



OPEN

Whole genome and in-silico analyses of G1P[8] rotavirus strains from pre- and post-vaccination periods in Rwanda

Sebotsana Rasebotsa^{1,6}, Peter N. Mwangi^{1,6}, Milton T. Mogotsi¹, Saheed Sabiu¹, Nonkululeko B. Magagula², Kebareng Rakau², Jeannine Uwimana³, Leon Mutesa³, Narcisse Muganga³, Didier Murenzi³, Lisine Tuyisenge³, Jose Jaimes⁴, Mathew D. Esona⁴, Michael D. Bowen⁴, M. Jeffrey Mphahlele², Mapaseka L. Seheri², Jason M. Mwenda⁵ & Martin M. Nyaga¹✉

Rwanda was the first low-income African country to introduce RotaTeq vaccine into its Expanded Programme on Immunization in May 2012. To gain insights into the overall genetic make-up and evolution of Rwandan G1P[8] strains pre- and post-vaccine introduction, rotavirus positive fecal samples collected between 2011 and 2016 from children under the age of 5 years as part of ongoing surveillance were genotyped with conventional RT-PCR based methods and whole genome sequenced using the Illumina MiSeq platform. From a pool of samples sequenced (n = 158), 36 were identified as G1P[8] strains (10 pre-vaccine and 26 post-vaccine), of which 35 exhibited a typical Wa-like genome constellation. However, one post vaccine strain, RVA/Human-wt/RWA/UFS-NGS:MRC-DPRU442/2012/G1P[8], exhibited a RotaTeq vaccine strain constellation of G1-P[8]-I2-R2-C2-M2-A3-N2-T6-E2-H3, with most of the gene segments having a close relationship with a vaccine derived reassortant strain, previously reported in USA in 2010 and Australia in 2012. The study strains segregated into two lineages, each containing a paraphyletic pre- and post-vaccine introduction sub-lineages. In addition, the study strains demonstrated close relationship amongst each other when compared with globally selected group A rotavirus (RVA) G1P[8] reference strains. For VP7 neutralization epitopes, amino acid substitutions observed at positions T91A/V, S195D and M217T in relation to the RotaTeq vaccine were radical in nature and resulted in a change in polarity from a polar to non-polar molecule, while for the VP4, amino acid differences at position D195G was radical in nature and resulted in a change in polarity from a polar to non-polar molecule. The polarity change at position T91A/V of the neutralizing antigens might play a role in generating vaccine-escape mutants, while substitutions at positions S195D and M217T may be due to natural fluctuation of the RVA. Surveillance of RVA at whole genome level will enhance further assessment of vaccine impact on circulating strains, the frequency of reassortment events under natural conditions and epidemiological fitness generated by such events.

Group A rotavirus is a significant viral etiological agents of acute gastroenteritis (AGE) resulting in approximately 125,000 deaths annually in children under 5 years worldwide¹. Rwanda rolled out the RotaTeq (Merck and Co., Whitehouse Station, NJ, USA) vaccine into her Expanded Program on Immunization in May 2012 and the vaccination coverage rate has been consistently high over the years > 95% since 2013, resulting in significant

¹Next Generation Sequencing Unit, Division of Virology, Faculty of Health Sciences, University of the Free State, Bloemfontein 9300, South Africa. ²Diarrhoeal Pathogens Research Unit, Sefako Makgatho Health Sciences University, Medunsa, Pretoria 0204, South Africa. ³Kigali University Teaching Hospital, College of Medicine and Health Sciences, University of Rwanda, P.O. Box 4285, Kigali, Rwanda. ⁴Division of Viral Diseases, NCIRD, Centers for Disease Control and Prevention, Atlanta, GA 30333, USA. ⁵World Health Organization, Regional Office for Africa, P.O. Box 06, Brazzaville, Republic of Congo. ⁶These authors contributed equally: Sebotsana Rasebotsa and Peter N. Mwangi. ✉email: NyagaMM@ufs.ac.za

reduction in RVA-associated hospitalization among children younger than 5 years old². Specifically, in 2013 and 2014 post introduction of the vaccine, 49% and 48% decreases in gastroenteritis cases requiring hospitalization, respectively, and 61% and 70% decreases in RV-specific diagnosis were observed, respectively⁷.

The RVA are double-stranded RNA viruses with a segmented genome enclosed within a triple-protein layer. The 11-gene segments code for six structural proteins (VP1–VP4, VP6 and VP7) and depending on the strain, five or six non-structural proteins (NSP1–NSP5/6)⁴. A more complex RVA classification is in use based on the genotype properties of all the 11 gene segments as Gx-Px-Ix-Rx-Cx-Mx-Ax-Nx-Tx-Ex-Hx which encodes for VP7-VP4-VP6-VP1-VP2-VP3-NSP1-NSP2-NSP3-NSP4-NSP5/6 proteins (x represents the genotype number)⁵. The segmented nature of RVA enables reassortment events that drive RVA evolution along with genomic rearrangement, genetic drift, deletion of gene sequences and zoonotic transmission^{6–8}. Globally, majority of the human RVA strains possess either the Wa-like genotype constellation (I1-R1-C1-M1-A1-N1-T1-E1-H1) or the DS-1-like genotype constellation (I2-R2-C2-M2-A2-N2-T2-E2-H2)^{9–11}. Currently, 36 G, 51 P, 26 I, 22 R, 20 C, 20 M, 31 A, 22 N, 22 T, 27 E and 22 H genotypes have been approved by the Rotavirus Classification Working Group (<https://rega.kuleuven.be/cev/viralmetagénomics/virus-classification>). Of the six common RVA genotypes (G1P[8], G2P[4], G3P[8], G4P[8], G9P[8] and G12P[8]) circulating globally, the G1P[8] genotype constitutes most of the human RVA infections¹². In Africa, the G1P[8] genotype accounts for approximately 29% of all the circulating RVA strains¹².

To date, four oral live attenuated RVA vaccines, Rotarix (GlaxoSmithKline, Rixensart, Belgium), RotaTeq (Merck and Co., Whitehouse Station, NJ, USA), ROTAVAC (Bharat Biotech, India) and ROTASIL (Serum Institute of India Pvt. Ltd., India) have been prequalified for global use by World Health Organization (WHO)¹³. All these vaccines are reported to be safe and highly efficacious based on clinical trial evaluations. In 2009, WHO recommended routine immunization of RVA vaccine to all infants to reduce the high mortality associated with RVA infections¹⁴. Rwanda was the first low-income country to introduce RotaTeq vaccine in May 2012 but switched to Rotarix in April 2017¹⁵. RotaTeq is an oral live attenuated RV vaccine comprising 5 bovine-human RVA reassortant strains (G1P[5], G2P[5], G3P[5], G4P[5] and G6P[8]) with a WC3 bovine backbone. The genotypes G1, G2, G3, G4 and P[8] are human rotavirus derived, while genotypes G6 and P[5] are bovine RVA derived¹⁶. RotaTeq is co-administered with other routine childhood vaccines in a three-dose schedule recommended at 2, 4 and 6 months of age¹⁷.

The rationale for RVA vaccine introduction was to combat the high morbidity and mortality of RVA-associated disease burden in Rwanda¹⁸. Rotavirus strain G1P[8] is the most prevalent genotype globally, several countries have analyzed the full genome of G1P[8] strains, however, Rwanda lacks such reports^{8,19–24}. Amongst these studies, one investigated the whole gene sequences of G1P[8] RVA strains collected from the pre- and post-vaccination era in South Africa and the second investigated only G1P[8] strains collected during the post-vaccination era in Brazil^{20,22}. Magagula and colleagues concluded that all the G1P[8] strains analyzed exhibited a Wa-like genetic constellation and shared a moderate nucleotide identity of 89–96% and 93–95% to Rotarix and RotaTeq G1P[8] vaccine components, respectively. In the second study, Santos and colleagues, reported double reassortments between Wa- and DS-1-like genogroups and also, Wa- and AU-1-like genogroups. A comparison of the amino acid residues present in the antigenic epitopes of VP7 and VP4, indicated a differences in the electrostatic charge distribution, between the Brazilian G1P[8] wild-types strains and Rotarix and RotaTeq G1P[8] vaccine strains. The availability of RVA strains collected from 2011 to 2016 in Rwanda during the pre- and post-vaccination periods with the RotaTeq vaccine presented an opportunity to assess vaccine impact on the molecular diversity of circulating strains. To gain insight into the overall genetic makeup and evolution of the G1P[8] strains from Rwanda, whole genome analyses was conducted and the data was compared with RVA strains from other parts of the world. To our knowledge, this study is the first to analyze circulating RVA strains in Rwanda on a whole genome level as compared to the traditional binary classification into G/P genotypes.

Results

Whole genome constellation analyses. A pool of 158 randomly selected Rwandan stool samples were sequenced and only 36 were genotype G1P[8], which was the inclusion criteria for this study. The genome constellations of the 35 pre- and post-vaccine G1P[8] strains described in this study was the typical G1-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1. One strain was a reassortant G1P[8] with genetic backbone of I2-R2-C2-M2-A3-N2-T6-E2-H3 (Fig. 1). The reassortant G1P[8] Rwandan strain was detected from a 2 months old infant collected 4 days after vaccination with the first dose of RotaTeq vaccine. Almost all gene segments of the reassortant G1P[8] strain shared a near absolute nucleotide and amino acid identity to cognate gene sequences of the RotaTeq vaccine G1P[8] strain. This suggested that the reassortant G1P[8] strain detected in a Rwandan child with diarrhoea was a RotaTeq vaccine derived G1P[8] (vdG1P[8]) strain. The lengths of the open reading frame (ORFs) for segments 1–11 for all 35 typical G1P[8] strains were 3,264 bp, 2,637 bp, 2,505 bp, 2,325 bp, 1,458 bp, 1,194 bp, 930 bp, 951 bp, 978 bp, 525 bp, and 594 bp, respectively. The ORF sequences for all 11 genes of these 35 Rwandan G1P[8] were sequenced and deposited in GenBank under accession numbers MN632673–MN633067 for all the gene segments.

Phylogenetic and sequence analyses. The concatenated ORF sequence analyses revealed the clustering of the Rwandan pre- and post-vaccine RVA strains into two lineages (Fig. 1). Within each lineage there were two paraphyletic sub-lineages with pre-vaccine introduction strains occupying one of the sub-lineages and post-vaccine introduction strains occupying the other (Fig. 1). Overall, the 35 typical pre- and post-vaccine introduction G1P[8] strains demonstrated a close genetic relationship amongst themselves as compared with other RVA G1P[8] strains circulating on a global scale (Fig. 2A–K). Even though Rotarix vaccine was not used in Rwanda at the time of the study, we included the Rotarix vaccine gene sequences in all the analyses.

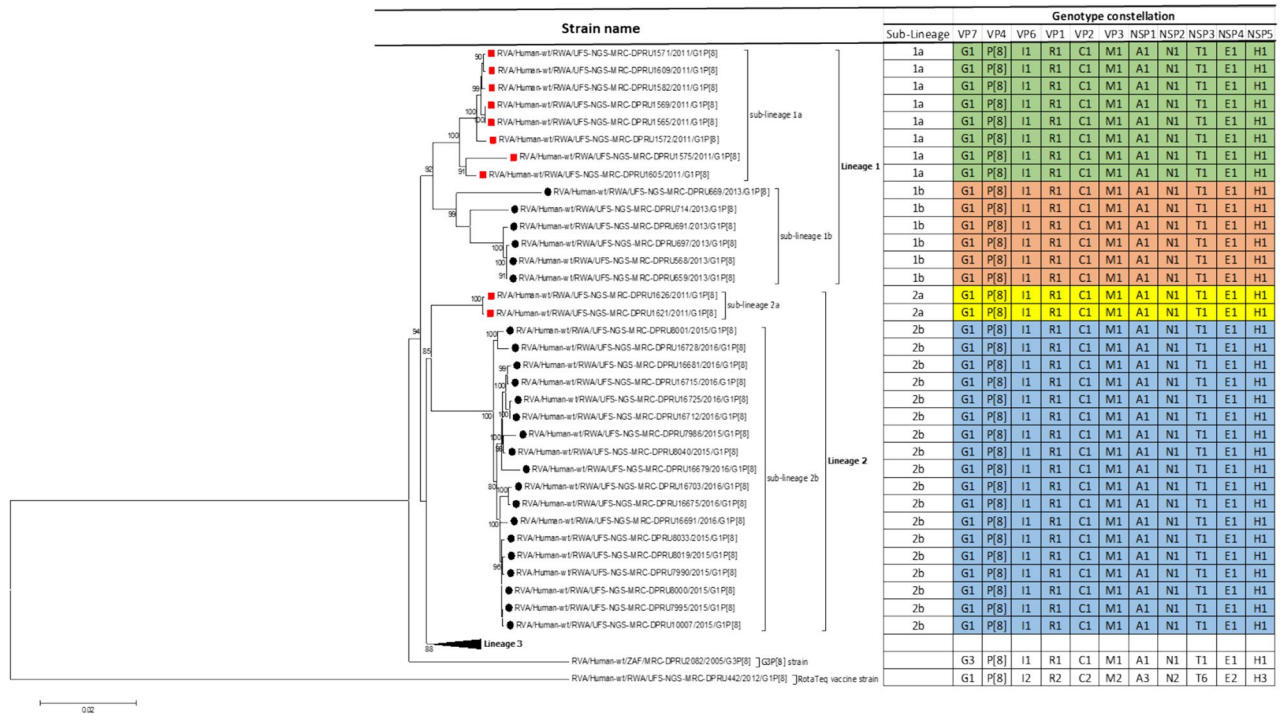


Figure 1. Maximum likelihood phylogram revealing the genetic relatedness of the concatenated, whole-gene ORF sequences (17,492 base pairs) for the 25 post-vaccine G1P[8] and 10 pre-vaccine G1P[8] RVA study strains characterized by whole-gene analysis. The strain-specific genotype constellations are included to the right of each strain and were identified using the assembled ORF sequences. Black circle denotes post-vaccine strains, while red squares denotes pre-vaccine strains. Bootstrap values $\geq 70\%$ are indicated at branch nodes where applicable and the scale bar indicates the number of nucleotide substitutions per site. On the table, colored in yellow and sky blue are lineage 2 pre- and post-vaccine strains belonging in sub-lineage 2a and sub-lineage 2b, respectively. Colored in green and brown are lineage 1 pre- and post-vaccine strains belonging in sub-lineage 1a and sub-lineage 1b, respectively.

