HHS Public Access

Author manuscript

Infect Genet Evol. Author manuscript; available in PMC 2015 October 26.

Published in final edited form as:

Infect Genet Evol. 2014 December ; 28: 513–523. doi:10.1016/j.meegid.2014.09.021.

Comparative Genomic Analysis of Genogroup 1 (Wa-like) Rotaviruses circulating in the USA, 2006 – 2009

Sunando Roy[a], **Mathew D. Esona**[a], **Ewen F. Kirkness**[b], **Asmik Akopov**[b], **J. Kyle McAllen**[b], **Mary Wikswo**[a], **Margaret M. Cortese**[a], **Daniel C. Payne**[a], **Umesh Parashar**[a], **Jon R. Gentsch**[a], **Michael D. Bowen**[a,*], and **the National Rotavirus Strain Surveillance System; the New Vaccine Surveillance System**

[a]Division of Viral Diseases, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, USA.

[b]The J. Craig Venter Institute, Rockville Maryland, USA

Abstract

Group A Rotaviruses (RVA) are double stranded RNA viruses that are a significant cause of acute pediatric gastroenteritis. Beginning in 2006 and 2008, respectively, two vaccines, Rotarix™ and RotaTeq®, have been approved for use in the USA for prevention of RVA disease. The effects of possible vaccine pressure on currently circulating strains in the USA and their genome constellations are still under investigation. In this study we report 33 complete RVA genomes (ORF regions) collected in multiple cities across USA during 2006 – 2009, including 8 collected from children with verified receipt of 3 doses of rotavirus vaccine. The strains included 16 G1P[8], 10 G3P[8], and 7 G9P[8]. All 33 strains had a Wa like backbone with the consensus genotype constellation of G(1/3/9)-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1. From maximum likelihood based phylogenetic analyses, we identified 3 to7 allelic constellations grouped mostly by respective G types, suggesting a possible allelic segregation based on the VP7 gene of RVA

[*]**Corresponding author** Mailing address: Gastroenteritis and Respiratory Viruses Laboratory Branch, DVD, NCIRD, Centers for Disease Control and Prevention, 1600 Clifton Rd., NE, Mailstop G04, Atlanta, GA, 30333; Phone: 404-639-4922; Fax: 404-639-3645; mkb6@cdc.gov..

Participants in the National Rotavirus Strain Surveillance System include the following: Kathy Dugaw, Seattle Children's Hospital, Seattle, WA; Gail Bloom, Clarian Health Partners, Indianapolis, IN; Paul A. Yam and Sandra Jameson, Children's Memorial Hospital of Omaha, Omaha, NE; Barbara McKee, Long Beach Memorial Medical Center, Long Beach, CA; Ann Marie Riley, Boston Children's Hospital, Boston, MA; Kenneth Thompson, University of Chicago Medical Center, Chicago, IL; Carolyn Wright and W. Lawrence Drew, University of California, San Francisco, UCSF Medical Center at Mount Zion, San Francisco, CA; Jim Dunn, Cook Children's Medical Center, Fort Worth, TX; and Valerie Hoover, Orlando Regional Medical Center, Orlando, FL.

Participants in the New Vaccine Surveillance Network include the following: Peter G. Szilagyi and Geoffrey A. Weinberg, Department of Pediatrics, University of Rochester School of Medicine and Dentistry, Rochester, NY; Mary Allen Staat, Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, OH; Kathryn M. Edwards and James Chappell, Department of Pediatrics, Vanderbilt University Medical Center, Nashville, TN.

Article summary line: Complete genomes of 33 Genogroup 1 (Wa-like) rotavirus strains collected in multiple cities across the USA during 2006 – 2009 were determined and analyzed.

Conflict of Interest Statement

The authors of this study declare that they have no conflict of interest, financial or otherwise, related to this article.

Disclaimer

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention. Names of specific vendors, manufacturers, or products are included for public health and informational purposes; inclusion does not imply endorsement of the vendors, manufacturers, or products by the Centers for Disease Control and Prevention or the US Department of Health and Human Services.

primarily for the G3 and G9 strains. The vaccine failure strains showed similar grouping for all genes in G9 strains and most genes of G3 strains suggesting that these constellations were necessary to evade vaccine-derived immune protection. Substitutions in the antigenic region of VP7 and VP4 genes were also observed for the vaccine failure strains which could possibly explain how these strains escape vaccine induced immune response. This study helps elucidate how RVA strains are currently evolving in the population post vaccine introduction and supports the need for continued RVA surveillance.

Keywords

Rotavirus; vaccine; failure; allele; VP4; VP7

1. Introduction

Group A Rotaviruses (RVA) are the major cause of acute gastroenteritis in children under 5 years of age and the leading cause of gastroenteritis related deaths (~450,000) in developing countries every year (Estes and Kapikian, 2007; Parashar et al., 2009; Tate et al., 2012). In industrialized countries, RVA associated mortality is minimal yet the financial cost of treatment associated with disease burden is enormous (Payne et al., 2011). The RVA genome is composed of 11 double-stranded RNA (dsRNA) segments which code for 11or 12 proteins in total, six structural proteins VP1-VP4, VP6 and VP7, and five or six nonstructural proteins NSP1-NSP5/6 (Estes and Kapikian, 2007). The VP7 and VP4 proteins form the outer layer of the viral capsid and have been historically used to classify RVA serotypes and genotypes into respective G and P types (Estes and Kapikian, 2007). These proteins have been extensively studied and a number of antigenic epitopes have been characterized. An extended classification system based on all 11 gene segments has been introduced by the Rotavirus Classification Working Group (RCWG) (Matthijnssens et al., 2008b) and this classification system designates genotypes in the format Gx-P[x]-Ix-Rx-Cx-Mx-Ax-Nx-Tx-Ex-Hx for the genes VP7, VP4, VP6, VP1-3, NSP1-5 respectively. Presently there are at least 27 G, 37 P, 16 I, 9 R, 9 C, 8 M, 16 A, 9 N, 12 T, 14 E, and 11 H types (Matthijnssens et al., 2011; Trojnar et al., 2013). The most common genogroup constellations worldwide based on the new classification are Wa like Genogroup 1 (Gx-P[x]-I1-R1-C1-M1-A1-N1-T1-E1-H1) and DS-1 like Genogroup 2 (Gx-P[x]-I2-R2-C2-M2-A2-N2-T2-E2-H2) (Matthijnssens et al., 2008b; Matthijnssens et al., 2012a). The genotype classification for an unknown strain is based on sequence identity cutoffs established by the RCWG and is currently implemented in the RotaC webserver (Maes et al., 2009).

In humans, the most common RVA G/P genotypes worldwide are G1P[8], G2P[4], G3P[8], G4P[8], and G9P[8] (Banyai et al., 2012; Gentsch et al., 2005). Two vaccines RotaTeq® (Merck) and Rotarix™ (GlaxoSmithKline) were introduced in the U.S. in 2006 and 2008, respectively (Cortese et al., 2009). RotaTeq® is a pentavalent human bovine reassortant vaccine which contains four G types (G1, G2, G3 and G4, VP7 gene) along with the P[8] VP4 type on a bovine WC3 (G6P[5]) backbone (Matthijnssens et al., 2010b). On the other hand, Rotarix™ is a human RVA derived G1P[8] strain (Ward, 2009). These vaccines have shown high efficacy in reducing the RVA burden in developed countries after their

introduction in various immunization programs (Gray, 2011). Unfortunately the same vaccines have much lower efficacy in some developing countries and this has caused some concern (Armah et al., 2010; Phua et al., 2009; Zaman et al., 2010). In addition it has been proposed that RVA vaccination in several countries may have driven the selection of new predominant genotypes through immune pressure (Carvalho-Costa et al., 2009; Hull et al., 2011; Matthijnssens et al., 2012b; Zeller et al., 2010) though current evidence remains inconclusive. Globally, multiple surveillance networks have been established to study RVA prevalence in various countries (Carvalho-Costa et al., 2011; Hull et al., 2011; Iturriza-Gomara et al., 2009; Kirkwood et al., 2010; Payne et al., 2008). In United States the two main RVA surveillance networks, National Rotavirus Strain Surveillance System (NRSSS) and the New Vaccine Surveillance Network (NVSN) have been established to monitor RVA strain prevalence in multiple parts of the country (Hull et al., 2011; Payne et al., 2008).

RVA genomes show high genomic diversity similar to other known RNA viruses. The diversity is generated by point mutations, reassortment, rearrangement and recombination events (Estes, 2007; Kirkwood et al., 2010). The proteins of RVA are known to evolve at substantially different rates with VP1 and VP2 being highly conserved whereas NSP1 is known to be highly divergent (Matthijnssens et al., 2008a). A recent study measured evolutionary rates of all 11 gene segments for the SA11-H96 strain and found nearly a 100 fold difference in evolutionary substitution rates among the 11 gene segments (Mlera et al., 2013). The evolutionary rates for the VP7 gene of G9 and G12 strains have also been calculated by Mathijnnesens et al to be 1.87×10^{-3} and 1.66×10^{-3} substitutions/site/year respectively (Matthijnssens et al., 2010a). Reassortment occurs when gene segments are exchanged between two different strains of RVA during a co-infection event. Multiple reassortments have been documented between two genogroups and also between different host species, albeit at a lower frequency (Varghese et al., 2004). Animal-to-human cross-transmission and reassortment events between human and animal strains have been suggested as one of the possible reasons for the low vaccine efficacy in developing countries (Martella et al., 2010; Palombo, 2002). Reassortment has also been recently reported between sub-genotypic clusters (Maunula and Von Bonsdorff, 2002; McDonald et al., 2012). Rearrangements events, although infrequent, have been observed in the NSP1 and NSP3 genes of cell culture passaged strains (Kojima et al., 2000). The short electropherotype of DS-1 like RVA strains is the result of naturally-occurring duplication and insertion events in the NSP5 genes (Giambiagi et al., 1994; Gonzalez et al., 1989). Intragenic recombination events are extremely rare and only a few cases have been reported in the VP7 and NSP2 genes (Donker et al., 2011; Phan et al., 2007). There is one report of intergenic recombination in the VP7 gene (Martinez-Laso et al., 2009).

In this study we sequenced complete open reading frames (ORFs) for 33 genogroup 1 strains (16 G1P[8], 10 G3P[8], and 7 G9P[8]) collected from multiple cities across the United States. Maximum likelihood based phylogenetic analysis along with other sequence analysis tools helped identify multiple sub-genotype allelic constellations recently circulating in USA. These sub allelic constellations may enhance our understanding of RVA evolution under vaccine pressure and help identify possible mechanisms of immune escape which result in RVA gastroenteritis in vaccinated individuals.

2. Materials and Methods

2.1 Surveillance Testing and Genotyping

Fecal specimens were collected from children with acute gastroenteritis from 12 sites in the United States during the 2006-2007, 2007-2008, and 2008-2009 RVA seasons. All samples were tested by enzyme immunoassay (EIA) using the Premier® Rotaclone® Rotavirus Detection Kit (Meridian Diagnostics Inc., Cincinnati, OH). At the Centers for Disease Control and Prevention, (CDC), RVA dsRNA was extracted and VP7 and VP4 genotyping was carried out using a two-step amplification method as described previously (Hull et al., 2011).

2.2 Sample Selection

A total of 33 samples were selected for genomic characterization of which 25 were from the NRSSS surveillance network and 8 from the NVSN. The NRSSS samples were selected based upon previous EIA and VP4/VP7 genotyping results and were collected from sites in Boston MA (1 sample), Chicago IL (1) Orlando FL (1), Fort Worth TX (6), Indianapolis IN (2), Long Beach CA (3), Omaha NE (3), San Francisco CA (1), and Seattle WA (7) whereas the NVSN samples were from study sites in Cincinnati OH (1), Nashville TN (1), and Rochester NY (6). Vaccination histories were available only for the children enrolled at NVSN sites. All 8 samples collected at NVSN sites were from children that received 3 doses of RotaTeq® vaccine.

2.3 Sanger Sequencing

Total RNA was extracted from 33 samples using MagNA Pure Compact extraction system with the RNA Isolation kit (Roche Applied Science, Indianapolis IN) and were sent to J. Craig Venter Institute for high-throughput Sanger sequencing. Oligonucleotide primers were designed using an automated primer design tool [PMID 18405373, 23131097]. Primers, with M13 tags added, were designed at intervals along both the sense and antisense strands, and provided amplicon coverage of at least 4-fold (see supplementary material). RT-PCRs were performed with 1 ng of RNA using OneStep RT-PCR kits (Qiagen, Valencia, CA) according to manufacturer's instructions with minor modifications: 1) reactions were scaled down to 1/5 the recommended volumes; 2) the RNA templates were denatured at 95°C for 5 min; and 3) 1.6 U RNase Out (Invitrogen, Carlsbad, CA) was used. The RT-PCR products were sequenced with an ABI Prism BigDye v3.1 terminator cycle sequencing kit (Applied Biosystems, Carlsbad, CA). Raw sequence traces were trimmed to remove any primer-derived sequence as well as low quality sequence, and gene sequences were assembled using Minimus, part of the open-source AMOS project [The AMOS project. http://amos.sourceforge.net]. The gene sequences were then manually edited using ClOE (Closure Editor; JCVI) and ambiguous regions were resolved by additional sequencing when necessary.

