



Emergence of Double- and Triple-Gene Reassortant G1P[8] Rotaviruses Possessing a DS-1-Like Backbone after Rotavirus Vaccine Introduction in Malawi

✉ Khuzwayo C. Jere,^{a,b} ✉ Chrispin Chaguza,^{a,b} Naor Bar-Zeev,^{a,b,c} Jenna Lowe,^a Chikondi Peno,^b Benjamin Kumwenda,^b Osamu Nakagomi,^{a,d} Jacqueline E. Tate,^e Umesh D. Parashar,^e Robert S. Heyderman,^{b,f} Neil French,^{b,c} Nigel A. Cunliffe,^c Miren Iturriza-Gomara^{c,g} for the VACSURV Consortium

^aInstitute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom

^bMalawi-Liverpool-Wellcome Trust Clinical Research Programme/Department of Medical Laboratory Sciences, College of Medicine, University of Malawi, Blantyre, Malawi

^cCentre for Global Vaccine Research, Institute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom

^dGraduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan

^eEpidemiology Branch, Division of Viral Diseases, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

^fDivision of Infection and Immunity, University College London, London, United Kingdom

^gNIHR Health Protection Research Unit in Gastrointestinal Infections, University of Liverpool, Liverpool, United Kingdom

ABSTRACT To combat the high burden of rotavirus gastroenteritis, multiple African countries have introduced rotavirus vaccines into their childhood immunization programs. Malawi incorporated a G1P[8] rotavirus vaccine (Rotarix) into its immunization schedule in 2012. Utilizing a surveillance platform of hospitalized rotavirus gastroenteritis cases, we examined the phylodynamics of G1P[8] rotavirus strains that circulated in Malawi before (1998 to 2012) and after (2013 to 2014) vaccine introduction. Analysis of whole genomes obtained through next-generation sequencing revealed that all randomly selected prevaccine G1P[8] strains sequenced ($n = 32$) possessed a Wa-like genetic constellation, whereas postvaccine G1P[8] strains ($n = 18$) had a DS-1-like constellation. Phylodynamic analyses indicated that postvaccine G1P[8] strains emerged through reassortment events between human Wa- and DS-1-like rotaviruses that circulated in Malawi from the 1990s and hence were classified as atypical DS-1-like reassortants. The time to the most recent common ancestor for G1P[8] strains was from 1981 to 1994; their evolutionary rates ranged from 9.7×10^{-4} to 4.1×10^{-3} nucleotide substitutions/site/year. Three distinct G1P[8] lineages chronologically replaced each other between 1998 and 2014. Genetic drift was the likely driver for lineage turnover in 2005, whereas replacement in 2013 was due to reassortment. Amino acid substitution within the outer glycoprotein VP7 of G1P[8] strains had no impact on the structural conformation of the antigenic regions, suggesting that it is unlikely that they would affect recognition by vaccine-induced neutralizing antibodies. While the emergence of DS-1-like G1P[8] rotavirus reassortants in Malawi was therefore likely due to natural genotype variation, vaccine effectiveness against such strains needs careful evaluation.

IMPORTANCE The error-prone RNA-dependent RNA polymerase and the segmented RNA genome predispose rotaviruses to genetic mutation and genome reassortment, respectively. These evolutionary mechanisms generate novel strains and have the potential to lead to the emergence of vaccine escape mutants. While multiple African countries have introduced a rotavirus vaccine, there are few data describing the evolution of rotaviruses that circulated before and after vaccine introduction. We re-

Received 20 July 2017 Accepted 3 November 2017

Accepted manuscript posted online 15 November 2017

Citation Jere KC, Chaguza C, Bar-Zeev N, Lowe J, Peno C, Kumwenda B, Nakagomi O, Tate JE, Parashar UD, Heyderman RS, French N, Cunliffe NA, Iturriza-Gomara M, VACSURV Consortium. 2018. 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:e01246-17. <https://doi.org/10.1128/JVI.01246-17>.

Editor Tom Gallagher, Loyola University Medical Center

Copyright © 2018 Jere et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Khuzwayo C. Jere, Khuzwayo.Jere@liverpool.ac.uk.

N.A.C. and M.I.-G. contributed equally.

port the emergence of atypical DS-1-like G1P[8] strains during the postvaccine era in Malawi. Three distinct G1P[8] lineages circulated chronologically from 1998 to 2014; mutation and reassortment drove lineage turnover in 2005 and 2013, respectively. Amino acid substitutions within the outer capsid VP7 glycoprotein did not affect the structural conformation of mapped antigenic sites, suggesting a limited effect on the recognition of G1-specific vaccine-derived antibodies. The genes that constitute the remaining genetic backbone may play important roles in immune evasion, and vaccine effectiveness against such atypical strains needs careful evaluation.

KEYWORDS rotavirus, phylodynamics, genome reassortment, lineage turnover, whole-genome sequencing, Malawi

Diarrhea is a leading cause of mortality in children under the age of 5 years globally (1, 2). The majority of hospitalizations and deaths of infants due to severe dehydrating diarrhea are caused by group A rotaviruses (RVAs) (3). The World Health Organization (WHO) recommended the universal introduction of rotavirus vaccines in 2009, particularly in countries where the diarrhea mortality rate is high (4). A global decline from 528,000 to 215,000 rotavirus-associated deaths per year among children <5 years of age between 2000 and 2013 has been reported, and live-attenuated oral rotavirus vaccines (Rotarix [RV1] and RotaTeq [RV5]) have now been incorporated into the national immunization programs of over 60 countries worldwide (5).

RVAs are members of the *Reoviridae* virus family. They are enveloped icosahedral viruses that contain a triple-layered capsid encasing 11 genome segments of double-stranded RNA (dsRNA). The rotavirus genome encodes six structural proteins (VP1 to VP4, VP6, and VP7) and five to six nonstructural proteins (NSP1 to NSP5/NSP6) (6). Nucleotide homology cutoff values for the open reading frame (ORF) for each genome segment are used to classify rotavirus strains on the basis of the whole-genome composition (7, 8). To date, 35 G (VP7), 50 P (VP4), 26 I (VP6), 21 R (VP1), 19 C (VP2), 19 M (VP3), 30 A (NSP1), 20 N (NSP2), 21 T (NSP3), 26 E (NSP4), and 21 H (NSP5) genotypes have been described (8–12). (<http://rega.kuleuven.be/cev/viralmetagonomics/virus-classification>).

Genotypes G1 to G4, G9, and G12 in association with genotype P[4], P[6], or P[8] are the predominant genotypes associated with human rotavirus infection worldwide (6, 13–15). Although several G and P genotype combinations have been detected among human rotaviruses, the genotypes for the other nine genes are limited to predominantly genotype 1 (I1-R1-C1-M1-A1-N1-T1-E1-H1; Wa-like) and genotype 2 (I2-R2-C2-M2-A2-N2-T2-E2-H2; DS-1-like) (16). For instance, typically, G1P[8], G3P[8], G4P[8], G9P[8], and G12P[8] RVAs have a Wa-like genotype constellation, whereas G2P[4] and G8P[4] or G8P[6] strains usually possess a DS-1-like constellation (16–18). The segmented RNA genome of rotaviruses and their error-prone RNA-dependent RNA polymerase, which lacks proofreading activity (6), allow various evolutionary mechanisms, including genetic mutation, recombination, and reassortment. This leads to the emergence of distinct lineages within individual genotypes or reassortant viruses containing segments from different progenitor strains (6, 19, 20).

Novel double-reassortant DS-1-like G1P[8] rotaviruses have recently emerged in Southeast Asia. These atypical G1 strains were initially detected in outbreaks of gastroenteritis among Japanese children (21–23), followed by reports from Thailand (24, 25) and then Vietnam (26). To date, there is no evidence that these atypical G1 strains are widespread. Rotavirus strain surveillance conducted in Blantyre, Malawi, since 1997, and the introduction of the monovalent, Wa-like G1P[8] rotavirus vaccine (Rotarix or RV1) into Malawi's childhood immunization program in 2012 allow the study of the evolution of G1P[8] strains before and after vaccine introduction. We have, for the first time, detected DS-1-like G1P[8] rotavirus strains in Africa, which became predominant following vaccine introduction. The evolutionary forces that were associated with the emergence of the atypical G1P[8] rotavirus strains were also determined.