Phylogenetic and sequence analyses of VP7 gene. Phylogenetically, RVA G1 reference genotypes from human and animal RVA strains utilized in this analyses mapped into the seven known lineages (I–VII) and 35 of 36 Rwandan pre- and post-vaccine introduction G1P[8] strains clustered into lineage I, which consisted of a global collection of G1 strains detected during 2003–2015 RVA seasons^{20,25}. The lineage I strains were further segregated into sub-lineages usually by year and in some cases by vaccination status. Within this lineage I, the G1 strains segregated into two sub-lineages: one consisted of 19 G1 strains detected in the post-vaccination era and the other was a mixture of both pre-vaccine (ten G1) and post-vaccine (five G1) introduction strains (Fig. 2A). The nucleotide (nt) and amino acid (aa) identities of the 19 G1 post-vaccine introduction strains amongst themselves was in the range of 99.2–100% and 98.0–100%, respectively, whereas the nt (aa) similarities of the five post- and ten pre-vaccine introduction mixed G1 sub-lineage was 94.7–100% (92.7–100%). Further comparison of the G1 strains in the typical post-vaccine sub-lineage with those in the mixed pre- and post-vaccine sub-lineage revealed a nt (aa) similarities of 92.7–96.3% (90.7–96.0%). The VP7 gene of Rwandan post-vaccination introduction strain RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU669/2013/G1P[8] was an orphan gene that did not cluster in either the typical post-vaccine sub-lineage or mixed pre-and post-vaccine introduction sub-lineage (Fig. 2A). The VP7 gene sequence of the RotaTeq vdG1P[8] strain, RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU442/2012/G1P[8], clustered into lineage III (Fig. 2A) alongside cognate gene sequence of the RotaTeq vaccine strain RVA/Vaccine/USA/RotaTeq-WI79-9/1992/G1P7[5] (Fig. 2A). The VP7 gene sequences of strain RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU442/2012/G1P[8] displayed absolute identity with the VP7 gene of RotaTeq reassortant strain, RVA/Vaccine/USA/RotaTeq-WI79-9/1992/G1P7[5] and appeared to be a RotaTeq VP7 gene.

Comparative analyses of neutralizing antigenic epitopes in the VP7 proteins of Rwandan G1P[8] and vaccine strains of RVA. Structurally, the VP7 gene contains two defined neutralization epitopes: 7-1 and 7-2²⁶. Aoki and colleagues further subdivided 7-1 epitope into 7-1a and 7-1b²⁶. These three antigenic epitopes comprise 29 amino acid residues (14 residues in 7-1a, 6 residues in 7-1b and 9 residues in 7-2). Using the G1 alignment for the VP7 gene, we identified amino acid differences in these neutralization epitopes between the wild-type Rwandan G1 RVA strains and the RotaTeq and Rotarix vaccine G1 strain (Fig. 3A). Out of these 29 amino acid residues, 20 (amino acids G96, W98, K99, Q104, V129 and D130 in 7-1a region, all positions in 7-1b region and K143, D145, Q146, N221, G264 in the 7-2 region) were completely conserved among all the Rwandan G1 strains. Relative to the G1 component of the RotaTeq vaccine, the Rwandan G1 strains showed up

A VP7

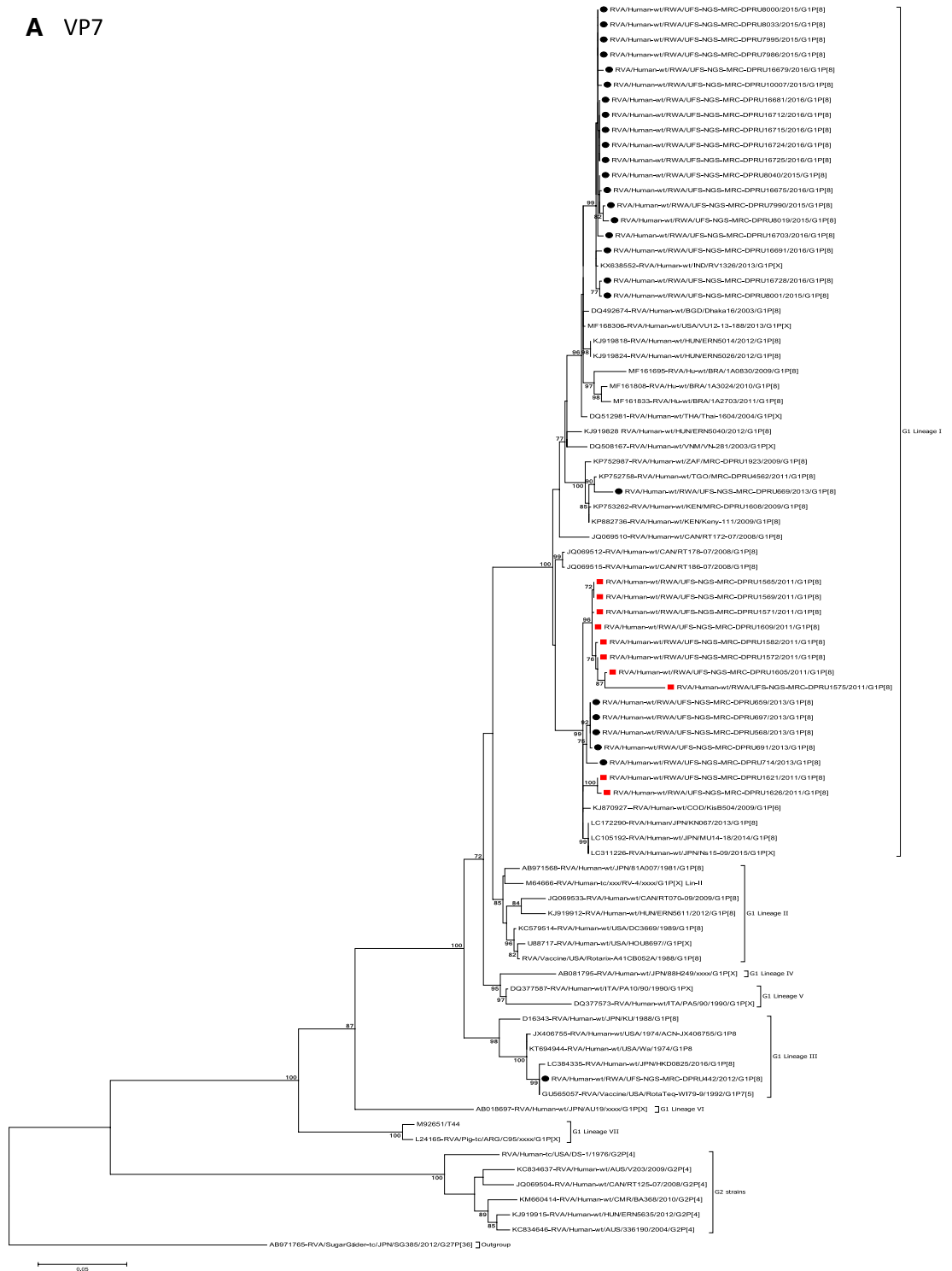


Figure 2. (A)–(K) Phylogenetic relatedness of rotavirus group A species base on (A) VP7, (B) VP4, (C) VP6, (D) VP1, (E) VP2, (F) VP3, (G) NSP1, (H) NSP2, (I) NSP3, (J) NSP4 and (K) NSP5 of the study strains from Rwanda with representatives of known human and animal rotavirus genotypes. Pre- and post-vaccine strains are indicated with a red squares and black circles, respectively. Bootstrap values $\geq 70\%$ are indicated at each branch node. Scale bars represents substitutions per nucleotide site.

to 10–11 differences (T87I, T91A/V, N94S, D97E, D100E, S123N, K291R, S147N, L148F and M217T) located on the surface of the protein structure RotaTeq (Fig. 3A,B). Among these changes seen relative to RotaTeq vaccine G1 strain, only the T91A/V, S190D and M217T substitution at positions 91, 190 and 217 involve a change from

B VP4

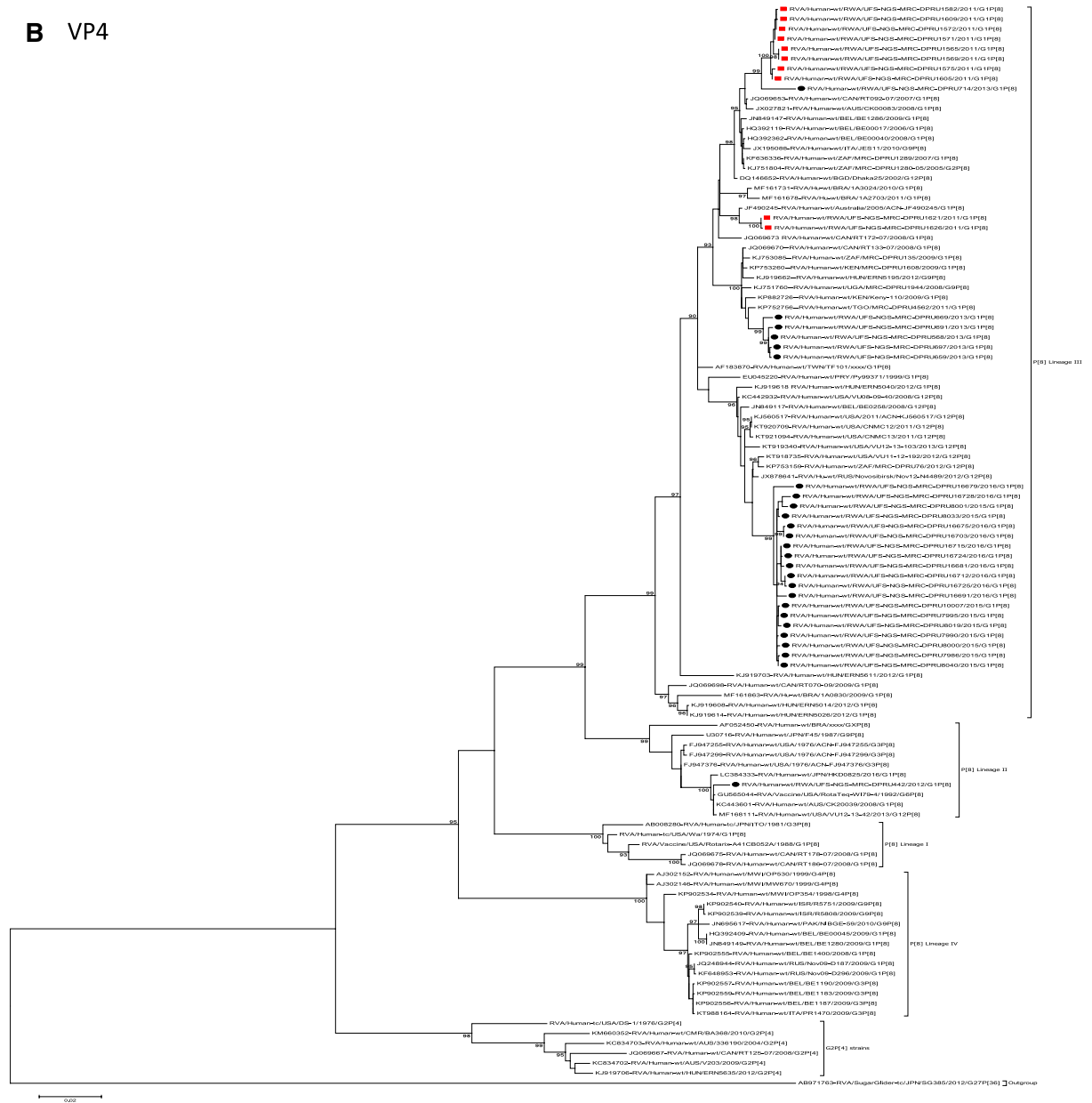


Figure 2. (continued)

a polar to non-polar molecule, polar to negative molecule and non-polar to polar molecule, respectively, were radical in nature. The remainder of the changes seen were neutral or conservative in nature.

