Comparative genomic analysis of genogroup 1 and genogroup 2 rotaviruses circulating in seven US cities, 2014–2016

Mathew D. Esona,1,* Rashi Gautam,1 Eric Katz,2 Jose Jaime,1 M. Leanne Ward,1 Mary E. Wikswo,1 Naga S. Betrapally,1 Slavica M. Rustempasic,1 Rangaraj Selvarangan,3 Christopher J. Harrison,3 Julie A. Boom,4 Jan Englund,5 Eileen J. Klein,5 Mary Allen Staat,6 Monica M. McNeal,6 Natasha Halasa,7 James Chappell,7 Geoffrey A. Weinberg,8 Daniel C. Payne,1 Umesh D. Parashar,1 and Michael D. Bowen1

1Division of Viral Diseases, Centers for Disease Control and Prevention, Viral Gastroenteritis Branch, Atlanta, GA, USA, 2Cherokee Nation Assurance, Contracting Agency to the Division of Viral Diseases, Centers for Disease Control and Prevention, Arlington, VA, USA, 3Kansas City Children’s Mercy Hospitals and Clinics, Kansas City, MO, USA, 4Texas Children’s Hospital, Houston, TX, USA, 5Seattle Children’s Hospital, Seattle, WA, USA, 6Division of Infectious Diseases, Department of Pediatrics, University of Cincinnati, Cincinnati Children’s Hospital Medical Center, Cincinnati, OH, USA, 7Vanderbilt University Medical Center, Nashville, TN, USA and 8University of Rochester School of Medicine and Dentistry, Rochester, NY, USA

*Corresponding author: E-mail: mdi4@cdc.gov

Abstract

For over a decade, the New Vaccine Surveillance Network (NVSN) has conducted active rotavirus (RVA) strain surveillance in the USA. The evolution of RVA in the post-vaccine introduction era and the possible effects of vaccine pressure on contemporary circulating strains in the USA are still under investigation. Here, we report the whole-gene characterization (eleven ORFs) for 157 RVA strains collected at seven NVSN sites during the 2014 through 2016 seasons. The sequenced strains included 52 G1P[8], 47 G12P[8], 18 G9P[8], 24 G2P[4], 5 G3P[6], as well as 7 vaccine strains, a single mixed strain (G9G12P[8]), and 3 less common strains. The majority of the single and mixed strains possessed a Wa-like backbone with consensus genotype constellation of G1/G3/G9/G12-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1, while the G2P[4], G3P[6], and G2P[8] strains displayed a DS-1-like genetic backbone with consensus constellation of G2/G3-P[4]/P[6]/P[8]-I2-R2-C2-M2-A2-N2-T2-E2-H2. Two intergenogroup reassortant G1P[8] strains were detected that appear to be progenies of reassortment events between Wa-like G1P[8] and DS-1-like G2P[4] strains. Two Rotarix® vaccine (RV1) and two RV5 derived (vd) reassortant strains were detected. Phylogenetic and similarity matrices analysis revealed 2–11 sub-genotypic allelic clusters among the genes of Wa- and DS-1-like strains. Most study strains clustered into previously defined alleles. Amino acid (AA) substitutions occurring in the neutralization epitopes of the VP7 and VP4 proteins characterized in this study were mostly neutral in nature, suggesting that these RVA proteins were possibly under strong negative or purifying selection in order to maintain competent and actual functionality, but fourteen radical (AA changes that occur between groups) AA substitutions were noted that may allow RVA strains to gain a selective advantage through immune escape. The tracking of RVA strains at the sub-genotypic allele constellation level will enhance our understanding of RVA
evolution under vaccine pressure, help identify possible mechanisms of immune escape, and provide valuable information for formulation of future RVA vaccines.

Key words: RVA; whole-genome sequencing; alleles; neutralization epitopes; NVSN; USA.