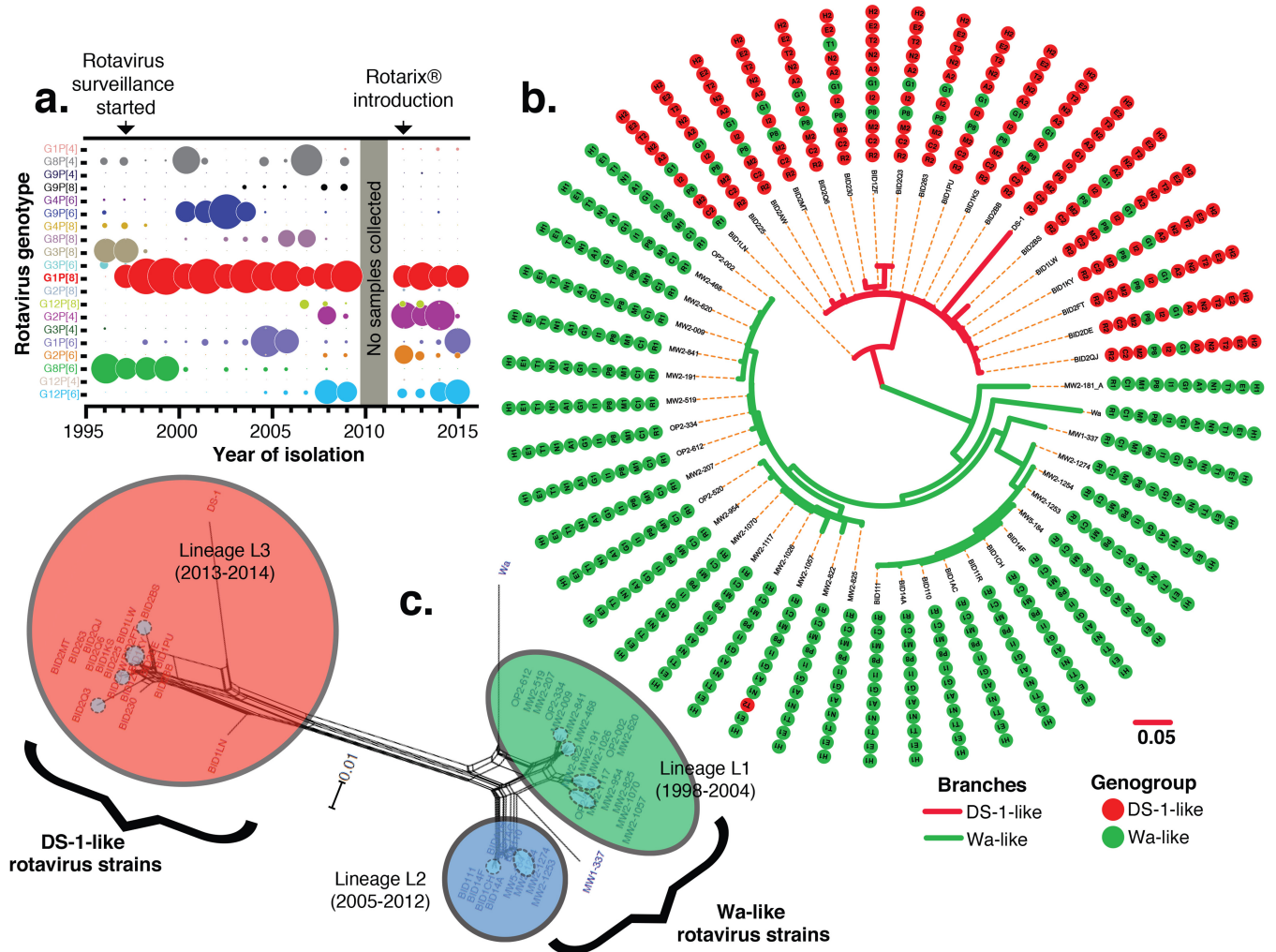


FIG 1 Rotavirus strains, genetic constellations, and phylogenetic networks of G1P[8] rotaviruses detected in Malawian children at QECH from 1997 to 2015. (a) Schematic representation of the proportions of all genotypes detected in rotavirus-positive stool samples. The size of the circle is directly proportional to the frequency of detection of G and P genotypes. There were no rotavirus surveillance activities in 2010; hence, samples were not collected at this time. (b) Bayesian maximum clade credibility tree for concatenated whole-rotavirus-genome sequences illustrating the genetic constellation and reassortment patterns for 32 pre- and 18 postvaccine G1P[8] strains from Malawi as well as the prototype Wa and DS-1 strains for comparison. Genotype and genogroup assignments for each segment were based on nucleotide sequence identities, assigned by using RotaC. Genome segments that were assigned the same genotype are shown with the same color and genotype numbers. (c) Phylogenetic network of complete concatenated whole-genome sequences of G1P[8] rotavirus strains detected in Malawi from 1998 to 2014. Branches are drawn to scale, and splits in the network indicate reassortments. Network clusters are color-coded and named in accordance with their phylogenetic lineages (L1 to L3) that correlated with the time of strain isolation before or after rotavirus vaccine introduction. Clusters L1 (green) and L2 (blue) contain prevaccine G1P[8] strains, whereas L3 (red) contains postvaccine strains. Network subclusters within each main cluster are shaded in blue (L1), red (L2), or orange (L3).

RESULTS

Emergence of reassortant DS-1-like G1P[8] rotavirus strains. All prevaccine (33.4%; 1,634/4,888) and postvaccine (22.6%; 477/2,109) rotavirus-positive stool samples collected from children presenting with acute severe diarrhea at Queen Elizabeth Central Hospital (QECH) in Blantyre, Malawi, were genotyped as part of ongoing rotavirus surveillance (27–30). Multiple strains were characterized, and G1P[8] rotaviruses were consistently predominant strains that were detected each year before (39.4%; 554/1406) and after (31.4%; 95/303) vaccine introduction (Fig. 1a; see also Fig. S1 in the supplemental material). Whole-genome sequences of 32 prevaccine G1P[8] strains (collected between 1998 and 2012) and 18 postvaccine G1P[8] strains (collected from 2013 to 2014) were successfully generated (see Table S1 in the supplemental material for the yearly distribution). Interruption of surveillance in 2010 meant that no G1P[8] strains were available between 2010 and 2011. Among the prevaccine G1P[8] strains, 31

TABLE 1 Evolutionary selective forces and recombination in all 11 proteins of the Malawian G1P[8] rotavirus strains^a

Protein	dN/dS by SLAC	ω (β/α) by FUBAR
VP1	0.0502	2.22
VP2	0.0538	2.67
VP3	0.0965	2.47
VP4	0.1052	3.15
VP6	0.0310	4.66
VP7	0.2033	6.01
NSP1	0.2239	4.21
NSP2	0.0920	5.62
NSP3	0.0870	5.44
NSP4	0.1097	7.10
NSP5	0.1097	7.47

^aShown are posterior distributions of synonymous (α) and nonsynonymous (β) substitution rates over sites as well as mean posterior probabilities for a ω (β/α) value of <1 at a site. The consensus selective force in all cases was purifying selection, and there was no recombination determined by GARD or SBP.

had the G1-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1 genotype constellation and hence were designated Wa-like G1P[8] strains, whereas 1 had the G1-P[8]-I1-R1-C1-M1-A1-N1-T2-E1-H1 genotype constellation and hence was designated a monoreassortant Wa-like G1P[8] strain. In contrast, 16/18 of the postvaccine G1P[8] strains had a DS-1-like genotype constellation (G1-P[8]-I2-R2-C2-M2-A2-N2-T2-E2-H2) and hence were designated atypical double-reassortant DS-1-like G1P[8] strains. The remaining two postvaccine G1P[8] strains (BID1LN and BID230) had Wa-like VP1 (R1) and Wa-like NSP3 (T1) genes, respectively, and hence were designated atypical triple-reassortant DS-1-like G1P[8] strains (Fig. 1b and Table S1).

Atypical postvaccine G1P[8] strains emerged through genome reassortment between Wa-like and DS-1-like human rotaviruses. While the concatenated sequences of all 11 genome segments of pre- and postvaccine G1P[8] rotaviruses clustered with prototype Wa- and DS-1-like strains, respectively (Fig. 1b), phylogenetic networks constructed by using characterized concatenated whole-genome sequences of G1P[8] strains revealed frequent reassortment events between strains from the same or different network clusters (Fig. 1c). The Malawi G1P[8] strains were distributed into three main phylogenetic network clusters (L1, L2, and L3), which also contained multiple subclusters (shaded within main clusters in Fig. 1c). The L1 phylogenetic cluster contained strains that circulated from 1998 to 2004, and strains detected from 2005 to 2012 were grouped into the L2 cluster, while the L3 cluster consisted only of postvaccine G1P[8] strains that circulated from 2013 to at least 2014 in Blantyre.

The phylogenetic relationship of Malawi G1P[8] strains was inferred by using the maximum likelihood method, for which complete nucleotide sequences of RVAs available in GenBank and sequences of Wa-like (VP4 and VP7 only) and DS-1-like (the other nine genes) strains from Malawi were used. The DS-1-like genes of the postvaccine G1P[8] reassortants clustered closely with those of G2 strains that circulated simultaneously in Malawi (see Fig. S2 in the supplemental material) and not with DS-1-like G1P[8] strains identified in Southeast Asia and Japan. Sequences within each cluster were 95 to 99.8% similar (calculated by using nucleotide identity matrices [data not shown]).

All 11 genome segments of the G1P[8] strains were undergoing purifying selection, thus potentially resulting in stabilizing selection following the purging of deleterious variants arising during error-prone rotavirus replication due to its RNA genome (Table 1). The genetic algorithm for recombination detection (GARD) (31) and single-breakpoint recombination (SBP) (32) did not identify any recombination events within each genome segment of the study strains (Table 1). Thus, the changes in the genetic constellation of postvaccine G1P[8] strains were likely generated through the exchange of a whole gene segment (reassortment) between circulating Wa-like and DS-1-like human rotaviruses.

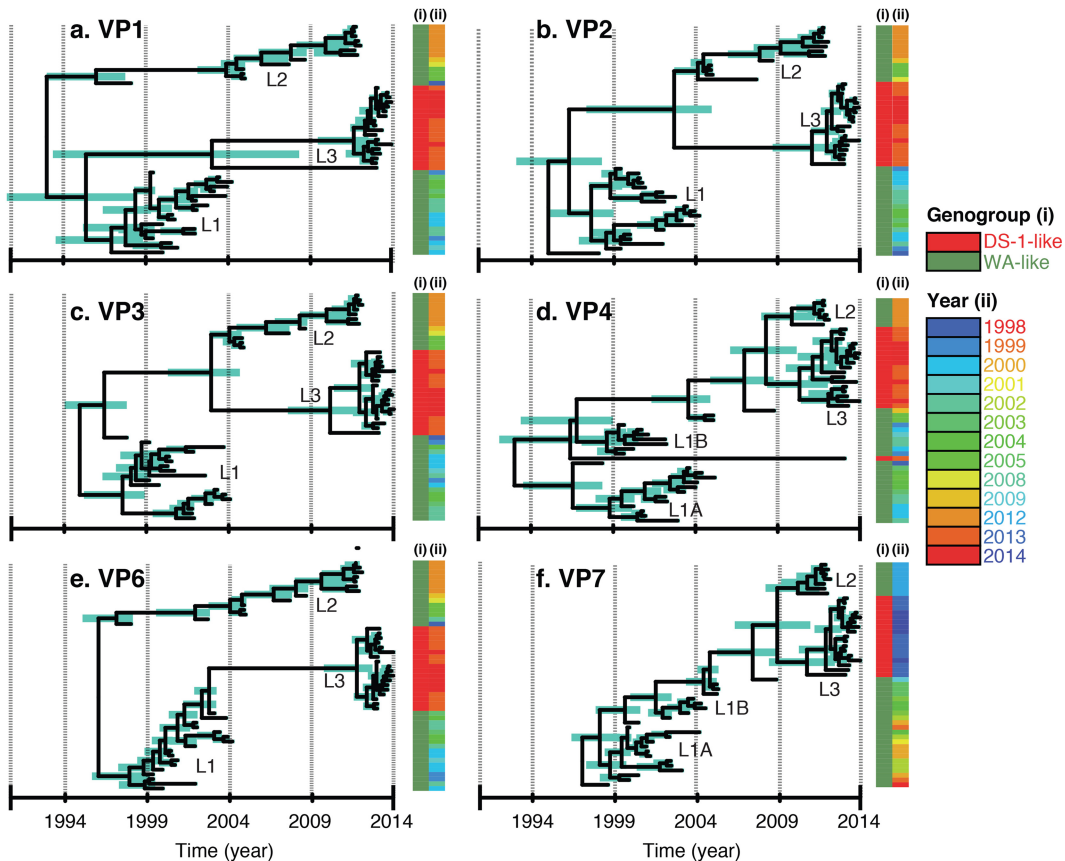


FIG 2 Bayesian maximum clade credibility (MCC) time tree based on complete nucleotide sequences illustrating lineage replacement within the genome segments encoding structural proteins of the G1P[8] strains that circulated in Malawi from 1998 to 2014. With the exception of VP4 and VP7 genes that had L2 and L3 genes sharing close ancestry, the rest had three distinct G1P[8] lineages. L1, L2, and L3 represent lineages 1, 2, and 3, respectively.