Comparative analyses of cytotoxic T lymphocytes epitopes of the G1 proteins of Rwandan and vaccine strains of RVA. Cytotoxic T lymphocytes (CTL) have been linked to clearance of RV infection resulting in complete short-term and partial long-term protection against re-infection²⁷. Two linear CTL epitopes responsible for this activity have been reported on VP7 protein at amino acid positions 16–28 and 40–52^{28,29}. An analyses of the Rwandan G1 pre- and post-vaccine introduction strains showed three amino acid differences at positions T41F/S, V42M and A46V with G1 component of the RotaTeq and Y41F/S, V42M and A46V to Rotarix vaccines (Fig. 4). Amongst these amino acid differences, the change at position T41F relative to the RotaTeq G1 gene sequence was radical in nature and resulted in a change in polarity from polar to nonpolar, while the other two substitutions were conservative in nature. Twenty-three out of the 26 amino acids that comprises the T-cell antigen epitopes were completely conserved.

Phylogenetic and sequence analyses of VP4 genes. The VP4 gene sequences of the 35 Rwandan strains and a single RotaTeq vdG1P[8] strain collected in the pre- and post-vaccination periods were compared with representative human RVA from the four established VP4 P[8] genotype lineages (I to IV)^{25,30}. With the

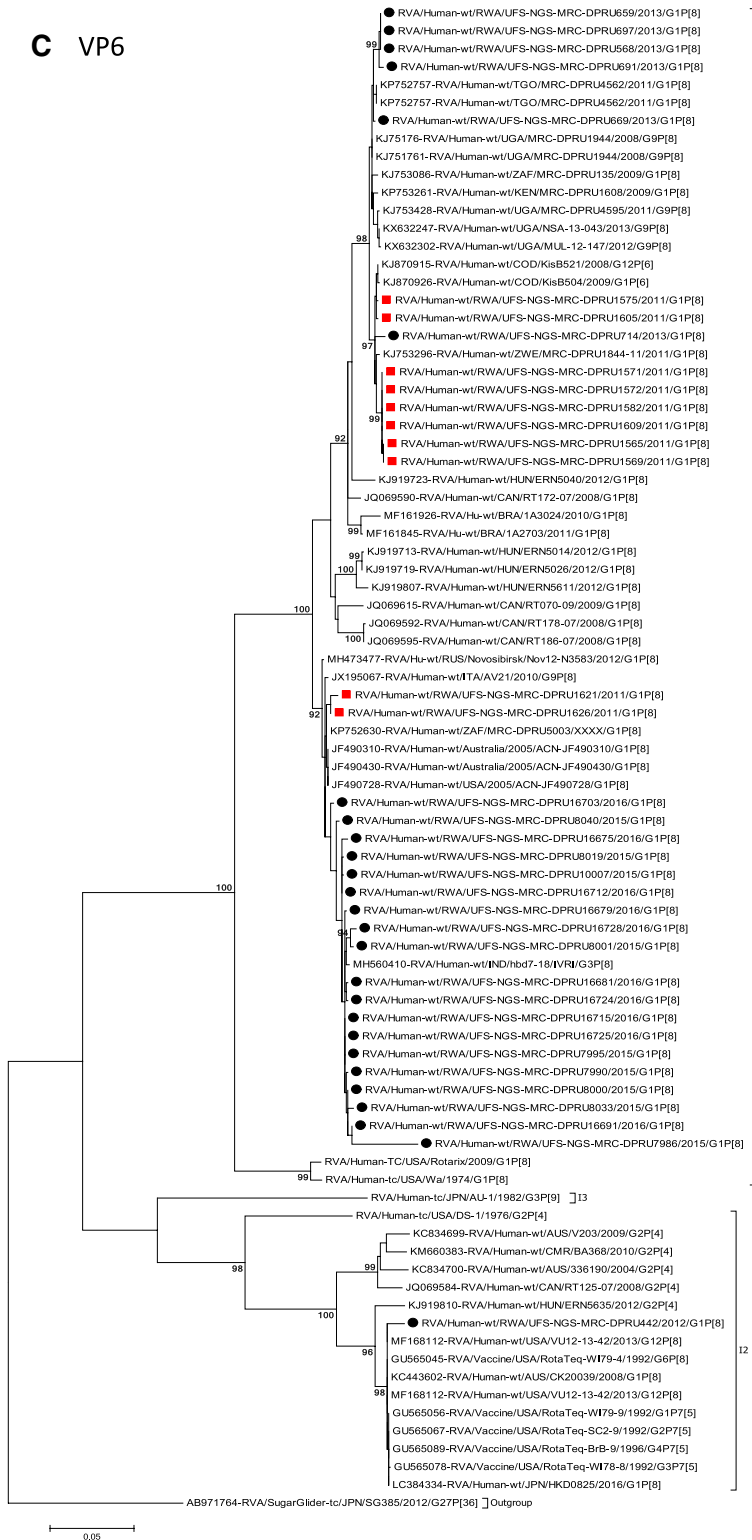


Figure 2. (continued)

exception of RotaTeq vdG1P[8] strain RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU442/2012/G1P[8] which clustered in P[8]-lineage II with the P[8] component of RotaTeq vaccine strain RVA/vaccine/USA/RotaTeq-W-179-4/1992/G6P1A[8], the other 35 study strains clustered in P[8]-lineage III (Fig. 2B). Within P[8]-lineage III cluster, the Rwandan strains segregated into 3 typical and one mixed sub-lineages (Fig. 2B). The first sub-lineage consisted of 8 pre-vaccine introduction P[8] and a single post-vaccine introduction P[8] strains and they shared nt (aa) identities in the range of 99.7–100% (99.4–100%) amongst themselves (Fig. 2B). The second sub-lineage comprised of only 2 pre-vaccine introduction strains detected in 2011 RVA season and shared an abso-

D VP1



Figure 2. (continued)

lute identity with each other (Fig. 2B). The third sub-lineage was characterised of 5 post-vaccine introduction strains circulating in the 2013 RVA season and shared a nt (aa) similarities amongst themselves ranging from 99.4–99.9% (98.6–100%) (Fig. 2B) and the fourth sub-lineage comprised of 19 post-vaccine introduction P[8] strains and shared nt (aa) identities amongst themselves in the range of 98.8–100% (98.0–100%) (Fig. 2B). Comparison of pre-vaccine introduction strains in the mixed sub-lineage to the post-vaccine introduction strains in the third and fourth sub-lineages revealed a moderate nt (aa) similarities in the range of 95.7–97.8% (89.5–94.5%). The VP4 P[8] component of the RotaTeq vdG1P[8] strain RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU442/2012/G1P[8] clustered distinctly in P[8]-lineage II together with the RotaTeq vaccine strain RVA/

E VP2

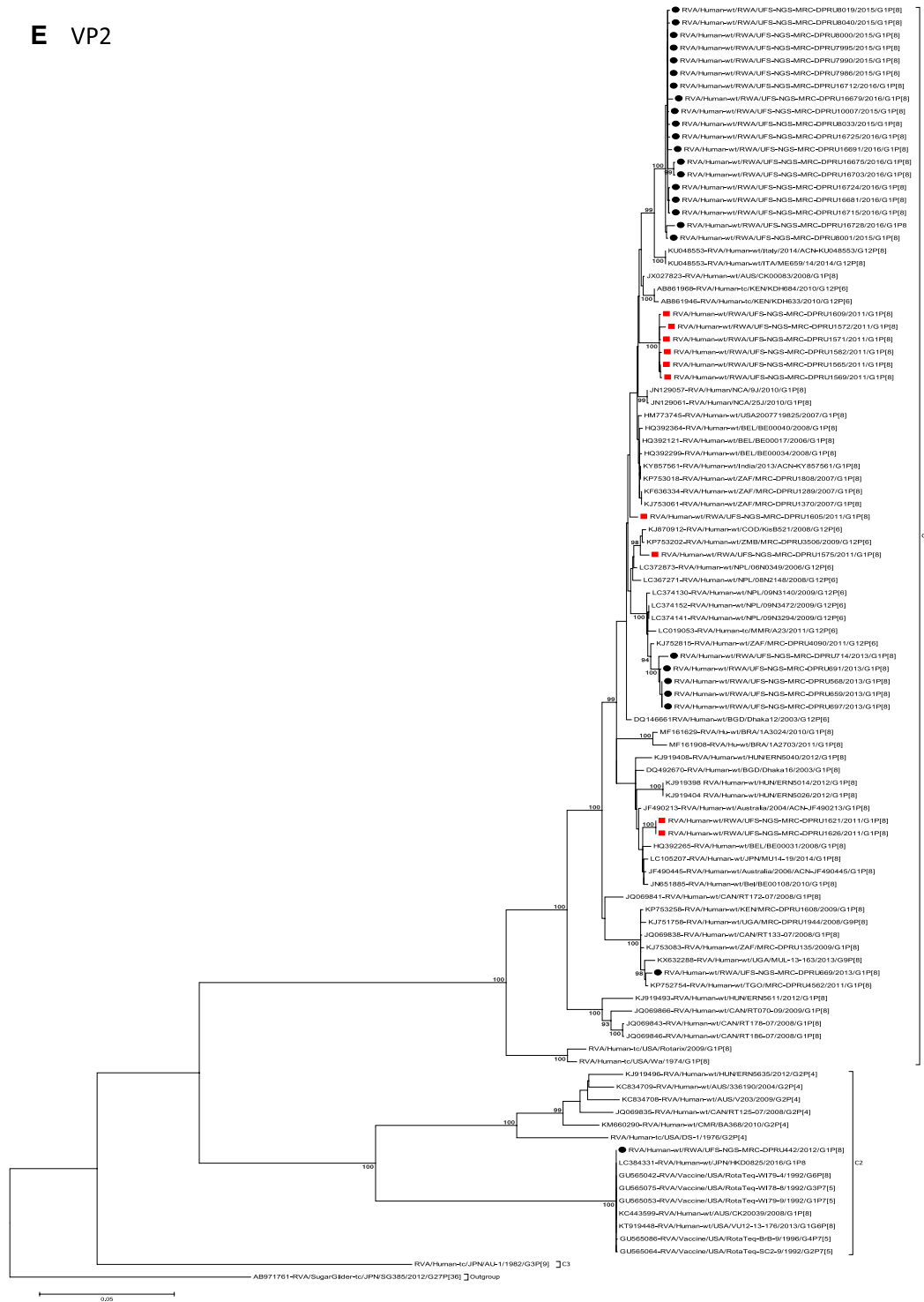


Figure 2. (continued)

Vaccine/USA/RotaTeq-WI79-4/1992/G6P1A8 and G3P[8] strains detected in 1976 in the US, G1P[8] strains detected in 2008 in Australia and G1P[8] detected in 2016 in Japan (Fig. 2B).

Comparative analyses of neutralizing antigenic epitopes in the VP4 protein of Rwandan G1P[8] and vaccine strains of RVA. The VP4 protein, which in this study represented P[8] strains was divided into the VP8* and VP5* regions for comparison. The VP8* region contains four (8-1 to 8-4) neutralizing antigenic epitopes, while the VP5* region has five (5-1 to 5-5)³¹. These two epitopes contain 37 amino acid residues, 25 in the VP8* and 12 in the VP5* antigenic epitope regions. Out of these 37 amino acid residues that spans the neutralization epitopes, those at positions 100, 146, 148, 150, 188, 190 and 194 (8-1 region), 180 and 183 (8-2 region), 114, 116, 132, 133, and 135 (8-3 region), 87, 88, and 89 (8-4 region), 384, 386, 388, 393, 394,

