pathogens

MDPI

Article

Whole Genome Characterization and Evolutionary Analysis of G1P[8] Rotavirus A Strains during the Pre- and Post-Vaccine Periods in Mozambique (2012–2017)

Benilde Munlela [1,2,*,†], Eva D. João [1,3,*,†], Celeste M. Donato [4,5,6], Amy Strydom [7], Simone S. Boene [1,2], Assucênio Chissaque [1,3], Adilson F. L. Bauhofer [1,3], Jerónimo Langa [1], Marta Cassocera [1,3], Idalécia Cossa-Moiane [1,8], Jorfélia J. Chilaúle [1], Hester G. O'Neill [7] and Nilsa de Deus [1,9]

[1] Instituto Nacional de Saúde (INS), Distrito de Marracuene, Maputo 3943, Mozambique; simone.boene@ins.gov.mz (S.S.B.); assucenio.chissaque@ins.gov.mz (A.C.); adilson.bauhofer@ins.gov.mz (A.F.L.B.); Jeronimo.Langa@ins.gov.mz (J.L.); marta.Cassocera@ins.gov.mz (M.C.); idalecia.moiane@ins.gov.mz (I.C.-M.); jorfelia.chilaule@ins.gov.mz (J.J.C.); nilsa.dedeus@ins.gov.mz (N.d.D.)
[2] Centro de Biotecnologia, Universidade Eduardo Mondlane, Maputo 3453, Mozambique
[3] Instituto de Higiene e Medicina Tropical (IHMT), Universidade Nova de Lisboa, UNL, Rua da Junqueira 100, 1349-008 Lisbon, Portugal
[4] Enteric Diseases Group, Murdoch Children's Research Institute, 50 Flemington Road, Parkville, Melbourne 3052, Australia; celeste.donato@mcri.edu.au
[5] Department of Paediatrics, the University of Melbourne, Parkville 3010, Australia
[6] Biomedicine Discovery Institute and Department of Microbiology, Monash University, Clayton 3800, Australia
[7] Department of Microbial, Biochemical and Food Biotechnology, University of the Free State, 205 Nelson Mandela Avenue, Bloemfontein 9301, South Africa; strydoma@ufs.ac.za (A.S.); oneillhg@ufs.ac.za (H.G.O.)
[8] Institute of Tropical Medicine (ITM), Kronenburgstraat 43, 2000 Antwerp, Belgium
[9] Departamento de Ciências Biológicas, Universidade Eduardo Mondlane, Maputo 3453, Mozambique
[*] Correspondence: benilde.munlela@ins.gov.mz or benildeantnio@gmail.com (B.M.); evadora1@hotmail.com (E.D.J.); Tel.: +258-848814087 (B.M.); +258-827479229 (E.D.J.)
[†] These authors contributed equally in this work.

check for updates

Abstract: Mozambique introduced the Rotarix® vaccine (GSK Biologicals, Rixensart, Belgium) into the National Immunization Program in September 2015. Although G1P[8] was one of the most prevalent genotypes between 2012 and 2017 in Mozambique, no complete genomes had been sequenced to date. Here we report whole genome sequence analysis for 36 G1P[8] strains using an Illumina MiSeq platform. All strains exhibited a Wa-like genetic backbone (G1-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1). Phylogenetic analysis showed that most of the Mozambican strains clustered closely together in a conserved clade for the entire genome. No distinct clustering for pre- and post-vaccine strains were observed. These findings may suggest no selective pressure by the introduction of the Rotarix® vaccine in 2015. Two strains (HJM1646 and HGM0544) showed varied clustering for the entire genome, suggesting reassortment, whereas a further strain obtained from a rural area (MAN0033) clustered separately for all gene segments. Bayesian analysis for the VP7 and VP4 encoding gene segments supported the phylogenetic analysis and indicated a possible introduction from India around 2011.7 and 2013.0 for the main Mozambican clade. Continued monitoring of rotavirus strains in the post-vaccine period is required to fully understand the impact of vaccine introduction on the diversity and evolution of rotavirus strains.

1. Introduction

Rotavirus is one of the leading causes of diarrheal disease in children under five years of age [1,2]. Worldwide, the number of deaths due to rotavirus infection in children under five years of age in 2016 was estimated to be 128,500, of which 104,733 occurred in sub-Saharan Africa [2]. Rotavirus is a member of the *Reoviridae* family. The genome is comprised of 11 double-stranded ribonucleic acid (dsRNA) segments. The mature virus has an icosahedral capsid formed by three concentric protein layers. The 11 segments of the rotavirus genome encode 12 viral proteins: 6 structural proteins VP1-VP4, VP6 and VP7and 6 non-structural proteins (NSP1-NSP6) [3–6].

The gene segments encoding the external capsid proteins, VP7 and VP4, are used in a binary classification system defining G and P genotypes, respectively [5,7]. Currently, 36 G and 51 P genotypes have been described in humans and various animal species [7–10]. At least 73 combinations of human rotavirus group A (RVA) G/P genotypes have been described, of which the most common combinations are G1P[8], G2P[4], G3P[8], G4P[8], G9P[8] and G12P[8] [10,11]. However, the implementation of whole genome sequencing has led to comprehensive sequence-based classification of all RVA genes into genotypes, which are identified and differentiated according to particular cut-off values of nucleotide sequence identities [9,11]. Currently, 26 I (VP6), 22 R (VP1), 20 C (VP2), 20 M (VP3), 31 A (NSP1), 22 N (NSP2), 22 T (NSP3), 27 E (NSP4) and 22 H (NSP5) genotypes have been described [8]. The whole genome constellation of a strain can be described following the nomenclature Gx-P[x]-Ix-Rx-Cx-Mx-Ax-Nx-Tx-Ex-Hx. Two major genotype constellations have been designated for strains that commonly infect humans: Wa-like (I1-R1-C1-M1-A1-N1-T1-E1-H1) and DS-1-like (I2-R2-C2-M2-A2-N2-T2-E2-H2). A third constellation also observed in human strains, called AU-1-like (I3-R3-C3-M3-A3-N3-T3-E3-H3), has been shown to have a feline/canine origin [9,11].

Four live oral vaccines, namely Rotarix® (GlaxoSmithKline Biologics, Rixensart, Belgium), RotaTeq® (Merck & Co., Kenilworth, NJ, USA), Rotavac® (Bharat Biotech, Hyderabad, India) and Rotasiil® (Serum Institute of India Pvt. Ltd., Pune, India) have been prequalified by the World Health Organization [12,13]. Rotarix® and RotaTeq® have been introduced into the immunization programs of more than 100 countries [13]. Rotarix® is a monovalent vaccine containing a single human G1P[8] strain and is administered from the age of six weeks [13]. Prior to vaccine introduction in Mozambique, a high burden of rotavirus disease was reported in children under five years old. The rate of rotavirus infection in urban (Maputo City) and rural (Manhiça District) areas between 2012 and 2013 was 42.4% [6]. In 2011, a 24.0% infection rate was reported in the Gaza province, another rural area in southern Mozambique [14]. In both studies G1P[8] was detected at a low frequency [14,15]. Data from the National Surveillance of Diarrhea (ViNaDia) revealed a high rotavirus infection rate of 40.2% and 38.3% in 2014 and 2015, respectively [16]. The Rotarix® vaccine was introduced in Mozambique in 2015 with increasing vaccine coverage of 70% and 80% in 2016 and 2017, respectively [16,17]. Post vaccine introduction, the rotavirus infection rate was reduced to 12.2% and 13.5% in 2016 and 2017, respectively [16]. During ViNaDia surveillance, G1P[8] strains were consistently observed in the pre- (2012–2015) and post-vaccination period (2016–2019). However, in the post-vaccine period a decrease in G1P[8] strains was observed which coincided with the emergence of other non-G1P[8] genotypes such as G3P[4] and G3P[8] [18].