G1P[8] strains that circulated in Malawi between 1998 and 2014 exhibit distinct replacement dynamic patterns.

Bayesian inferences of time-measured trees were individually constructed for each of the 11 genome segments of the G1P[8] strains to further determine their evolutionary dynamics. As illustrated by phylogenetic networks, prior to the emergence of the reassortant L3 lineage, both structural and nonstructural genes segregated into at least 2 distinct lineages with a common ancestor in the mid-1990s or before (Fig. 2; see also Fig. S3 in the supplemental material). For NSP1, NSP2, NSP3, VP1, VP2, VP3, and VP6, a single lineage (L1) was predominant until the mid-2000s (2003 to 2005), and following its disappearance, strains forming a second lineage, L2, circulated until the emergence of the DS-1-like G1P[8] reassortant strains. For the VP4, VP7, and NSP4 genes, two clusters cocirculated until 2004, and replacement with strains in L2 did not occur until later, between 2010 and 2012. In 2013, VP7- and VP4-encoding genes of the emergent reassortant strains also formed a third cluster with a likely common ancestor with strains in L2. Both maximum clade credibility (MCC) coalescent framework and maximum likelihood phylogenetic approaches showed that G1P[8] strains acquired DS-1-like genome segments (genotype 2) in the nine non-VP4 and non-VP7 genes in 2013, outcompeting typical Wa-like L2 variants (genotype 1) during the postvaccine era (Fig. 2 and Fig. S2 to S4). These analyses also showed that the post-2013 Malawi DS-1-like G1P[8] reassortant strains clustered with or were derived from Wa-like G1P[8] strains and DS-1-like G2 strains that cocirculated in Malawi and hence emerged through reassortment among local strains. The genes of Malawi DS-1-like G1P[8] strains clustered away from those of DS-1-like reassortants that emerged recently in southeast Asian countries and Japan, indicating that they were likely not imported.

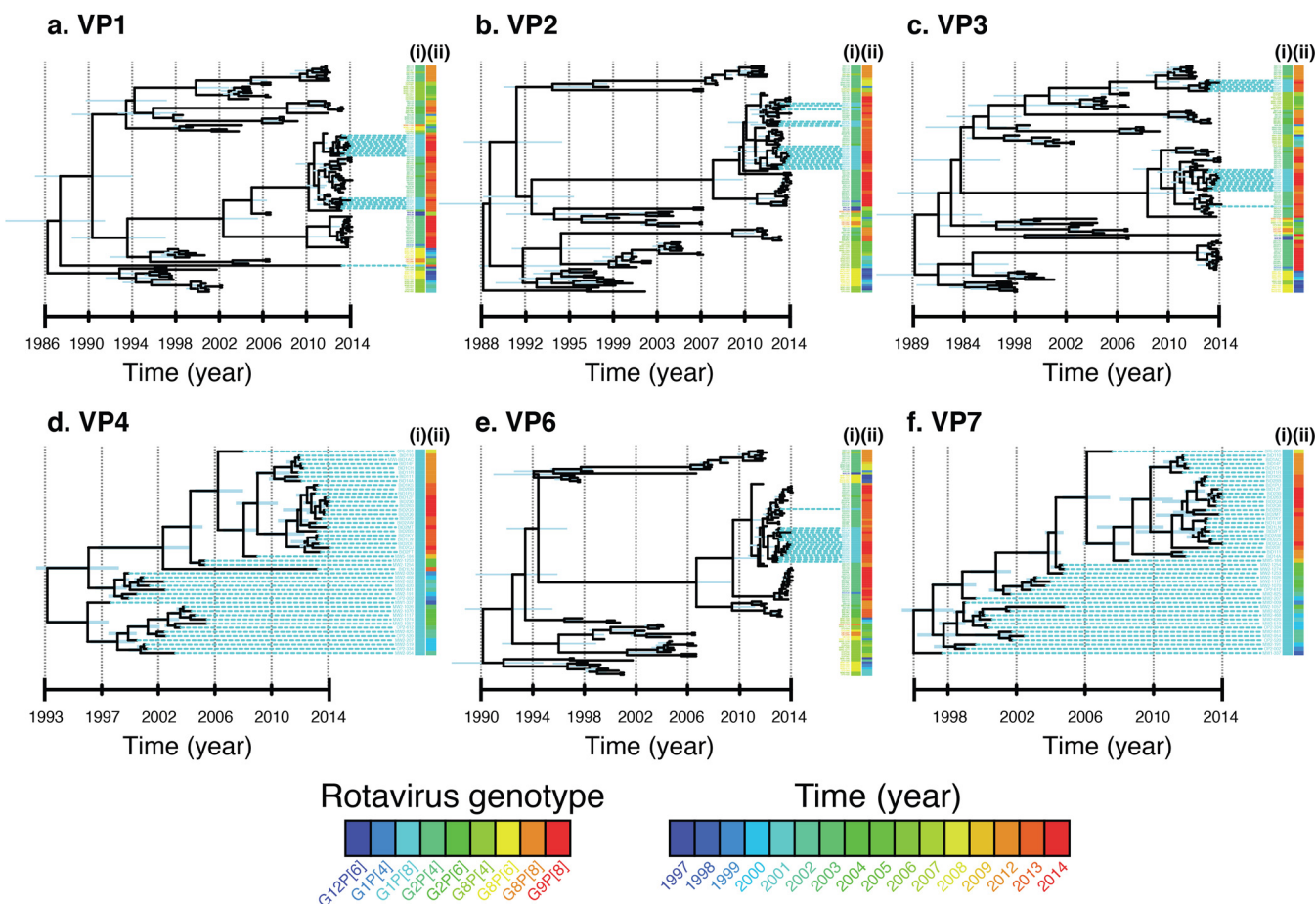


FIG 3 Bayesian MCC time tree based on complete nucleotide sequences of the structural proteins for G1P[8] strains from Malawi. Only DS-1-like genome segments for typical DS-1-like strains that were assigned G2P[4], G2P[6], G8P[4], G8P[6], and G12P[6] outer capsid genotypes from Malawi were included to calculate evolutionary dynamics for VP1- to VP4-, VP6-, and VP7-encoding genome segments for the atypical DS-1-like G1P[8] strains (L3 cluster).

Since the postvaccine Malawi G1P[8] strains contained a DS-1-like genetic backbone, only sequences for the Wa-like genome segments for G1P[8] strains generated in this study were used for phylogenetic analysis of VP4 and VP7 genes, whereas cognate genes of DS-1-like strains that were assigned various G and P types (Table S1), which also circulated in Malawi from 1997 to 2014, were used to estimate evolutionary dynamics for non-VP4 and non-VP7 genes for postvaccine G1P[8] strains (Fig. 3 and Fig. S4). The calculated mean times per gene to the most recent common ancestor (tMRCAs) ranged from the years 1986 to 1996 (Fig. 4a). When only non-VP4 and non-VP7 genes (NSP1 to NSP4, VP1 to VP3, and VP6) for reassortant DS-1-like G1P[8] and cognate genes of DS-1-like strains collected from Malawi between 1997 and 2014 were used to infer Bayesian time-measured trees, the tMRCAs for the atypical G1P[8] strains (L3 cluster) was estimated to range from 2009 to 2011, which was similar to predominantly DS-1-like G2 strains that were detected after vaccine introduction (Fig. 3 and Fig. S4). Marginal differences between the mean evolutionary rates for each genome segment were observed, which ranged from 9.7×10^{-4} to 4.1×10^{-3} nucleotide substitutions per site per year (Fig. 4b). VP2 had the lowest mutation rates (9.7×10^{-4} nucleotide substitutions per site per year; 95% highest posterior density [HPD] interval, 7.4×10^{-4} to 1.2×10^{-3} substitutions per site per year), whereas VP3 had the highest mutation rates (4.1×10^{-3} substitutions per site per year; HPD, 3.1×10^{-3} to 5×10^{-3} substitutions per site per year) (Fig. 3 and 4b and Fig. S4).