F VP3

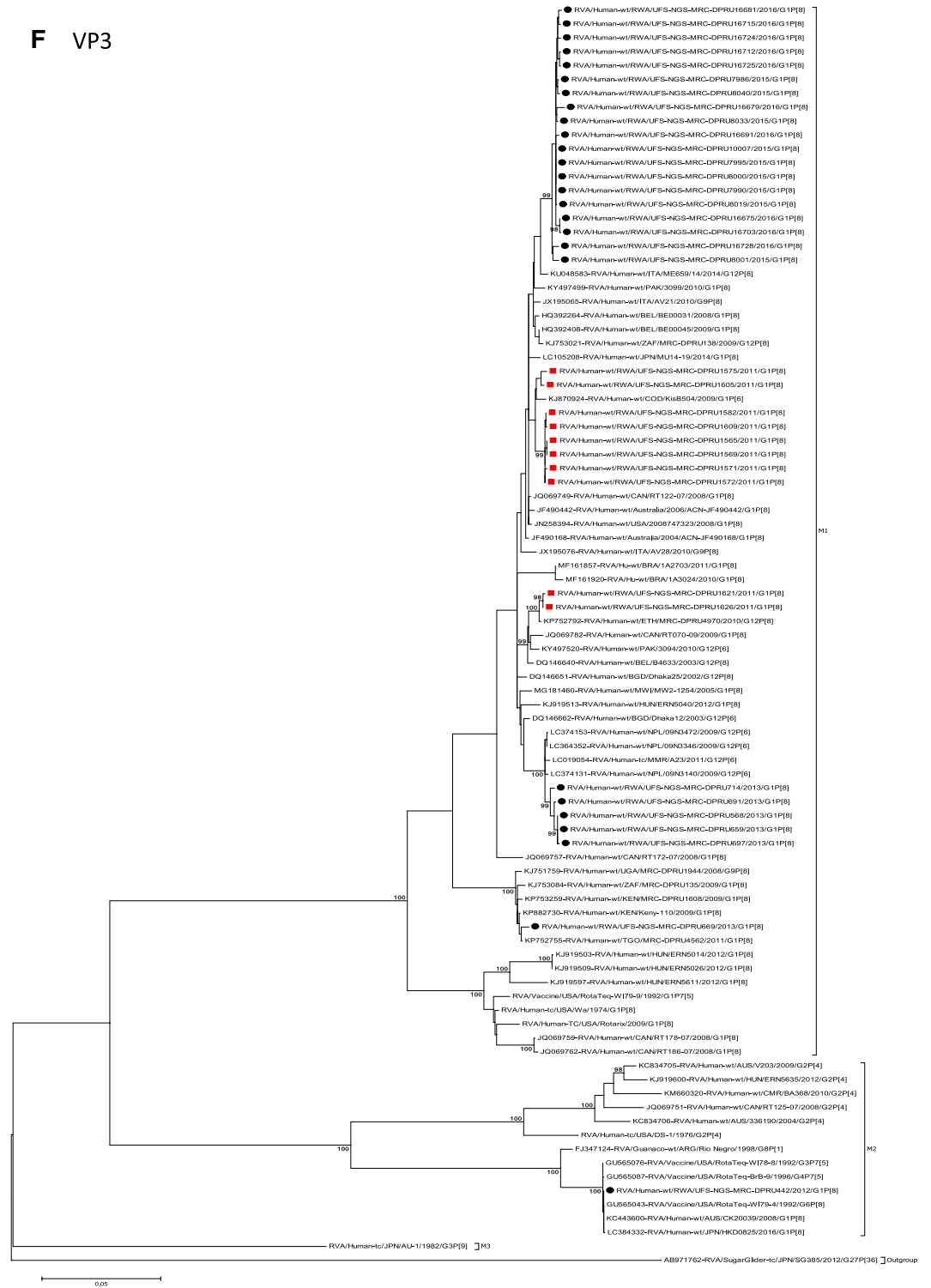


Figure 2. (continued)

398, 440, and 441 (5-1 region), 434 (5-2 region), 459 (5-3 region), 429 (5-4 region), and 306 (5-5 region), are known neutralization escape mutation sites (Fig. 5A)^{32,33}. A comparison of the Rwandan P[8] strains to the P[8] component of RotaTeq and Rotarix vaccines showed 29 identical amino acid residues distributed through the antigenic epitopes of the VP4 gene (Fig. 5A). The differences between the P[8] study strains and vaccine P[8] component of RotaTeq and Rotarix were mostly contained in VP8* epitopes 8-1 and 8-3. Relative to the RotaTeq, the P[8] component of the Rwandan G1P[8] strains exhibited 3–4 amino acid differences. The changes identified relative to the RotaTeq P[8] vaccine strain were seen at positions E150D and D195G in neutralization epitope 8-1 region (Fig. 5A,B) and were all located on the surface of the protein structure. Overall, the amino acid changes identified at positions N113D, S131R and N135D, resulted in a change in charge from a polar molecule with

G NSP1

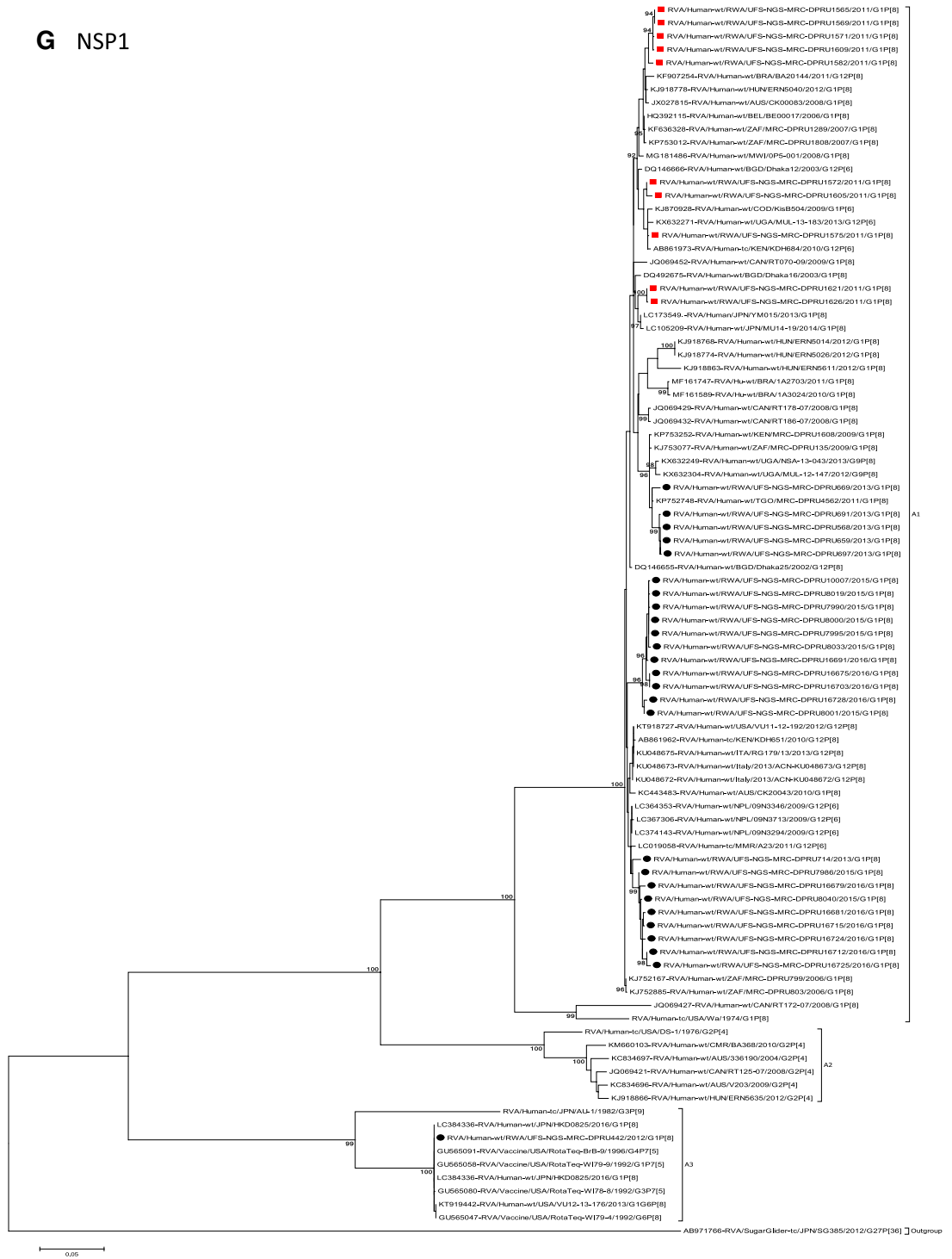


Figure 2. (continued)

potential side chain with five hydrogen bonds to a negatively charged molecule with four hydrogen bonds, and finally, amino acid substitutions at position N195G for both vaccines, resulted in a change in polarity.

Phylogenetic and sequence analyses of VP1–VP3, NSP1, NSP4 and NSP5. These six genes segregated into typical pre- and post-vaccine introduction clusters, hence presented together. Phylogenetic analyses based on VP1–VP3, NSP1, NSP4 and NSP5 nucleotide sequences of each Rwandan G1P[8] strain showed that each gene (exception of RotaTeq vG1P[8] strain) segregated and clustered with cognate gene sequences of Wa-like strains belonging to genotype 1 (Fig. 2D–G, J, K). The VP1 gene sequences of the 35 typical Rwandan G1P[8] strains segregated into a single pre-vaccine introduction sub-lineage exhibiting nt (aa) similarities in the range of 99.9–100% (99.8–100%) and 2 post-vaccine introduction sub-lineages with nt (aa) identities in the range of

H NSP2

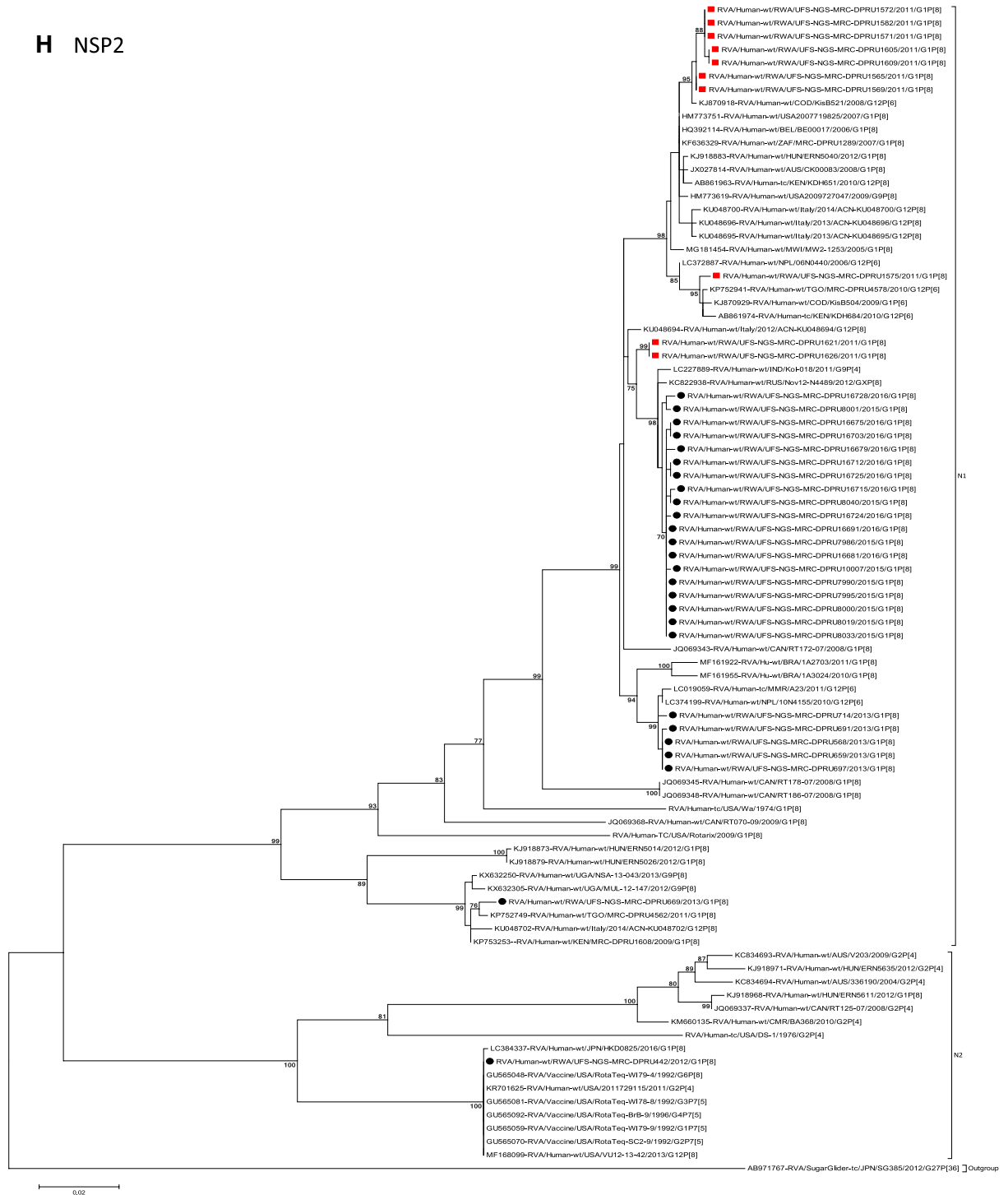


Figure 2. (continued)

96.8–98.0% (98.7–99.5%) (Fig. 2D). Four pre-vaccine introduction orphan strains that did not cluster in either of the three sub-lineages were also observed (Fig. 2D). For VP2, the C1 genes segregated into three typical sub-lineages (one typical pre- and two post-vaccine introduction) and five orphan strains which did not cluster into either of the three sub-lineages (Fig. 2E). The two post-vaccine introduction sub-lineage consisting of five and 19 strains shared nt (aa) similarities in range of 99.6–100% (99.6–100%) and 99.3–100% (99.7–100%), respectively. Furthermore, the typical pre-vaccine introduction sub-lineage incorporated strains (n = 6) detected in the 2011 RVA season and with nt (aa) sequence identities amongst themselves in the range of 99.9–100% (99.8–100%). Phylogenetically, the VP3 gene sequences of the Rwandan pre- and post-vaccine introduction strains separated into three typical sub-lineages and three orphan strains were observed (Fig. 2F). The first post-vaccine sub-