1. Introduction

Diarrhea is a leading cause of morbidity and mortality in young children under the age of 5 years globally (Estes 2013; Liu et al. 2016). A significant number of hospitalizations and deaths due to severe dehydrating diarrhea amongst this group are attributed to group A rotaviruses (RVAs) (Liu et al. 2016). RVAs are members of the Reoviridae family and are a universal pathogen that causes acute gastroenteritis (AGE) resulting in morbidity and mortality in young children (Estes 2013). Globally, RVA associated infections have been linked to an estimated 128, 500 deaths among children under 5 years of age in 2016 (Troeger et al. 2018). In high-income countries like the USA, the mortality associated with RVA AGE is low, but the morbidity is still significant. In the pre-vaccination era in the USA, ≥55,000 hospitalizations due to RVA AGE occurred annually among children ≤60 months of age, with twenty to sixty deaths annually (Cortese and Parashar 2009). This number has decreased by 40,000–50,000 since vaccine licensure (Tate et al. 2009).

The RVA virion consists of a tri-layered, non-enveloped icosahedron with an eleven segment genome of double-stranded (ds) RNA, encoding six structural (VP1–VP4, VP6, and VP7) and depending on the strain, five or six non-structural (NSP1-NSP5/NSP6) proteins (Estes and Greenberg 2013; Esmon and Gautam 2015). Prior to 2008, classification of RVA used a binomial typing system, GxP[x], based on antigenic or genetic characterization of the outer capsid proteins VP7 (G-type) and VP4 (P-type) (Estes and Greenberg 2013). The traditional binomial system has been replaced by an extended classification system that assigns genotypes to all eleven genes (Matthijnssens et al. 2008a) and defines the RVA genotype constellation as Gx-P[x]-I3-R3-C3-M3-A3-N3-T3-E3-H3, representing the genetic diversity of RVA. In recent years, whole-genome-based phylogenetic analyses using next-generation sequencing (NGS) data have been used to define the genotype constellations of RVA strains and most human RVA strains show high similarity in all segments to either the Wa-like genogroup 1 constellation (Gx-P[x]-I1-R1-C1-M1-A1-T1-E1-H1) or the DS-1-like genogroup 2 constellation (Gx-P[x]-I2-R2-C2-M2-A2-T2-E2-H2) (Matthijnssens et al. 2008a; McDonald et al. 2009a). In addition, a small group of human RVA strains belong to the AU-1-like genogroup 3 constellation (Gx-P[x]-I3-R3-C3-M3-A3-T3-E3-H3) (Matthijnssens et al. 2008b). Based on this classification system, 36 G, 51 P, 26 I, 22 R, 20 C, 20 M, 31 A, 22 N, 22 T, 27 E, and 22 H genotypes have been identified in human and animal hosts to date and classified based on differences in the nucleotide sequence identities of each encoding gene segment (Esmon et al. 2010; Abe et al. 2013; Matthijnssens et al. 2011; Papp et al. 2012; Estes and Greenberg 2013; Trojnar et al. 2013; Jere et al. 2014; Yinda et al. 2016; Esmon et al. 2018), https://regaskuleuven.be/cev/viralmetagenomics/virus-classification/newgenotypes; last accessed on 02 March 2021. Although a variety of G and P genotypes have been reported in humans, the most common VP7 (G) and VP4 (P) genotypes are G1, C2, G3, C4, G9, C12, P[4], P[6], and P[8] and have been associated with an estimated 80–90 per cent of RVA associated diarrhea infections in the post-vaccination era (Banyai et al. 2012; Estes and Greenberg 2013; Doro et al. 2014; Esmon and Gautam 2015).

The segmented RNA genome of RVA enables reassortment in vivo during co-infection with multiple RVAs, permitting the occurrence of intra- and inter-genogroup reassortment events (Estes 2013; Cowley et al. 2016; Esmon et al. 2017). This may lead to the emergence of distinct lineages within individual genotypes or reassortant viruses containing segments from different parent strains (Estes 2013; Ramig 1997). Lately, the detection of inter-genogroup reassortant strains possessing genetic constellations of G1/2/3/8–P[8]–I2–R2–C2–M2–A2–N2–T2–E2–H2 are on the increase among the human populations (Fujii et al. 2014; Yamamoto et al. 2014; Komoto et al. 2015; Arana et al. 2016; Cowley et al. 2016; Perkins et al. 2017; Jere et al. 2018).