The whole genomes of Mozambican G2P[4], G8P[4], G12P[6] and G12P[8] RVA strains from the pre-vaccination period have been described [19,20]. However, there are no reports of the whole genome analyses of G1P[8] strains from Mozambique. To address this, the consensus sequences of 36 G1P[8] strains collected between 2012–2017 from vaccinated and non-vaccinated children were analyzed to investigate the diversity and evolution of G1P[8] strains.

2. Results

2.1. Genome Constellation

A total of 36 G1P[8] (12 from the pre-vaccine period and 24 from the post-vaccine period) strains were successfully sequenced with an average coverage ranging from 450.0 to 46060.5 per sequence (Supplementary Table S1). Complete open reading frames (ORFs) were obtained for 393 of the 396 genome segments analyzed. A partial ORF (99.0%) for segment four of RVA/Human-wt/MOZ/HCN0690/2015/G1P[8] was obtained, while two genome segments (encoding VP2 and VP3, respectively) of RVA/Human-wt/MOZ/HGM0059/2014/G1P[8] could not be determined as insufficient data were generated for these two segments (Supplementary Table S1). The genotype constellations were determined and all strains exhibited a Wa-like genetic backbone (G1-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1). The nucleotide (nt) identities among Mozambican strains varied from 92.5–100.0% and the comparison between Rotarix® and the 11 genes of the Mozambican strains revealed 84.0–97.9% nt identity (Supplementary Table S2).

2.2. Phylogenetic Analyses

2.2.1. Sequence Analyses of VP7 and VP4

The VP7 encoding sequences of the 36 Mozambican G1P[8] strains, collected between 2012 to 2017 from non-vaccinated and vaccinated children (Supplementary Table S3), were compared with human rotavirus sequences representing VP7 G1 lineages (I-VII) [21–28]. The Mozambican strains clustered into two distinct lineages, I and II (Figure 1a). The majority of Mozambican strains formed a highly conserved clade, and were closely related to various Indian strains circulating between 2012 and 2013. HGM0544 was moderately divergent to the rest of the strains in the clade sharing 99.2–99.7% nucleotide (nt) identity and 98.8–99.4% amino acid (aa) identity. Strains from the pre- and post-vaccine era were intermingled in lineage II. Only two Mozambican strains from this study clustered in the VP7 lineage I and were more diverse than the 34 strains clustering in lineage II. MAN0033, collected in a rural area in southern Mozambique before vaccine introduction, was closely related to Malawian strains from 2012 and to previously characterized Mozambican strains detected in 2011 [14]. HJM1646, collected in southern Mozambique after vaccine introduction, clustered distinctly and only shared 92.5–92.9% nt and 92.7–93.3% aa identity to the other Mozambican strains. HJM1646 clustered with contemporary Indian strains in a sub-lineage of African and global strains (Figure 1a).

The P[8] encoding sequences of the 36 Mozambican strains were compared with human rotavirus sequences representing the four lineages (I–IV) [21–27] (Figure 1b). The Mozambican strains clustered in the major P[8] lineage III. Similar to the VP7 tree, the majority of Mozambican P[8] sequences formed a highly conserved clade, and were closely related to various Indian strains circulating between 2012 and 2013. The P[8] encoding sequence of HJM1646 clustered with HGM0544, despite clustering in different VP7 lineages. These two strains were moderately divergent to the rest of the study strains in the Mozambican clade and clustered close to another Mozambican strain, RVA/Human-wt/MOZ/0060a/2012/G12P[8]P[14], which was previously detected in the Manhiça district in southern Mozambique [19]. MAN0033 clustered distinctly to the rest of the Mozambican strains sharing 95.7–96.2% nt and 98.3–98.7% aa identity and was closely related to contemporary Malawian strains isolated in 2012 (Figure 1b, Supplementary Table S1).
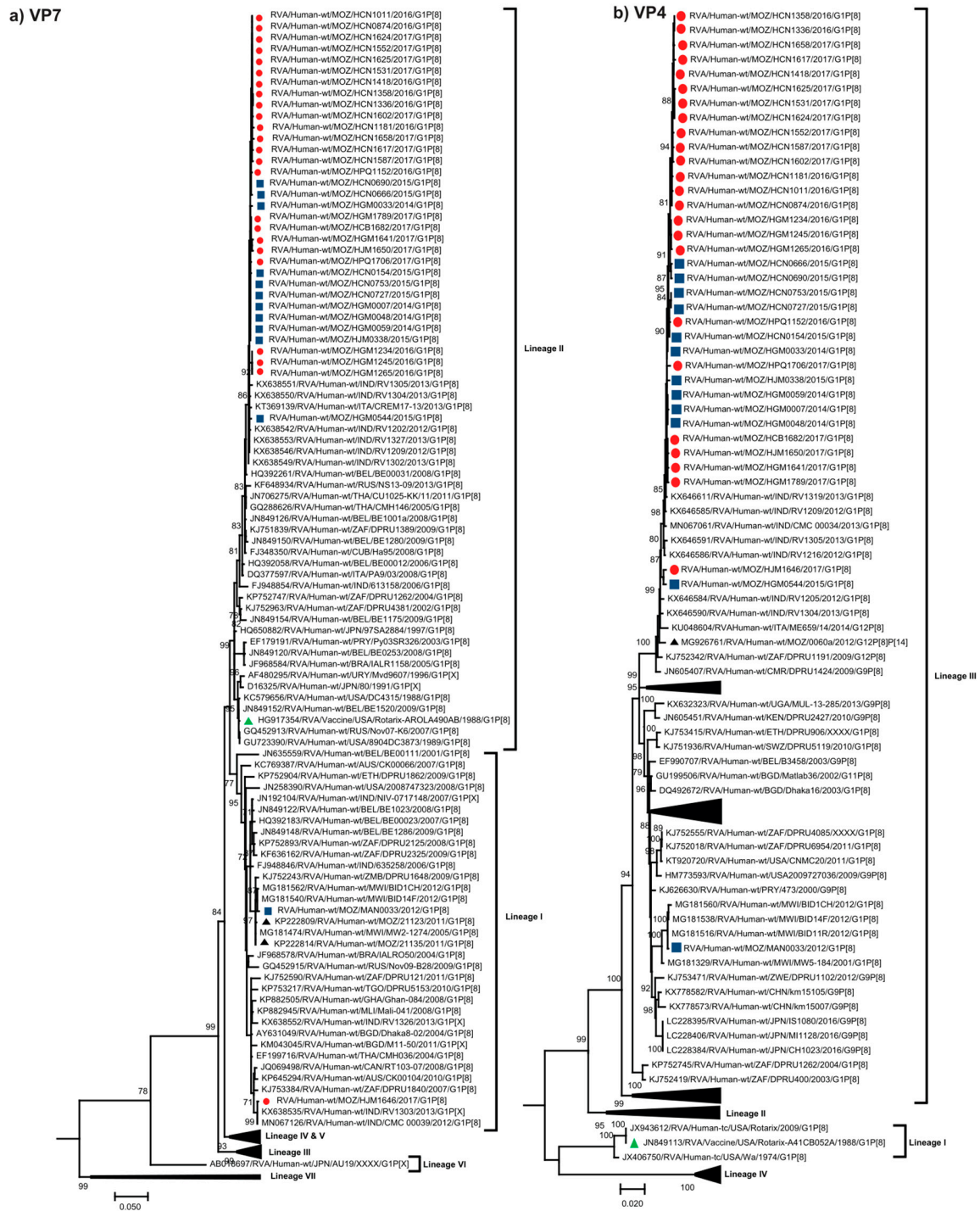
Figure 1. Phylogenetic trees based on the ORF (open reading frame) nucleotide sequence of the (**a**) VP7 and (**b**) VP4 genes of G1P[8] strains circulating in Mozambique and global strains obtained from GenBank. The trees were constructed based on the maximum likelihood method implemented in MEGA X [29], applying the best-fit nucleotide substitution model Tamura-3-parameter (T92+G+I) for VP7 and General Time Reversible (GTR-G) for VP4, determined by JModelTest [30]. Bootstrap values (1000 replicates) ≥70% are shown with DS-1 serving as an out-group (not shown in the final tree). Scale bar indicates genetic distance expressed as the number of nucleotide substitutions per site. Pre-vaccine Mozambican strains are indicated by blue squares, post-vaccine by red circles, the Rotarix® vaccine strain by a green triangle and Mozambican strains from previous studies [14,19] are indicated by black triangles. Lineages are defined from I-VIII for VP7 and I-IV for VP4 [21–23,25–28].