2.4 Whole Genome Phylogenetic Analysis

Genotypes for each gene segment were determined using RotaC v2.0 webserver (Maes et al., 2009). For each gene, multiple alignments were made using the MUSCLE algorithm

implemented in MEGA 5.1 (Tamura et al., 2011). Maximum likelihood trees were constructed for each gene in PhyML 3.0 using the optimal model for each alignment as identified by jModeltest 2 and approximate Likelihood Ratio Test (aLRT) statistics computed for branch support (Anisimova and Gascuel, 2006; Darriba et al., 2012; Guindon et al., 2010). The best models were selected based on the corrected Akaike Information Criterion and were General Time Reversible (GTR)+I+G (NSP1), GTR+G (NSP2, VP4), Transition model (TIM2)+I (NSP3), Tamura-Nei (TrN)+G (NSP4), Hasegawa-Kishino-Yano (HKY)+G (NSP5), TIM1+I (VP1,VP3), GTR+I (VP2), TrN+I (VP6), HKY+I+G (VP7). . Sub-genotypic clusters were identified as tight phylogenetic clusters with aLRT support greater >75%. Sequences were tested for possible recombination using the Genetic Algorithm Recombination Detection (GARD) algorithm implemented in Datamonkey (Delport et al., 2010; Kosakovsky Pond and Frost, 2005; Kosakovsky Pond et al., 2006; Pond and Frost, 2005). Selection analysis was performed using a combination of Single Likelihood Ancestor Counting (SLAC), Fixed Effects Likelihood (FEL) and Random Effects Likelihood (REL) analysis in Datamonkey (Kosakovsky Pond and Frost, 2005; Pond and Frost, 2005). Substitutions in the VP7, VP8* and VP5* regions were mapped on crystal structures available in PDB (www.pdb.org) using VMD 1.9.1 (Berman et al., 2000; Humphrey et al., 1996). For VP7 we used the RRV crystal 3FMG, for VP8* the Wa crystal 2DWR and for VP5* the RRV crystal 2B4I was used (Aoki et al., 2009; Blanchard et al., 2007; Yoder and Dormitzer, 2006).

2.5 Accession numbers

The nucleotide sequences for the ORFs of the eleven gene segments for each strain were submitted to GenBank (total of 363 sequences). The accession numbers are listed in Table 1.

3. Results

Complete ORF sequences were obtained for all 11 genes of 33 strains. Out of the 33 strains sequenced, 16 were G1P[8], 10 were G3P[8], and 7 were G9P[8] RVA strains.. Using RotaC 2.0 webserver, we assigned genotypes to the 11 different gene segments for all 33 RVA samples. All strains were found to be Wa-like genogroup 1 strains and the consensus 11-gene genotype constellation for the 33 RVA samples was G(1/3/9)-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1.

To this dataset we added 58 previously published sequences (McDonald et al., 2012) that were from samples collected during 2005 – 2008 from the NVSN site in Nashville (see supplementary material). The samples were all genogroup 1 strains with G1P[8], G3P[8] and G12P[8] genotypes (McDonald et al., 2012). Using the alignments of ORF sequences for each gene segment, we determined the optimal model based upon AICc values. We tested for subgenotype clustering of strains based on maximum likelihood analyses. Sub-genotype clustering was identified by nodes with aLRT support ≥ 75%. The phylogenetic trees for all 11 gene segments are shown in Fig 1(A-K) and the color coding of alleles shown in Figure 2. The VP7 gene clusters are based on the various individual genotypes G1, G3, G9, G12 (Fig 1A). The G1 strains were further divided into three distinct clusters indicating 3 alleles [AI (red), AII (maroon), and AIII (olive)]. The strains in the AI Allele

clustered with the prototype G1P[8] strain Wa and the G1 genes of the two vaccines, RotaTeq® and Rotarix™, whereas the AII and AIII strains formed distinct, supported clusters.

For the other 10 gene segments the clusters were colored based on association with VP7 genotypes (G1, G3, G9). The Wa-like G1 cluster was Allele A (red), the majority G3 cluster was Allele B (green), and the G9-like cluster was Allele C (blue). Strains that formed separate clusters outside the three designated alleles were separately clustered into Alleles D (purple), E (lime), F (pink), G (teal) and H (aqua). Doublets (pairs of strains) are indicated in gray and singleton and outgroup strains were labelled in black (Fig 1B-K, Fig 2). The analysis (of all 91 sequences) identified 3 to 7 subgenotypic clusters along with a few doublets and single outgroup strains for all 11 gene segments. Similar clustering was observed for most of the G3 and G9 RVA strains in all 11 gene segments (Figure 2). The G1 strains were broadly divided into 3 alleles. Strains 2008747288, 2007719698, 2007744509, 2007744510, except in the NSP3, VP6 and VP7 genes, clustered with the Wa like strains whereas strains 2008747100 and 2008747106 differed in their clustering pattern in the VP6, NSP3 and NSP1 genes only. The other G1 strains from this study did not exhibit a fixed clustering pattern across the 11 gene segments suggesting possible past intra-genotype reassortment events between the alleles.

Out of the 33 strains, 8 were from children known to have received at least one dose of RotaTeq® vaccine and these strains were referred to as vaccine failure (VF) strains. Six VF strains were from Rochester (G9P[8]) and one each (G3P[8]) from Nashville and Cincinnati. Four more G3P[8] VF strains from the McDonald et al. dataset were also included in this analysis. The G9 VF strains showed clustering across all 11 gene segments into Allele C whereas the G3 VF strains clustered primarily with Allele B, with a few exceptions (Fig 2). For strain 2009726997, the NSP4 gene did not cluster with other G3P[8] strains whereas strain VU08-09-22 clustered with Allele C for VP6, NSP1 and NSP4 genes and, for the NSP3 gene, strains not assigned to alleles (Fig 2).

We further tested for positive selection in the ORFs of all strains based on combined SLAC, FEL and REL tests. The number of codons analyzed per ORF is shown in Table 2. No positively selected sites were identified with high confidence ($p \leq 0.1$) among the 91 RVA samples (Table 2). All the genes were found to contain sites under strong purifying selection, however, ranging from 38 of 175 sites in NSP4 (21.7%) to 187 of 326 sites (57.1%) in VP7 (Table 2). VP6 exhibited the second highest percentage of sites (49.6%) under purifying selection. The percentage of sites under strong purifying selection ranged from 25.8% (VP2) to 57.1% (VP7) in the structural proteins and 21.7% (NSP4) to 44.2% (NSP1) among the non-structural proteins (Table 2). Tests for recombination were also carried out, but recombination was not detected in any of the 11 genes based on results of GARD analysis.

Using alignments for the VP4 and VP7 genes, we identified amino acid substitutions in the antigenic regions between the wild type RVA strains and the RotaTeq® and Rotarix™ vaccine strains. For the VP7 protein, the VF strains in our data set were genotype G3/G9 so we compared them to the G3 component of RotaTeq® as no known G9 vaccine component

was available. The conservative or non-conservative nature of substitutions was based on the classification proposed by Zhang et al (Zhang, 2000). In the G3 VF strains, a single substitution at position 242 was detected (T242N) that was also present in the G9 VF strains (Fig 3). In G9 VP7 VF strains, substitutions were observed at positions 94, 96, 146, 189, 208, 238 and 242 in antigenic sites A, B, E, C, and F (Fig 3). Positions 94, 96, 189, and 238 are known neutralization escape mutation site (Aoki et al., 2009). In the G9 VP7 protein, we observed a N/S94G substitution, G96T substitution, Q146S substitution, S189Q substitution, Q208I substitution and a N238D substitution . Most of these substitutions are conservative in nature except the Q208I substitution (position 208), which causes change in polarity and thus possibly affecting epitope structure. None of the G1 strains analyzed in this study were VF strains so we could not perform the substitution comparison with the G1 genes of the vaccine strains.

The VP4 protein, which in the study represented all P[8] strains, was divided into the VP8* and VP5* regions for comparison. In VP8*, positions 106, 108, 113, 120, 145, 150, and 195 show amino acid substitutions out of which positions 113, 145 and 195 are known neutralization escape mutation sites (Fig 4) (Dormitzer et al., 2004; Kobayashi et al., 1990; McKinney et al., 2007; Monnier et al., 2006; Zhou et al., 1994). In the VP8* region, we observed V106I substitution, I108V substitution, N113D substitution, T/M120N substitution, S145G substitution, E150D substitution, and N/D195G substitution in in all the G3 and G9 strains, and some G1 strains. The G1 strains that possessed the above substitutions were the ones that did not cluster with Allele A strains for most of the 11 genes. Most of these changes are conservative in nature except a N113D substitution at position 113 involving a change in charge, M120N at position 120 involving a change in polarity and D195G substitution at position 195 involving a change in charge. Three sites, positions 281, 385 and 604, were observed to have substitutions in the VP5* region (Fig 5). The substitutions were V281I, H/Y385D and L604V respectively. Out of these three substitutions only H/Y385D substitution at position 385 was non-conservative with a change in charge and this position is also a known neutralization escape mutation site (Dormitzer et al., 2004; Kobayashi et al., 1990; Larralde et al., 1991; Matsui et al., 1989). These substitutions were also present in all G3 strains and a few G1 strains in this study suggesting possible vaccine failure phenotypes.

The 12 VF strains in this study shared a common clustering pattern across all 10 (except VP7) genes. The G9P[8] VF strains always clustered with Allele C strains whereas the G3P[8] VF strains primarily clustered with Allele B strains. The 11 gene segments of 4 RVA samples from Fort Worth (2008747329, 2008747332, 2008747336, 2008747337) also shared similar clustering with respect to the G3P[8] VF strains (Fig 2). Vaccination histories were not available for those 4 samples but two of the 4 children were known to be age-eligible for vaccination. The substitutions in the VP7 proteins of the G9P[8] VF strains were unique at 6 of 7 sites when compared with the G3P[8] VF strains and the vaccine strains. Only 1 of 7 substitutions was shared by both the G9 and G3 VF strains. All of the ten substitutions found in the VP4 protein were shared by both G9P[8] and G3P[8] vaccine failure strains. These substitutions were also present, however, in all of the other G3 and G9 strains analyzed in this study and ~73% of the G1 strains.

4. Discussion

In this study we sequenced 33 RVA samples collected post vaccine introduction (2006-2009) from sites across the continental United States. The G1 samples collected were evenly distributed across the 2006 – 2007 and the 2007 – 2008 seasons (none in the 2008 – 2009 season). The majority of the G3 strains were from the 2007 – 2008 season whereas most G9 strains were from the 2008 – 2009 season. Maximum likelihood based phylogenetic clustering identified 3 to 7 distinct allelic clusters across all 11 gene segments. Allelic clusters for most part were defined by the VP7 gene with a few exceptions in some strains. Most G3 and G9 strains formed their own clusters across all 11 genes that also included the G1 alleles AI and AII in most cases. We hypothesize that the G1 strains from this study cluster on the tree with known G3 and G9 strains due to inter-allelic reassortment events as reported previously by McDonald et al. (McDonald et al., 2012). Such reassortment events may lead to constellations that are more favorable to the virus evolution. From the McDonald et al. dataset, which contain a large number of samples from the pre vaccination era (2005 – 2006), we identified 3 G1 alleles circulating in the population. The presence of such mixed allelic G1 reassortants pre vaccine introduction and their rise in numbers post vaccination may suggest that these mixed alleles are possibly selected for under vaccine pressure over older Wa like G1 alleles and can evolve to give rise to potential VF strains (Arista et al., 2006). Arista et al. suggested that the older Wa like lineages became extinct due to immune pressure but this study detected strains circulating in US (Allele AI) which are similar to older Wa-like strains. These strains may provide a reservoir for selection of constellations that may lead to formation of vaccine failure strains. A larger dataset both pre and post vaccination is required to understand the evolution of these mixed allelic G1 reassortants during the pre-vaccine era and their rise in numbers during the post-vaccine era. In this study our allelic calls do not always match with previously reported studies (McDonald et al 2012). This is primarily due to Likelihood based measures that are heavily dependent on the dataset and the models used and hence these strains need to be re-defined based on the current phylogenetic trees.

In this study, we found that the vaccine failure strains have a fixed constellation across the 11 genes. It is possible that these constellations are present due to preferred interactions between proteins coded by these alleles which help in evading vaccination-induced protection in children (McDonald et al., 2009). All the vaccine failure strains also shared one substitution in VP7 gene and 10 substitutions in VP4 gene, which may lead to potential evasion of vaccine induced immunity. Although such fixed constellations and substitutions were observed in other G3 and G9 samples (e.g., G3P[8] strains from Fort Worth), we did not have the vaccination histories to determine whether these strains were also potential VF strains and we know that at least some of the children were too old to have been vaccinated (i.e., born more than 6 months before RotaTeq® vaccine was licensed in the US).