Genetic analysis of RVA at the whole-genome level using NGS has provided a better understanding of the diversity of RVA strains post-vaccine licensure era, resulting from the accumulation of point mutations, gene rearrangement, reassortment as well as recombination (Estes 2013; Ramig 1997). Also, RVA whole-genome analysis simplifies accurate elucidations of the origin of a given strain and similarly assists in outlining its evolutionary patterns (Ghosh and Kobayashi 2011).

For over a decade, the New Vaccine Surveillance Network (NVSN) has been conducting active RVA AGE strain surveillance in the USA since 2006 (Payne et al. 2008). Several RVA surveillance studies conducted by NVSN have revealed a changing pattern of circulating RVA genotypes over time since the licensure of RotaTeq (G1, G2, G3, G4, P[8] pentavalent vaccine) in 2006 and Rotarix (G1P[8] monovalent vaccine) in 2008 (Payne et al. 2009; Bowen et al. 2016). In the USA, post-vaccine surveillance data from the 2007 through 2013 RVA seasons revealed dramatic changes in genotype predominance from G1P[8] to G3P[8] in 2009 and then from G3P[8] to G1P[8] in 2012 (Bowen et al. 2016). To better understand the origin and evolutionary patterns of circulating RVA strains and to monitor changes in gene sequences that may affect vaccine performance, continued surveillance at the whole-genome level is necessary. This report provides whole-genome ORF data and analysis of RVA strains from the NVSN for the 2014 through 2016 RVA seasons.

2. Materials and methods

2.1 Ethics statement

Institutional review board approvals were obtained from the Centers for Disease Control and Prevention (CDC) and from each NVSN surveillance site mentioned below. In this study, written informed consent was obtained from the parent or guardian of each child at the time of enrollment.

2.2 Sample selection for whole-genome analysis

The VP7 and VP4 genotype results (Esmon et al. 2020) were used to select a subset of stool specimens for Sanger and NGS. Stool samples included in this study were selected from all seven NVSN surveillance sites (Fig. 1) who participated in the
following RVA seasons, 2013–2014 (hereafter referred to as 2014), 2014–2015 (hereafter referred to as 2015), and 2015–2016 (hereafter referred to as 2016). In the Rotavirus Surveillance and Molecular Epidemiology Laboratory, we have found that samples with NSP3 qRT-PCR Ct values of ≤23 will yield high-quality NGS sequence data and the sequence quality decreases as viral load decreases (unpublished). Hence, using Sanger sequence data for VP7 and VP4 and the NSP3 qRT-PCR Ct values ≤23 criterion, 157 randomly selected, common and unusual samples from the 2014 to 2016 seasons, which genotyped as 52 G1P[8], 47 G12P[8], 18 G9P[8], 24 G2P[4], 5 G3P[6], 3 RV1 containing samples, 4 RV5 samples which genotyped as G6P[8] (n = 2), G6P[5] (n = 1), G1P[5] (n = 1), and 4 WT strains; G2P[8], G3P[8], G4P[8], and G9G12P[8], were included in our NGS analysis.

2.3 Viral dsRNA extraction for NGS
RVA dsRNA for whole-genome sequencing was extracted from each fecal sample and purified using the commercial guanidinium isothiocyanate reagent TRI-REAGENT-LS (Molecular Research Centre), according to the manufacturer’s protocol or using the MagNA Pure Compact RNA Isolation Kit on the automated MagNA Pure Compact Instrument (Roche Applied Science, Indianapolis, IN, USA) as described previously (Esona et al. 2013a).

2.4 cDNA library synthesis, amplification, and NGS
Sequencing templates and libraries were prepared using either a sequence-independent whole-genome RT-PCR amplification method (Potgieter et al. 2009) or NEBNext Ultra RNA Library Prep Kit for Illumina v1.2 and NEBNext Multiplex Oligos for Illumina (New England Biolabs, Ipswich, MA, USA), according to manufacturer’s protocol. cDNA libraries generated using NEBNext Ultra RNA Library Prep Kit were purified using Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA, USA) according to manufacturer’s specifications. NGS was carried out on an Illumina MiSeq sequencer (Miseq Sequencer, San Diego, CA, USA) using the MiSeq reagent kit v.2 with 500 cycles and the standard 250 bp paired-end reads method.