We then utilized a Gaussian Markov random field (GMRF) tree prior to investigate whether Rotarix introduction had an impact on the relative population size of the

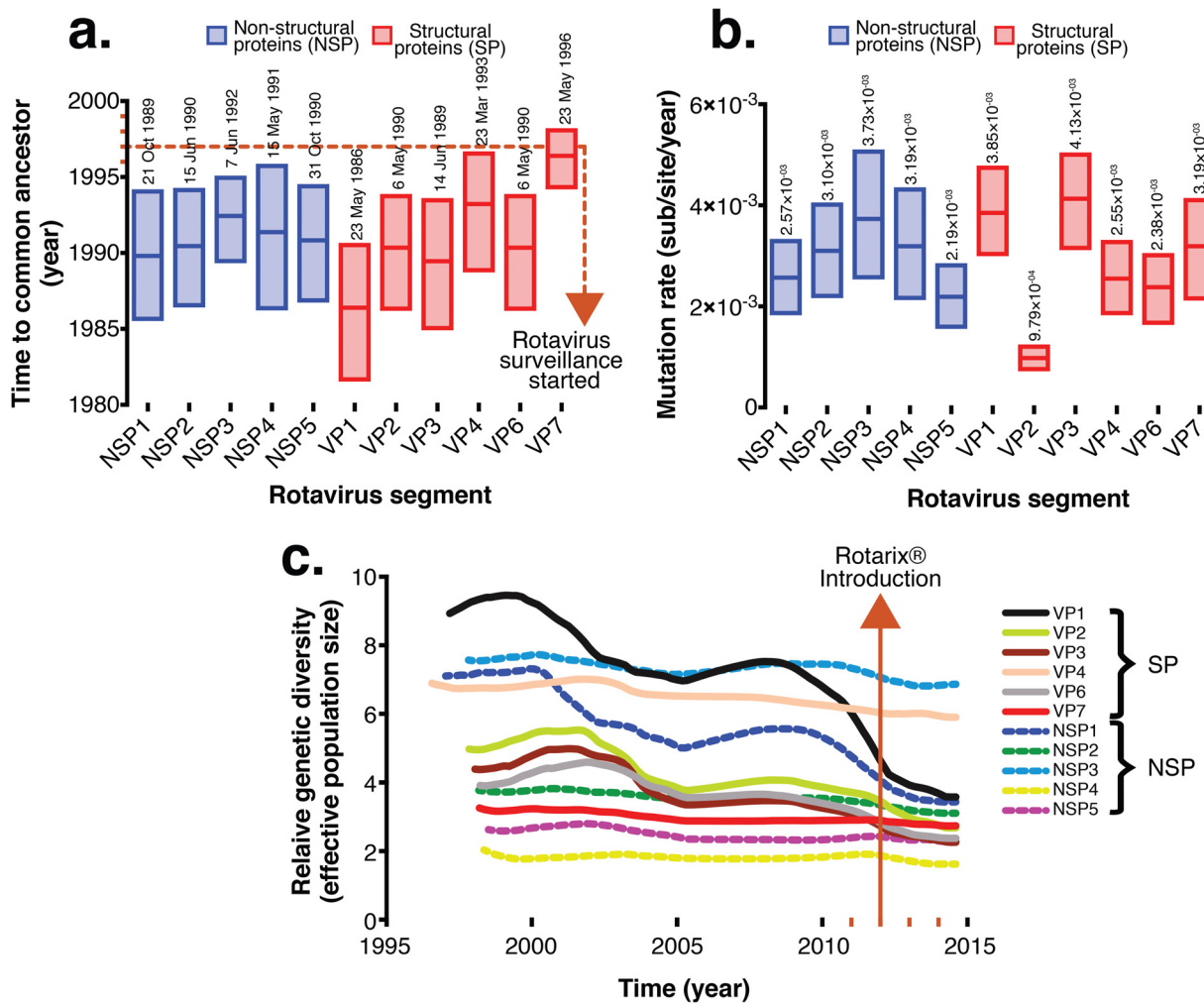


FIG 4 Time to the most recent common ancestor (tMRCAs), evolutionary rates for each genome segment of Malawian G1P[8] strains, and comparative population dynamics of G1P[8] rotavirus strains circulating in Malawi, 1998 to 2014. (a) Evolutionary rates and tMRCAs for each genome segment of the atypical Malawian G1P[8] strains shown together with their 95% highest posterior density (HPD) intervals. (b) Absolute values for the means and ranges of evolutionary rates and tMRCAs at 95% HPD intervals. (c) Phylogenies and relative genetic diversities estimated by using the Gaussian Markov random field (GMRF) model represent Bayesian Skygrid plots for VP1- to VP4-, VP6-, and NSP1- to NSP5-encoding genome segments. Lines in the GMRF plot represent the mean relative genetic diversity through time.

circulating G1P[8] strain. There was no evidence to suggest that vaccine introduction affected either G1P[8] genetic diversity or population size for the VP4, VP7, NSP2, NSP3, and NSP5 genes, as the peaks and troughs of their Skygrid plots exhibited similar stable profiles just before (2005 to 2012) and after (2013 to 2014) vaccine use. In contrast, genes encoding VP1 to VP3, VP6, and NSP1 had relatively stable profiles and also small effective population sizes during the postvaccine era compared to those before vaccine introduction, which could be natural, as similar downward trends were already occurring before vaccine introduction (Fig. 4c).

Mutations within VP7 antigenic regions did not affect the structural conformation of neutralizing epitopes essential for antigenic recognition by neutralizing antibodies. Mapped amino acid motifs that constitute neutralizing epitopes on the outer capsid glycoprotein were compared between Rotarix and Malawian G1P[8] strains. In total, 15 lineage-defining amino acid substitutions were identified across the entire VP7 sequence, with only 5 of these being located at mapped antigenic regions 7-1a (S123N and K291R) and 7-2 (AR C [M202T, M212T, and N221S]) (Fig. 5a). A single amino acid substitution, N221S, located at one of these antigenic regions, differentiated L3 from L2 cluster strains (Fig. 5).

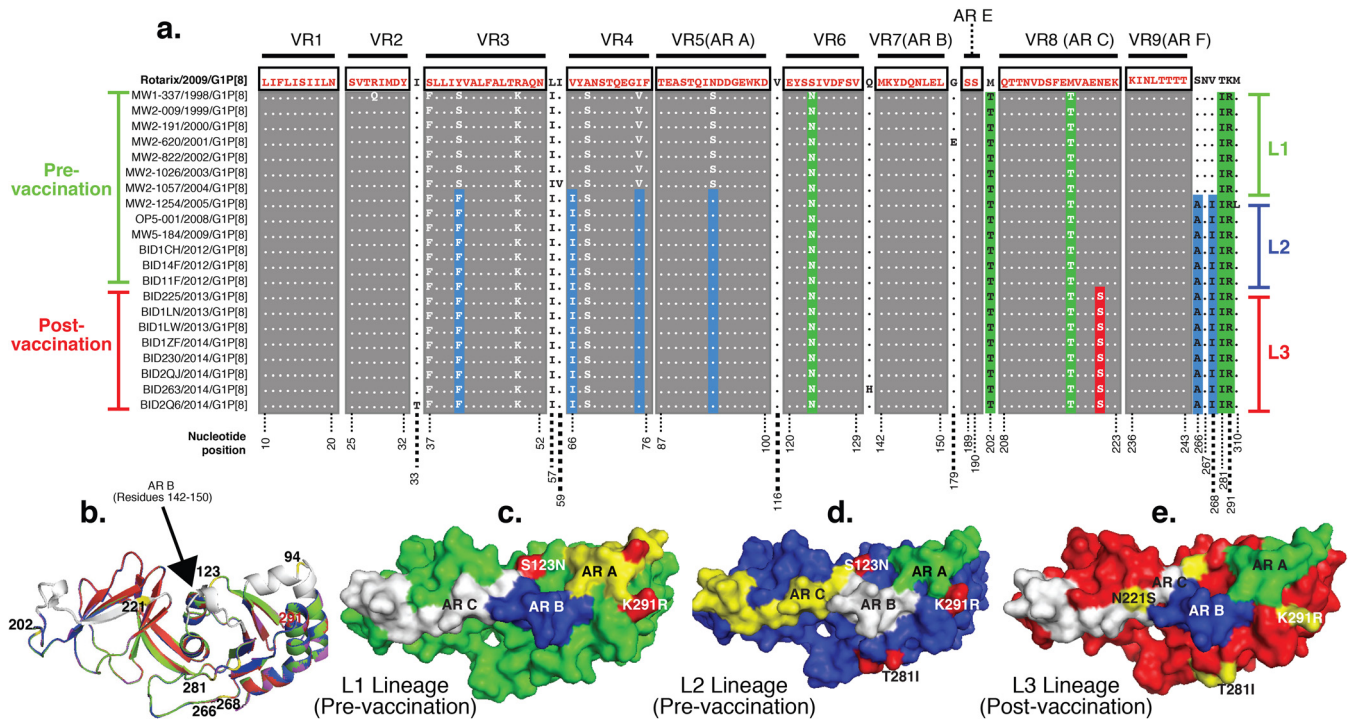


FIG 5 Amino acid substitutions and structural conformation of the outer capsid glycoprotein of Malawian G1P[8] strains. (a) Complete VP7 sequences of representative pre- and postvaccine G1P[8] strains aligned to that of RV1 exhibiting amino acid substitutions that occurred within the variable regions (VR) and mapped antigenic regions (AR) over time. Lineage-defining amino acid amino acid substitutions are highlighted in green, blue, and yellow for the L1, L2, and L3 lineages, respectively. Pre- and postvaccine strains are shown with vertical green and red bars on the right, respectively. Strains belonging to the L1, L2, and L3 phylogenetic clusters are shown with green, blue, and red bars, respectively, on the right. (b) Perfect alignment of superimposed VP7 structures exhibiting few differences between RV1 and L1 to L3 strains. Antigenic regions A, B, and C are shown in white. L1 to L3 and RV1 strains are shown in yellow, green, blue, and red, respectively. (c to e) Surface visualization of VP7 from outside the virion on the 3-fold axis displaying amino acid differences when structures for L1 (c), L2 (d), and L3 (e) G1P[8] strains were superimposed on the structure of the outer capsid glycoprotein of RV1. Numbers correspond to the positions where mutations occurred.

The VP7 structures of L1 to L3 strains and that of the Rotarix strain aligned perfectly when superimposed (Fig. 5b), implying a conserved conformation consistent with the conservation in chemical properties of the replacing amino acids. The replacing amino acids did not appear to impact the structural conformation of the antigenic regions of the glycoprotein of the G1P[8] strains circulating before and after vaccine introduction (Fig. 5c to e).

DISCUSSION

Malawi was one of the first African countries to introduce a rotavirus vaccine into its infant immunization schedule in October 2012. By 2015, vaccine coverage had exceeded 95% (30). We enhanced rotavirus surveillance activities in the postvaccine period in Malawi primarily to assess the impact of vaccine introduction on the burden of rotavirus disease (27, 30). The availability of a collection of rotavirus strains isolated before (1997 to 2012) (28) and after Rotarix introduction offered a rare opportunity to assess the early impact of vaccine introduction on the genetic diversity of circulating strains and their evolutionary patterns over time. In the present study, phylogenetic analysis and evolutionary history were inferred for G1P[8] rotaviruses, the most prevalent rotavirus strains globally (6). In Malawi, G1P[8] strains were the only rotaviruses consistently detected year to year, from 1998 to 2014, hence enabling a systematic inference of evolutionary patterns over time. Furthermore, G1P[8] strains are homologous, with respect to VP7 and VP4 genotypes, to the Rotarix vaccine that is in use in Malawi (33), hence permitting homologous genomic comparison.