VP4 is an outer capsid component and in this study we found that 42.8% of the amino acid residues in this protein are under strong purifying selection. A previous study found the VP8* region of the P[8] protein associated with genotype G12 contains 103 amino acids under strong purifying selection (Mijatovic-Rustempasic et al., 2014). The VP4 contains epitopes involved in antibody-mediated neutralization (Dormitzer et al., 2004; Kobayashi et

al., 1990; Matsui et al., 1989) as well as functional domains involved in virus attachment and entry into host cells (Estes and Kapikian, 2007). It is likely that multiple selective pressures are influencing the evolution of VP4. VP6 also appears to be under strong purifying selection but the processes involved in the evolution of this protein remain to be determined. Previous studies have reported positive selection in the NSP2, NSP4, and VP7 genes (Donker and Kirkwood, 2012; Mijatovic-Rustempasic et al., 2014; Song and Hao, 2009) but in this study we did not find evidence of this process in our data sets for these 3 genes

The VP7 and VP4 proteins form the outer surface structure of the viral capsid, hence their antigenic regions have been well characterized. RVA immunity is known to be both homotypic, heterotypic and polygenic (Desselberger and Huppertz, 2011). To identify sites that could possibly help escape vaccine derived immunity, we looked at the alignments of the antigenic regions of the RVA strains in this study. For the VP7 gene, seven substitutions were observed in all G9 strains out of which four were at known neutralization escape sites (Hoshino et al., 2005; Hoshino et al., 2004). When compared with Rotarix™ and RotaTeq® only one change Q208I was accompanied by a change in polarity observed in the G9 strains. Change in polarity at this site may possibly help the G9 strains escape neutralization although we did not find similar substitution in the G3 vaccine failure strain. Zeller et al. (Zeller et al., 2012) described a N238D change in G3P[8] strains from Belgium that creates a potential N-linked glycosylation site that is absent in the G3 strain of RotaTeq®. All strains in this study, except the G9 strains, have an N at position 238 indicating the presence of this predicted glycosylation site. For the G9 strains, a D at position 238 gives rise to a predicted integrin binding site instead of a glycosylation site. Whereas the glycosylation site is thought to facilitate viral growth in cell culture systems, the function of the integrin binding site is unknown (Graham et al., 2005). The N238D mutation, however, appears in virulent murine RVA strains produced by serial passage of avirulent RVA strains in mice (Tsugawa et al., 2014). It is interesting to note that even the Rotarix™ VP7 protein contains the same glycosylation site at position. Most of the changes at the VP7 neutralization escape sites identified in this study were conservative in nature. The mutations in the G9 strains were unique to the group but the mutation present in G3 strains were present in all G3 and even some G1 strains in this study. In the VP8* region of the VP4 gene, 7 substitutions were observed out of which 4 were non conservative in nature when compared with the Rotarix™ and the RotaTeq® consensus sequences. Other conserved epitope changes at multiple sites have previously been reported recently in strains circulating in the US and Belgium (Mijatovic-Rustempasic et al., 2014; Zeller et al., 2012). The functional relevance of these changes at known epitopes is yet to be determined. With regard to the non-conservative amino acid substitutions in the VP8* region, the N113D substitution at position 113 and D195G substitution at position 195 involve a change in charge, from neutral to negative and negative to neutral, respectively. Both of these sites lie in known neutralization escape sites and differ from the sequences of Rotarix™ and RotaTeq®. The other non-conservative change was at position 120 where an M120N change was accompanied with change in polarity. In the VP5* region there were three positions were substitutions were observed. Out of the 3, H/Y385D substitution at position 385, a neutralization escape site, was non-conservative with a change in charge. Amino acid residues 382 – 400 in the VP5* region

also contain a potential membrane interaction loop and may be essential for viral virulence, especially in facilitating viral attachment or penetration (Dormitzer et al., 2004; Trask et al., 2010), In a cell culture-murine model system, a charge change at this position has been shown to be associated with reversion of an avirulent strain to a virulent one (Tsugawa et al., 2014). It is interesting to note that these substitutions are present in all sequences except the Allele A strains. If these substitutions are responsible for neutralization escape, it would suggest that the Allele A strains would gradually be replaced over time by the other reassortant G1 types due to vaccine derived evolutionary pressures. A more detailed analysis with multiple strains from pre vaccination years is necessary to assess whether these substitutions are actively arising under vaccine pressure.

The main limitation of the genetic analysis performed in this study was the closely related nature of the strains that were sampled. All strains shared a common Wa like genogroup 1 backbone with the only difference being in the VP7 gene (G1/G3/G9). The low genetic distance between the genes of the strains sampled may interfere in understanding better the effect of vaccine pressure, geographical distance, and time on the evolution of these virus strains currently circulating in continental US. A more detailed sampling of multiple genotypes with different backbones is currently underway to determine better the effect of vaccine pressures on circulating strains. Also to test the hypothesis that specific amino acid substitutions can lead to immune evasion of vaccine induced immunity, VF strains should be compared to strains from unvaccinated children that are matched by age, site, and date of illness to control for these variables. Such studies are currently underway at CDC.

5. Conclusion

Large scale full genome sequencing initiatives are necessary to understand the changes in circulating strains post introduction of the vaccine. We were able to capture changes in recent circulating strains that could possibly lead to the rise of vaccine failure strains. We were also able to identify inter-allelic reassortment events in currently circulating G1 strains. Such reassortment events have been reported previously, albeit infrequently, (McDonald et al., 2011; McDonald et al., 2012) and can help in producing possibly strains better suited to escape vaccine derived immune pressure. This study highlights the need for further surveillance and other full genome initiatives to study in detail the evolution of RVA strains currently in circulation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

We wish to thank Tara Kerin, Slavica Mijatovic-Rustempasic, David Spiro, and Elizabeth Teel for their contributions to this study. We also wish to thank Rashi Gautam for her critical review of the manuscript.

This project has been funded in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services under contract number HHSN272200900007C.

References

Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Systematic biology. 2006; 55:539–552. [PubMed: 16785212]

Aoki ST, Settembre EC, Trask SD, Greenberg HB, Harrison SC, Dormitzer PR. Structure of rotavirus outer-layer protein VP7 bound with a neutralizing Fab. Science. 2009; 324:1444–1447. [PubMed: 19520960]

Arista S, Giammanco GM, De Grazia S, Ramirez S, Lo Biundo C, Colomba C, Cascio A, Martella V. Heterogeneity and temporal dynamics of evolution of G1 human rotaviruses in a settled population. Journal of virology. 2006; 80:10724–10733. [PubMed: 16928744]

Armah GE, Sow SO, Breiman RF, Dallas MJ, Tapia MD, Feikin DR, Binka FN, Steele AD, Laserson KF, Ansah NA, Levine MM, Lewis K, Coia ML, Attah-Poku M, Ojwando J, Rivers SB, Victor JC, Nyambane G, Hodgson A, Schodel F, Ciarlet M, Neuzil KM. Efficacy of pentavalent rotavirus vaccine against severe rotavirus gastroenteritis in infants in developing countries in sub-Saharan Africa: a randomised, double-blind, placebo-controlled trial. Lancet. 2010; 376:606–614. [PubMed: 20692030]

Banyai K, Laszlo B, Duque J, Steele AD, Nelson EA, Gentsch JR, Parashar UD. 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. 2012; 30(Suppl 1):A122–130. [PubMed: 22520121]

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic acids research. 2000; 28:235–242. [PubMed: 10592235]

Blanchard H, Yu X, Coulson BS, von Itzstein M. Insight into host cell carbohydrate-recognition by human and porcine rotavirus from crystal structures of the virion spike associated carbohydrate-binding domain (VP8*). Journal of molecular biology. 2007; 367:1215–1226. [PubMed: 17306299]

Carvalho-Costa FA, Araujo IT, Santos de Assis RM, Fialho AM, de Assis Martins CM, Boia MN, Leite JP. Rotavirus genotype distribution after vaccine introduction, Rio de Janeiro, Brazil. Emerging infectious diseases. 2009; 15:95–97. [PubMed: 19116062]

Carvalho-Costa FA, Volotao Ede M, de Assis RM, Fialho AM, de Andrade Jda S, Rocha LN, Tort LF, da Silva MF, Gomez MM, de Souza PM, Leite JP. Laboratory-based rotavirus surveillance during the introduction of a vaccination program, Brazil, 2005-2009. The Pediatric infectious disease journal. 2011; 30:S35–41. [PubMed: 21048523]

Cortese MM, Parashar UD, Centers for Disease, C., Prevention. Prevention of rotavirus gastroenteritis among infants and children: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR. Recommendations and reports : Morbidity and mortality weekly report. Recommendations and reports / Centers for Disease Control. 2009; 58:1–25.

Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nature methods. 2012; 9:772. [PubMed: 22847109]

Delport W, Poon AF, Frost SD, Kosakovsky Pond SL. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics. 2010; 26:2455–2457. [PubMed: 20671151]

Desselberger U, Huppertz HI. Immune responses to rotavirus infection and vaccination and associated correlates of protection. The Journal of infectious diseases. 2011; 203:188–195. [PubMed: 21288818]

Donker NC, Boniface K, Kirkwood CD. Phylogenetic analysis of rotavirus A NSP2 gene sequences and evidence of intragenic recombination. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2011; 11:1602–1607. [PubMed: 21689784]

Donker NC, Kirkwood CD. Selection and evolutionary analysis in the nonstructural protein NSP2 of rotavirus A. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2012; 12:1355–1361.

Dormitzer PR, Nason EB, Prasad BV, Harrison SC. Structural rearrangements in the membrane penetration protein of a non-enveloped virus. Nature. 2004; 430:1053–1058. [PubMed: 15329727]

Estes, M.; Kapikian, A. Rotaviruses. In: Knipe, DM.; H.P.; Griffin, DE.; Lamb, RA.; Martin, MA.; Roizman, B.; Straus, SE., editors. Fields Virology. 5th Edition ed.. Kluwer/Lippincott; Williams and Wilkins; Philadelphia, PA: 2007. p. 1917-1974.

Estes, MK.; Kapikian, A. Rotaviruses. In: Knipe, DM.; Howley, PM.; Griffin, DE.; Lamb, RA.; Martin, MA.; Roizman, B.; Straus, SE., editors. Fields Virology. 5th ed.. Kluwer/Lippincott; Williams and Wilkins; Philadelphia, PA: 2007. p. 1917-1974.

Gentsch JR, Laird AR, Bielfelt B, Griffin DD, Banyai K, Ramachandran M, Jain V, Cunliffe NA, Nakagomi O, Kirkwood CD, Fischer TK, Parashar UD, Bresee JS, Jiang B, Glass RI. Serotype diversity and reassortment between human and animal rotavirus strains: implications for rotavirus vaccine programs. The Journal of infectious diseases. 2005; 192(Suppl 1):S146–159. [PubMed: 16088798]

Giambiagi S, Gonzalez Rodriguez I, Gomez J, Burrone O. A rearranged genomic segment 11 is common to different human rotaviruses. Archives of virology. 1994; 136:415–421. [PubMed: 8031245]

Gonzalez SA, Mattion NM, Bellinzoni R, Burrone OR. Structure of rearranged genome segment 11 in two different rotavirus strains generated by a similar mechanism. The Journal of general virology. 1989; 70(Pt 6):1329–1336. [PubMed: 2543781]

Graham KL, Fleming FE, Halasz P, Hewish MJ, Nagesha HS, Holmes IH, Takada Y, Coulson BS. Rotaviruses interact with alpha4beta7 and alpha4beta1 integrins by binding the same integrin domains as natural ligands. The Journal of general virology. 2005; 86:3397–3408. [PubMed: 16298987]

Gray J. Rotavirus vaccines: safety, efficacy and public health impact. Journal of internal medicine. 2011; 270:206–214. [PubMed: 21668823]

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic biology. 2010; 59:307–321. [PubMed: 20525638]

Hoshino Y, Honma S, Jones RW, Ross J, Santos N, Gentsch JR, Kapikian AZ, Hesse RA. A porcine G9 rotavirus strain shares neutralization and VP7 phylogenetic sequence lineage 3 characteristics with contemporary human G9 rotavirus strains. Virology. 2005; 332:177–188. [PubMed: 15661150]

Hoshino Y, Jones RW, Ross J, Honma S, Santos N, Gentsch JR, Kapikian AZ. Rotavirus serotype G9 strains belonging to VP7 gene phylogenetic sequence lineage 1 may be more suitable for serotype G9 vaccine candidates than those belonging to lineage 2 or 3. Journal of virology. 2004; 78:7795–7802. [PubMed: 15220453]

Hull JJ, Teel EN, Kerin TK, Freeman MM, Esona MD, Gentsch JR, Cortese MM, Parashar UD, Glass RI, Bowen MD, National Rotavirus Strain Surveillance, S. United States rotavirus strain surveillance from 2005 to 2008: genotype prevalence before and after vaccine introduction. The Pediatric infectious disease journal. 2011; 30:S42–47. [PubMed: 21183839]

Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. Journal of molecular graphics. 1996; 14:33–38. 27–38. [PubMed: 8744570]

Iturriza-Gomara M, Dallman T, Banyai K, Bottiger B, Buesa J, Diedrich S, Fiore L, Johansen K, Korsun N, Kroneman A, Lappalainen M, Laszlo B, Maunula L, Matthinjnssens J, Midgley S, Mladenova Z, Poljsak-Prijatelj M, Pothier P, Ruggeri FM, Sanchez-Fauquier A, Schreier E, Steyer A, Sidaraviciute I, Tran AN, Usonis V, Van Ranst M, de Rougemont A, Gray J. Rotavirus surveillance in europe, 2005-2008: web-enabled reporting and real-time analysis of genotyping and epidemiological data. The Journal of infectious diseases. 2009; 200(Suppl 1):S215–221. [PubMed: 19821712]

Kirkwood CD, Boniface K, Bishop RF, Barnes GL. Australian Rotavirus Surveillance Program: annual report, 2009/2010. Communicable diseases intelligence. 2010; 34:427–434.