I NSP3

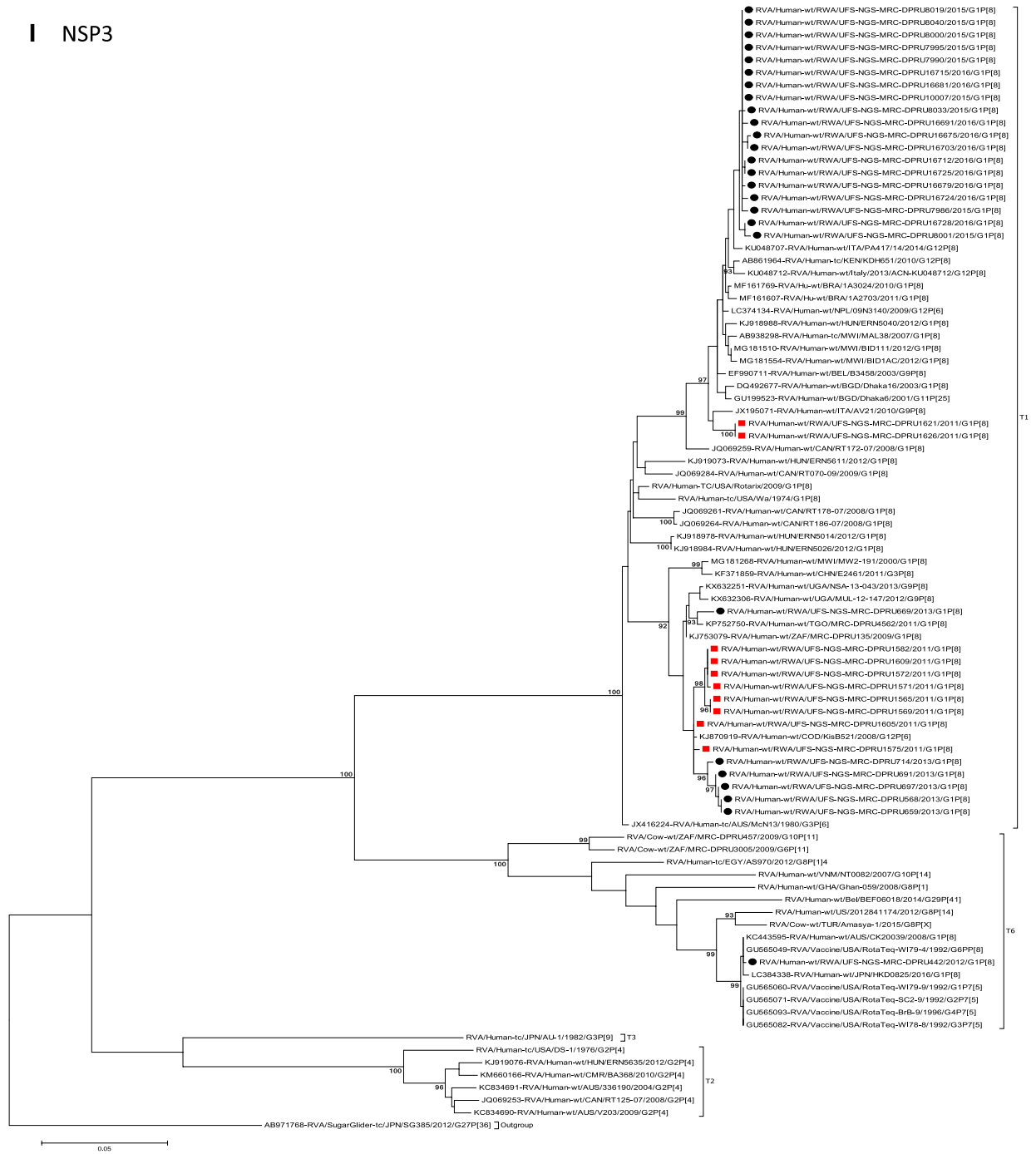


Figure 2. (continued)

lineage consisted of 19 strains and the second one had five strains that shared nt (aa) similarities ranging from 99.2–100% (98.9–100%) and 99.4–100% (99.3–100%), respectively. The pre-vaccine introduction sub-lineage contained strains that shared nt (aa) identities of 99.3–100% (98.5–100%). The NSP1 gene sequences of the typical Rwandan pre- and post-vaccine introduction G1P[8] strains were grouped into a single sub-lineage consisted of all ten pre-vaccine introduction strains and three typical post-vaccination introduction sub-lineages (Fig. 2G). The pre-vaccine strains shared nt (aa) similarities that ranged from 99.0–99.9% (97.5–99.8%). The three post-vaccine introduction NSP1 genes that segregated into three sub-lineages revealed nt (aa) identities of 98.7–99.9% (98.1–99.8%), 99.0–99.9% (98.4–99.8%) and 98.2–99.9% (98.1–99.8%), respectively. The NSP4 gene sequences of the Rwandan strains also segregated into two typical post-vaccination sub-lineages and a single pre-vaccination sub-lineage (Fig. 2J). Six orphan Rwandan E1 genotypes did not cluster into either the pre- or post-vaccine introduction sub-lineages (Fig. 2J). The first typical post-vaccine introduction strain sub-lineage comprised of 19 NSP4 genes and the second one had five E1 strains detected in 2011 RVA season which exhibited

J NSP4

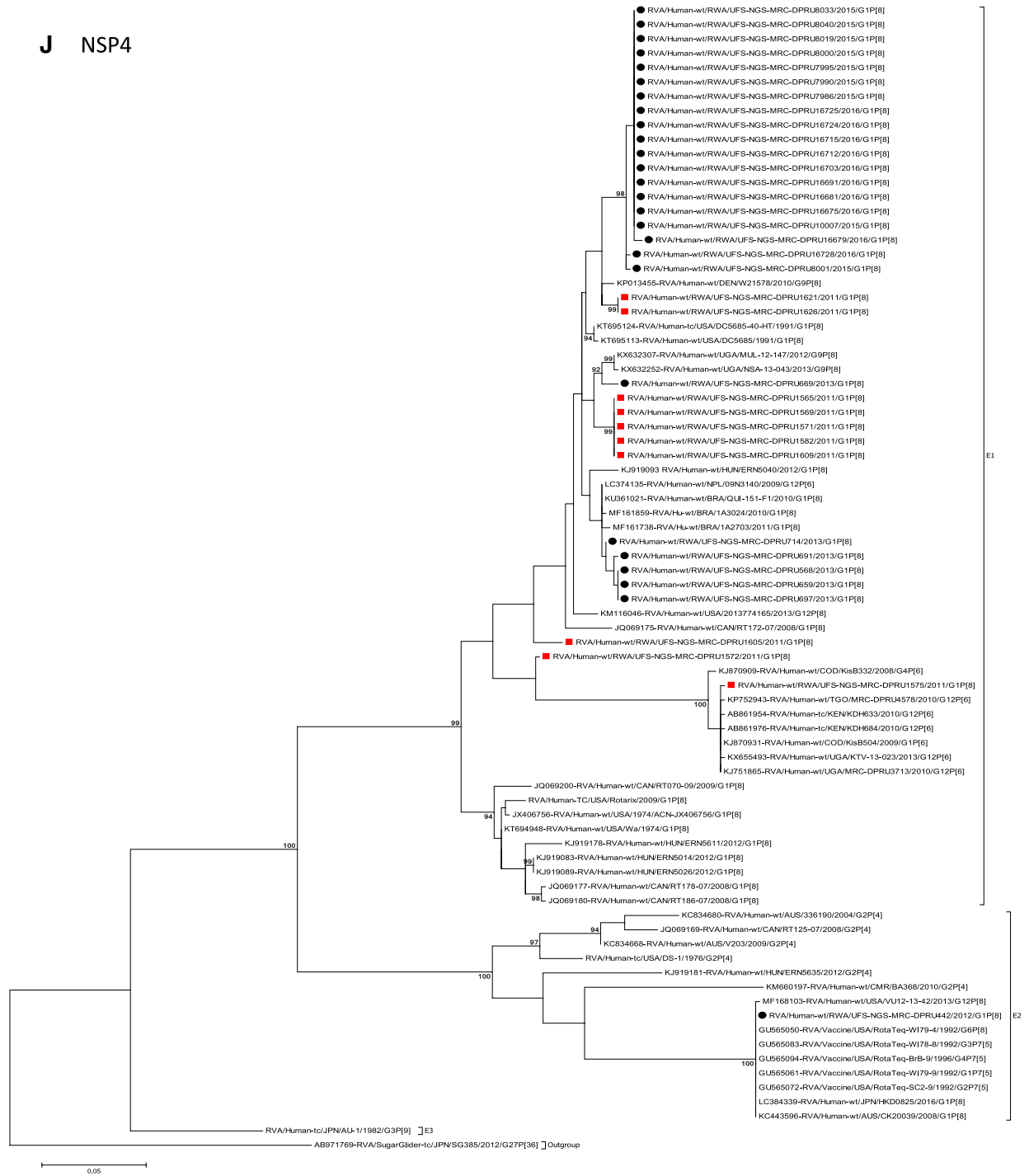


Figure 2. (continued)

a nt (aa) identities in the range of 99.4–100% (99.4–100%) and 99.4–100% (99.4–100%), respectively. The pre-vaccination sub-lineage comprised of five NSP4 strains with absolute gene identities among themselves. Phylogenetically, the NSP5 gene did not resolve and the bootstrap support was not enough to separate the strains into distinct sublineages. On the other hand, a homologous relationship was observed for the VP1, VP2, VP3, NSP1, NSP4 and NSP5 gene segments of strain RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU442/2012/G1P[8] with cognate gene sequences of RVA/Vaccine/USA/Rotateq-W179-4/1992/G6P1A8.

Phylogenetic and sequence analyses of VP6, NSP2, and NSP3. Phylogenetically, the VP6, NSP2 and NSP3 gene segments of each Rwandan G1P[8] strain, with the exception of RotaTeq vdG1P[8] strain, segregated or grouped with cognate gene sequences of Wa-like strains belonging to genotype 1 (I1, N1 and T1, respectively) (Fig. 2C,H,I), while those of the RotaTeq vdG1P[8] clustered together with genotype I2, N2 and

K NSP5

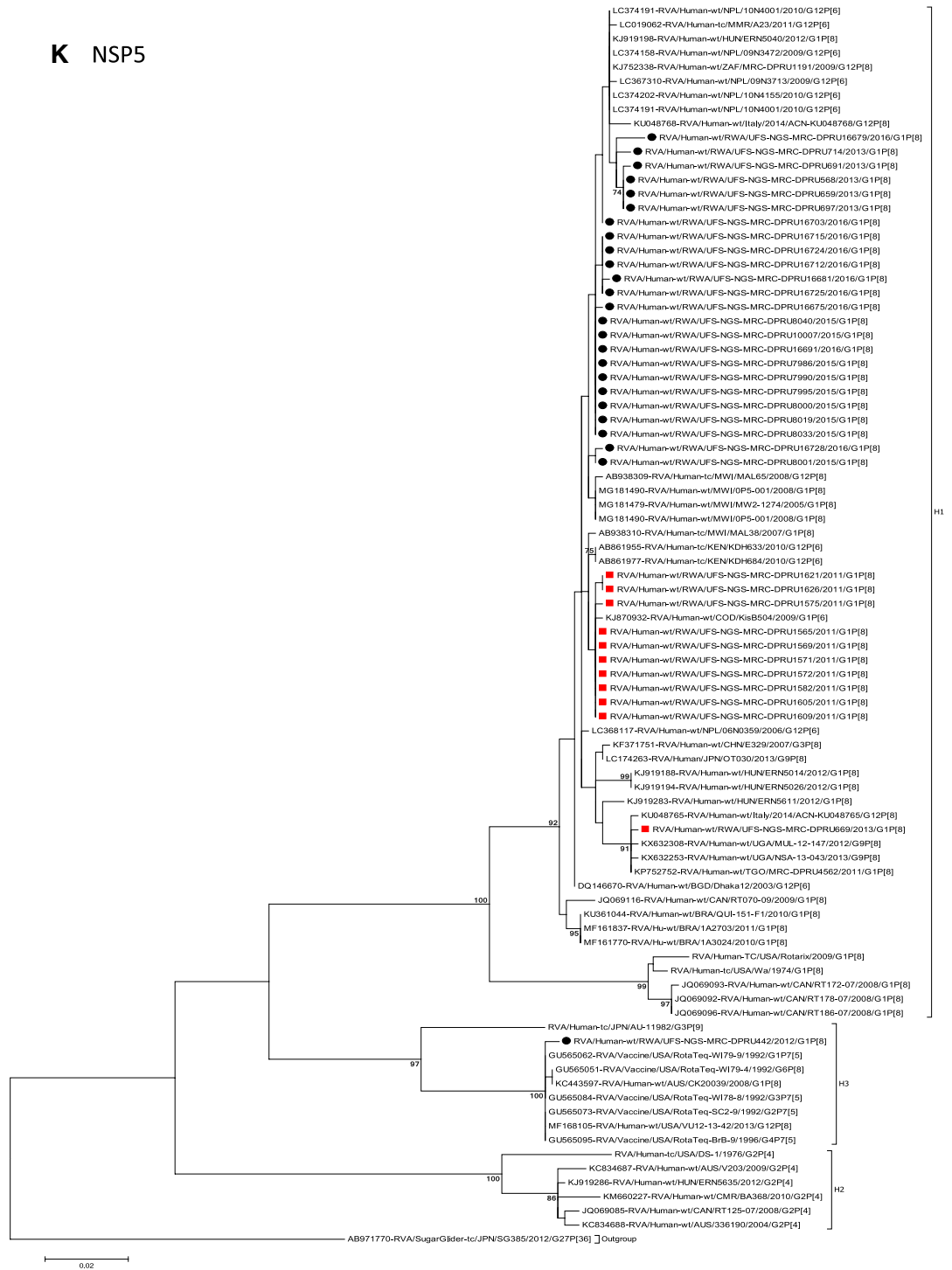


Figure 2. (continued)