2.5 Whole-genome sequence analysis
Contigs were assembled from individual reads by de novo assembly and guided assembly with default parameters in CLC Genomics Workbench 7.0.4 or 11.0 software (http://www.clcbio.com/products/clc-genomics-workbench/). Genotypes were determined according to guidelines of the Rotavirus Classification Working Group (Matthijnssens et al. 2008b) using the online genotyping tool, RotaC (available at: http://rotac.regatools.be/; last accessed on 10 January 2018) (Maes et al. 2009), and the NCBI’s BLASTN program. Nucleotide (NT) sequences of whole genes of study strains were deposited in GenBank under accession numbers listed in Supplementary Table S1.

2.6 Phylogenetic, sequence, and structural analyses
Prior to phylogenetic and sequence analyses, all duplicate sequences for each genome segment were identified by ElimDupes (Abecasis, Vandamme, and Lemey 2006-2007) and removed. Unique whole ORF NT sequences were aligned and subjected to sequence comparisons as described previously (Ward et al. 2016; Esona et al. 2017, 2018). Briefly, multiple alignments were made using the MUSCLE algorithm implemented in MEGA 6 software (Tamura et al. 2013). Once aligned, the DNA Model Test program implemented in MEGA version 6 was used to identify optimal evolutionary models that best fit sequence datasets using the Corrected Akaike Information Criterion (AICc). Maximum-likelihood trees were constructed using MEGA 6 with 1000 bootstrap replicates to estimate branch support. NT and deduced AA sequence identities among strains were calculated for each gene based on distance matrices prepared using the p-distance algorithm in MEGA 6 software (Tamura et al. 2013).

Figure 1. Location of NVSN site in the USA from which rotavirus positive specimens were obtained.
For allele assignments within each gene, multiple alignments of strains exhibiting monophyletic clustering in the phylogenetic analyses were prepared using MEGA6. NT sequence identities within each monophyletic cluster and between clusters were calculated using the p-distance algorithm as described above. Monophyletic groupings with appropriate bootstrap values were collapsed using MEGA6. Maximum likelihood trees indicating the relationship of study strains to globally circulating strains were constructed using sequences available in the GenBank database. Using the derived within- and between-cluster distances, sub-genotype alleles were assigned to sequence groups based on (1) phylogenetic clusters with bootstrap support of ≥70% and (2) NT identity of ≥95.8% within a cluster.

VF7 and VP4 structural modeling was performed with AA sequences for each study strain and reference strains using the Swiss-Model protein structure homology-modeling server (available at: https://swissmodel.expasy.org; last accessed on 22 August 2018) (Waterhouse et al. 2018). The designation and coloring of the amino acid residues on the protein structures were done using UCSF Chimera (Petterson et al. 2004). AA residue comparisons were performed using BLOSUM62 scoring in the NCBI Amino Acid Explorer tool (available at: https://www.ncbi.nlm.nih.gov/Class/Structure/aa/aa_explorer.cgi; last accessed on 20 October 2018) and was also used to classify AA changes as radical, conservative or neutral in nature.

3. Results

3.1 Whole-genome sequencing of RVA

The whole genome (ORF sequences) of 157 RVA strains 52 G1P[8], 47 G12P[8], 18 G9P[8], 24 G2P[4], 5 G3P[6], 3 RV1 strains, 4 RV5 strains (2 G6P[8], 1 G6P[5] and 1 G1P[5]), 1 G2P[6], 1 G3P[8], 1 G4P[8], and 1 G9P[12P] from the 2014-2016 seasons were determined by NGS. One hundred eighteen of these strains possessed a Wa-like backbone with consensus genotype constellation of G1/G3/C4/G9/G12-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1 and 24 G2P[4] and 5 G3P[6] strains possessed a DS-1-like genetic backbone with a consensus constellation of G2/G3-P[4]/P[6]-I2-R2-C2-M2-A2-N2-T2-E2-H2.