Kobayashi N, Taniguchi K, Urasawa S. Identification of operationally overlapping and independent cross-reactive neutralization regions on human rotavirus VP4. The Journal of general virology. 1990; 71(Pt 11):2615–2623. [PubMed: 1701477]

Kojima K, Taniguchi K, Kawagishi-Kobayashi M, Matsuno S, Urasawa S. Rearrangement generated in double genes, NSP1 and NSP3, of viable progenies from a human rotavirus strain. Virus research. 2000; 67:163–171. [PubMed: 10867195]

Kosakovsky Pond SL, Frost SD. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Molecular biology and evolution. 2005; 22:1208–1222. [PubMed: 15703242]

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. GARD: a genetic algorithm for recombination detection. Bioinformatics. 2006; 22:3096–3098. [PubMed: 17110367]

Larralde G, Li BG, Kapikian AZ, Gorziglia M. Serotype-specific epitope(s) present on the VP8 subunit of rotavirus VP4 protein. Journal of virology. 1991; 65:3213–3218. [PubMed: 1709699]

Maes P, Matthijnssens J, Rahman M, Van Ranst M. RotaC: a web-based tool for the complete genome classification of group A rotaviruses. BMC microbiology. 2009; 9:238. [PubMed: 19930627]

Martella V, Banyai K, Matthijnssens J, Buonavoglia C, Ciarlet M. Zoonotic aspects of rotaviruses. Veterinary microbiology. 2010; 140:246–255. [PubMed: 19781872]

Martinez-Laso J, Roman A, Rodriguez M, Cervera I, Head J, Rodriguez-Avial I, Picazo JJ. Diversity of the G3 genes of human rotaviruses in isolates from Spain from 2004 to 2006: cross-species transmission and inter-genotype recombination generates alleles. The Journal of general virology. 2009; 90:935–943. [PubMed: 19264637]

Matsui SM, Mackow ER, Greenberg HB. Molecular determinant of rotavirus neutralization and protection. Advances in virus research. 1989; 36:181–214. [PubMed: 2472045]

Matthijnssens 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. 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. Journal of virology. 2008a; 82:3204–3219. [PubMed: 18216098]

Matthijnssens 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, Parreno V, Rahman M, Ruggeri FM, Saif LJ, Santos N, Steyer A, Taniguchi K, Patton JT, Desselberger U, Van Ranst M. Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). Archives of virology. 2011; 156:1397–1413. [PubMed: 21597953]

Matthijnssens J, Ciarlet M, Rahman M, Attoui H, Banyai K, Estes MK, Gentsch JR, Iturriza-Gomara M, Kirkwood CD, Martella V, Mertens PP, Nakagomi O, Patton JT, Ruggeri FM, Saif LJ, Santos N, Steyer A, Taniguchi K, Desselberger U, Van Ranst M. Recommendations for the classification of group A rotaviruses using all 11 genomic RNA segments. Archives of virology. 2008b; 153:1621–1629. [PubMed: 18604469]

Matthijnssens J, Heylen E, Zeller M, Rahman M, Lemey P, Van Ranst M. Phylodynamic analyses of rotavirus genotypes G9 and G12 underscore their potential for swift global spread. Molecular biology and evolution. 2010a; 27:2431–2436. [PubMed: 20522727]

Matthijnssens J, Joelsson DB, Warakomski DJ, Zhou T, Mathis PK, van Maanen MH, Ranheim TS, Ciarlet M. Molecular and biological characterization of the 5 human-bovine rotavirus (WC3)-based reassortant strains of the pentavalent rotavirus vaccine, RotaTeq. Virology. 2010b; 403:111–127. [PubMed: 20451234]

Matthijnssens J, Mino S, Papp H, Potgieter C, Novo L, Heylen E, Zeller M, Garaicoechea L, Badaracco A, Lengyel G, Kisfali P, Cullinane A, Collins PJ, Ciarlet M, O'Shea H, Parreno V, Banyai K, Barrandeguy M, Van Ranst M. Complete molecular genome analyses of equine rotavirus A strains from different continents reveal several novel genotypes and a largely conserved genotype constellation. The Journal of general virology. 2012a; 93:866–875. [PubMed: 22190012]

Matthijnssens J, Nakagomi O, Kirkwood CD, Ciarlet M, Desselberger U, Van Ranst M. Group A rotavirus universal mass vaccination: how and to what extent will selective pressure influence prevalence of rotavirus genotypes? Expert review of vaccines. 2012b; 11:1347–1354. [PubMed: 23249234]

Maunula L, Von Bonsdorff CH. Frequent reassortments may explain the genetic heterogeneity of rotaviruses: analysis of Finnish rotavirus strains. Journal of virology. 2002; 76:11793–11800. [PubMed: 12414921]

McDonald SM, Davis K, McAllen JK, Spiro DJ, Patton JT. Intra-genotypic diversity of archival G4P[8] human rotaviruses from Washington, DC. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2011; 11:1586–1594.

McDonald SM, Matthijnssens J, McAllen JK, Hine E, Overton L, Wang S, Lemey P, Zeller M, Van Ranst M, Spiro DJ, Patton JT. Evolutionary dynamics of human rotaviruses: balancing reassortment with preferred genome constellations. PLoS pathogens. 2009; 5:e1000634. [PubMed: 19851457]

McDonald SM, McKell AO, Rippinger CM, McAllen JK, Akopov A, Kirkness EF, Payne DC, Edwards KM, Chappell JD, Patton JT. Diversity and relationships of cocirculating modern human rotaviruses revealed using large-scale comparative genomics. Journal of virology. 2012; 86:9148–9162. [PubMed: 22696651]

McKinney BA, Kallewaard NL, Crowe JE Jr. Meiler J. Using the natural evolution of a rotavirus-specific human monoclonal antibody to predict the complex topography of a viral antigenic site. Immunome research. 2007; 3:8. [PubMed: 17877819]

Mijatovic-Rustempasic S, Teel EN, Kerin TK, Hull JJ, Roy S, Weinberg GA, Payne DC, Parashar UD, Gentsch JR, Bowen MD. Genetic analysis of G12P[8] rotaviruses detected in the largest U.S. G12 genotype outbreak on record. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2014; 21:214–219.

Mlera L, O'Neill HG, Jere KC, van Dijk AA. Whole-genome consensus sequence analysis of a South African rotavirus SA11 sample reveals a mixed infection with two close derivatives of the SA11-H96 strain. Archives of virology. 2013; 158:1021–1030. [PubMed: 23263646]

Monnier N, Higo-Moriguchi K, Sun ZY, Prasad BV, Taniguchi K, Dormitzer PR. High-resolution molecular and antigen structure of the VP8* core of a sialic acid-independent human rotavirus strain. Journal of virology. 2006; 80:1513–1523. [PubMed: 16415027]

Palombo EA. Genetic analysis of Group A rotaviruses: evidence for interspecies transmission of rotavirus genes. Virus genes. 2002; 24:11–20. [PubMed: 11928984]

Parashar UD, Burton A, Lanata C, Boschi-Pinto C, Shibuya K, Steele D, Birmingham M, Glass RI. Global mortality associated with rotavirus disease among children in 2004. The Journal of infectious diseases. 2009; 200(Suppl 1):S9–S15. [PubMed: 19817620]

Payne DC, Humiston S, Opel D, Kennedy A, Wikswo M, Downing K, Klein EJ, Kobayashi A, Locke D, Albertin C, Chesley C, Staat MA. A multi-center, qualitative assessment of pediatrician and maternal perspectives on rotavirus vaccines and the detection of Porcine circovirus. BMC pediatrics. 2011; 11:83. [PubMed: 21943237]

Payne DC, Staat MA, Edwards KM, Szilagyi PG, Gentsch JR, Stockman LJ, Curns AT, Griffin M, Weinberg GA, Hall CB, Fairbrother G, Alexander J, Parashar UD. Active, population-based surveillance for severe rotavirus gastroenteritis in children in the United States. Pediatrics. 2008; 122:1235–1243. [PubMed: 19047240]

Phan TG, Okitsu S, Maneekarn N, Ushijima H. Evidence of intragenic recombination in G1 rotavirus VP7 genes. Journal of virology. 2007; 81:10188–10194. [PubMed: 17609273]

Phua KB, Lim FS, Lau YL, Nelson EA, Huang LM, Quak SH, Lee BW, Teoh YL, Tang H, Boudville I, Oostvogels LC, Suryakiran PV, Smolenov IV, Han HH, Bock HL. Safety and efficacy of human rotavirus vaccine during the first 2 years of life in Asian infants: randomised, double-blind, controlled study. Vaccine. 2009; 27:5936–5941. [PubMed: 19679216]

Pond SL, Frost SD. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics. 2005; 21:2531–2533. [PubMed: 15713735]

Song XF, Hao Y. Adaptive evolution of rotavirus VP7 and NSP4 genes in different species. Computational biology and chemistry. 2009; 33:344–349. [PubMed: 19665933]

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular biology and evolution. 2011; 28:2731–2739. [PubMed: 21546353]

Tate JE, Burton AH, Boschi-Pinto C, Steele AD, Duque J, Parashar UD, Network W.H.-c.G.R.S. 2008 estimate of worldwide rotavirus-associated mortality in children younger than 5 years before the introduction of universal rotavirus vaccination programmes: a systematic review and meta-analysis. The Lancet infectious diseases. 2012; 12:136–141. [PubMed: 22030330]

Trask SD, Kim IS, Harrison SC, Dormitzer PR. A rotavirus spike protein conformational intermediate binds lipid bilayers. Journal of virology. 2010; 84:1764–1770. [PubMed: 20007281]

Trojnar E, Sachsenroder J, Twardziok S, Reetz J, Otto PH, Johne R. Identification of an avian group A rotavirus containing a novel VP4 gene with a close relationship to those of mammalian rotaviruses. The Journal of general virology. 2013; 94:136–142. [PubMed: 23052396]

Tsugawa T, Tatsumi M, Tsutsumi H. Virulence-associated genome mutations of murine rotavirus identified by alternating serial passages in mice and cell cultures. Journal of virology. 2014; 88:5543–5558. [PubMed: 24599996]

Varghese V, Das S, Singh NB, Kojima K, Bhattacharya SK, Krishnan T, Kobayashi N, Naik TN. Molecular characterization of a human rotavirus reveals porcine characteristics in most of the genes including VP6 and NSP4. Archives of virology. 2004; 149:155–172. [PubMed: 14689281]

Ward R. Mechanisms of protection against rotavirus infection and disease. The Pediatric infectious disease journal. 2009; 28:S57–59. [PubMed: 19252425]

Yoder JD, Dormitzer PR. Alternative intermolecular contacts underlie the rotavirus VP5* two- to three-fold rearrangement. The EMBO journal. 2006; 25:1559–1568. [PubMed: 16511559]

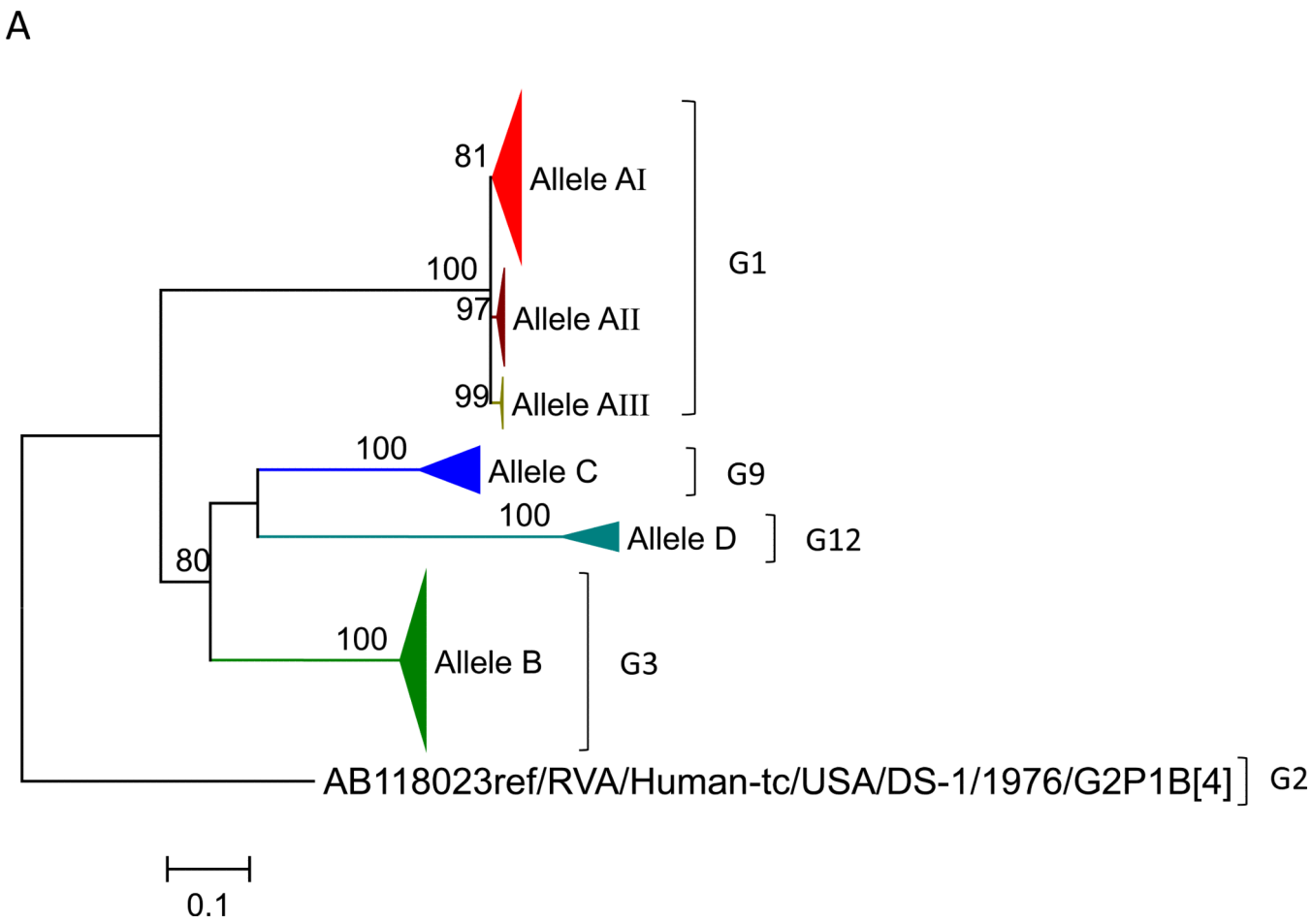
Zaman K, Dang DA, Victor JC, Shin S, Yunus M, Dallas MJ, Podder G, Vu DT, Le TP, Luby SP, Le HT, Coia ML, Lewis K, Rivers SB, Sack DA, Schodel F, Steele AD, Neuzil KM, Ciarlet M. Efficacy of pentavalent rotavirus vaccine against severe rotavirus gastroenteritis in infants in developing countries in Asia: a randomised, double-blind, placebo-controlled trial. Lancet. 2010; 376:615–623. [PubMed: 20692031]

Zeller M, Patton JT, 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. Journal of clinical microbiology. 2012; 50:966–976. [PubMed: 22189107]