2.2.2. Sequence Analyses of VP1-VP3 and VP6

Thirty-three Mozambican strains formed conserved, monophyletic clades that were observed in the VP1, VP2 and VP3 trees, closely related to RVA/Human-wt/IND/CMC00034/2013/G1P[8], and within lineages comprising of contemporary African and global strains (Figure 2a–c). In the VP6 tree, the 35 Mozambican strains clustered together but did not form a discrete monophyletic clade, and were closely related to contemporary Indian G1P[8] strains (Figure 2d). HJM1646 and HGM0544 clustered together in the VP1 tree were moderately divergent from the main Mozambican clade (Figure 2a). In the VP2 tree, HJM1646 clustered close to the monophyletic clade while HGM0544 fell within the Mozambican clade (Figure 2b). In the VP3 tree, these strains clustered together, distinct from the Mozambican clade, adjacent to previously characterized G12P[6] Mozambican strains (Figure 2c) [19]. MAN0033 fell within the same lineage as the main Mozambican clade, showing minor divergence in the VP1 and VP2 tree and more pronounced divergence in the VP3 tree, closely related to contemporary Malawian G1P[8] strains (Figure 2a–c). This strain clustered within a different lineage in the VP6 tree, closely related to the same group of Malawian G1P[8] strains and adjacent to the Mozambican G12P[6] strains (Figure 2d).

2.2.3. Sequence Analyses of NSP1-NSP5/6

The conserved monophyletic clade, comprised of 33 Mozambican strains, was observed in the NSP1–NSP4 trees, with RVA/Human-wt/IND/CMC00034/2013/G1P[8] interspersed within the clade in the NSP2 tree (Figure 2e–h). HJM1646 and HGM0544 continued to show varied clustering patterns across the trees. In the NSP3 tree these strains clustered together and were divergent from the main Mozambican clade, clustering with Indian strains including RVA/Human-wt/IND/CMC00034/2013/G1P[8], and close to Mozambican G12P[6] strains (Figure 2g). HJM1646 was divergent to the Mozambican clade in the NSP1 and NSP4 trees, but clustered close to the monophyletic clade in the NSP2 tree. HGM0544 clustered with the main Mozambican clade in the NSP1, NSP2 and NSP4 trees. In the NSP5 tree, HJM1646 and HGM0544, along with 33 other Mozambican strains, formed a monophyletic clade that was interspersed with global strains (Figure 2i). MAN0033 clustered distinctly to the rest of the Mozambican strains and was closely related to contemporary Malawian strains isolated in 2012 across these trees (Figure 2e–i). The five G12P[6] [19] Mozambican strains fell within neighboring clusters to MAN0033 in the VP6 and NSP2 trees (Figure 2d,f).

2.3. Evolutionary Analysis of VP7 and VP4 Genes

A randomly subsampled dataset of 378 G1 genes that were representative of global strains temporally and genetically were analyzed (Supplementary Figure S1). The Mozambican strains detected between 2012 and 2017 shared a common ancestral strain circulating in 2009.9 (95% HPD 2008.1–2010.9). Of the Mozambican strains characterized in this study, 34 clustered within the same lineage and shared a most recent common ancestor in 2011.7 (95% HPD 2011.1–2012.0) and diverged from the closest related Indian strains around the same time. MAN0033 and HJM1646 clustered in the other major lineage present in the tree. MAN0033 and closely related Malawian strains shared a common ancestor in 2010.1 (95% HPD 2008.5–2010.9). These variants, circulating in Malawi, Zambia and Mozambique, diverged from a group of Indian G1P[8] strains in 2001.3 (95% HPD 1998.4–2003.5). HJM1646 was divergent to the other G1 strains from Mozambique in this lineage and shared its most recent common ancestor with Indian strains in 2012.9 (95% HPD 2012.4–2013.0) (Supplementary Figure S1).
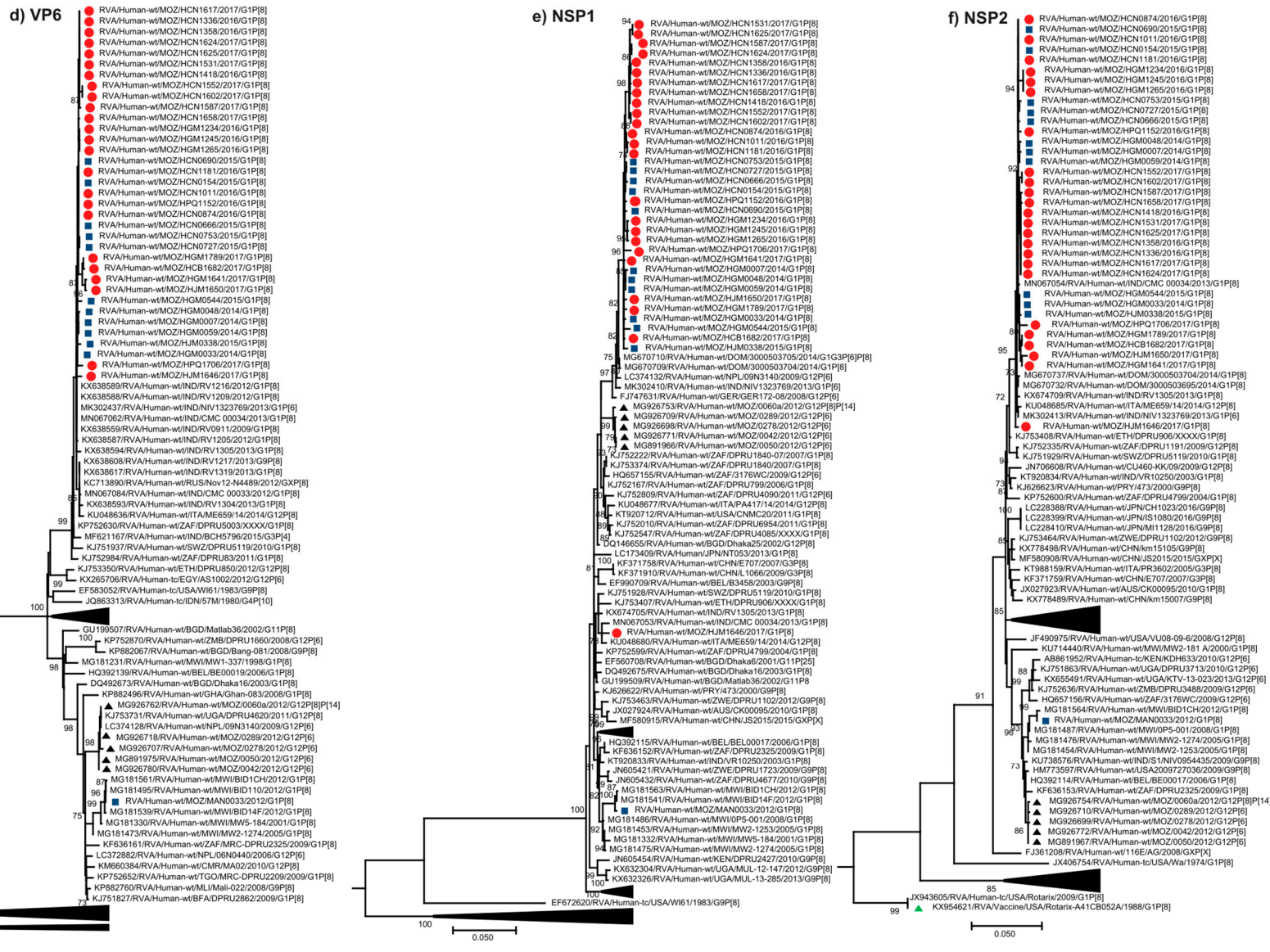
Figure 2. *Cont.*