Among three intergenogroup reassortant strains, one G2P[8] strain possessed a DS-1-like genetic backbone with a consensus genotype constellation of G2-F[8]-I2-R2-C2-M2-A2-N2-T2-E2-H2, and two G1P[8] strains on a DS-1-like backbone exhibited a genetic constellation of G1-F[8]-I2-R2-C2-M2-A2-N2-T2-E2-H2. Two Rotarix vaccine (RV1) reassortant strains were detected: one apparently produced by reassortment events between a RV1 (VP7, VP2, NSP1, and NSP3 genes), a RotaTeq (RV5) strain donating the VP4 gene and a wildtype (WT) genotype 1 Wa-like strain donating the VP1, VP3, VP6, NSP2, NSP4, and NSP5 genes, hence generating a strain with a genetic constellation of G1(RV1)-P[8] (RV5)-I1(WT)-R1(WT)-C1(RV1)-M1(WT)-A1(RV1)-N1(WT)-T1(RV1)-E1(WT)-H1(WT) and the other generated by reassortment between a RV1 (VP7, VP4 genes) and a WT DS-1-like strain (VP1, VP2, VP3, VP6, NSP1, NSP2, NSP3, NSP4, and NSP5 genes) that displayed a genetic constellation of G1(RV1)-P[8] (RV1)-I2(WT)-R2(WT)-C2(WT)-M2(WT)-A2(WT)-N2(WT)-T2(WT)-E2(WT)-H2(WT). The third RV1 strain was determined to be a pure vaccine strain. Two RV5-derived reassortant strains were also identified: a RV5 vdG6P[8] strain with genetic constellation of G6-P[5]-I2-R2-C2-M2-A3-N2-T6-E2-H3, generated by an apparent reassortment event between vaccine component strains, RVA/Vaccine/USA/RotaTeq-WI79-4/1992/G6P1A[8] and either RVA/Vaccine/USA/RotaTeq-WI78-8/1992/G3P7[5] or RVA/Vaccine/USA/RotaTeq-BrB8-9/1996/G4P7[5]. The second was RV5 vdG6P[8], generated through a reassortment event between a WT G6 strain and a RV5 strain RVA/Vaccine/USA/RotaTeq-WI79-4/1992/G6P1A[8]. The remaining two RV5 G6P[8] and G1P[5] vaccine strains that were sequenced were pure vaccine component strains.

3.2 Sub-genotype allele designation for the study strain genes

To increase the robustness of these analysis, we added 104 previously published sequences from US samples collected during the 2005–2011 RVA seasons (McDonald et al. 2012; Dennis et al. 2014; Roy et al. 2014) including both genogroup 1 strains (G1P[8], G3P[8], G12P[8]) and genogroup 2 strains (G2P[4], G12P[6]) with previously designated sub-genotype alleles (McDonald et al. 2012; Dennis et al. 2014; Roy et al. 2014). In addition, reference RVA strains from around the world possessing previously assigned genotypes/lineages/sub-lineages were included. Using the ORF sequence alignments and the optimal models based upon AICc values, GTR + G + I (VP1, VP2, VP3, VP4, VP6, NSP1, NSP2, and NSP3), and GTR + G (NSP4), HKY + G (NSP5), and HKY + G + I (VP7), we performed maximum likelihood analyses and identified candidate alleles based upon genotype clades with bootstrap support ≥70%. We found that the NT identity within each candidate allele for each of the eleven genes was ≥95.8% and between-allele distances were in the range of 89–93.2% per cent. Using these criteria, we defined a distinct sub-genotypic allele cluster as a monophyletic group with ≥70% bootstrap support and ≥95.8% within-group sequence identity and applied these criteria to the phylogenetic estimates. The resulting phylogenetic trees and allele color-coding for all eleven genes are shown in Fig. 2A–K.