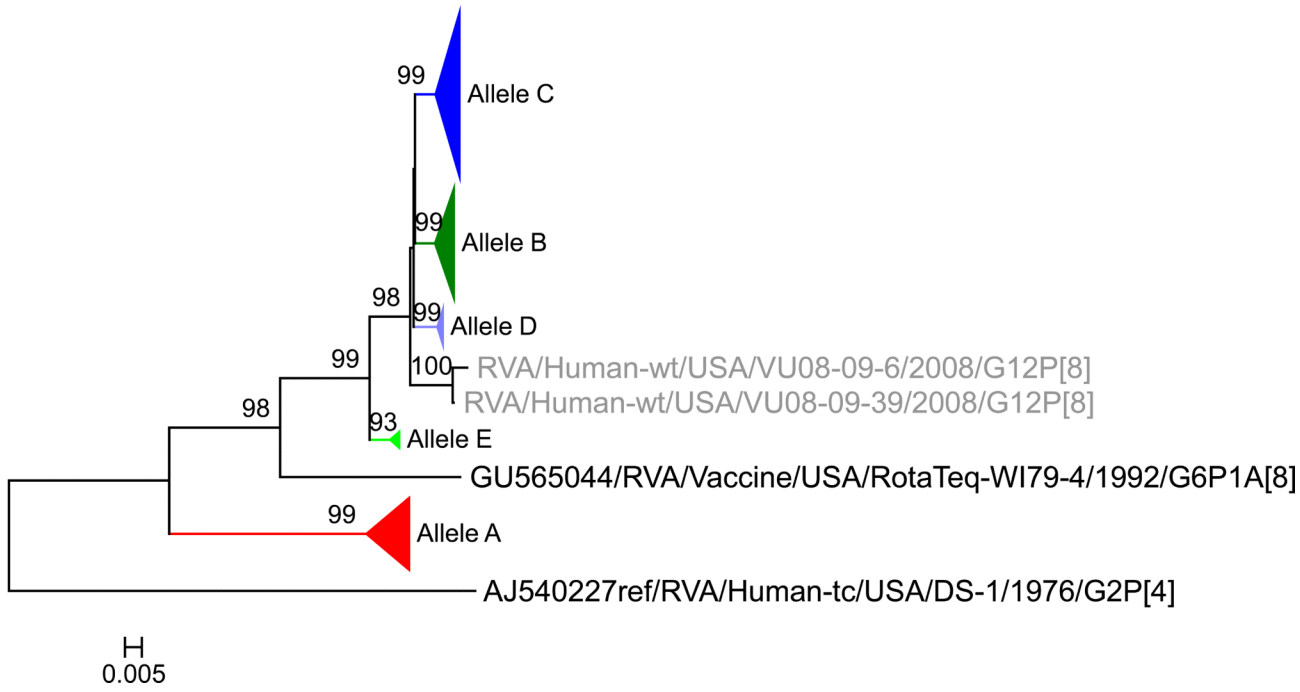
Zeller M, Rahman M, Heylen E, De Coster S, De Vos S, Arijs I, Novo L, Verstappen N, Van Ranst M, Matthijnssens J. Rotavirus incidence and genotype distribution before and after national rotavirus vaccine introduction in Belgium. Vaccine. 2010; 28:7507–7513. [PubMed: 20851085]

Zhang J. Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. Journal of molecular evolution. 2000; 50:56–68. [PubMed: 10654260]

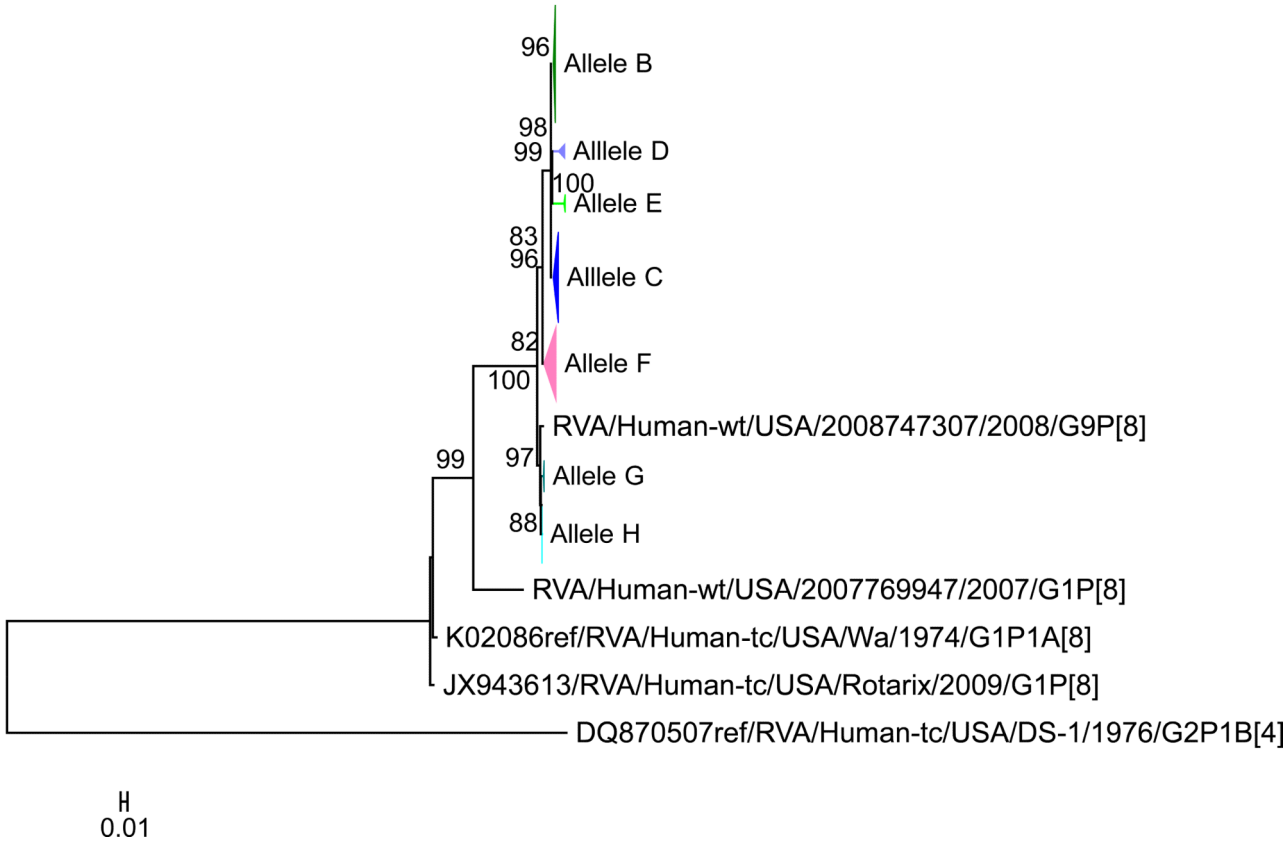
Zhou YJ, Burns JW, Morita Y, Tanaka T, Estes MK. Localization of rotavirus VP4 neutralization epitopes involved in antibody-induced conformational changes of virus structure. Journal of virology. 1994; 68:3955–3964. [PubMed: 7514681]