While the prevaccine G1P[8] strains had a typical Wa-like genetic constellation (8, 16, 34, 35), the backbone of the G1P[8] strains detected after vaccine introduction pos-

essed a DS-1-like genotype constellation, which is associated most frequently with G2P[4] strains, occasionally with G9 and G12 strains, and with G8P[4] or G8P[6] strains in Africa (8, 17, 18, 36). The detection of atypical reassortant DS-1-like G1P[8] strains had not been documented during 15 years of prevaccine rotavirus strain surveillance in Malawi or elsewhere in Africa. Similar atypical reassortant DS-1-like G1P[8] strains emerged recently in Southeast Asian countries where rotavirus vaccine use is limited (21–26). Atypical DS-1-like G1P[8] strains from Malawi and Asia exhibited distinct phylogenetic clustering patterns in all 11 genome segments, indicating that these atypical DS-1-like G1P[8] strains most likely emerged independently in Malawi and Asia and did not arise through importation. In addition, (i) tight clustering between the G and P genes of the atypical Malawian G1P[8] strains with those of other G1P[8] Wa-like strains, (ii) monophyletic clustering between non-G and non-P genes of atypical Malawian G1P[8] strains and those of other Malawian DS-1-like strains, and (iii) evidence of frequent reassortment events among DS-1- and Wa-like strains circulating in Malawi during the study period as revealed by phylogenetic networks suggest that the atypical Malawian G1P[8] strains emerged locally through genome reassortment among cocirculating Wa- and DS-1-like strains. This is further supported by the detection of a high prevalence of G2 strains in Malawi from 2012 (27, 30), thus providing the required circulating strains to allow the emergence of Wa/DS-1-like reassortant strains.

The Malawian G1 mutation rate falls in the same range as those of rotavirus G9 and G12 strains (37–41). Zeller et al. (35) recently analyzed the phylodynamics of typical Wa-like G1P[8] strains that circulated before and after vaccine introduction in Belgium and Australia. Unlike the Malawian G1P[8] strains, all G1P[8] strains from Belgium and Australia had a Wa-like genetic constellation, and their tMRCA ranged from 1846 to 1945, whereas their evolutionary rates, which ranged from 6.05×10^{-4} to 1.01×10^{-3} nucleotide substitutions per site per year, were similar to those for Malawian G1P[8] strains. These rates are relatively lower than the known rates of RNA viruses (42), possibly due to the double-stranded nature of the rotavirus genome.

The emergence of the atypical reassortant DS-1-like G1P[8] strains in Malawi coincided with the programmatic rollout of Rotarix. Our data suggest that the atypical Malawian DS-1-like G1P[8] strains were derived from a combination of reassortment and drift of rotaviruses that were circulating locally since the 1990s. We found that at least three G1P[8] lineages were circulating in Malawi from 1998 to 2014. The diversity of the circulating G1P[8] variants exhibited periodic lineage replacements, similarly to influenza virus (43), dengue virus (44), and other enteric viruses such as noroviruses (45), where lineage replacement also appears to be an important evolutionary mechanism in response to herd immunity (35, 46, 47). Lineage diversity and replacement coincided temporally for blocks of genes and can be explained by drift and reassortment events occurring hand in hand.

Although the detection of DS-1-like G1P[8] strains coincided with the widespread use of a G1P[8] Rotarix rotavirus vaccine in Malawi in 2013, it is difficult to ascertain the role that the vaccine had in the emergence of these atypical strains considering the short postvaccine period. The data from phylodynamic analyses suggested that these strains were derived from those strains circulating in the human population well before vaccine introduction. It was difficult to determine the effect of vaccine introduction on the effective population size of the circulating G1P[8] strains, as only postvaccine G1 strains from 2013 and 2014 were analyzed, a time which was too early to detect genetic variations in the virus population size. The frequent detection of DS-1-like G2 strains just before Rotarix introduction and during the postvaccine era in Malawi (27, 30) may indicate a natural surge of DS-1-like rotaviruses during this period, similarly to G2 cyclic seasonal epidemic patterns that have been observed in many countries, including Africa (48, 49). However, the predominance of reassortant DS-1-like G1P[8] strains after vaccine introduction could suggest that there is positive selection for atypical G1P[8] strains and that such selection may not be driven exclusively by the VP7 and VP4 specificities. The DS-1-like genotype constellation was, however, found in rotavirus diarrhea cases regardless of the vaccination status of the children (48% were vaccinated

[data not shown]). The relatively fewer rotavirus G1P[8] genomes that were analyzed are a potential limitation of the present study; as such, sequencing of additional G1P[8] strains that circulated in Malawi post-2014 is under way to determine whether emerging DS-1-like G1P[8] strains remain in circulation.

It was shown previously that mutations along the three main mapped neutralizing epitopes of VP7 can generate vaccine escape mutants (50). For instance, amino acid substitutions at positions 94, 97, 147, and 291 significantly affect the antigenic recognition of human G1 strains (50). When VP7 sequences of postvaccine G1P[8] strains (L3) were compared to that of Rotarix, only the K291R (7-1a) substitution occurred within the sites associated with antibody escape mutants, and this substitution was already present among the strains circulating before vaccine introduction since 1998 and hence was not selected due to potential vaccine pressure. Furthermore, both lysine (K) and arginine (R) are positively charged polar proteins; hence, this substitution is unlikely to produce significant changes in the biochemical properties of VP7. The only substitution present among the postvaccine G1 strains was N221S. As both asparagine (N) and serine (S) are small noncharged residues, this substitution does not appear to have a significant impact on the overall conformation and structure of the protein surface. However, the loss of an asparagine residue may have resulted in the loss of a potential glycosylation site, which could affect VP7 antigenic determinants (51). This change does not appear to be a universal glycosylation position, as serine also occurs naturally at position 221 for some non-G1 strains like S2 (G2), RV-5 (G3), and ST5 (G4) strains (52). While this change is outside the currently proven glycosylation sites (positions 69 to 71, 238 to 240, and 318 to 320), the N221S change occurred within neutralizing epitope C (antigenic region C) of VP7 (52). In order to exclude the potential for this amino acid substitution, further functional studies may be warranted, bearing in mind that changes in immunogenicity and neutralization patterns have been attributed to different glycosylation patterns using mutated laboratory strains (51, 53).

The only hydrophobic-to-hydrophilic change, which would potentially affect the structural conformation and stability of proteins significantly, occurred outside the mapped antigenic regions at position 266 (alanine[A] to S) and was also present among strains circulating pre-2009, those of L2. However, it is possible that substitutions in the non-AR could affect the stability of the viral particle or protein assortment specificities, since VP6 serves as an anchor for the outer capsid VP4 and VP7 proteins, where the 260 trimers of VP7 lie directly on top of the VP6 trimers (54). Contact with VP6 is facilitated by the arm-like extensions formed by the VP7 N termini that also form a lattice with other VP7 trimers. This interaction allows the gripping of VP7 to the intermediate VP6 layer and reinforces the integrity of the outer shell (55). Such interactions may drive the selection of particular VP7 and/or VP4 lineages in reassortant strains and explain lineage replacements that may not necessarily be explained exclusively in terms of immune pressure. In a recent analysis, vaccine effectiveness against all G1P[8] strains 3 years after Rotarix introduction in Malawi was 82% (30), suggesting a high degree of protection against atypical DS-1-like G1P[8] strains, given that these G1P[8] strains were detected in randomly selected stool samples collected between 2013 and 2014. Further analysis is under way in order to assess the extent of the spread of the DS-1-like G1P[8] strains and to calculate vaccine effectiveness against various G and P types possessing Wa- and DS-1-like genetic backbones.

In conclusion, genome reassortment and mutation are the major evolutionary mechanisms that influenced the genetic diversity of G1P[8] strains that circulated in Malawi from 1998 to 2014. Atypical DS-1-like G1P[8] strains emerged in 2013 through genome reassortment events between Wa- and DS-1-like human strains that can be traced back to the 1990s in Malawi. Mutations within outer capsid VP7 of Malawian G1P[8] strains compared to RV1 had no impact on the structural conformation of antigenic regions, suggesting little or no effect on the recognition of vaccine-induced antibodies. Thus, the remaining genome segments (non-G or -P) might also play an important role in immune evasion. It is likely that the atypical DS-1-like G1P[8] strains emerged through natural strain evolutionary pressure, which is unrelated to vaccine

use. However, the predominance of atypical reassortant DS-1-like G1P[8] strains, which coincided with vaccine introduction, could suggest the positive selection of atypical G1P[8] strains that were undergoing purifying selection. Vaccine effectiveness against such atypical strains needs careful investigation.

MATERIALS AND METHODS

Rotavirus strains. Stool samples collected from children aged <5 years presenting with acute gastroenteritis at QECH in Blantyre, Malawi, from 1998 to 2014 were utilized (27–29). Diarrhea case definition, rotavirus screening, VP4 and VP7 genotyping, and descriptions of strains that circulated in Malawi were reported previously (27–29). In total, 4,888 stool specimens were collected before vaccine introduction (1997 to 2012), whereas 2,109 were collected after vaccine introduction (2013 to 2014). Rotaviruses with G1P[8] outer capsid proteins were the only strains that were detected every year from 1998 to 2015 (comprising 554 strains isolated before and 95 strains isolated after vaccine introduction). Therefore, only G1P[8] strains from each surveillance year, and from samples that had sufficient fecal material for dsRNA reextraction, were selected for further examination. Where available, a single fecal sample from each month of each year was randomly selected for whole-genome sequencing; only samples from which whole-genome data were obtained were included in the analysis. Ethical approval was obtained from the National Health Sciences Research Committee, Lilongwe, Malawi (approval number 867), and the Research Ethics Committee of the University of Liverpool, Liverpool, United Kingdom (approval number 000490).