Figure 2. *Cont.*
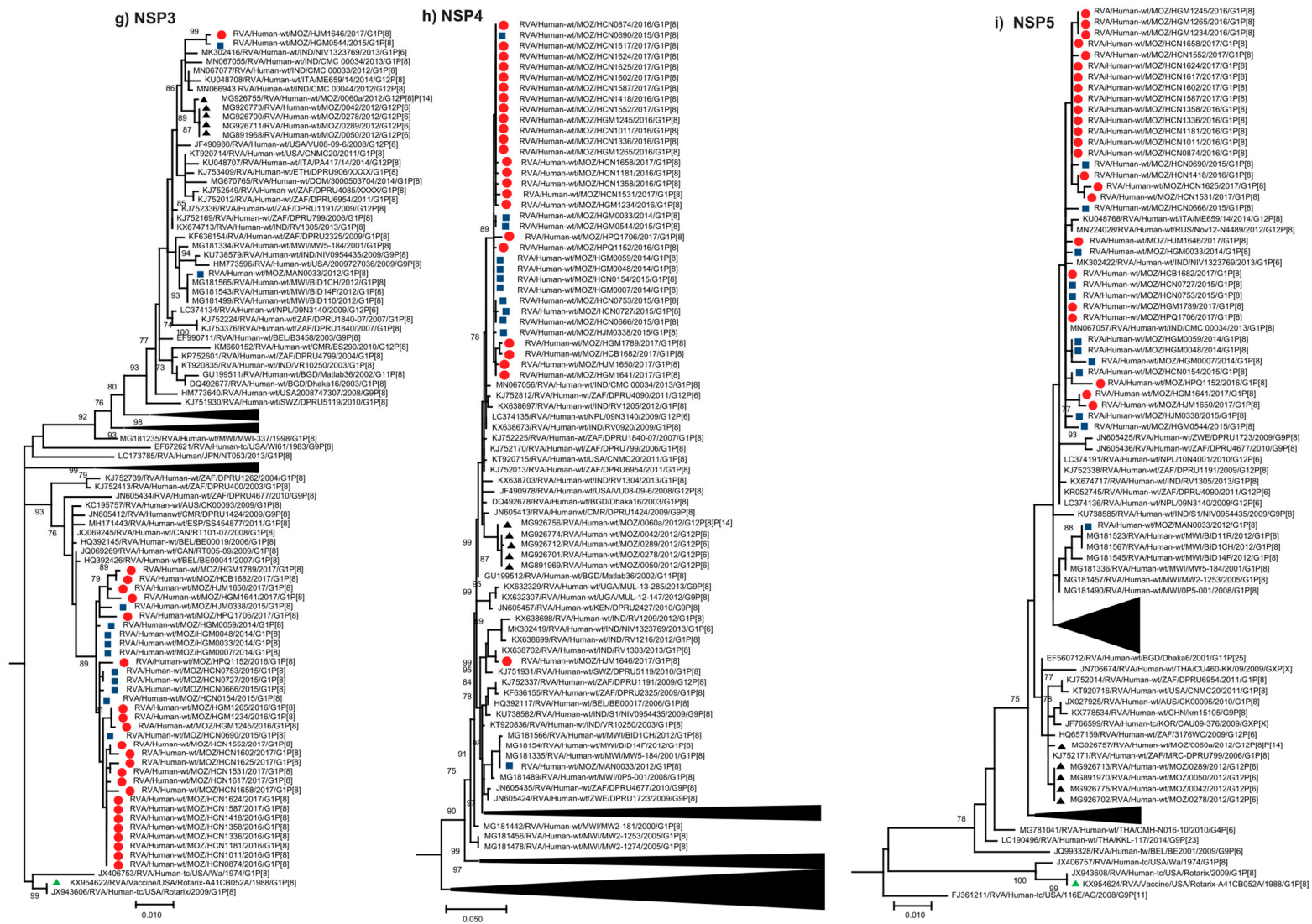
Figure 2. Phylogenetic trees based on the ORF nucleotide sequences of the (**a**) VP1, (**b**) VP2, (**c**) VP3, (**d**) VP6, (**e**) NSP1, (**f**) NSP2, (**g**) NSP3, (**h**) NSP4 and (**i**) NSP5

genes of G1P[8] strains circulating in Mozambique and global strains obtained from GenBank. The trees were constructed based on the maximum likelihood method implemented in MEGA X [29], using the best-fit nucleotide substitution model General Time Reversible (GTR+G+I) for VP3, GTR+G for VP2, NSP2 and NSP3, Hasegawa Kishino Yano (HKY+G+I) for VP6 and NSP1, HKY+G for VP1, NSP4 and NSP5/6, determined by JModelTest [30]. Bootstrap values (1000 replicates) ≥70% are shown with DS-1 serving as an out-group (not shown in the final tree). Scale bar indicates genetic distance expressed as the number of nucleotide substitutions per site. Pre-vaccine Mozambican strains are indicated by blue squares, post-vaccine by red circles, the Rotarix® vaccine strain by a green triangle and Mozambican strains from a previous study [19] are indicated by black triangles.

A subsampled dataset of 235 P[8] genes, representative of global strains temporally and genetically, was also analyzed. Thirty-three Mozambican strains characterized in this study clustered within the same lineage and shared a common ancestor in 2013.0 (95% HPD 2012.1–2013.6) and diverged from the closest related Indian strains around 2011.6 (95% HPD 2011.2–2011.9) (Supplementary Figure S2). The most recent common ancestor of HGM0544 and HJM1646 (that was moderately divergent to the rest of the Mozambican strains in the major clade), was estimated to be 2014.1 (95% HPD 2013.0–2014.9). Clustering in a separate lineage to the other Mozambican G1P[8] strains, MAN0033 diverged from the closest Malawian strain in 2010.9 (95% HPD 2009.5–2011.7) (Supplementary Figure S2).