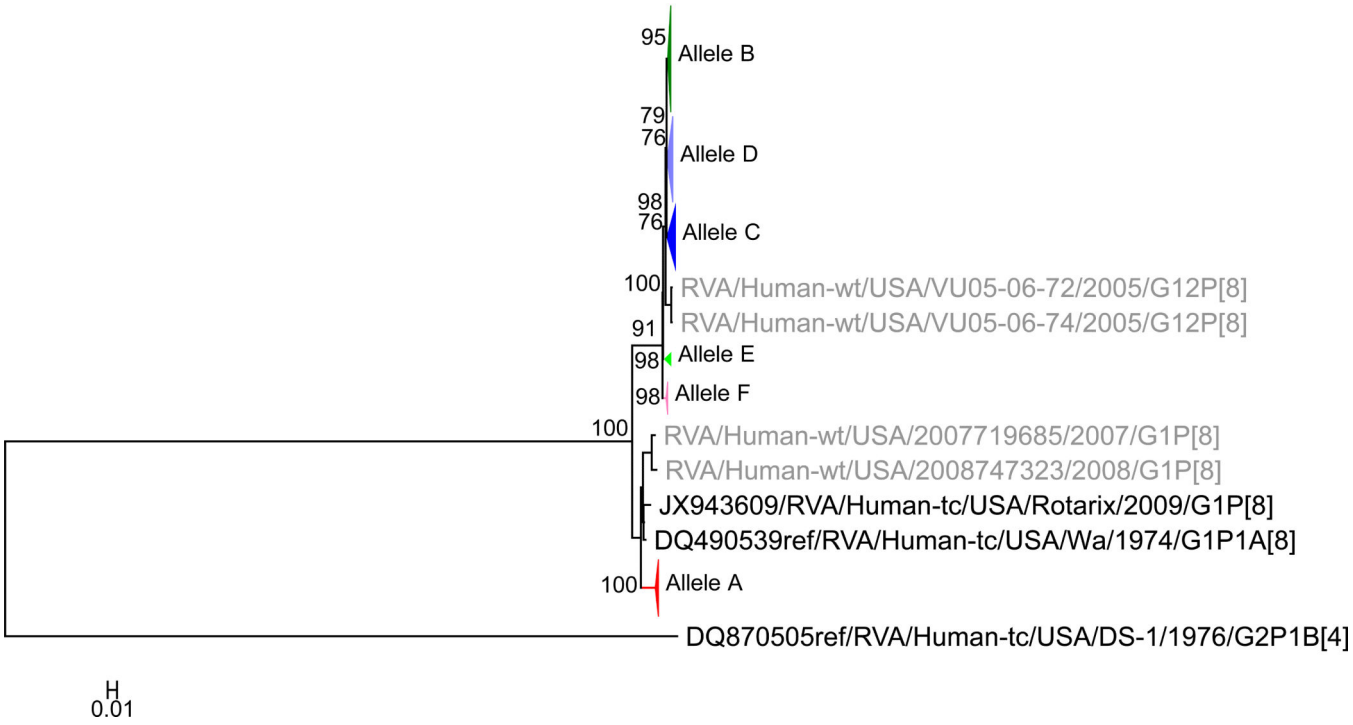
B



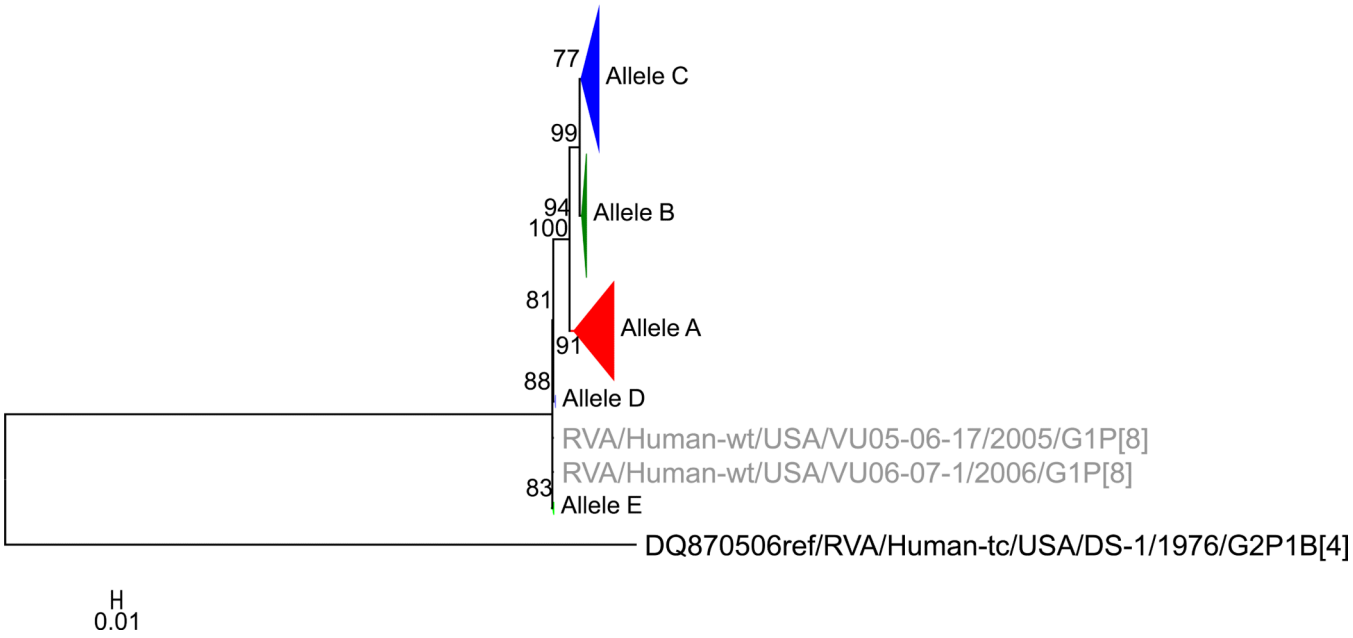
C



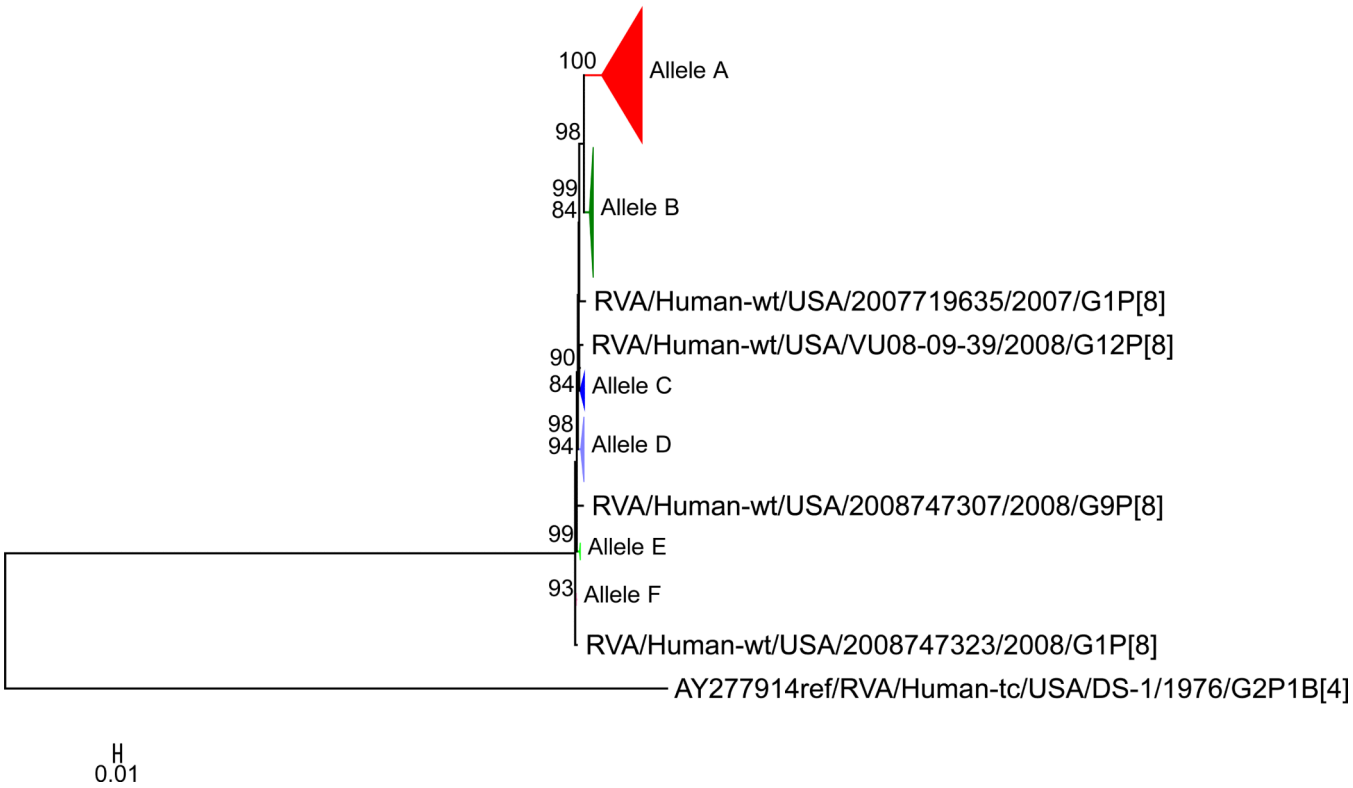
D



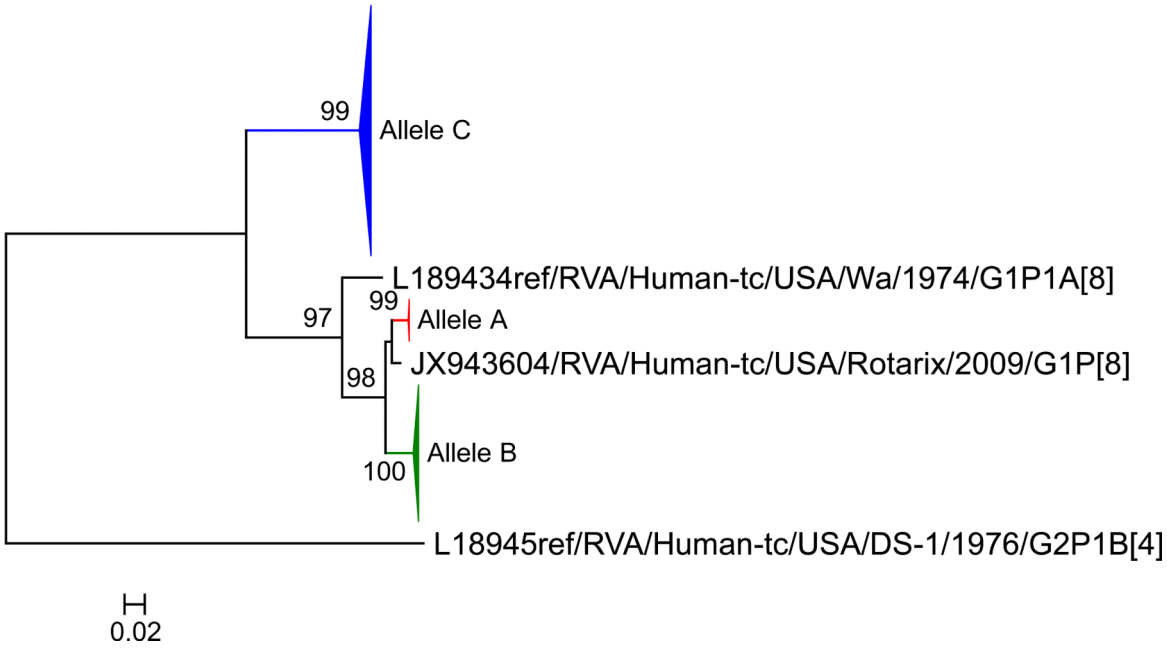
E



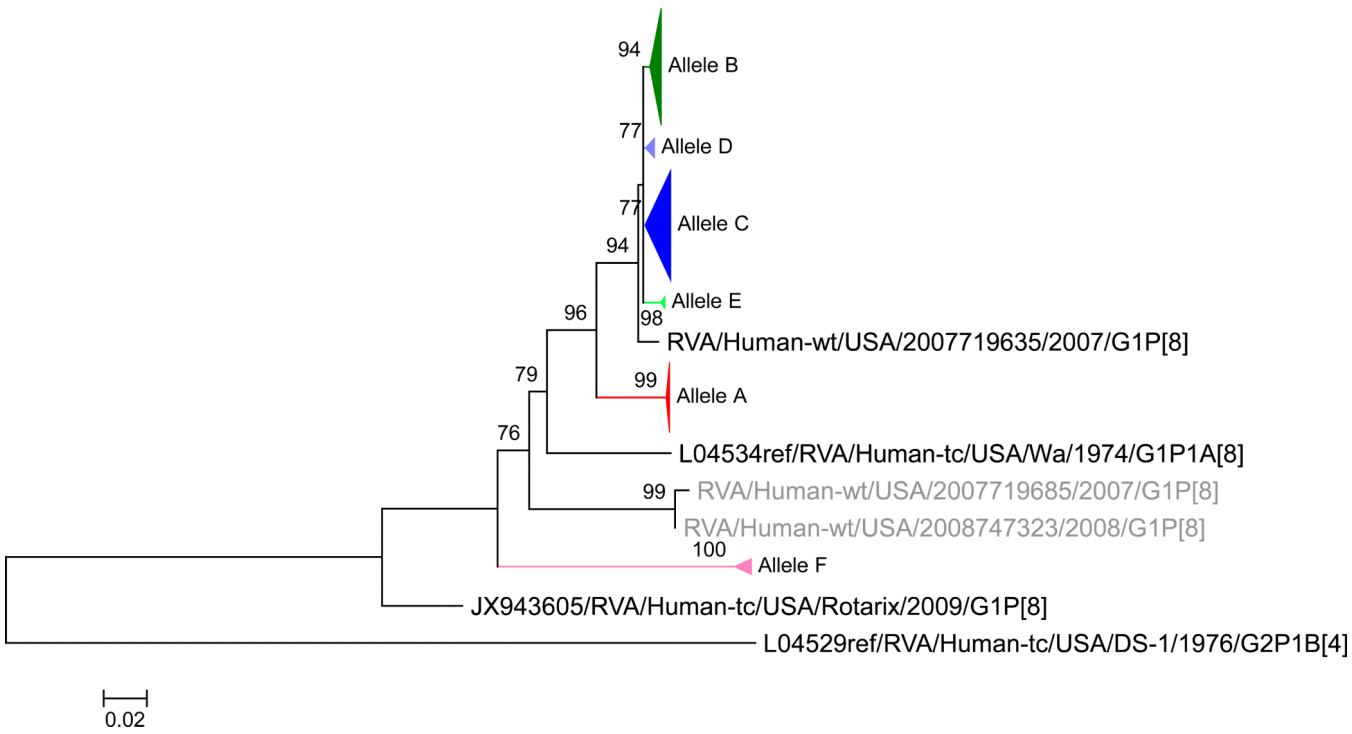
F

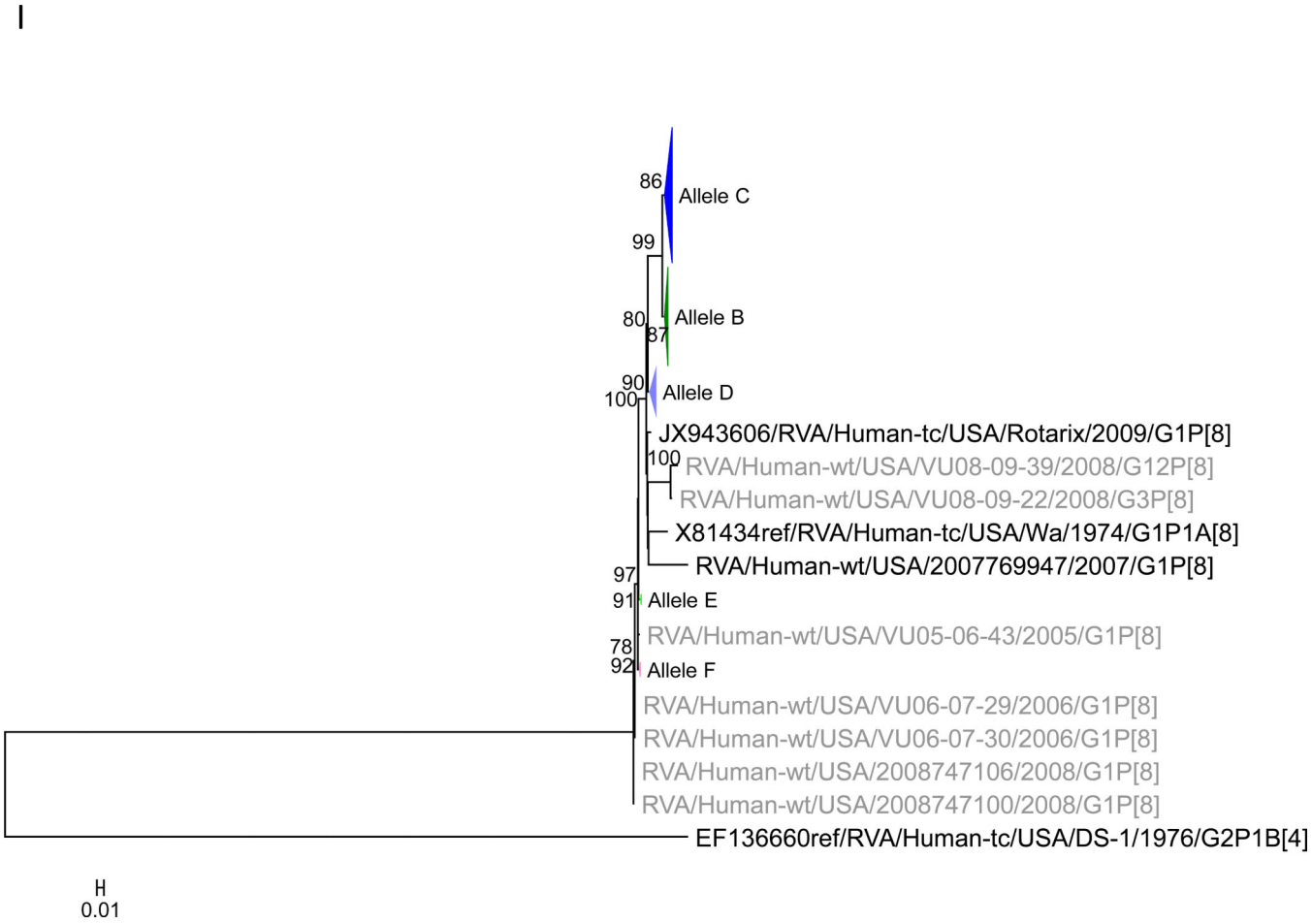


G

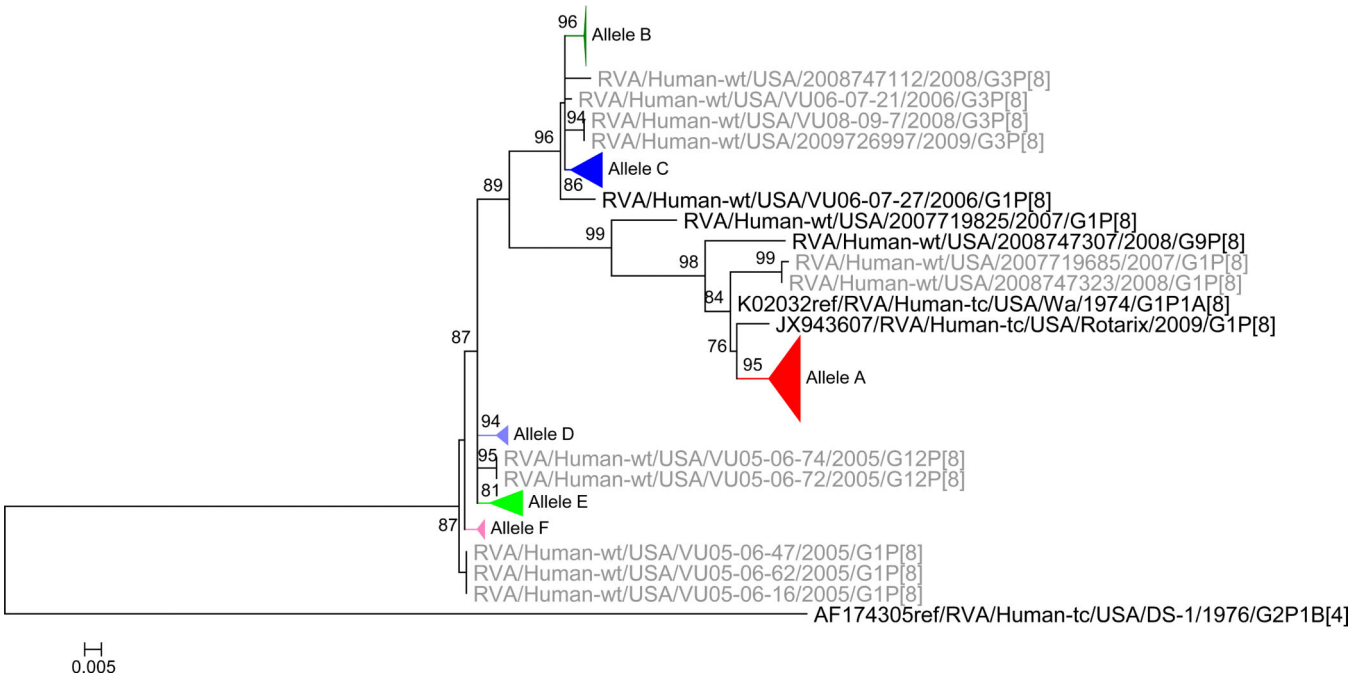


H





J



K

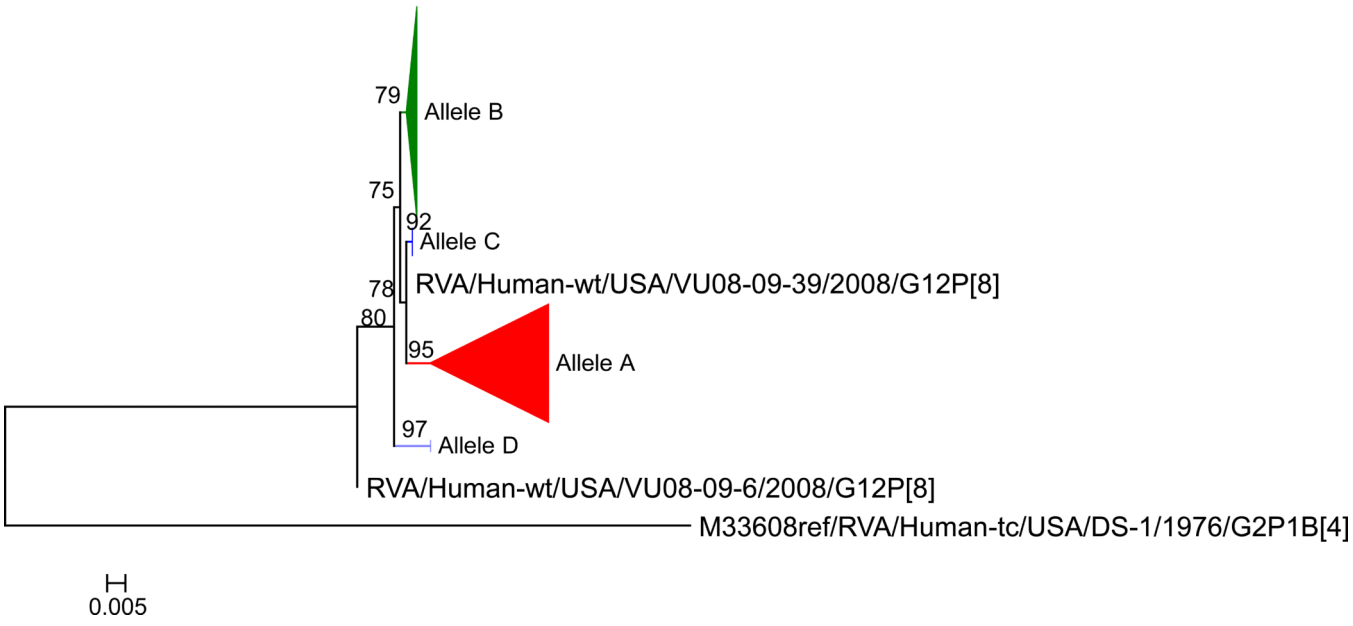
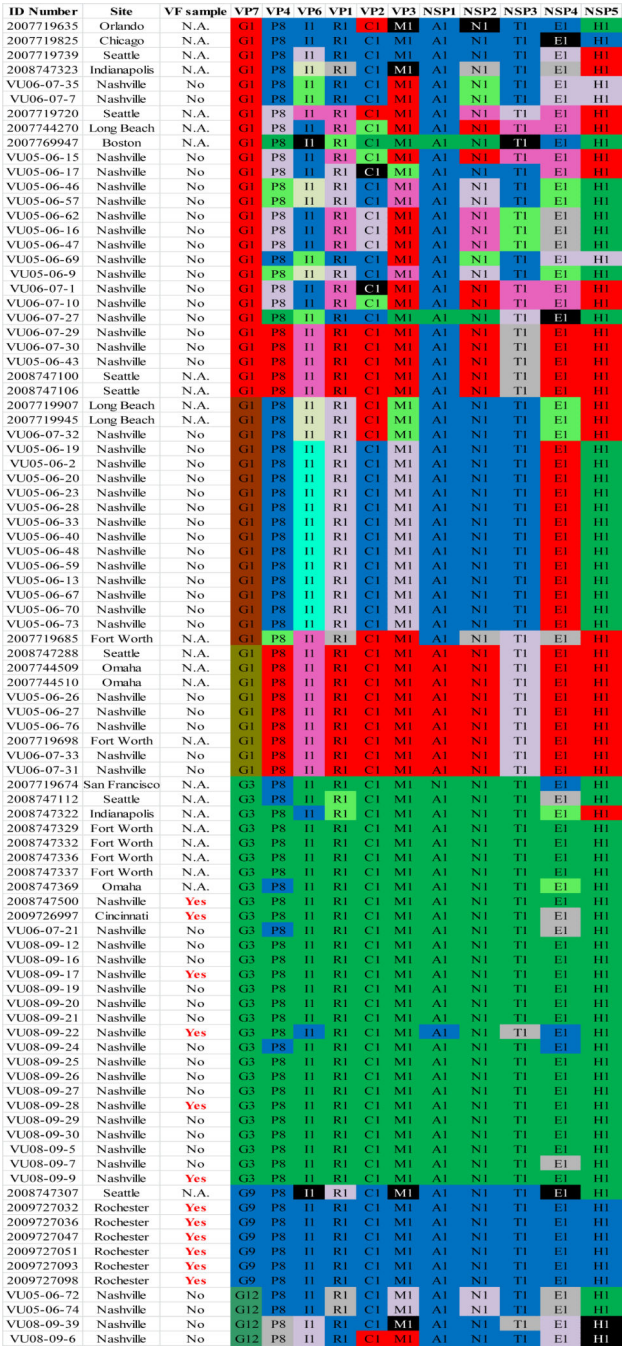


Figure 1.

Maximum likelihood trees with aLRT values showing branch support for the 11 RVA genes. The different alleles are colored in red, green, blue, purple, lime, pink, teal and aqua for Alleles A, B, C, D, E, F, G and H respectively. Doublets and singletons are shown in grey and black, respectively (Fig. 1A) VP7. The Allele A in further divided into three clusters: red strains that mostly cluster with G1 (Wa) strains and maroon and olive strains that forms a distinct clusters from the reference Wa Strain. (1B) VP4; (1C) VP6; (1D) VP1; (1E) VP2; (1F) VP3; (1G) NSP1; (1H) NSP2; (1I) NSP3; (1J) NSP4; (1K) NSP5. To see the individual strains comprising each colored triangle, consult the supplemental figures (see supplementary material).



ID Number	Site	VF sample	VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5
2007719635	Orlando	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2007719825	Chicago	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2007719739	Seattle	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2008747323	Indianapolis	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU06-07-35	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU06-07-7	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2007719720	Seattle	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2007744270	Long Beach	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2007769947	Boston	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-15	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-17	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-46	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-57	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-62	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-16	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-47	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-69	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-9	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU06-07-1	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU06-07-10	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU06-07-27	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU06-07-29	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU06-07-30	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-43	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2008747100	Seattle	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2008747106	Seattle	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2007719907	Long Beach	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2007719945	Long Beach	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU06-07-32	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-19	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-2	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-20	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-23	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-28	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-33	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-40	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-48	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-59	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-13	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-67	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-70	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-73	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2007719685	Fort Worth	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2008747288	Seattle	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2007744509	Omaha	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2007744510	Omaha	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-26	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-27	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-76	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2007719698	Fort Worth	N.A.	