T6 strains, respectively. Phylogenetic analyses of the VP6, NSP2 and NSP3 nucleotide sequences of the pre- and post-vaccine introduction G1P[8] strains from Rwandan showed a segregation into typical sub-lineages and mixed sub-lineages (consisting of both pre- and post-vaccine introduction strains) (Fig. 2C,H,I). For the VP6 gene, the study strains separated into a post-vaccine sub-lineage and two mixed sub-lineages each consisting of pre- and post-vaccine introduction VP6 genes (Fig. 2C). The nt (aa) similarities amongst the five VP6 genes in the typical post-vaccine introduction strains sub-lineage ranged from 99.3–100% (99.7–100%). The nine I1 genes that grouped in the first mixed pre- and post-vaccine introduction strains sub-lineage revealed gene identities of 98.9–100% (99.5–100%), while the nt (aa) identities of the 21 strains that belonged to the second mixed pre- and post-vaccine introduction strains sub-lineage was in the ranged of 95.1–100% (93.7–100%). The NSP2 genes segregated into three sub-lineages, a pre-vaccine sub-lineage, a mixed pre- and post-vaccine sub-lineage and a post-

	Lineage	T-cell epitope 16-28 aa																T-cell epitope 40-52 aa											
		16	17	18	19	20	21	22	23	24	25	26	27	28	40	41	42	43	44	45	46	47	48	49	50	51	52		
RVA/Vaccine/USA/Rotarik-A41CB052A/1988/G1P1A[8]	II	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	Y	V	A	L	F	A	L	T	R	A	Q	N		
GU566057-RVA/Vaccine/USA/RotaTeq-W179-9/1992/G1P7[5]	III	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	T	V	A	L	F	A	L	T	R	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU16675/2016/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU16679/2016/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU16681/2016/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU16691/2016/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU16703/2016/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU16712/2016/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU16715/2016/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU16724/2016/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU16725/2016/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU16728/2016/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU10007/2015/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU7986/2015/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU7990/2015/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU7995/2015/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU8000/2015/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU8001/2015/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU8019/2015/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU8033/2015/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU8040/2015/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	M	A	L	F	V	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU442/2012/G1P[8]	III	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	T	V	A	L	F	A	L	T	R	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU568/2013/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU659/2013/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU669/2013/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU691/2013/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU697/2013/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU714/2013/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	L	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU1565/2011/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU1569/2011/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU1571/2011/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU1572/2011/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU1575/2011/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU1582/2011/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU1605/2011/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU1609/2011/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU1621/2011/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	L	A	L	T	K	A	Q	N		
RVA/Human-wt/T GO/MRC-DPRU4562/2011/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	F	V	A	L	F	A	L	T	K	A	Q	N		
RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU1626/2011/G1P[8]	I	I	I	L	L	N	Y	I	L	K	S	V	T	R	I	S	V	A	L	L	A	L	T	K	A	Q	N		

Figure 4. Alignment of antigenic residues in T-cell antigen epitopes of the G1 vaccine component contained in Rotarix and RotaTeq compared to Rwanda G1 wild type strains circulating from 2011–2016 RVA seasons. Amino acid changes are highlighted in sky blue.

NGS-MRC-DPRU1626/2011/G1P[8] and RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU669/2011/G1P[8] that did not cluster in either of the typical or mixed sub-lineages were observed (Fig. 2I). The VP6, NSP2 and NSP3 gene sequences of the single RotaTeq vdG1P[8] strains shared nt (aa) homology of 98.0%, ≥99.7% and 100% with cognate gene sequence of RotaTeq vaccine strains, respectively.

Discussion

The introduction of the rotavirus vaccines has resulted in the reduction of the global burden of RVA associated diarrhoea diseases and hospitalization¹. Hence, continuous RVA surveillance in different settings remains vital to document and characterize post-vaccine licensure era RVA strains, possible vaccine derived strains and intergenogroup reassortant strains as part effort to document impact of rotavirus vaccination. Therefore, in the present study, the whole gene analyses of 35 Rwandan G1P[8] and a single RotaTeq vdG1P[8] strains collected from children less than 5 years of age during pre- and post-vaccination with RotaTeq showed that all 35 of the strains possessed a typical Wa-like genotype constellation, while the single RotaTeq vdG1P[8] strain RVA/Human-wt/RWA/UF S-NGS-MRC-DPRU442/2012/G1P[8], was apparently produced by a reassortment event between vaccine strains, RVA/Vaccine/USA/RotaTeq-W179-4/1992/G6P1A[8] and RVA/Vaccine/USA/Rotateq-W179-9/1992/G1P7[5]. However, the VP4 and VP6 genes may have originated from other circulating vaccine strains as their homologies with vaccine strain RVA/Vaccine/USA/RotaTeq-W179-4/1992/G6P1A[8] and RVA/Vaccine/USA/Rotateq-W179-9/1992/G1P7[5] were lower than other genes. Whole gene analyses confirmed that the VP7 and VP4 genes were vaccine derived and the RotaTeq vaccine WC3 bovine genetic backbone was confirmed to be G1-P[8]-I2-R2-C2-M2-A3-N2-T6-E2-H3. Though reassortment events between RotaTeq vaccine strains which generated vaccine derived G1P[8] reassortant have been reported previously in Australia, South Korea, Finland and USA, this is the first reported RotaTeq vdG1P[8] from the African continent^{34–37}. Previous studies have shown that reassortment may occur more frequently between the VP7 gene of RotaTeq vaccine strain RVA/Vaccine/USA/Rotateq-W179-9/1992/G1P7[5] and VP4 gene of RotaTeq vaccine strain RVA/Vaccine/USA/RotaTeq-W179-4/1992/G6P1A[8] than between other genotypes^{35,36}. The likelihood that this vdG1P[8] may possess increased virulence and cause AGE exists³⁶. The combination of the two human outer capsid proteins could potentially enhance cell binding and entry into enterocytes³⁶, hence an increase in virulence of this vdG1P[8] strain.

Although inter-genogroup reassortant G1P[8] strains have been reported globally, generally the G1P[8] strains are usually associated with the Wa-like backbone^{5,38–41}. Phylogenetically, the 11 gene segments of the Rwanda pre- and post-vaccine introduction G1P[8] strains, with exception of the RotaTeq vdG1P[8], were highly similar and clustered together with cognate gene segments of Wa-like strains (Figs. 1, 2A–K) and further segregated into either typical sub-lineages consisted of only pre-vaccine or post-vaccine introduction strains or mixtures of pre- and post-vaccine introduction strains. The segregation of pre- and post-vaccine introduction G1P[8] strains into different lineages and/or sub-lineages have been previously described in South Africa²⁰. In each gene, we observed segregation of strains in terms of vaccination status and year of detection. With the exception of the NSP1 gene which showed the 2013, 2015 and 2016 post-vaccine strains in 3 different sub-lineages, the

A

	Lineage	VP8*																VP5*																				
		100	146	148	150	188	190	192	193	194	195	196	180	183	113	114	115	116	125	131	132	133	135	87	88	89	384	386	388	393	394	398	440	441	434	459	429	306
RVA/Vaccine/USA/Rotarix-A41CB052A/1988/G1P1A[8]	P[8]-I	D	S	Q	E	S	T	N	L	N	N	I	T	A	N	P	V	D	S	N	D	N	N	T	N	Y	F	I	W	P	G	R	T	P	E	L	R	
RVA/Vaccine/USA/RotaTeq-W179-4/1992/G6P1A[8]	P[8]-II	D	S	Q	E	S	T	N	L	N	D	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU442/2012/G1P[8]	P[8]-III	D	S	Q	E	S	T	N	L	N	D	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU16675/2016/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU16679/2016/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU16681/2016/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	S	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU16691/2016/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU16703/2016/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU16712/2016/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	S	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU16715/2016/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	S	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU16724/2016/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	S	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU16725/2016/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	S	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU16728/2016/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU10007/2015/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU7986/2015/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU7990/2015/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU7995/2015/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU8000/2015/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU8001/2015/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU8019/2015/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU8033/2015/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU8040/2015/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU568/2013/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	D	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU659/2013/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	D	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU669/2013/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	D	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU691/2013/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	D	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU697/2013/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	D	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU714/2013/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	D	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU1565/2011/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU1569/2011/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU1571/2011/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU1572/2011/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU1575/2011/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU1582/2011/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU1605/2011/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU1609/2011/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU1621/2011/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	D	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/RWA/UFS-NGS-MRC-DPRU1626/2011/G1P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	D	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-tc/USA/Wa/1974/G1P[8]	P[8]-I	D	S	Q	E	S	T	N	L	N	N	I	T	A	N	P	V	D	S	S	N	D	N	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-tc/JPN/IT0/1981/G3P[8]	P[8]-I	D	S	Q	E	S	T	N	L	N	H	H	I	T	A	N	P	V	D	S	S	N	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/JPN/IF45/1987/G9P[8]	P[8]-II	D	S	Q	E	S	T	N	L	N	N	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/ZAF/MRC-DPRU76/2012/G12P[8]	P[8]-III	D	S	Q	D	S	T	N	L	N	G	I	T	A	N	P	V	D	N	R	N	D	D	N	T	N	Y	F	L	W	P	G	R	T	P	E	L	R
RVA/Human-wt/MWI/OP354/1998/G4P[8]	P[8]-IV	D	S	Q	E	S	T	D	L	S	I	T	A	D	P	V	D	S	R	N	D	N	N	T	N	Y	F	L	W	P	G	R	T	P	D	L	R	

B

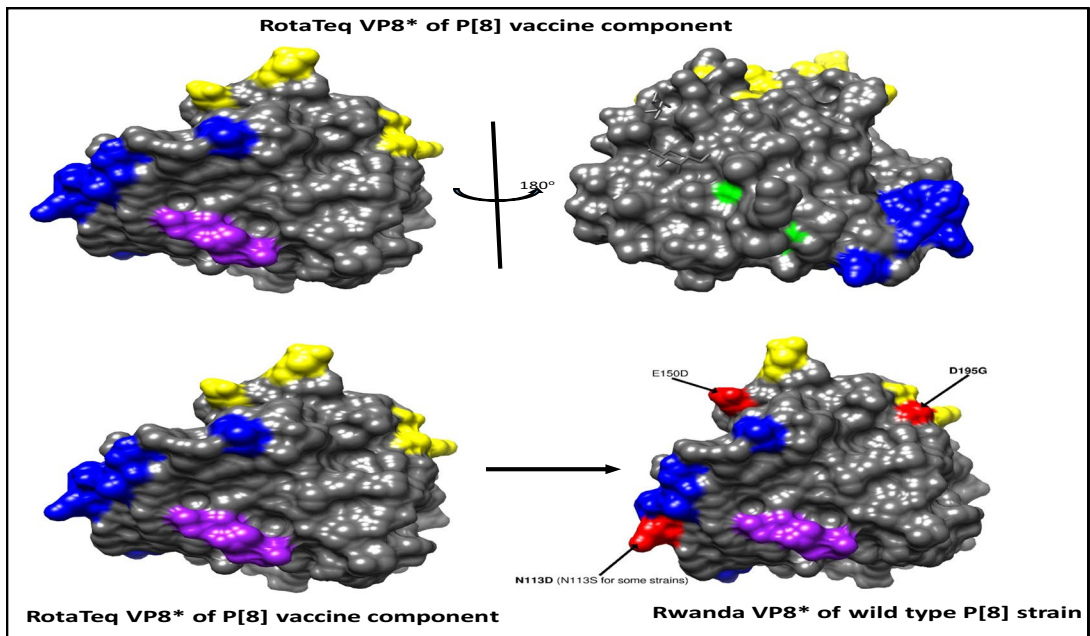


Figure 5. (A) The alignment of the P[8] component of Rotarix and RotaTeq vaccines and wild type P[8] strains circulating in Rwanda from 2011–2016 RVA seasons, based on the antigenic residues in VP4. VP4 is divided into two regions VP8* (antigenic epitope 8-1, 8-2, 8-3, and 8-4) and VP5* (antigenic epitope 5-1, 5-2, 5-3, 5-4, and 5-5). The black dot indicates changes in the residues associated with escape neutralization with monoclonal antibodies. Amino acids differences between Rotarix and RotaTeq are shown at positions 195 (antigenic epitope 8-1), 125, 131 and 135 (antigenic epitope 8-3) and are indicated in boldface. Study strains amino acid residues highlighted in green and brown are residues that are different from Rotarix and RotaTeq, respectively. Also, amino acids differences between study strains and Rotarix and RotaTeq are highlighted in sky blue. Post-vaccine and pre-vaccine G1P[8] study strains are bolded in black and red, respectively. (B) Location of surface-exposed amino acid differences between VP8* protein of RotaTeq P[8] vaccine component versus a wild-type P[8] strain from Rwanda (indicated in red). Antigenic epitopes in (B) are colored in yellow (8-1), green (8-2), blue (8-3) and purple (8-4).

post-vaccine introduction G1P[8] strains detected during the 2015–2016 RVA seasons always clustered together within the post-vaccine introduction sublineages. The VP1–VP3, NSP1, NSP4 and NSP5 gene sequences of the pre-vaccine introduction strains always clustered together. Analyses of the VP7, VP4, VP6, NSP2 and NSP3 gene sequences, however, showed mixed sub-lineages consisting of both the pre-vaccine strains detected in the 2011 RVA season and post-vaccine introduction strains detected in the 2013 (VP7, VP4 and NSP3), and 2015–2016 (VP6 and NSP2) seasons. In addition, within the typical pre- and post-vaccine introduction sub-lineages of all 11 gene segments, the sequence similarity was extremely high $\geq 99.9\%$ and between sub-lineages the nt (aa) similarities were moderately high in the range of 96–98% (98.1–99.2%). The high genetic relationship between gene sequences of the Rwandan pre- and post-vaccine introduction G1P[8] strains is an indication that the same strains that were circulating before vaccine introduction in Rwanda, are the same strains or their progeny that are causing RVA-associated diarrheal diseases and hospitalization after vaccination.