2.4. Comparative Analysis of Neutralizing Antigenic Epitopes of the VP7 and VP4 Genes of Mozambican Strains and the Rotarix® Vaccine Strain

The rotavirus VP7 protein consists of two antigenic epitopes, 7-1 and 7-2, with 7-1 subdivided into 7-1a and 7-1b [31]. The comparative analysis of the VP7 antigenic epitopes between Mozambican strains and the Rotarix® vaccine strain revealed amino acid substitutions in all three antigenic sites. However, most of the amino acid substitutions were observed in antigenic region 7-2. A total of 30 strains shared conserved amino acid differences at positions N147D and 25 strains at M217I. Sporadic mutations were observed in MAN0033 (unvaccinated) and HJM1646 (fully vaccinated) (S123N, K291R and M217T). The HJM1646 strain contained an additional amino acid substitution at N96S (Figure 3).

Antigenic regions	7-1a														7-1b						7-2								
	87	91	94	96	97	98	99	100	104	123	125	129	130	291	201	211	212	213	238	242	143	145	146	147	148	190	217	221	264
Rotarix®	T	T	N	G	E	W	K	D	Q	S	V	V	D	K	Q	N	V	D	N	T	K	D	Q	N	L	S	M	N	G
HCB1682	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.
HCN0154	.	.	.	.	.	.	.	.	.	.	.	N	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN0666	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN0690	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN0727	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN0753	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN0874	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN1011	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN1181	.	.	.	.	.	.	E	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN1336	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN1358	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN1418	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN1531	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN1552	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN1587	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	S
HCN1602	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN1617	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN1624	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN1625	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HCN1658	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HGM0007	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.
HGM0033	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.
HGM0048	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.
HGM0059	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.
HGM0544	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	I	.	.	.	.	.	.	.	.	.	.	.	.
HGM1234	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	I	.	.
HGM1245	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	I	.	.
HGM1265	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	I	.	.
HGM1641	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.
HGM1789	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.
HJM0338	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.
HJM1646	.	.	S	.	.	.	.	.	.	N	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.
HJM1650	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.
HPQ1152	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	I	.	.
HPQ1706	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	D	.	.	.	.	.
MAN0033	.	.	.	.	.	.	.	.	.	N	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.

Figure 3. The alignment of amino acids corresponding to three VP7 antigenic epitopes (7-1a, 7-1b and 7-2). The amino acid sequence of Rotarix® is the reference strain and the conserved residues between the Rotarix® to Mozambican strains are indicated by dots (.) and residues that differ are in bold.

Activation of the protein VP4 requires proteolytic cleavage to produce the VP8* and VP5* subunits. These regions contain four (8-1 to 8-4) and five (5-1 to 5-5) antigenic epitopes, respectively [32,33]. The amino acid substitutions between the Rotarix® vaccine strain and Mozambican P[8] strains were concentrated in the 8-1 and 8-3 epitopes. There were five conserved amino acid substitutions, at positions E150D, N195D/G, S125N, S131R and N135D. Sporadic mutations were observed in MAN0033 (N195S and N113D), HJM0338 (S146N) and HGM1789 (P114T) (Figure 4).

Antigenic region	8-1											8-2		8-3									8-4			5-1								5-2	5-3	5-4	5-5
	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
Rotarix®	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	I	W	P	G	R	T	P	E	L	R
HCN1181	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	.	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN1011	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN0874	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN1418	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN1531	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN1624	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN1625	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN0753	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN1358	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN1336	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN1552	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN1587	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN1602	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN1617	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN1658	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HGM1234	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HGM1245	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HGM1265	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN0666	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN0690	.	.	.	D	.	.	.	.	.	.	D	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN0727	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCN0154	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HPQ1152	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HGM1641	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HGM0007	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HGM0048	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HGM0059	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HJM1650	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HCB1682	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HGM0033	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HGM1789	.	.	.	D	.	.	.	.	.	.	G	.	.	.	T	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HJM0338	.	N	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HPQ1706	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HJM1646	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HGM0544	.	.	.	D	.	.	.	.	.	.	G	.	.	.	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MAN0033	.	.	.	D	.	.	.	.	.	.	S	.	.	D	.	.	.	N	R	.	.	D	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Figure 4. The alignment of the amino acids corresponding to the VP4 antigenic epitopes (8-1, 8-2, 8-3, 8-4 for VP8* and 5-1, 5-2, 5-3, 5-4, 5-5 for VP5). The amino acid sequence of Rotarix® is the reference strain and the conserved residues between the Rotarix® to Mozambican strains are indicated by dots (.) and residues that differ are in bold.

3. Discussion

In the present study, whole genome sequencing was performed for 36 G1P[8] RVA strains obtained from Mozambican children with gastroenteritis between 2012–2017 (12 from the pre-vaccine period and 24 from the post-vaccine period). This is the first study to perform whole genome analysis of G1P[8] strains in Mozambique, facilitating the description of genetic diversity and the origins of Mozambican strains.

Of the 36 strains characterized, 33 clustered within the same conserved Mozambican clade across all trees. Two strains, HGM0544 and HJM1646, showed varied patterns by clustering within and were distinct from the Mozambican clade across trees, suggesting these strains had undergone reassortment events. The strain, MAN0033, clustered distinctly from the rest of the Mozambican strains in all trees. This strain was closely related to a conserved group of Malawian G1P[8] strains suggesting that this strain may have been recently introduced from a neighboring country. No distinct clustering patterns were observed based on the year of isolation or vaccination status, which suggests that strains with limited sequence diversity may have circulated among children in the country over the five year period investigated (2012–2017). The homogeneous population of G1P[8] strains suggests that the introduction of Rotarix®has not resulted in a dramatic shift in the diversity of G1P[8] strains circulating in Mozambique. A similar finding was reported in South Africa where no distinct clustering was observed for strains from the pre- and post-vaccine introduction period [26]. Analysis of G1P[8] strains in Brazil over a 27 year period also did not detect any evidence of a selective pressure exerted by the mass introduction of Rotarix® [34]. In contrast, Australia and Belgium reported some unique clusters of G1P[8] strains following vaccine introduction, which may have been due to natural fluctuation or the first signs of vaccine-driven evolution [35]. In Rwanda, unique clusters of G1P[8] strains were identified following RotaTeq introduction [36]. Although neighboring countries reported the widespread (Malawi) and sporadic (South Africa) detection of G1P[8] strains that had undergone reassortment with DS-1 like strains [28,37], all Mozambican strains characterized in this study exhibited a typical Wa-like genetic backbone. Despite some reports of vaccine-derived G1P[8] strains detected in Australia and England, none of the strains identified in this study were derived from the Rotarix® vaccine [38,39].

Although there are seven recognized lineages described for global G1 sequences [40], the majority of Mozambican strains from this study clustered in lineage II, with only two strains clustering in lineage I. However, one strain that clustered in lineage I represented the oldest Mozambican strain sequenced in this study from 2012, which clustered with previously characterized G1P[8] Mozambican strains from 2011 [14]. This may suggest that lineage I strains were replaced in later years by G1 strains associated with lineage II [26,41]. The VP7 lineage I strains were detected in the south of Mozambique which may suggest geographical restriction in the circulation of strains. However, these results can be in part due to the short sampling period of this study. Of the four established lineages of the P[8] genotype [21], all Mozambican strains clustered in lineage III and shared a high level of genetic similarity, except strain MAN0033 which clustered in a distinct sub-lineage.