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU06-07-33	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU06-07-31	Nashville	No	G1	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2007719674	San Francisco	N.A.	G3	P8	I1	R1	C1	M1	N1	N1	T1	E1	H1
2008747112	Seattle	N.A.	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2008747322	Indianapolis	N.A.	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2008747329	Fort Worth	N.A.	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2008747332	Fort Worth	N.A.	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2008747336	Fort Worth	N.A.	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2008747337	Fort Worth	N.A.	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2008747369	Omaha	N.A.	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2008747500	Nashville	Yes	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2009726997	Cincinnati	Yes	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU06-07-21	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-12	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-16	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-17	Nashville	Yes	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-19	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-20	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-21	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-22	Nashville	Yes	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-24	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-25	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-26	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-27	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-28	Nashville	Yes	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-29	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-30	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-5	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-7	Nashville	No	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-9	Nashville	Yes	G3	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2008747307	Seattle	N.A.	G9	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2009727032	Rochester	Yes	G9	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2009727036	Rochester	Yes	G9	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2009727047	Rochester	Yes	G9	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2009727051	Rochester	Yes	G9	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2009727093	Rochester	Yes	G9	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
2009727098	Rochester	Yes	G9	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-72	Nashville	No	G12	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU05-06-74	Nashville	No	G12	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-39	Nashville	No	G12	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1
VU08-09-6	Nashville	No	G12	P8	I1	R1	C1	M1	A1	N1	T1	E1	H1

Figure 2.

Allelic clusters for the 91 RVA strains. Genes from strains that clustered with Allele A (G1) are shown in red, maroon and olive. Allele B (G3) genes are shown in green and Allele C (G9) strains are shown in blue, Independent clusters are shown in purple (Allele D), lime (Allele E), pink (Allele F), teal (Allele G) and aqua (Allele H). Doublet strains are shown in grey and singleton strains in black.

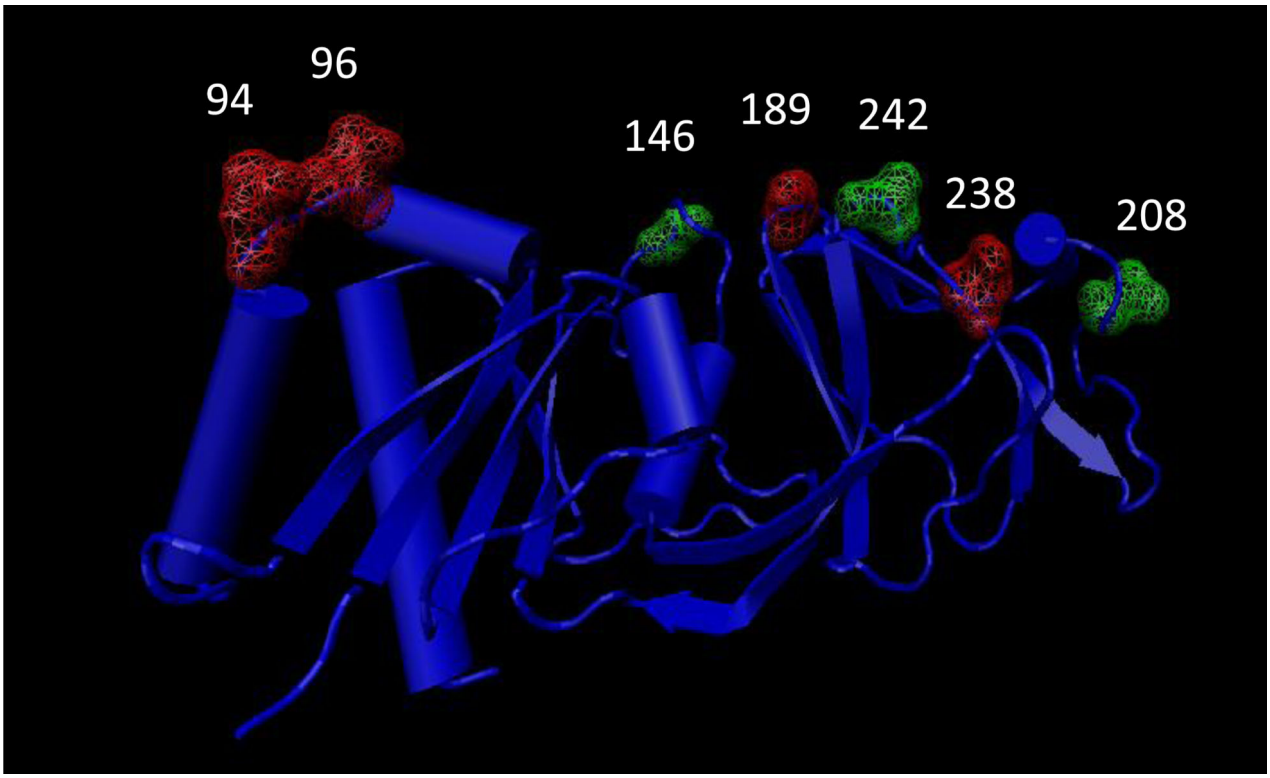


Figure 3.
Crystal structure of RVA VP7 protein (3FMG). Substitutions at the antigenic site are marked in green and substitutions at reported neutralization escape mutation sites are marked in red.