Preparation of purified dsRNA and cDNA for rotavirus whole-genome sequencing. Rotavirus dsRNA was extracted and purified as previously described (17, 56). An additional DNase I treatment step following a lithium chloride precipitation step was included to remove contaminating DNA (Sigma-Aldrich, Dorset, UK). Purified dsRNA was quantified by using a Qubit 3.0 fluorometer (Life Technologies, CA, USA). Sequence-independent cDNA synthesis and PCR amplification procedures described previously (17, 56) were used to amplify cDNA for rotavirus whole genomes from samples with ≥ 2 ng/ μ l dsRNA.

RNA and cDNA library construction and Illumina HiSeq sequencing. After denaturation of dsRNA at 95°C for 5 min, ScriptSeq RNA-Seq Library Preparation kit V2 was used to generate Illumina sequencing libraries for samples that had <2 ng/ μ l dsRNA (Illumina Inc., CA, USA). Purified cDNA generated from samples with > 2 ng/ μ l dsRNA was subjected to standard bar-coding and library construction for Illumina sequencing using a Nextera XT DNA library preparation kit (Illumina Inc., CA, USA). Rotavirus VP6-specific quantitative PCR (qPCR) (57) and a 2100 bioanalyzer (Agilent Technologies Inc., CA, USA) were used to perform quality control on the DNA libraries, followed by sequencing using the HiSeq 2000 platform (Illumina Inc., CA, USA) at the Centre for Genomic Research (CGR), University of Liverpool, United Kingdom.

Sequence assembly and determination of rotavirus genotypes. Illumina adapter sequences were trimmed from the raw Fastq sequence data by using Cutadapt v1.2.1 (58) and Sickle v1.2 software (59). Complete consensus nucleotide sequences were generated by mapping trimmed Illumina sequence reads to various complete nucleotide sequences of prototype rotavirus genogroup strains by using both *de novo* and reference DNA sequence assembler algorithms implemented in Geneious software v8 (60). Rotavirus genotypes were assigned to each of the 11 genome segments by using the RotaC v2.0 Web-based automated rotavirus genotyping tool (<http://rotac.regatools.be/>) (7).

Sequence alignments and maximum likelihood phylogeny construction. Reference nucleotide sequences for each rotavirus genome segment were retrieved from the rotavirus resource in the GenBank database (61). This was followed by multiple-sequence alignment of the assembled sequences for the study strains by using Muscle v3.8.31 (62), included in MEGA v6.0 (63). Initial phylogenetic trees for each segment were inferred by using the maximum likelihood approach implemented in MEGA by selecting the DNA model that best fit the data according to the corrected Akaike information criterion (AICc), as described previously (12). We used a generalized time-reversible (GTR) model with Gamma (γ) heterogeneity across nucleotide sites, while the reliability of the branching order and partitioning were assessed by performing 1,000 bootstrap replicates (64).

Bayesian inference of phylogenies and population dynamics. Coalescence analyses were performed by using BEAUTi v1.7.5 and BEAST v1.8 (65, 66) with the following parameter specifications: lognormal relaxed (uncorrelated) clock model (67), a constant-size coalescent tree prior, a Hasegawa-Kishino-Yano (HKY85) nucleotide substitution model with estimated base frequencies (68), and a γ site heterogeneity model with 4 rate categories (69) and a prior mutation rate (μ) of $\sim 1.0 \times 10^{-3}$ nucleotide substitutions/site/year, as previously reported by Zeller et al. (35). The maximum likelihood trees generated as described above were used as starting trees for the Bayesian analysis in BEAST. We used a GMRF tree prior, which also allows the investigation of the population dynamics, i.e., effective population size ($N_e \tau$) or relative genetic diversity, over time. A total of 200 million Markov chain Monte Carlo (MCMC) iterations were performed and sampled every 40,000th generation. The first 20 million iterations (10% of the total) from the MCMC analysis, burn-in time, were discarded since these iterations may represent states that the chain explored before reaching the equilibrium state of the target distribution. The mean values and 95% HPDs of the mutation rates and the tMRCAs for each rotavirus segment were calculated from the BEAST output by using Tracer v1.6.0 (<https://tree.bio.ed.ac.uk/software/tracer/>). The MCC tree for each viral segment was generated by using Tree Annotator v2.1.2 (<http://beast.bio.ed.ac.uk/treeannotator>) and visualized by using FigTree v1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>) and BioPython scripts (70).

Detection of natural selection in proteins encoded by rotavirus segments. ORFs that encode both structural and nonstructural proteins were identified and extracted from the multiple-sequence alignments of each rotavirus genomic segment. We used BioPython (70) scripts to extract and manipulate the sequence alignments. The ORF alignments for each protein were then converted to the corresponding codon alignments by using CodonAlign (<http://www.hiv.lanl.gov/cgi-bin/CodonAlign>). The alignments were then used to calculate the global ratio of synonymous to nonsynonymous substitutions (dN/dS) for each ORF by using single-likelihood ancestor counting (SLAC) (71) and a fast, unconstrained Bayesian approximation for inferring selection (FUBAR) (72). To identify specific sites under selection in the ORFs, we also used the mixed-effects model of episodic selection (MEME) (73), random-effects likelihood (REL) (74), and fixed-effects likelihood (FEL) (75) methods implemented in DataMonkey, a Web server for the HyPhy package (76). The occurrence of genetic recombination was checked by using GARD (31) and SBP (32). We used the following default significance levels: P , Bayes factor, or posterior probability values of 0.1 for SLAC, MEME, and FEL; 0.9 for FUBAR; and 50 for REL. The HKY85 nucleotide substitution model, beta-gamma site-to-site rate variation, and neighbor-joining trees were used for the selection analysis. All analyses were done by using the DataMonkey Web server (71).

Structure comparison between the outer capsid glycoprotein of RV1 and G1P[8] strains to predict changes in antibody binding. To investigate the likely impact of amino acid substitutions on anti-RV1 antibody recognition due to mutations that occurred over time within the antigenic regions of Malawian G1P[8] rotaviruses, the structural conformations of the outer capsid glycoprotein of pre- and postvaccine strains were compared to that of RV1. Representative VP7 sequences for each G1P[8] lineage were utilized for protein structure modeling using Modeler version 9.17 (77). Templates were searched in the Protein Data Bank (78), using an integrated Web-based HHpred program (79). The best model with the highest zdope score was selected for analysis from the 100 models that were generated for each sequence.

Accession number(s). All complete nucleotide and deduced amino acid sequences generated in this study were deposited in GenBank (61) under accession numbers [MG181227](#) to [MG181941](#).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/JVI.01246-17>.

SUPPLEMENTAL FILE 1, XLS file, 0.1 MB.

SUPPLEMENTAL FILE 2, PDF file, 3.9 MB.

ACKNOWLEDGMENTS

We thank the collaborating members of the VACSURV Consortium (Anthony Costello, Charles Mwansambo, James Beard, Amelia C. Crampin, Carina King, Sonia Lewycka, Hazzie Mvula, Tambosi Phiri, Jennifer R. Verani, Cynthia G. Whitney, Louisa Pollock, and Aisleen Bennett). We acknowledge the support of the laboratory staff at the Malawi-Liverpool-Wellcome Trust Clinical Research Programme and the sequencing and informatics teams at the Centre for Genomic Research (CGR), University of Liverpool, United Kingdom.

This work was supported by a research grant from GlaxoSmithKline Biologicals and the Wellcome Trust (program grant number 091909/Z/10/Z and the MLW Programme Core award). K.C.J. is a Wellcome Trust training fellow (grant number 201945/Z/16/Z). C.C. acknowledges funding from the Commonwealth Scholarship Commission (Ph.D. studentship). M.I.-G. is partly supported by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Gastrointestinal Infections. The research was funded by the NIHR HPRU in Gastrointestinal Infections at the University of Liverpool in partnership with Public Health England (PHE), the University of East Anglia, the University of Oxford, and the Institute of Food Research. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health, or Public Health England. The funders had no role in the study design, data collection and interpretation, or the decision to submit the work for publication. GlaxoSmithKline Biologicals SA was provided the opportunity to review a preliminary version of the manuscript for factual accuracy, but the authors are solely responsible for final content and interpretation. The authors received no financial support or other form of compensation related to the development of the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the U.S. Centers for Disease Control and Prevention.

K.C.J., N.B.-Z., and N. F. have received investigator-initiated research grant support from GlaxoSmithKline Biologicals. M.I.-G. has received investigator-initiated research

grant support from GlaxoSmithKline Biologicals and Sanofi Pasteur Merck Sharpe & Dohme. O.N. has received research grant support and honoraria from Japan Vaccine and MSD for delivering lectures on rotavirus vaccines. N.A.C. has received research grant support and honoraria for participation in rotavirus vaccine advisory board meetings from GlaxoSmithKline Biologicals. All other authors report no potential conflicts.

K.C.J., J.E.T., U.D.P., N.F., R.S.H., N.A.C. and M.I-G. conceived and designed the study. N.A.C. and M.I-G. supervised the study. K.C.J., N.B-Z., C.P., and N.A.C. collected the stool samples. K.C.J., J.L., and C.P. performed the laboratory work. K.C.J. and C.C. carried out the bioinformatic and statistical analyses and drafted the manuscript. B.K. and K.C.J. conducted protein modeling. K.C.J., C.C., N.B-Z., J.L., B.K., C.P., J.E.T., U.D.P., O.N., R.S.H., N.F., N.A.C., and M.I-G. contributed to the discussions, interpretation of the results, and reviewing of the manuscript.