Globally, at least seven lineages have been previously described for G1 strains collected from different geographical locations^{25,30,42,43}. The emergence of distinct lineages or sub-lineages is attributed to the diverse evolutionary mechanisms such as mutation, recombination and reassortments⁷. Based on the VP7 gene, the 35 Rwandan G1 strains which consisted of both pre- and post-vaccine introduction strains, clustered in lineage I and were comparable to observation reported previously^{20, 22, 44}. On the other hand, the VP4 plays a role as an antibody-neutralization protein^{45,46}. Diversity of the P[8] genotypes has been demonstrated through four described distinct lineages^{30,47}. The clustering of the 35 P[8] strains in lineage III in this study is consistent with previous observations^{19, 20, 22, 44}. However, despite the clustering of all VP7 and VP4 genes of the study strains in Lineage I (Fig. 2A) and lineage III (Fig. 2B), respectively, antigenic variation seen at the neutralization epitope sites of the VP7 and VP4 proteins of the study strains have been previously described in Belgium⁴⁸ and Australia⁴⁹. Amino acid substitutions with effect on polarity changes were detected on Rwandan VP7 neutralization epitopes at positions T91A/V, S190D and M217. Amino acid substitutions on positions 94, 96, 147, 148, 190, 208, 211, 213 and 217 are critical and have been reported to alter rotavirus antigenicity^{50,51}. Therefore, the detected substitutions may be involved in antigenic drift in Rwandan G1 strains. The resultant alteration in charge and polarity observed at position T91A/V on the VP7 neutralization epitope may play a role in escape of host immunity as it was distinctively observed post vaccine introduction^{24,48}. Furthermore, the change in polarity at position S190D and M217 could be due to natural fluctuation of RVA as they were observed during the pre-vaccine era and consistently before and after vaccine introduction, respectively³⁸. The presence of amino acid differences in cytotoxic T lymphocyte epitope positions Y41F/S, V42M and A46V in Rwandan G1 strains possibly can result in greater host-immunity escape effects²⁸. In the intestine, trypsin like proteases cleave the VP4 spike protein into two structural domains (VP8* and VP5*)⁴. Four surface-exposed antigenic epitopes (8-1 to 8-4) have been described in the VP8* region, while five antigenic epitopes (5-1 to 5-5) in the VP5* region have been documented⁴⁸. The amino acid changes identified at positions N113D, S131R and N135D in this study, resulted in polarity changes and play a role in escape of host-immunity^{32,33}.

Overall, full genomic analyses of 35 Rwandan G1P[8] strains revealed the predominance of G1-I and P[8]-III, which is consistent with what has been reported previously in Belgium⁴⁸. The global prevalence of Wa-like human strains is hypothesized to be due to the ease of dissemination of this genetic backbone in the human host^{21,52}. Although this study was insightful in reporting the whole gene composition of the circulating G1P[8] strains in Rwanda, it was limited to a few years (just 1 year) pre- and (5 years) post-vaccination samples were included. Another limitation was that, for the sample where RotaTeq vaccine shedding was observed, other plausible alternate aetiologies such as adenovirus and/or norovirus that might have led to the diarrhoea symptoms and possible hospitalization, were not evaluated. Cell culture was not used to confirm the reassortment events as the project was mainly focused on in-silico work.

In conclusion, this is the first study to describe full genomic analyses of G1[P8] RVA strains in Rwanda. The detection of RotaTeq vdG1P[8] strain from RVA positive child hospitalized with AGE symptoms was unexpected. RVA strain surveillance at whole gene level will enhance further assessment of vaccine impact on circulating RVA strains, the frequency of reassortment events under natural conditions and epidemiological fitness of RVA generated by such events.

Methods

Sample collection. RVA positive fecal samples (n = 158) were obtained from children less than 5 years old who were hospitalized with AGE in Rwanda as part of the ongoing WHO AFRO RVA surveillance program. The samples were conventionally genotyped into G and P types at the South African Medical Research Council-Diarrhoeal Pathogens Research Unit (SAMRC/DPRC), a WHO Rotavirus Reference Laboratory in South Africa (WHO RRL-SA). The samples were collected during the pre- (2011–2012) and post- (2012–2016) RVA vaccination periods with RotaTeq vaccine in Rwanda. Collectively, all the samples were stored at $-20\text{ }^{\circ}\text{C}$ at SAMRC/DPRU and G1P[8] strains were selected for whole genome sequencing at the University of the Free State-Next Generation Sequencing (UFS-NGS) Unit.

Double-stranded RNA extraction and purification. The extraction of dsRNA was performed at the MRC-DPRU and involved a method previously described by Nyaga et al. (2018)⁵³. Briefly, a 100 mg stool sample was added to 200 μL freshly made phosphate buffered saline (PBS) (Sigma-Aldrich, Germany). A 900 μL volume of TRI-REAGENT-LS (Molecular Research Center, Inc, Cincinnati, OH, USA) was added to the suspended stool sample to homogenize as well as lyse the cells and cell components. A volume of 300 μL chloroform (Sigma-Aldrich, Germany) was then added and centrifugation (13,000 rpm for 20 min at $4\text{ }^{\circ}\text{C}$) was done in a temperature-controlled microcentrifuge (Eppendorf centrifuge 5427R, Hamburg, Germany). The supernatant containing the total RNA was precipitated by addition of 700 μL isopropanol (Sigma-Aldrich, Germany) and

by centrifugation at 16,000×g for 30 min at room temperature. The resulting pellet was re-dissolved by addition of 90 µL of ddH₂O (Merck KGaA, Germany). A concentration of 8 M LiCl₂ (Sigma, St. Louis, MO, USA) was used to remove ssRNA through precipitation for 16 h after which further centrifugation was done for 30 min at 16,000×g. The extracted dsRNA was purified by utilizing the MinElute gel extraction kit (Qiagen, Hilden, Germany) and the integrity and enrichment of the dsRNA was verified via agarose gel electrophoresis.

cDNA synthesis. Complementary DNA (cDNA) was generated from the extracted viral RNA utilizing Maxima H Minus Double-Stranded cDNA Synthesis Kit with minor modifications (Thermo Fischer Scientific, Waltham, MA, USA). Briefly, the extracted total RNA was denatured at 95 °C for five minutes and then 1 µL random hexamer primers were added. The hexamer primers were allowed to anneal at 65 °C for 5 min. A volume of 5 µL of First Strand Reaction Mix and 1 µL of First Strand Enzyme Mix was then added. The solution was then incubated at 25 °C, 50 °C and 85 °C for 10, 120 and 5 min, respectively. The tubes were removed from the thermocycler and second-strand synthesis was performed by adding 55 µL nuclease-free water, followed by addition of 20 µL of 5× Second Strand Reaction Mix and 5 µL of Second Strand Enzyme Mix. Subsequently, the solution was incubated at 16 °C for one hour and the reaction was stopped with 6 µL 0.5 M EDTA. Residual RNA was removed with 10 µL RNase I and the synthesized cDNA was incubated at room temperature for 5 min.

DNA library preparations and whole genome sequencing. DNA libraries were prepared using the Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA, USA) following the manufacturer's instructions. Briefly, DNA library preparation entailed fragmentation of the generated DNA, indexing using unique barcodes and amplification of tagged DNA and clean-up of the amplified DNA. The library quality and size was assessed using an Agilent 2100 BioAnalyzer (Agilent Technologies, Waldbronn, Germany) according to the manufacturer's specified protocol. The Illumina custom protocol was utilized to normalize the libraries to 4 nM. All the normalized libraries were then pooled together into a single tube by combining 5 µL of each bar-coded library. The pooled libraries were subjected to chemical denaturation using 0.2 N sodium hydroxide. After denaturation, 990 µL of pre-chilled hybridization buffer HT1 (Illumina) was added to the 10 µL of the 4 nM denatured DNA library to dilute to 20 pM. A further dilution of the denatured library was performed to get the desired final concentration of 8 pM. A PhiX control spike-in of 20% was used. Whole genome sequencing was performed for 600 cycles (301×2 paired-end) on a MiSeq benchtop sequencer (Illumina, San Diego, California, USA) using Illumina V3 reagent kit at the UFS-NGS Unit, Bloemfontein, South Africa.

Genome assembly. Illumina sequence reads were analyzed using Geneious software v11 (<https://www.geneious.com>)⁵⁴ and CLC Genomics Workbench v11 (CLC Bio, Qiagen) which entailed genome assembly and mapping the reads to reference-based sequences to obtain the full-length genomes.

Identification of genotype constellations. Genotyping was performed by utilizing RotaC, v 2.0⁵⁵, an automated online genotyping tool for group A RV strains. Genome constellations were generated by assigning genotypes to each genome segments.

GenBank accession numbers. The sequences were deposited into GenBank under the accession number MN632673-MN633067.

Phylogenetic, sequence analyses and protein modelling. For each gene segment, the ORF were aligned and subjected to sequence comparisons as described previously^{56–58}. Briefly, multiple alignments were made using the MUSCLE algorithm implemented in MEGA 6 software (<https://www.megasoftware.net/>)⁵⁹. Once aligned, the DNA Model Test program implemented in MEGA version 6 was used to identify the optimal evolutionary models that best fit the sequence datasets. Using the Corrected Akaike Information Criterion (AICc), the model GTR-G-I was found to best fit the sequence data for each gene segment. With this model, maximum likelihood trees were constructed using MEGA 6 with 1000 bootstrap replicates to estimate branch support. The nucleotide and deduced amino acid sequence identities among strains were calculated for each gene using distance matrices prepared using the p-distance algorithm in MEGA 6 software⁵⁹. Protein modelling was performed on amino acid sequences for each strain and reference strain using the Swiss-Model protein structure homology-modelling server (<https://swissmodel.expasy.org>)⁶⁰. The structures were modified using UCSF Chimera⁶¹.

Ethics. The protocol used in this study was reviewed and approved by the Health Sciences Research Ethics Committee (HSREC) of the University of the Free State under the following clearance number: UFS-HSD2019/1601/2810.

Disclaimer. The findings and conclusions in this report are those of the authors and do not represent the official position of the World Health Organization and of the Centers for Disease Control and Prevention. Names of specific vendors, manufacturers, or products are included for public health and informational purposes; inclusion does not imply endorsement of the vendors, manufacturers, or products by the World Health Organization, Centers for Disease Control and Prevention or the US Department of Health and Human Services^{38,62–64}.

