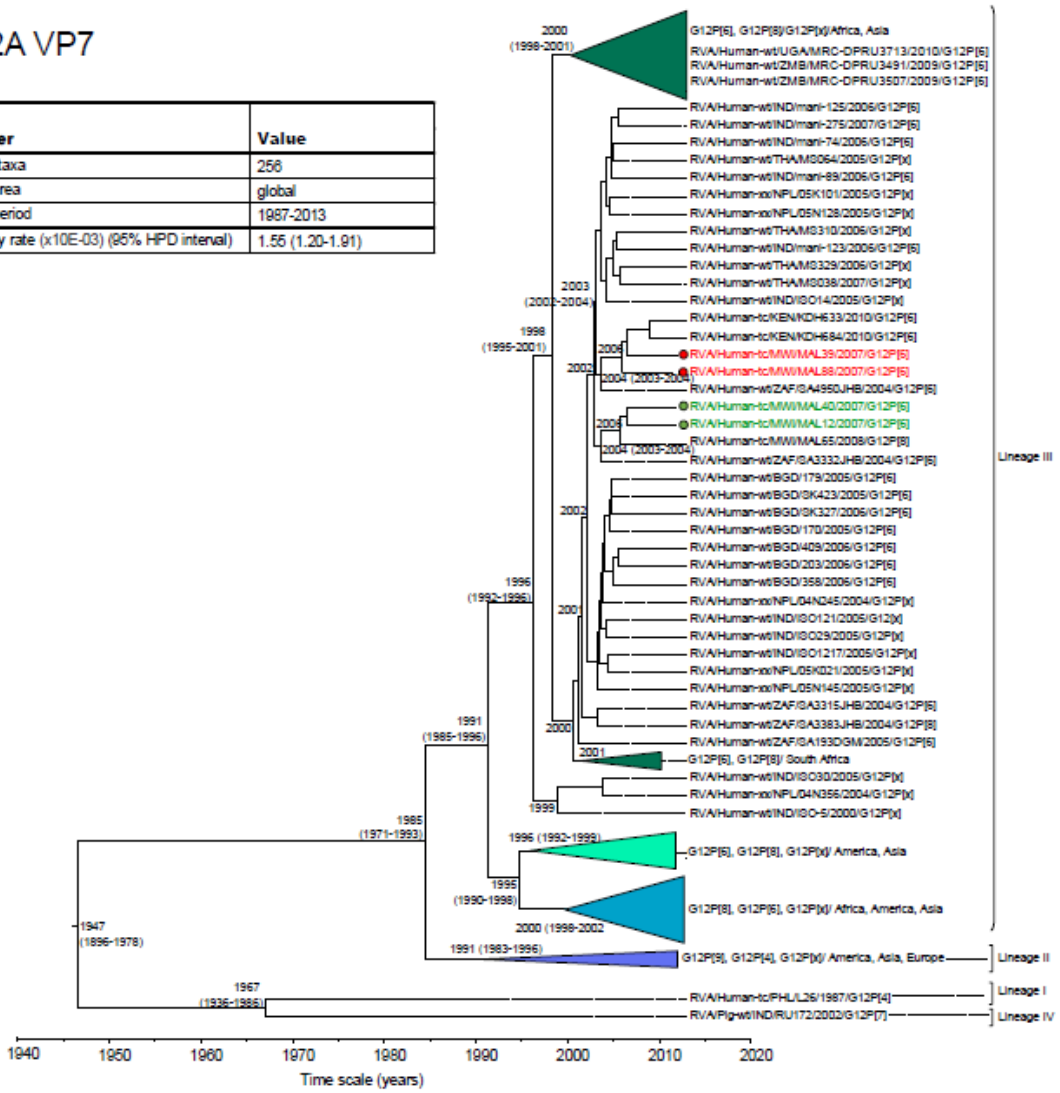

Fig. 2B VP4

Parameter	Value
Number of taxa	53
Sampling area	global
Sampling period	1997-2011
Evolutionary rate (x10E-03) (95% HPD interval)	0.99 (0.78-1.21)

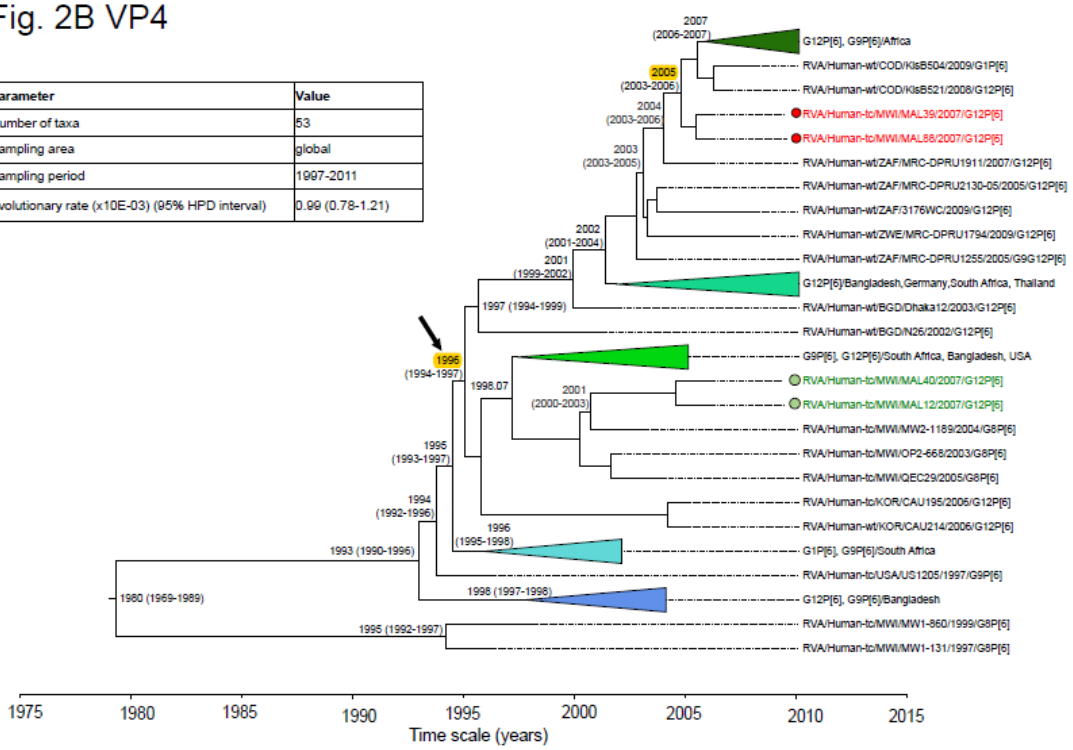

Fig. 2C VP3

Parameter	Value
Number of taxa	107
Sampling area	global
Sampling period	2000-2012
Evolutionary rate (x10E-03) (95% HPD interval)	0.94 (0.77-1.13)