REFERENCES

- Bryce J, Boschi-Pinto C, Shibuya K, Black RE. 2005. WHO estimates of the causes of death in children. *Lancet* 365:1147–1152. [https://doi.org/10.1016/S0140-6736\(05\)71877-8](https://doi.org/10.1016/S0140-6736(05)71877-8).
- Liu L, Oza S, Hogan D, Chu Y, Perin J, Zhu J, Lawn JE, Cousens S, Mathers C, Black RE. 2016. Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the sustainable development goals. *Lancet* 388:3027–3035. [https://doi.org/10.1016/S0140-6736\(16\)31593-8](https://doi.org/10.1016/S0140-6736(16)31593-8).
- Estes MK, Greenberg HB. 2013. Rotaviruses, p 1347–1401. *In* Knipe DM, Howley PM, Cohen JI, Griffin DE, Lamb RA, Martin MA, Racaniello VR, Roizman B (ed), *Fields virology*, 6th ed. Lippincott Williams & Wilkins, Philadelphia, PA.
- WHO. 2007. Rotavirus vaccines. *Wkly Epidemiol Rec* 82:285–295.
- Tate JE, Burton AH, Boschi-Pinto C, Parashar UD. 2016. Global, regional, and national estimates of rotavirus mortality in children <5 years of age, 2000–2013. *Clin Infect Dis* 62(Suppl 2):S96–S105. <https://doi.org/10.1093/cid/civ1013>.
- Estes MK, Kapikian AZ. 2007. Rotaviruses, p 1917–1974. *In* Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE (ed), *Fields virology*, 5th ed. Lippincott Williams & Wilkins, Philadelphia, PA.
- Maes P, Matthijnsens J, Rahman M, Van Ranst M. 2009. RotaC: a Web-based tool for the complete genome classification of group A rotaviruses. *BMC Microbiol* 9:238. <https://doi.org/10.1186/1471-2180-9-238>.
- Matthijnsens J, Ciarlet M, McDonald SM, Attoui H, Banyai K, Brister JR, Buesa J, Esona MD, Estes MK, Gentsch JR, Iturriza-Gomara M, Johne R, Kirkwood CD, Martella V, Mertens PP, Nakagomi O, Parrero V, Rahman M, Ruggeri FM, Saif LJ, Santos N, Steyer A, Taniguchi K, Patton JT, Desselberger U, Van Ranst M. 2011. Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). *Arch Virol* 156:1397–1413. <https://doi.org/10.1007/s00705-011-1006-z>.
- Papp H, Al-Mutairi LZ, Chehadeh W, Farkas SL, Lengyel G, Jakab F, Martella V, Szucs G, Banyai K. 2012. Novel NSP4 genotype in a camel G10P[15] rotavirus strain. *Acta Microbiol Immunol Hung* 59:411–421. <https://doi.org/10.1556/AMicr.59.2012.3.11>.
- Trojnar E, Sachsenroder J, Twardziok S, Reetz J, Otto PH, Johne R. 2013. Identification of an avian group A rotavirus containing a novel VP4 gene with a close relationship to those of mammalian rotaviruses. *J Gen Virol* 94:136–142. <https://doi.org/10.1099/vir.0.047381-0>.
- Guo D, Liu J, Lu Y, Sun Y, Yuan D, Jiang Q, Lin H, Li C, Si C, Qu L. 2012. Full genomic analysis of rabbit rotavirus G3P[14] strain N5 in China: identification of a novel VP6 genotype. *Infect Genet Evol* 12:1567–1576. <https://doi.org/10.1016/j.meegid.2012.06.010>.
- Jere KC, Esona MD, Ali YH, Peenze I, Roy S, Bowen MD, Saeed IK, Khalafalla AI, Nyaga MM, Mphahlele J, Steele D, Seheri ML. 2014. Novel NSP1 genotype characterised in an African camel G8P[11] rotavirus strain. *Infect Genet Evol* 21:58–66. <https://doi.org/10.1016/j.meegid.2013.10.002>.
- Santos N, Hoshino Y. 2005. Global distribution of rotavirus serotypes/genotypes and its implication for the development and implementation of an effective rotavirus vaccine. *Rev Med Virol* 15:29–56. <https://doi.org/10.1002/rmv.448>.
- Banyai K, Laszlo B, Duque J, Steele AD, Nelson EA, Gentsch JR, Parashar UD. 2012. Systematic review of regional and temporal trends in global rotavirus strain diversity in the pre rotavirus vaccine era: insights for understanding the impact of rotavirus vaccination programs. *Vaccine* 30(Suppl 1):A122–A130. <https://doi.org/10.1016/j.vaccine.2011.09.111>.
- Esona MD, Gautam R. 2015. Rotavirus. *Clin Lab Med* 35:363–391. <https://doi.org/10.1016/j.cll.2015.02.012>.
- Matthijnsens J, Van Ranst M. 2012. Genotype constellation and evolution of group A rotaviruses infecting humans. *Curr Opin Virol* 2:426–433. <https://doi.org/10.1016/j.coviro.2012.04.007>.
- Jere KC, Mlera L, O'Neill HG, Potgieter AC, Page NA, Seheri ML, van Dijk AA. 2011. Whole genome analyses of African G2, G8, G9, and G12 rotavirus strains using sequence-independent amplification and 454(R) pyrosequencing. *J Med Virol* 83:2018–2042. <https://doi.org/10.1002/jmv.22207>.
- Matthijnsens J, Ciarlet M, Heiman E, Arijs I, Delbeke T, McDonald SM, Palombo EA, Iturriza-Gomara M, Maes P, Patton JT, Rahman M, Van Ranst M. 2008. Full genome-based classification of rotaviruses reveals a common origin between human Wa-like and porcine rotavirus strains and human DS-1-like and bovine rotavirus strains. *J Virol* 82:3204–3219. <https://doi.org/10.1128/JVI.02257-07>.
- Blackhall J, Fuentes A, Magnusson G. 1996. Genetic stability of a porcine rotavirus RNA segment during repeated plaque isolation. *Virology* 225:181–190. <https://doi.org/10.1006/viro.1996.0586>.
- Gouvea V, Brantly M. 1995. Is rotavirus a population of reassortants? *Trends Microbiol* 3:159–162. [https://doi.org/10.1016/S0966-842X\(00\)88908-8](https://doi.org/10.1016/S0966-842X(00)88908-8).
- Kuzuya M, Fujii R, Hamano M, Kida K, Mizoguchi Y, Kanadani T, Nishimura K, Kishimoto T. 2014. Prevalence and molecular characterization of G1P[8] human rotaviruses possessing DS-1-like VP6, NSP4, and NSP5/6 in Japan. *J Med Virol* 86:1056–1064. <https://doi.org/10.1002/jmv.23746>.
- Fujii Y, Nakagomi T, Nishimura N, Noguchi A, Miura S, Ito H, Doan YH, Takahashi T, Ozaki T, Katayama K, Nakagomi O. 2014. Spread and predominance in Japan of novel G1P[8] double-reassortant rotavirus strains possessing a DS-1-like genotype constellation typical of G2P[4] strains. *Infect Genet Evol* 28:426–433. <https://doi.org/10.1016/j.meegid.2014.08.001>.
- Yamamoto SP, Kaida A, Kubo H, Iritani N. 2014. Gastroenteritis outbreaks caused by a DS-1-like G1P[8] rotavirus strain, Japan, 2012–2013. *Emerg Infect Dis* 20:1030–1033. <https://doi.org/10.3201/eid2006.131326>.
- Komoto S, Tacharoenuang R, Guntapong R, Ide T, Haga K, Katayama K, Kato T, Ouchi Y, Kurahashi H, Tsuji T, Sangkitporn S, Taniguchi K. 2015. Emergence and characterization of unusual DS-1-like G1P[8] rotavirus strains in children with diarrhea in Thailand. *PLoS One* 10:e0141739. <https://doi.org/10.1371/journal.pone.0141739>.
- Komoto S, Tacharoenuang R, Guntapong R, Ide T, Tsuji T, Yoshikawa T, Tharmaphornpilas P, Sangkitporn S, Taniguchi K. 2016. Reassortment of human and animal rotavirus gene segments in emerging DS-1-like G1P[8] rotavirus strains. *PLoS One* 11:e0148416. <https://doi.org/10.1371/journal.pone.0148416>.
- Nakagomi T, Nguyen MQ, Gauchan P, Agbemabiese CA, Kaneko M, Do LP, Vu TD, Nakagomi O. 2017. Evolution of DS-1-like G1P[8] double-gene reassortant rotavirus A strains causing gastroenteritis in children in