Maximum likelihood phylogenetic analysis showed that the majority of the Mozambican strains, with the exception of MAN0033, were most closely related to a conserved group of Indian strains across most genes. Even the two reassortant strains (HJM1646 and HGM0544) were most closely related to Indian strains. This suggests that there may have been multiple, contemporary introductions of diverse strains from a similar origin, perhaps India, into Mozambique. This was further supported by the results of the Bayesian analysis, where the time to the most recent common ancestor for the VP7 and VP4 genes of the main Mozambican clade were 2011.7 and 2013.0, respectively, and which had diverged from the closest Indian strain in 2011.7 and 2011.6, respectively. This suggests that the strains became endemic shortly after being introduced and became the dominant variant circulating in the population. These Indian strains were submitted directly to the GenBank database and no associated manuscripts were found, so it is unclear if these strains were associated with any particular outbreak or were detected as part of routine surveillance.

Overall, the VP7 and VP4 antigenic epitopes exhibited conserved substitutions among the Mozambican strains when compared to Rotarix®. The substitutions in the 7-2 VP7 epitope at position M217I, N147D and 8-1, 8-3 VP4 at positions E150D, N195D and S125N, S131R, N135D were observed in pre- and post-vaccine introduction strains, suggesting that these substitutions are not due to the vaccine introduction.

The main limitation of the study was the limited number of strains successfully sequenced, and that fewer strains were sequenced from the pre-vaccine period.

There is a need to expand the whole genome analysis to other strains detected in Mozambique such as G3, G9 in combination with P[4], P[6] and P[8] genotypes reported previously [18] in order to evaluate the possible influence of vaccine introduction on other rotavirus genotypes.

Mozambique has introduced the Rotarix® vaccine, however cases of rotavirus infection associated with G1P[8] strains resulting in hospitalization of children are still being reported. The present analysis showed that G1P[8] strains detected in the post-vaccine period did not undergo significant mutations in the epitope regions that could result in vaccine escape. However, the short post vaccine period analyzed (three years) may have influenced these results, as it may be too early to see major genetic changes associated with vaccine pressure. These results highlight the need for future studies to understand host factors such as the role of histo-blood group antigen status, nutritional status and enteric co-infections that can influence the vaccine effectiveness in Mozambique.

4. Materials and Methods

4.1. Ethics Approval

The ViNaDia Protocol was approved by the National Health Bioethics Committee of Mozambique (CNBS) under number (IRB00002657, reference Nr: 348/CNBS/13). Participants' anonymity and confidentiality were guaranteed.

4.2. Sample Collection

Forty-three fecal samples, collected between 2012 and 2017 that were positive for RVA by ELISA (Prospect EIA rotavirus, Basingstoke, UK) and identified as genotype G1P[8] by multiplex RT-PCR according to described protocols [42,43], were selected for sequencing according to year of isolation, location of collection (region of Mozambique) and the child's vaccination status. The samples were obtained from children <5 years of age, hospitalized with acute gastroenteritis, and collected at five sentinel sites of the National Diarrheal Surveillance (ViNaDia), which are Hospital Geral de Mavalane (HGM), Hospital Geral Jose Macamo (HJM), Hospital Central da Beira (HCB), Hospital Geral de Quelimane (HGQ) and Hospital Central de Nampula (HCN), and from a previous study of Centro de Investigação em Saúde da Manhiça (CISM) in Mozambique between 2012 and 2013 (Supplementary Figure S3) [15]. Clinical information was collected through a structured questionnaire from ViNaDia which included metadata such as age, gender, site and vaccination status.

4.3. RNA Extraction and cDNA Synthesis

Total RNA was extracted from stool samples with TRI-reagent (Sigma, Darmstadt, Germany) and single-stranded RNA was precipitated with lithium chloride. The self-priming PC3-T7 loop primer (Integrated DNA Technologies, Coralville, IA, USA) was ligated to dsRNA in order to obtain full-length sequences and cDNA was synthesized using the Maxima H Minus double-stranded cDNA kit (Thermo Fisher Scientific, Massachusetts, MA, USA) as previously described [20,44].

4.4. Next Generation Sequencing

The whole genome sequencing was performed using an Illumina MiSeq sequencing platform (Illumina, Inc. San Diego, CA, USA) at the Next Generation Sequencing Unit at the University of the

Free State (NGS-UFS). Sequencing was completed using the Nextera XT DNA Library Preparation Kit (Illumina, Inc., San Diego, CA, USA) using protocols previously describe [19].

4.5. Data Analyses

A de novo assembly was performed for all samples using CLC Bio Genomics Workbench (12.0.3; Qiagen, Aarhus, Denmark); all contigs with an average coverage above 100 were identified on the Nucleotide Basic Local Alignment Search Tool (BLASTn at the National Center for Biotechnology information (NCBI). References were chosen based on the Blastn results for reference mapping and extraction of consensus sequences for each segment. The genotyping tools, Virus Pathogenic database and analysis resource (ViPR) [45] and RotaC v2.0 [46], were used to determine the genotype of each gene. The sequences were submitted to GenBank and accession numbers MT737379-MT737772 were assigned.

4.6. Phylogenetic Analysis

Multiple nucleotide sequence alignments with strains obtained from GenBank [24] were made with multiple sequence alignment program (MAFFT v7.450) [47] on Geneious prime v2020.0.3 and Multiple Sequence Comparison by Log Expectation (MUSCLE) [48] alignment available in Molecular Evolutionary Genetic Analysis X (MEGA X) [29]. The optimal nucleotide substitution model for phylogenetic analysis was selected based upon the Akaike information criterion (corrected) (AICc) ranking implemented in the model selection algorithm available on JModelTest [30] and the models selected for each segment were: Tamura-3 parameter (T92+G+I) [49] for VP7; General Time Reversible (GTR+G+I) [50] for VP3; GTR+G for VP2, VP4, NSP2 and NSP3; Hasegawa Kishino Yano (HKY+G+I) for VP6 and NSP1; and HKY+G for VP1, NSP4 and NSP5/6 [51]. The maximum-likelihood trees were generated using MEGA X [29] using 1000 bootstrap replicates to estimate branch support. Pairwise distance matrix nucleotides were obtained in MEGA X using the p-distance algorithm [29]. Amino acid sequences of the VP7 and VP4 Mozambican strains were aligned and epitopes were identified and compared to those of the vaccine strain Rotarix® (A41CB052A, with accession numbers JN849114 and JN849113 for VP7 and VP4) using MEGA X [23,31].

4.7. Evolutionary Analysis

Maximum likelihood trees were generated using the Randomized Accelerated Maximum Likelihood (RAxML) program (v2.0.0) [52], applying the nucleotide substitution model GTR+G. The trees were used as the input for TempEst 1.5.3 to plot root-to-tip genetic distances, and sequences not conforming to a linear evolutionary pattern were discarded [53]. Time-measured evolutionary histories were reconstructed using the Bayesian Evolutionary Analysis Sampling Trees (BEAST) Program package (v 1.7.5) [54]. The nucleotide substitution model Hasegawa-Kishino-Yano model (HKY+G) for VP7 and GTR+G for VP4 were selected based on AICc raking in jModelTest [30].

The parameters applied included a relaxed uncorrelated lognormal molecular clock to account for varied evolutionary rates among lineages and a coalescent Gaussian Markov random field (GMRF) Bayesian Skyride tree prior. Three independent Markov chain Monte Carlo (MCMC) chains were run for 200 million generations with sampling every 20,000 generations, with the first 10% discarded as burn-in. Convergence and mixing of the chains was assessed using Tracer (v1.7.1) and all parameters yielded effective sample sizes ≥ 200 [55]. The Maximum Clade Credibility (MCC) trees were summarized using TreeAnnotator (v1.10.4) [54]. The time-ordered MCC trees were visualized in FigTree (v1.4.4) (http://tree.bio.ed.ac.uk/software/figtree/).