Received: 9 March 2020; Accepted: 17 July 2020

Published online: 10 August 2020

References

1. Troeger, C. *et al.* Rotavirus vaccination and the global burden of rotavirus diarrhea among children younger than 5 years. *JAMA Pediatr.* **172**(10), 958–965 (2018).
2. Ngabo, F. *et al.* Effect of pentavalent rotavirus vaccine introduction on hospital admissions for diarrhoea and rotavirus in children in Rwanda: a time-series analysis. *Lancet Glob. Health* **4**(2), e129–e136 (2016).
3. Vesikari, T. Success of rotavirus vaccination in Africa: good news and remaining questions. *Lancet Glob. Health* **4**(2), e76–e77 (2016).
4. Estes, M. K. & Greenberg, H. B. Rotaviruses. In *Fields Virology* (eds Knipe, D. M. & Howley, P. M.) (Wolters Kluwer Health/Lippincott, Williams and Wilkins, Philadelphia, 2013).
5. Matthijnssens, J. *et al.* Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). *Adv. Virol.* **156**(8), 1397–1413 (2011).
6. Dóro, R., Farkas, S. L., Martella, V. & Banyai, K. Zoonotic transmission of rotavirus: surveillance and control. *Expert Rev. Anti-infect. Ther.* **13**(11), 1337–1350 (2015).
7. Kirkwood, C. D. Genetic and antigenic diversity of human rotaviruses: potential impact on vaccination programs. *J. Infect. Dis.* **202**, S43–S48 (2010).
8. Banyai, K. *et al.* Sequencing and phylogenetic analysis of the coding region of six common rotavirus strains: evidence for intragroup reassortment among co-circulating G1P [8] and G2P [4] strains from the United States. *J. Med. Virol.* **83**(3), 532–539 (2011).
9. Banyai, K. *et al.* 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**, A122–A130 (2012).
10. Gentsch, J. R. *et al.* Serotype diversity and reassortment between human and animal rotavirus strains: implications for rotavirus vaccine programs. *J. Infect. Dis.* **192**, S146–S159 (2005).
11. Matthijnssens, J. *et al.* Full genome-based classification of rotaviruses reveals a common origin between human Wa-Like and porcine rotavirus strains and human DS-1-like and bovine rotavirus strains. *J. Virol.* **82**(7), 3204–3219 (2008).
12. Dóro, R. *et al.* Review of global rotavirus strain prevalence data from six years post vaccine licensure surveillance: is there evidence of strain selection from vaccine pressure? *Infect. Genet. Evol.* **28**, 446–461 (2014).
13. Kirkwood, C. D. & Steele, A. D. Rotavirus Vaccines in China. *JAMA Netw. Open* **1**(4), e181579. <https://doi.org/10.1001/jamanetworkopen.2018.1579> (2018).
14. World Health Organization. Rotavirus Vaccination. Meeting of the immunization Strategic Advisory Group of Experts. *Wkly. Epidemiol. Rec.* **84**(23), 220–236 (2009).
15. Seheri, L. M. *et al.* Rotavirus strain diversity in Eastern and Southern African countries before and after vaccine introduction. *Vaccine* **36**(47), 7222–7230 (2018).
16. Dennehy, P. H. Rotavirus vaccines: an overview. *Clin. Microbiol. Rev.* **21**(1), 198–208 (2008).
17. Heaton, P. M. & Ciarlet, M. The pentavalent rotavirus vaccine: discovery to licensure and beyond. *Clin. Infect. Dis.* **45**(12), 1618–1624 (2007).
18. Ngabo, F. *et al.* Can routinely collected national data on childhood morbidity and mortality from diarrhea be used to monitor health impact of rotavirus vaccination in Africa? Examination of pre-vaccine baseline data from Rwanda. *Pediatr. Infect. Dis. J.* **33**, S89–S93 (2014).
19. Arora, R. & Chitambar, S. D. Full genomic analysis of Indian G1P [8] rotavirus strains. *Infect. Genet. Evol.* **11**(2), 504–511 (2011).
20. Magagula, N. B. *et al.* Whole genome analyses of G1P [8] rotavirus strains from vaccinated and non-vaccinated South African children presenting with diarrhea. *J. Med. Virol.* **87**(1), 79–101 (2015).
21. Rahman, M. *et al.* Complete genomic analysis of a Bangladeshi G1P [8] rotavirus strain detected in 2003 reveals a close evolutionary relationship with contemporary human Wa-like strains. *Infect. Genet. Evol.* **10**(6), 746–754 (2010).
22. Santos, F. S. *et al.* G1P [8] Rotavirus in children with severe diarrhea in the post-vaccine introduction era in Brazil: Evidence of reassortments and structural modifications of the antigenic VP7 and VP4 regions. *Infect. Genet. Evol.* **69**, 255–266 (2019).
23. Shintani, T. *et al.* Whole genomic analysis of human G1P[8] rotavirus strains from different age groups in China. *Viruses* **4**, 1289–1304 (2012).
24. Zeller, M. *et al.* Genome-wide evolutionary analyses of G1P[8] strains isolated before and after rotavirus vaccine introduction. *Genome Biol Evol.* **7**, 2473–2483 (2015).
25. Le, V. P. *et al.* Genetic variation of prevalent G1P [8] human rotaviruses in South Korea. *J. Med. Virol.* **82**(5), 886–896 (2010).
26. Aoki, S. T. *et al.* Structure of rotavirus outer-layer protein VP7 bound with a neutralizing Fab. *Science* **324**(5933), 1444–1447 (2009).
27. Franco, M. A., Tin, C. & Greenberg, H. B. CD8+ T cells can mediate almost complete short-term and partial long-term immunity to rotavirus in mice. *J. Virol.* **71**(5), 4165–4170 (1997).
28. Wei, J. *et al.* A naturally processed epitope on rotavirus VP7 glycoprotein recognized by HLA-A2. 1-restricted cytotoxic CD8+ T cells. *Viral Immunol.* **22**(3), 189–194 (2009).
29. Morozova, O. V., Sashina, T. A., Fomina, S. G. & Novikova, N. A. Comparative characteristics of the VP7 and VP4 antigenic epitopes of the rotaviruses circulating in Russia (Nizhny Novgorod) and the Rotarix and RotaTeq vaccines. *Adv. Virol.* **160**(7), 1693–1703 (2015).
30. Arista, S. Heterogeneity and temporal dynamics of evolution of G1 human rotaviruses in a settled population. *J. Virol.* **80**(21), 10724–10733 (2006).
31. Dormitzer, P. R., Sun, Z. Y. J., Wagner, G. & Harrison, S. C. The rhesus rotavirus VP4 sialic acid binding domain has a galectin fold with a novel carbohydrate binding site. *EMBO J.* **21**(5), 885–897 (2002).
32. Dormitzer, P. R., Nason, E. B., Prasad, B. V. & Harrison, S. C. Structural rearrangements in the membrane penetration protein of a non-enveloped virus. *Nature* **430**(7003), 1053 (2004).
33. Monnier, N. *et al.* High-resolution molecular and antigen structure of the VP8* core of a sialic acid-independent human rotavirus strain. *J. Virol.* **80**(3), 1513–1523 (2006).
34. Markkula, J., Hemming, M. & Vesikari, T. Detection of vaccine-derived rotavirus strains in nonimmunocompromised children up to 3–6 months after RotaTeq vaccination. *Pediatr. Infect. Dis. J.* **34**(3), 296–298 (2015).
35. Than, V. T., Jeong, S. & Kim, W. Characterization of RotaTeq vaccine-derived rotaviruses in South Korean infants with rotavirus gastroenteritis. *J. Med. Virol.* **87**(1), 112–116 (2015).
36. Donato, C. M. *et al.* Identification of strains of rotateq rotavirus vaccine in infants with gastroenteritis following routine vaccination. *J. Infect. Dis.* **206**(3), 377–383. <https://doi.org/10.1093/infdis/jis361> (2012).
37. Payne, D. C. *et al.* Sibling transmission of vaccine-derived rotavirus (RotaTeq) associated with rotavirus gastroenteritis. *Pediatrics* **125**(2), e438–e441. <https://doi.org/10.1542/peds.2009-1901> (2010).
38. Jere, K. C. *et al.* Emergence of double- and triple-gene reassortant G1P [8] rotaviruses possessing a DS-1-like backbone after rotavirus vaccine introduction in Malawi. *J. Virol.* **92**(3), e01246–e1317 (2018).
39. Fujii, Y. *et al.* 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 (2014).
40. Kuzuya, M. *et al.* 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 (2014).

41. Luchs, A. *et al.* First Detection of DS-1-like G1P[8] Double-gene Reassortant Rotavirus Strains on The American Continent, Brazil, 2013. *Sci. Rep.* **9**, 2210 (2019).
42. Bányai, K. *et al.* Trends in the epidemiology of human G1P [8] rotaviruses: a Hungarian study. *J. Infect. Dis.* **200**, S222–S227 (2009).
43. Bucardo, F., Rippinger, C. M., Svensson, L. & Patton, J. T. Vaccine-derived NSP2 segment in rotaviruses from vaccinated children with gastroenteritis in Nicaragua. *Infect. Genet. Evol.* **12**(6), 1282–1294 (2012).
44. da Silva, M. F. M. *et al.* G1P [8] species A rotavirus over 27 years—pre-and post-vaccination eras—in Brazil: full genomic constellation analysis and no evidence for selection pressure by Rotarix vaccine. *Infect. Genet. Evol.* **30**, 206–218 (2015).
45. Ofit, P. A., Clark, H. F., Blavat, G. & Greenberg, H. B. Reassortant rotaviruses containing structural proteins vp3 and vp7 from different parents induce antibodies protective against each parental serotype. *J. Virol.* **60**(2), 491–496 (1986).
46. Ward, R. L., McNeal, M. M., Sander, D. S., Greenberg, H. B. & Bernstein, D. I. Immunodominance of the VP4 neutralization protein of rotavirus in protective natural infections of young children. *J. Virol.* **67**(1), 464–468 (1993).
47. Cho, M. K. *et al.* Full genomic analysis of a human rotavirus G1P [8] strain isolated in South Korea. *J. Med. Virol.* **85**(1), 157–170 (2013).
48. Zeller, M. *et al.* Genetic analyses reveal differences in the VP7 and VP4 antigenic epitopes between human rotaviruses circulating in Belgium and rotaviruses in Rotarix and RotaTeq. *J. Clin. Microbiol.* **50**(3), 966–976 (2012).
49. Diwakarla, C. S. & Palombo, E. A. Genetic and antigenic variation of capsid protein VP7 of serotype G1 human rotavirus isolates. *J. Gen. Virol.* **80**(2), 341–344 (1999).
50. Ahmed, K., Nakagomi, T. & Nakagomi, O. Molecular identification of a novel G1 VP7 gene carried by a human rotavirus with a super-short RNA pattern. *Virus Genes* **35**(2), 141–145 (2007).
51. Trinh, Q. D. *et al.* Sequence analysis of the VP7 gene of human rotavirus G1 isolated in Japan, China, Thailand, and Vietnam in the context of changing distribution of rotavirus G-types. *J. Med. Virol.* **79**(7), 1009–1016 (2007).
52. Ghosh, S. & Kobayashi, N. Whole-genomic analysis of rotavirus strains: current status and future prospects. *Future Microbiol.* **6**(9), 1049–1065 (2011).
53. Nyaga, M. M. *et al.* Whole-genome sequencing and analyses identify high genetic heterogeneity, diversity and endemicity of rotavirus genotype P [6] strains circulating in Africa. *Infect. Genet. Evol.* **63**, 79–88 (2018).
54. Kearse, M. *et al.* Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**(12), 1647–1649 (2012).
55. Maes, P., Matthijssens, J., Rahman, M. & Van Ranst, M. RotaC: a web-based tool for the complete genome classification of group A rotaviruses. *BMC Microbiol.* **9**(1), 238 (2009).
56. Esona, M. D. *et al.* Molecular characterization of a human G20P [28] rotavirus a strain with multiple genes related to bat rotaviruses. *Infect. Genet. Evol.* **57**, 166–170 (2018).
57. Esona, M. D. *et al.* Characterization of a triple-recombinant, reassortant rotavirus strain from the Dominican Republic. *J. Gen. Virol.* **98**(2), 134 (2017).
58. Ward, M. L. Molecular characterization of the first G24P [14] rotavirus strain detected in humans. *Infect. Genet. Evol.* **43**, 338–342 (2016).
59. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**(12), 2725–2729 (2013).
60. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**(W1), W296–W303 (2018).
61. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**(13), 1605–1612. <https://doi.org/10.1002/jcc.20084> (2004).
62. Anderson, E. J. Rotavirus vaccines: viral shedding and risk of transmission. *Lancet Infect. Dis.* **8**(10), 642–649. [https://doi.org/10.1016/S1473-3099\(08\)70231-7](https://doi.org/10.1016/S1473-3099(08)70231-7) (2008).
63. Hsieh, Y. C. *et al.* Comparison of virus shedding after lived attenuated and pentavalent reassortant rotavirus vaccine. *Vaccine* **32**(10), 1199–1204 (2014).
64. Smith, C. K., McNeal, M. M., Meyer, N. R., Haase, S. & Dekker, C. L. Rotavirus shedding in premature infants following first immunization. *Vaccine* **29**(45), 8141–8146 (2011).

Acknowledgements

The principal financial assistance of the WHO reference 2017/757922-0 is hereby acknowledged. The project was also funded in part by grants awarded to M.M.N. by the South Africa Medical Research Council (Self-Initiated Research grant), the Bill and Melinda Foundation (African Enteric Viruses Genome Initiative grant number BMGF OPP1180423 2017); the National Research Foundation (competitive programme for rated researchers, grant number 120814) and Poliomyelitis Research Foundation (grant number 19/16). The views and opinions expressed in this article are not those of the funding and affiliated institutions but are of the authors of this publication.

Author contributions

S.R., P.N.M., M.M. N.B.M., K.R. M.T.M. and M.M.N. performed laboratory experiments. S.R., P.N.M., M.T.M., J.J., M.D.E., and M.M.N. analysed and interpreted the data. S.R., P.N.M., M.M., S.S., N.B.M., K.R., J.U., L.M., N.M., D.M., L.T., M.D.E., M.D.B., M.J.M., M.L.S., J.M.M., M.T.M. and M.M.N. contributed to study design manuscript writing. All authors approved the final version of the manuscript. M.M.N. supervised the research.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.M.N.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020