- Vietnam in 2012/2013. *Arch Virol* 162:739–748. <https://doi.org/10.1007/s00705-016-3155-6>.
27. Bar-Zeev N, Kapanda L, Tate JE, Jere KC, Iturriza-Gomara M, Nakagomi O, Mwansambo C, Costello A, Parashar UD, Heyderman RS, French N, Cunliffe NA. 2015. Effectiveness of a monovalent rotavirus vaccine in infants in Malawi after programmatic roll-out: an observational and case-control study. *Lancet Infect Dis* 15:422–428. [https://doi.org/10.1016/S1473-3099\(14\)71060-6](https://doi.org/10.1016/S1473-3099(14)71060-6).
 28. Cunliffe NA, Gondwe JS, Broadhead RL, Molyneux ME, Woods PA, Bresee JS, Glass RI, Gentsch JR, Hart CA. 1999. Rotavirus G and P types in children with acute diarrhea in Blantyre, Malawi, from 1997 to 1998: predominance of novel P[6]G8 strains. *J Med Virol* 57:308–312. [https://doi.org/10.1002/\(SICI\)1096-9071\(199903\)57:3<308::AID-JMV15>3.0.CO;2-B](https://doi.org/10.1002/(SICI)1096-9071(199903)57:3<308::AID-JMV15>3.0.CO;2-B).
 29. Cunliffe NA, Gentsch JR, Kirkwood CD, Gondwe JS, Dove W, Nakagomi O, Nakagomi T, Hoshino Y, Bresee JS, Glass RI, Molyneux ME, Hart CA. 2000. Molecular and serologic characterization of novel serotype G8 human rotavirus strains detected in Blantyre, Malawi. *Virology* 274:309–320. <https://doi.org/10.1006/viro.2000.0456>.
 30. Bar-Zeev N, Jere KC, Bennett A, Pollock L, Tate JE, Nakagomi O, Iturriza-Gomara M, Costello A, Mwansambo C, Parashar UD, Heyderman RS, French N, Cunliffe NA. 2016. Population impact and effectiveness of monovalent rotavirus vaccination in urban Malawian children 3 years after vaccine introduction: ecological and case-control analyses. *Clin Infect Dis* 62(Suppl 2):S213–S219. <https://doi.org/10.1093/cid/civ1183>.
 31. Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. 2006. GARD: a genetic algorithm for recombination detection. *Bioinformatics* 22:3096–3098. <https://doi.org/10.1093/bioinformatics/btl474>.
 32. Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. *Mol Biol Evol* 23:1891–1901. <https://doi.org/10.1093/molbev/msl051>.
 33. Ward RL, Bernstein DI, Shukla R, McNeal MM, Sherwood JR, Young EC, Schiff GM. 1990. Protection of adults rechallenged with a human rotavirus. *J Infect Dis* 161:440–445. <https://doi.org/10.1093/infdis/161.3.440>.
 34. Magagala NB, Esona MD, Nyaga MM, Stucker KM, Halpin RA, Stockwell TB, Seheri ML, Steele AD, Wentworth DE, Mphahlele MJ. 2015. Whole genome analyses of G1P[8] rotavirus strains from vaccinated and non-vaccinated South African children presenting with diarrhea. *J Med Virol* 87:79–101. <https://doi.org/10.1002/jmv.23971>.
 35. Zeller M, Donato C, Trovao NS, Cowley D, Heylen E, Donker NC, McAllen JK, Akopov A, Kirkness EF, Lemey P, Van Ranst M, Matthijnsens J, Kirkwood CD. 2015. Genome-wide evolutionary analyses of G1P[8] strains isolated before and after rotavirus vaccine introduction. *Genome Biol Evol* 7:2473–2483. <https://doi.org/10.1093/gbe/evv157>.
 36. Heiman EM, McDonald SM, Barro M, Taraporewala ZF, Bar-Magen T, Patton JT. 2008. Group A human rotavirus genomics: evidence that gene constellations are influenced by viral protein interactions. *J Virol* 82:11106–11116. <https://doi.org/10.1128/JVI.01402-08>.
 37. Matthijnsens J, Heylen E, Zeller M, Rahman M, Lemey P, Van Ranst M. 2010. Phylodynamic analyses of rotavirus genotypes G9 and G12 underscore their potential for swift global spread. *Mol Biol Evol* 27:2431–2436. <https://doi.org/10.1093/molbev/msq137>.
 38. Nagaoka Y, Tatsumi M, Tsugawa T, Yoto Y, Tsutsumi H. 2012. Phylogenetic and computational structural analysis of VP7 gene of group A human rotavirus G1P[8] strains obtained in Sapporo, Japan from 1987 to 2000. *J Med Virol* 84:832–838. <https://doi.org/10.1002/jmv.23247>.
 39. Wentzel JF, Yuan L, Rao S, van Dijk AA, O'Neill HG. 2013. Consensus sequence determination and elucidation of the evolutionary history of a rotavirus Wa variant reveal a close relationship to various Wa variants derived from the original Wa strain. *Infect Genet Evol* 20:276–283. <https://doi.org/10.1016/j.meegid.2013.09.007>.
 40. Afrad MH, Matthijnsens J, Afroz SF, Rudra P, Nahar L, Rahman R, Hossain ME, Rahman SR, Azim T, Rahman M. 2014. Differences in lineage replacement dynamics of G1 and G2 rotavirus strains versus G9 strain over a period of 22 years in Bangladesh. *Infect Genet Evol* 28:214–222. <https://doi.org/10.1016/j.meegid.2014.10.002>.
 41. Trang NV, Yamashiro T, Anh LTK, Hau VT, Luan LT, Anh DD. 2012. Genetic variation in the VP7 gene of rotavirus G1P[8] strains isolated in Vietnam, 1998–2009. *Virus Res* 165:190–196. <https://doi.org/10.1016/j.virusres.2012.02.015>.
 42. Jenkins GM, Cowart LA, Signorelli P, Pettus BJ, Chalfant CE, Hannun YA. 2002. Acute activation of de novo sphingolipid biosynthesis upon heat shock causes an accumulation of ceramide and subsequent dephosphorylation of SR proteins. *J Biol Chem* 277:42572–42578. <https://doi.org/10.1074/jbc.M207346200>.
 43. Huang K, Zhu H, Fan X, Wang J, Cheung CL, Duan L, Hong W, Liu Y, Li L, Smith DK, Chen H, Webster RG, Webby RJ, Peiris M, Guan Y. 2012. Establishment and lineage replacement of H6 influenza viruses in domestic ducks in southern China. *J Virol* 86:6075–6083. <https://doi.org/10.1128/JVI.06389-11>.
 44. Martin E, Chirivella M, Co JKG, Santiago GA, Gubler DJ, Munoz-Jordan JL, Bennett SN. 2016. Insights into the molecular evolution of dengue virus type 4 in Puerto Rico over two decades of emergence. *Virus Res* 213:23–31. <https://doi.org/10.1016/j.virusres.2015.11.009>.
 45. Donaldson EF, Lindesmith LC, Lobue AD, Baric RS. 2010. Viral shape-shifting: norovirus evasion of the human immune system. *Nat Rev Microbiol* 8:231–241. <https://doi.org/10.1038/nrmicro2296>.
 46. Zeller M, Patton JT, Heylen E, De Coster S, Ciarlet M, Van Ranst M, Matthijnsens J. 2012. 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* 50:966–976. <https://doi.org/10.1128/JCM.05590-11>.
 47. Zhang S, McDonald PW, Thompson TA, Dennis AF, Akopov A, Kirkness EF, Patton JT, McDonald SM. 2014. Analysis of human rotaviruses from a single location over an 18-year time span suggests that protein coadaptation influences gene constellations. *J Virol* 88:9842–9863. <https://doi.org/10.1128/JVI.01562-14>.
 48. Page NA, Steele AD. 2004. Antigenic and genetic characterization of serotype G2 human rotavirus strains from the African continent. *J Clin Microbiol* 42:595–600. <https://doi.org/10.1128/JCM.42.2.595-600.2004>.
 49. Page NA, Steele AD. 2004. Antigenic and genetic characterization of serotype G2 human rotavirus strains from South Africa from 1984 to 1998. *J Med Virol* 72:320–327. <https://doi.org/10.1002/jmv.10571>.
 50. Coulson BS, Kirkwood C. 1991. Relation of VP7 amino acid sequence to monoclonal antibody neutralization of rotavirus and rotavirus monotype. *J Virol* 65:5968–5974.
 51. Caust J, Dyal-Smith ML, Lazdins I, Holmes IH. 1987. Glycosylation, an important modifier of rotavirus antigenicity. *Arch Virol* 96:123–134. <https://doi.org/10.1007/BF01320955>.
 52. Estes MK, Cohen J. 1989. Rotavirus gene structure and function. *Microbiol Rev* 53:410–449.
 53. Lazdins I, Coulson BS, Kirkwood C, Dyal-Smith M, Masendycz PJ, Sonza S, Holmes IH. 1995. Rotavirus antigenicity is affected by the genetic context and glycosylation of VP7. *Virology* 209:80–89. <https://doi.org/10.1006/viro.1995.1232>.
 54. Trask SD, McDonald SM, Patton JT. 2012. Structural insights into the coupling of virion assembly and rotavirus replication. *Nat Rev Microbiol* 10:165–177. <https://doi.org/10.1038/nrmicro2673>.
 55. Settembre EC, Chen JZ, Dormitzer PR, Grigorieff N, Harrison SC. 2011. Atomic model of an infectious rotavirus particle. *EMBO J* 30:408–416. <https://doi.org/10.1038/emboj.2010.322>.
 56. Potgieter AC, Page NA, Liebenberg J, Wright IM, Landt O, van Dijk AA. 2009. Improved strategies for sequence-independent amplification and sequencing of viral double-stranded RNA genomes. *J Gen Virol* 90:1423–1432. <https://doi.org/10.1099/vir.0.009381-0>.
 57. Iturriza-Gomara M, Elliot AJ, Dockery C, Fleming DM, Gray JJ. 2009. Structured surveillance of infectious intestinal disease in pre-school children in the community: 'the Nappy Study'. *Epidemiol Infect* 137:922–931. <https://doi.org/10.1017/S0950268808001556>.
 58. Chen C, Khaleel SS, Huang H, Wu CH. 2014. Software for pre-processing Illumina next-generation sequencing short read sequences. *Source Code Biol Med* 9:8. <https://doi.org/10.1186/1751-0473-9-8>.
 59. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet* 17:10–12. <https://doi.org/10.14806/ej.17.1.200>.
 60. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>.
 61. Benson D, Karsch-Mizrachi I, Lipman D, Ostell J, Rapp B, Wheeler D. 2002. GenBank. *Nucleic Acids Res* 30:17–20. <https://doi.org/10.1093/nar/30.1.17>.
 62. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
 63. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6:

- Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30:2725–2729. <https://doi.org/10.1093/molbev/mst197>.
64. Felsenstein J. 1988. Phylogenies from molecular sequences: inference and reliability. *Annu Rev Genet* 22:521–565. <https://doi.org/10.1146/annurev.ge.22.1.20188.002513>.
 65. Drummond A, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214. <https://doi.org/10.1186/1471-2148-7-214>.
 66. Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29:1969–1973. <https://doi.org/10.1093/molbev/mss075>.
 67. Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol* 4:e88. <https://doi.org/10.1371/journal.pbio.0040088>.
 68. Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 22:160–174. <https://doi.org/10.1007/BF02101694>.
 69. Yang Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J Mol Evol* 39:306–314. <https://doi.org/10.1007/BF00160154>.
 70. Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, de Hoon MJ. 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25:1422–1423. <https://doi.org/10.1093/bioinformatics/btp163>.
 71. Pond SL, Frost SD. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21:2531–2533. <https://doi.org/10.1093/bioinformatics/bti320>.
 72. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K. 2013. FUBAR: a fast, unconstrained Bayesian approximation for inferring selection. *Mol Biol Evol* 30:1196–1205. <https://doi.org/10.1093/molbev/mst030>.
 73. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. 2012. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet* 8:e1002764. <https://doi.org/10.1371/journal.pgen.1002764>.
 74. Kosakovsky Pond SL, Murrell B, Fourment M, Frost SDW, Delpont W, Scheffler K. 2011. A random effects branch-site model for detecting episodic diversifying selection. *Mol Biol Evol* 28:3033–3043. <https://doi.org/10.1093/molbev/msr125>.
 75. Yang Z, Nielsen R, Goldman N, Pedersen A. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155:431–449.
 76. Pond SL, Frost SD, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679. <https://doi.org/10.1093/bioinformatics/bti079>.
 77. Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779–815. <https://doi.org/10.1006/jmbi.1993.1626>.
 78. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. *Nucleic Acids Res* 28:235–242. <https://doi.org/10.1093/nar/28.1.235>.
 79. Soding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33:W244–W248. <https://doi.org/10.1093/nar/gki408>.