5. Conclusions

This study provides important insights into the whole genome sequences of G1P[8] strains in Mozambique. Whilst similar strains were detected prior to and following vaccine introduction, multiple introductions of diverse strains from India highlight the importance of continuously monitoring the

strains detected in Mozambique to determine if the strains are evolving by vaccine-induced selection or by natural evolutionary pressures.

References

1. Tate, J.E.; Burton, A.H.; Boschi-Pinto, C.; Parashar, U.D. Global, Regional, and National Estimates of Rotavirus Mortality in Children <5 Years of Age, 2000–2013. *Clin. Infect. Dis.* **2016**, *62* (Suppl. 2), S96–S105. [CrossRef] [PubMed]
2. Troeger, C.; Khalil, I.A.; Rao, P.C.; Cao, S.; Blacker, B.F.; Ahmed, T.; Armah, G.; Bines, J.E.; Brewer, T.G.; Colombara, D.V.; et al. Rotavirus Vaccination and the Global Burden of Rotavirus Diarrhea Among Children Younger Than 5 Years. *JAMA Pediatr.* **2018**, *172*, 958–965. [CrossRef] [PubMed]
3. Estes, M.K.; Cohen, J. Rotavirus gene structure and function. *Microbiol. Rev.* **1989**, *53*, 410–449. [CrossRef] [PubMed]
4. Jayaram, H.; Estes, M.; Prasad, B.V. Emerging themes in rotavirus cell entry, genome organization, transcription and replication. *Virus Res.* **2004**, *101*, 67–81. [CrossRef]
5. Desselberger, U. Rotaviruses. *Virus Res.* **2014**, *190*, 75–96. [CrossRef]
6. De Deus, N.; João, E.; Cuamba, A.; Cassocera, M.; Luís, L.; Acácio, S.; Mandomando, I.; Augusto, O.; Page, N. Epidemiology of Rotavirus Infection in Children from a Rural and Urban Area, in Maputo, Southern Mozambique, before Vaccine Introduction. *J. Trop. Pediatr.* **2017**, *64*, 141–145. [CrossRef]
7. Matthijnssens, J.; Ciarlet, M.; McDonald, S.M.; Attoui, H.; Banyai, K.; Brister, J.R.; Buesa, J.; Esona, M.D.; Estes, M.K.; Gentsch, J.R.; et al. Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). *Arch. Virol.* **2011**, *156*, 1397–1413. [CrossRef]
8. Rotavirus Classification Working Group. Virus Classification. Available online: https://rega.kuleuven.be/cev/viralmetagenomics/virus-classification (accessed on 5 June 2020).

9.　Matthijnssens, J.; Ciarlet, M.; Heiman, E.; Arijs, I.; Delbeke, T.; McDonald, S.M.; Palombo, E.A.; Iturriza-Gómara, M.; Maes, P.; Patton, J.T.; 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.* **2008**, *82*, 3204–3219. [CrossRef]

10.　Mwenda, J.M.; Tate, J.E.; Parashar, U.D.; Mihigo, R.; Agócs, M.; Serhan, F.; Nshimirimana, D. African Rotavirus Surveillance Network. *Pediatr. Infect. Dis. J.* **2014**, *33*, S6–S8. [CrossRef]

11.　Matthijnssens, J.; Van Ranst, M. Genotype constellation and evolution of group A rotaviruses infecting humans. *Curr. Opin. Virol.* **2012**, *2*, 426–433. [CrossRef]

12.　World Health Organization. WHO Prequalifies New Rotavirus Vaccine. Available online: http://www.who.int/medicines/news/2018/prequalified_new-rotavirus_vaccine/en/ (accessed on 5 June 2020).

13.　Burke, R.M.; Tate, J.E.; Kirkwood, C.D.; Steele, A.D.; Parashar, U.D. Current and new rotavirus vaccines. *Curr. Opin. Infect. Dis.* **2019**, *32*, 435–444. [CrossRef] [PubMed]

14.　Langa, J.S.; Thompson, R.; Arnaldo, P.; Resque, H.R.; Rose, T.; Enosse, S.M.; Fialho, A.; De Assis, R.M.S.; Da Silva, M.F.M.; Leite, J.P.G. Epidemiology of rotavirus A diarrhea in Chókwè, Southern Mozambique, from February to September, 2011. *J. Med. Virol.* **2016**, *88*, 1751–1758. [CrossRef] [PubMed]

15.　João, E.D.; Strydom, A.; O'Neill, H.G.; Cuamba, A.; Cassocera, M.; Acácio, S.; Mandomando, I.; Motanyane, L.; Page, N.; De Deus, N. Rotavirus A strains obtained from children with acute gastroenteritis in Mozambique, 2012–2013: G and P genotypes and phylogenetic analysis of VP7 and partial VP4 genes. *Arch. Virol.* **2017**, *163*, 153–165. [CrossRef] [PubMed]

16.　De Deus, N.; Chilaúle, J.J.; Cassocera, M.; Bambo, M.; Langa, J.S.; Sitoe, E.; Chissaque, A.; Anapakala, E.; Sambo, J.; Guimarães, E.L.; et al. Early impact of rotavirus vaccination in children less than five years of age in Mozambique. *Vaccine* **2018**, *36*, 7205–7209. [CrossRef]

17.　World Health Organization. WHO and UNICEF Estimates of National Immunization Coverage. Available online: http://www.who.int/immunization/monitoring_surveillance/routine/coverage/en/index4.html (accessed on 1 April 2020).

18.　João, E.; Munlela, B.; Chissaque, A.; Chilaúle, J.; Langa, J.S.; Augusto, O.; Boene, S.; Anapakala, E.; Sambo, J.; Guimarães, E.; et al. Molecular Epidemiology of Rotavirus A Strains Pre- and Post-Vaccine (Rotarix®) Introduction in Mozambique, 2012–2019: Emergence of Genotypes G3P[4] and G3P[8]. *Pathogens* **2020**, *9*, 671. [CrossRef]

19.　Strydom, A.; Motanyane, L.; Nyaga, M.M.; João, E.D.; Cuamba, A.; Mandomando, I.; Cassocera, M.; De Deus, N.; O'Neill, H.G. Whole-genome characterization of G12 rotavirus strains detected in Mozambique reveals a co-infection with a GXP[14] strain of possible animal origin. *J. Gen. Virol.* **2019**, *100*, 932–937. [CrossRef]

20.　Strydom, A.; João, E.D.; Motanyane, L.; Nyaga, M.M.; Potgieter, A.C.; Cuamba, A.; Mandomando, I.; Cassocera, M.; De Deus, N.; O'Neill, H.G. Whole genome analyses of DS-1-like Rotavirus A strains detected in children with acute diarrhoea in southern Mozambique suggest several reassortment events. *Infect. Genet. Evol.* **2019**, *69*, 68–75. [CrossRef]

21.　Le, V.P.; Chung, Y.-C.; Kim, K.; Chung, S.-I.; Lim, I.; Kim, W. Genetic variation of prevalent G1P[8] human rotaviruses in South Korea. *J. Med. Virol.* **2010**, *82*, 886–896. [CrossRef]

22.　Arista, S.; Giammanco, G.M.; De Grazia, S.; Ramirez, S.; Biundo, C.L.; Colomba, C.; Cascio, A.; Martella, V. Heterogeneity and Temporal Dynamics of Evolution of G1 Human Rotaviruses in a Settled Population. *J. Virol.* **2006**, *80*, 10724–10733. [CrossRef]

23.　Zeller, M.; Patton, J.T.; Heylen, E.; De Coster, S.; Ciarlet, M.; Van Ranst, M.; Matthijnssens, J. 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.* **2012**, *50*, 966–976. [CrossRef]