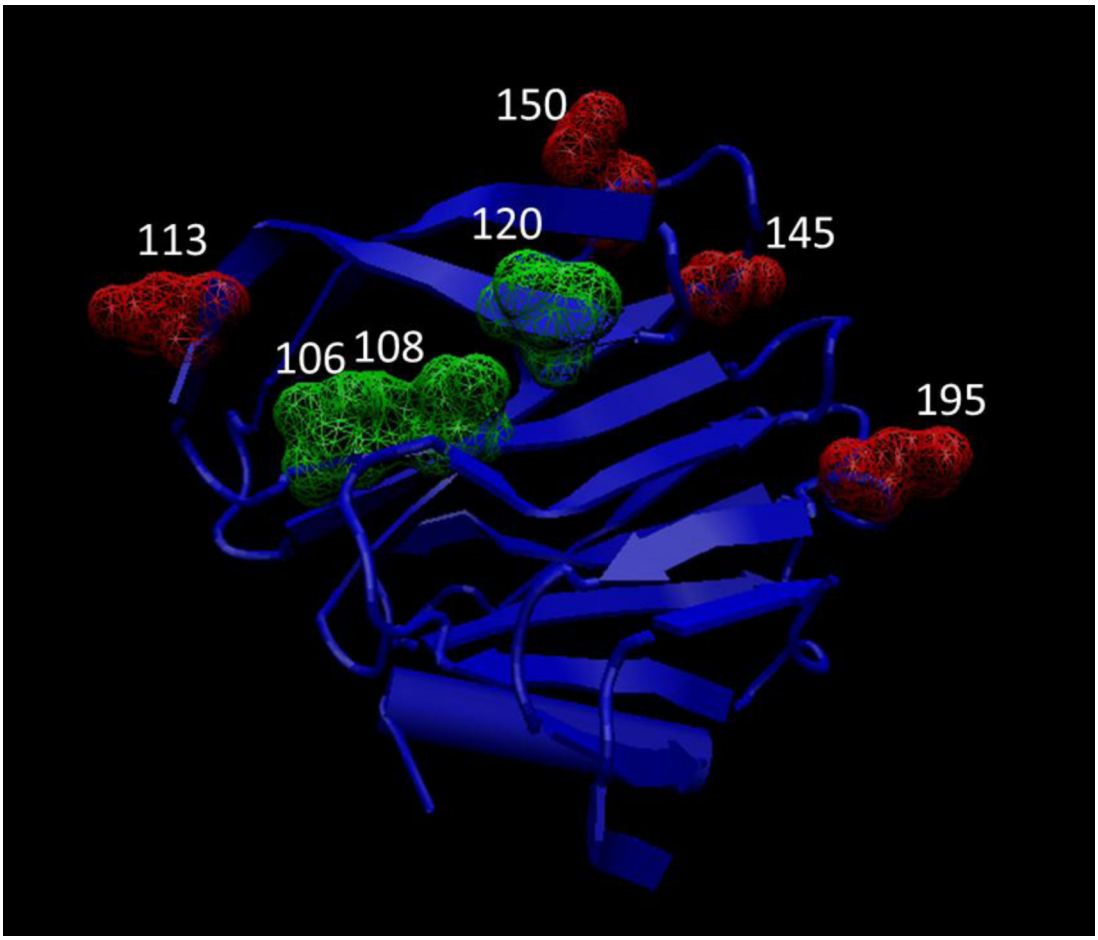


Figure 4.
Crystal structure of RVA VP8* region (2DWR). Substitutions at the antigenic site are marked in green and substitutions at reported neutralization escape mutation sites are marked in red.

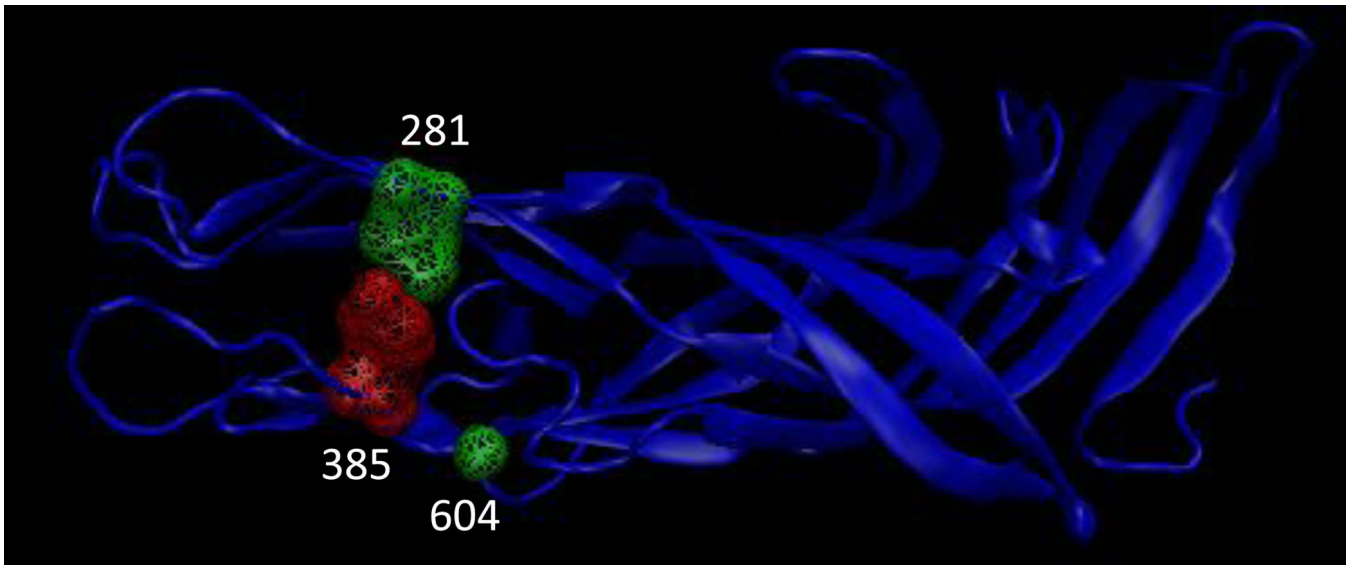


Figure 5.
Crystal structure of RVA VP5* region (2B4I). Substitutions at the antigenic site are marked in green and substitutions at reported neutralization escape mutation sites are marked in red.

Table 1

Accession Numbers for 33 RVA strains determined in this study.

Case ID	VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5
RVA/Human-wt/USA/2007719635/2007/G1P[8]	JN258368	JN258371	JN258370	JN258364	JN258374	JN258373	JN258372	JN258365	JN258369	JN258367	JN258366
RVA/Human-wt/USA/2007719674/2007/G3P[8]	JN258362	JN258360	JN258359	JN258353	JN258363	JN258361	JN258355	JN258354	JN258358	JN258357	JN258356
RVA/Human-wt/USA/2007719685/2007/G1P[8]	JN258346	JN258349	JN258348	JN258342	JN258350	JN258351	JN258352	JN258343	JN258347	JN258345	JN258344
RVA/Human-wt/USA/2007719698/2007/G1P[8]	HM773774	HM773769	HM773771	HM773766	HM773767	HM773768	HM773770	HM773773	HM773772	HM773775	HM773776
RVA/Human-wt/USA/2007719720/2007/G1P[8]	JN258338	JN258340	JN258335	JN258331	JN258341	JN258336	JN258337	JN258332	JN258334	JN258339	JN258333
RVA/Human-wt/USA/2007719739/2007/G1P[8]	HM773763	HM773758	HM773760	HM773755	HM773756	HM773757	HM773759	HM773762	HM773761	HM773764	HM773765
RVA/Human-wt/USA/2007719825/2007/G1P[8]	HM773752	HM773747	HM773749	HM773744	HM773745	HM773746	HM773748	HM773751	HM773750	HM773753	HM773754
RVA/Human-wt/USA/2007719907/2007/G1P[8]	HM773851	HM773846	HM773848	HM773843	HM773844	HM773845	HM773847	HM773850	HM773849	HM773852	HM773853
RVA/Human-wt/USA/2007719945/2007/G1P[8]	HM773840	HM773835	HM773837	HM773832	HM773833	HM773834	HM773836	HM773839	HM773838	HM773841	HM773842
RVA/Human-wt/USA/2007744270/2007/G1P[8]	HM773829	HM773824	HM773826	HM773821	HM773822	HM773823	HM773825	HM773828	HM773827	HM773830	HM773831
RVA/Human-wt/USA/2007744509/2007/G1P[8]	HM773818	HM773813	HM773815	HM773810	HM773811	HM773812	HM773814	HM773817	HM773816	HM773819	HM773820
RVA/Human-wt/USA/2007744510/2007/G1P[8]	HM773807	HM773802	HM773804	HM773799	HM773800	HM773801	HM773803	HM773806	HM773805	HM773808	HM773809
RVA/Human-wt/USA/2007769947/2007/G1P[8]	JN258401	JN258404	JN258403	JN258397	JN258406	JN258405	JN258407	JN258398	JN258402	JN258400	JN258399
RVA/Human-wt/USA/2008747100/2008/G1P[8]	HM773796	HM773791	HM773793	HM773788	HM773789	HM773790	HM773792	HM773795	HM773794	HM773797	HM773798
RVA/Human-wt/USA/2008747106/2008/G1P[8]	HM773785	HM773780	HM773782	HM773777	HM773778	HM773779	HM773781	HM773784	HM773783	HM773786	HM773787
RVA/Human-wt/USA/2008747112/2008/G3P[8]	HM773719	HM773714	HM773716	HM773711	HM773712	HM773713	HM773715	HM773718	HM773717	HM773720	HM773721
RVA/Human-wt/USA/2008747288/2008/G1P[8]	JN258380	JN258383	JN258382	JN258375	JN258385	JN258384	JN258377	JN258376	JN258381	JN258379	JN258378
RVA/Human-wt/USA/2008747307/2008/G9P[8]	HM773642	HM773637	HM773639	HM773634	HM773635	HM773636	HM773638	HM773641	HM773640	HM773643	HM773644
RVA/Human-wt/USA/2008747322/2008/G3P[8]	HM773741	HM773736	HM773738	HM773733	HM773734	HM773735	HM773737	HM773740	HM773739	HM773742	HM773743
RVA/Human-wt/USA/2008747323/2008/G1P[8]	JN258390	JN258393	JN258392	JN258386	JN258396	JN258394	JN258395	JN258387	JN258391	JN258389	JN258388
RVA/Human-wt/USA/2008747329/2008/G3P[8]	HM773708	HM773703	HM773705	HM773700	HM773701	HM773702	HM773704	HM773707	HM773706	HM773709	HM773710
RVA/Human-wt/USA/2008747332/2008/G3P[8]	HM773697	HM773692	HM773694	HM773689	HM773690	HM773691	HM773693	HM773696	HM773695	HM773698	HM773699
RVA/Human-wt/USA/2008747336/2008/G3P[8]	HM773686	HM773681	HM773683	HM773678	HM773679	HM773680	HM773682	HM773685	HM773684	HM773687	HM773688
RVA/Human-wt/USA/2008747337/2008/G3P[8]	HM773675	HM773670	HM773672	HM773667	HM773668	HM773669	HM773671	HM773674	HM773673	HM773676	HM773677
RVA/Human-wt/USA/2008747369/2008/G3P[8]	HM773664	HM773659	HM773661	HM773656	HM773657	HM773658	HM773660	HM773663	HM773662	HM773665	HM773666
RVA/Human-wt/USA/2008747500/2008/G3P[8]	HM773653	HM773648	HM773650	HM773645	HM773646	HM773647	HM773649	HM773652	HM773651	HM773654	HM773655
RVA/Human-wt/USA/2009726997/2009/G3P[8]	HM773730	HM773725	HM773727	HM773722	HM773723	HM773724	HM773726	HM773729	HM773728	HM773731	HM773732

Case ID	VP7	VP4	VP6	VP1	VP2	VP3	NSP1	NSP2	NSP3	NSP4	NSP5
RVA/Human-wt/USA/2009727032/2009/G9P[8]	HM773587	HM773582	HM773584	HM773579	HM773580	HM773581	HM773583	HM773586	HM773585	HM773588	HM773589
RVA/Human-wt/USA/2009727036/2009/G9P[8]	HM773598	HM773593	HM773595	HM773590	HM773591	HM773592	HM773594	HM773597	HM773596	HM773599	HM773600
RVA/Human-wt/USA/2009727047/2009/G9P[8]	HM773620	HM773615	HM773617	HM773612	HM773613	HM773614	HM773616	HM773619	HM773618	HM773621	HM773622
RVA/Human-wt/USA/2009727051/2009/G9P[8]	HM773631	HM773626	HM773628	HM773623	HM773624	HM773625	HM773627	HM773630	HM773629	HM773632	HM773633
RVA/Human-wt/USA/2009727093/2009/G9P[8]	HM534680	HM534675	HM534677	HM534672	HM534673	HM534674	HM534676	HM534679	HM534678	HM534681	HM534682
RVA/Human-wt/USA/2009727098/2009/G9P[8]	HM773609	HM773604	HM773606	HM773601	HM773602	HM773603	HM773605	HM773608	HM773607	HM773610	HM773611

Table 2

Results of selection analyses of 11 rotavirus proteins.

Protein	No. of sites	No. of sites under positive selection	No. (%) of sites under purifying selection
NSP1	491	0	217 (44.2)
NSP2	317	0	107 (33.8)
NSP3	315	0	132 (41.9)
NSP4	175	0	38 (21.7)
NSP5	198	0	46 (23.2)
VP1	1088	0	421 (38.9)
VP2	881	0	227 (25.8)
VP3	835	0	246 (29.5)
VP4	776	0	332 (42.8)
VP6	397	0	197 (49.6)
VP7	326	0	187 (57.1)