24.　Ianiro, G.; Delogu, R.; Fiore, L.; Ruggeri, F.M. Genetic variability of VP7, VP4, VP6 and NSP4 genes of common human G1P[8] rotavirus strains circulating in Italy between 2010 and 2014. *Virus Res.* **2016**, *220*, 117–128. [CrossRef] [PubMed]

25.　Almeida, T.N.V.; De Sousa, T.T.; Da Silva, R.A.; Fiaccadori, F.S.; Souza, M.; Badr, K.R.; de Paula Cardoso, D.d.D. Phylogenetic analysis of G1P[8] and G12P[8] rotavirus A samples obtained in the pre- and post-vaccine periods, and molecular modeling of VP4 and VP7 proteins. *Acta Trop.* **2017**, *173*, 153–159. [CrossRef] [PubMed]

26. Magagula, N.B.; Esona, M.D.; Nyaga, M.M.; Stucker, K.M.; Halpin, R.A.; Stockwell, T.B.; Seheri, M.L.; Steele, A.D.; Wentworth, D.E.; Mphahlele, M.J. Whole genome analyses of G1P[8] rotavirus strains from vaccinated and non-vaccinated South African children presenting with diarrhea. *J. Med. Virol.* **2015**, *87*, 79–101. [CrossRef] [PubMed]

27. Damanka, S.; Kwofie, S.; Dennis, F.E.; Lartey, B.L.; Agbemabiese, C.A.; Doan, Y.H.; Adiku, T.K.; Katayama, K.; Enweronu-Laryea, C.C.; Armah, G.E. Whole genome characterization and evolutionary analysis of OP354-like P[8] Rotavirus A strains isolated from Ghanaian children with diarrhoea. *PLoS ONE* **2019**, *14*, e0218348. [CrossRef] [PubMed]

28. Jere, K.C.; Chaguza, C.; Bar-Zeev, N.; Lowe, J.; Peno, C.; Kumwenda, B.; Nakagomi, O.; Tate, J.E.; Parashar, U.D.; Heyderman, R.S.; 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.* **2017**, *92*, e01246-17. [CrossRef] [PubMed]

29. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef]

30. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [CrossRef]

31. Aoki, S.T.; Settembre, E.C.; Trask, S.D.; Greenberg, H.B.; Harrison, S.C.; Dormitzer, P.R. Structure of Rotavirus Outer-Layer Protein VP7 Bound with a Neutralizing Fab. *Science* **2009**, *324*, 1444–1447. [CrossRef]

32. Dormitzer, P.R.; Sun, Z.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.* **2002**, *21*, 885–897. [CrossRef]

33. Dormitzer, P.R.; Nason, E.B.; Prasad, B.V.V.; Harrison, S.C. Structural rearrangements in the membrane penetration protein of a non-enveloped virus. *Nat. Cell Biol.* **2004**, *430*, 1053–1058. [CrossRef]

34. Da Silva, M.F.M.; Rose, T.L.; Gómez, M.M.; Carvalho-Costa, F.A.; Fialho, A.M.; De Assis, R.M.; Sde Andrade, J.d.S.R.; Volotão, E.D.M.; Leite, J.P.G. 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.* **2015**, *30*, 206–218. [CrossRef] [PubMed]

35. Zeller, M.; Donato, C.; Trovão, N.S.; Cowley, D.; Heylen, E.; Donker, N.C.; McAllen, J.K.; Akopov, A.; Kirkness, E.F.; Lemey, P.; et al. Genome-Wide Evolutionary Analyses of G1P[8] Strains Isolated Before and After Rotavirus Vaccine Introduction. *Genome Biol. Evol.* **2015**, *7*, 2473–2483. [CrossRef] [PubMed]

36. Rasebotsa, S.; Mwangi, P.N.; Mogotsi, M.T.; Sabiu, S.; Magagula, N.B.; Rakau, K.; Uwimana, J.; Mutesa, L.; Muganga, N.; Murenzi, D.; et al. Whole genome and in-silico analyses of G1P[8] rotavirus strains from pre- and post-vaccination periods in Rwanda. *Sci. Rep.* **2020**, *10*, 1–22. [CrossRef] [PubMed]

37. Mwangi, P.N.; Mogotsi, M.; Rasebotsa, S.P.; Seheri, M.L.; Mphahlele, M.J.; Ndze, V.N.; Dennis, F.E.; Jere, K.C.; Nyaga, M.M. Uncovering the First Atypical DS-1-like G1P[8] Rotavirus Strains That Circulated during Pre-Rotavirus Vaccine Introduction Era in South Africa. *Pathogens* **2020**, *9*, 391. [CrossRef]

38. Donato, C.M.; Ch'Ng, L.S.; Boniface, K.F.; Crawford, N.W.; Buttery, J.P.; Lyon, M.; Bishop, R.F.; Kirkwood, C.D. Identification of Strains of RotaTeq Rotavirus Vaccine in Infants With Gastroenteritis Following Routine Vaccination. *J. Infect. Dis.* **2012**, *206*, 377–383. [CrossRef]

39. Gower, C.M.; Dunning, J.; Nawaz, S.; Allen, D.; Ramsay, M.E.; Ladhani, S.N. Vaccine-derived rotavirus strains in infants in England. *Arch. Dis. Child.* **2019**, *105*, 553–557. [CrossRef]

40. Arora, R.; Chitambar, S. Full genomic analysis of Indian G1P[8] rotavirus strains. *Infect. Genet. Evol.* **2011**, *11*, 504–511. [CrossRef]

41. Kulkarni, R.; Arora, R.; Arora, R.; Chitambar, S.D. Sequence analysis of VP7 and VP4 genes of G1P[8] rotaviruses circulating among diarrhoeic children in Pune, India: A comparison with Rotarix and RotaTeq vaccine strains. *Vaccine* **2014**, *32*, A75–A83. [CrossRef]

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

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

44. Potgieter, A.C.; Page, N.A.; Liebenberg, J.; Wright, I.M.; Landt, O.; Van Dijk, A. Improved strategies for sequence-independent amplification and sequencing of viral double-stranded RNA genomes. *J. Gen. Virol.* **2009**, *90*, 1423–1432. [CrossRef] [PubMed]

45. Pickett, B.E.; Sadat, E.L.; Zhang, Y.; Noronha, J.M.; Squires, R.B.; Hunt, V.; Liu, M.; Kumar, S.; Zaremba, S.; Gu, Z.; et al. ViPR: An open bioinformatics database and analysis resource for virology research. *Nucleic Acids Res.* **2012**, *40*, D593–D598. [CrossRef] [PubMed]

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

47. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef] [PubMed]

48. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [CrossRef]

49. Hazkani-Covo, E.; Graur, D. A Comparative Analysis of numt Evolution in Human and Chimpanzee. *Mol. Biol. Evol.* **2006**, *24*, 13–18. [CrossRef]

50. Nei, M.; Kumar, S. *Molecular Evolution and Phylogenetics*; Oxford University Press: Oxford, UK, 2000.

51. Hasegawa, M.; Kishino, H.; Yano, T.-A. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **1985**, *22*, 160–174. [CrossRef]

52. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [CrossRef]

53. Rambaut, A.; Lam, T.T.; Carvalho, L.M.; Pybus, O.G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2016**, *2*, vew007. [CrossRef]

54. Suchard, M.A.; Lemey, P.; Baele, G.; Ayres, D.L.; Drummond, A.J.; Rambaut, A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **2018**, *4*, vey016. [CrossRef]

55. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [CrossRef] [PubMed]