Analysis of the study strains, reference strains (some with known alleles), and vaccine strains identified 2–11 sub-genotypic alleles. Most study strains clustered into previously defined alleles, but we defined 2–7 new alleles per gene (Fig. 2A–K). The individual genotypes C1, G2, C3, G9, and G12 form the basis of VP7 gene clusters (Fig. 2A). The G1 strains were further divided into six distinct alleles denoted by a spectrum of red colors. A majority (seventy-four per cent) of the G1 study strains were grouped together with published allele A strains (McDonald et al. 2012; Roy et al. 2014) which also contained previously reported G1P[8] reassortant strains detected in Malawi and Japan in 2013 (Yamamoto et al. 2014; Jere et al. 2018). Eleven (twenty-two per cent) of the remaining G1 study strains clustered in a newly defined allele B, together with RV1 strains. Previously defined alleles C and D, which contained only US G1 strains from the 2005–2008 seasons, were confirmed (McDonald et al. 2012; Roy et al. 2014). Two new alleles, E and F, were defined in this study. Allele E consisted of G1P[8] strains from Italy, Japan, and Korea, while F consisted of G1 reference strain Wa, contemporary US G1P[8] strains, a RV5 vdG1P[8] reassortant strain, and the G1 component of the RV5.

Similarly, the G2 strains were divided into six alleles denoted by a spectrum of sky-blue colors (Fig. 2A) with five new alleles, A, B, D, E, and F. Allele A consisted of the G2 reference strain DS-1 and other previously defined G2 lineage I strains (Mascarenhas et al. 2010). Because of a lack of bootstrap support, previously defined alleles A and B (Dennis et al. 2014) were combined to form a new allele designated as B. The G2P[8] reassortant strains clustered together with previously defined allele C strains (Dennis et al. 2014). The newly designated allele D contained all G2P[4] study strains and previously reported US strains detected in the 2013–2015 seasons (Dennis et al.
Figure 2. Maximum likelihood trees with bootstrap values showing branch support for the RVA VP7 (A), VP4 (B), VP6 (C), VP1 (D), VP2 (E), VP3 (F), NSP1 (G), NSP2 (H), NSP3 (I), NSP4 (J), and NSP5 (K) genes. Monophyletic groupings with bootstrap support of ≥70% and nucleotide similarity in the range of 99–100 per cent were collapsed and assigned an allele. VP7 genotypes G1, G2, G3, G9, and G12, alleles were colored with a spectrum of red, sky blue, blue, brown, and pink colors, respectively. VP4 genotypes P[4], P[6], and P[8], alleles were colored with a spectrum of aqua blue, chartreuse, and purple colors, respectively. For the internal gene segments, Wa-like genotypes were colored with a spectrum of red colors ranging from dark red (allele A) to light red (allele I). DS-1-like genotypes were colored with a spectrum of green colors ranging from dark green (allele A) to light green (allele I). Strains that formed separate clusters but without bootstrap support were designated and collapsed black and/or collapsed. Singlet or doublet strains were also colored black. RV1 and RV5 components were included in the analysis. Alleles, which consisted of the study strains, are indicated on the table to the right of each tree. Trees were drawn to scale as indicated by the scale bar.
Figure 2.
Figure 2.
Figure 2.
Figure 2.
along with G2P[4] strains from Vietnam (Trinh et al. 2010) and Thailand (Khamrin et al. 2007). In addition, two new G2 alleles, E and F, were defined. Allele E consisted of the G2 component of the RV5 and previously defined G2 lineage III strains, while F consisted of only porcine G2 strains belonging to lineage IV (Esteban et al. 2010).

G3 strains formed four distinct clusters of alleles denoted by a spectrum of blue colors (Fig. 2A). Previously defined allele A which contained only RVA strains from the 2007–2012 seasons was unchanged (McDonald et al. 2009b, 2012; Roy et al. 2014). Three new alleles, B, C, and D, were defined. Allele B consisted of mostly G3P[6] strains from sub-Saharan Africa, Europe, Asia, and those from the current study, while C comprised a group with mostly US G3P[6] strains detected in 1976 and two G3P[6] study strains. Allele D consisted of only equine-like G3P[8] reassortant strains from the USA, Japan, and Germany.

G9 strains were divided into four newly defined alleles, A, B, C, and D, indicated by a spectrum of brown colors (Fig. 2A). Allele A forms a cluster consisting of previously defined G9 lineage I strains (W61 and AU32) (Esona et al. 2013b). Allele B exclusively consisted of G9P[8] study strains detected in 2009 and strains belonging to the previously defined G9 lineage III (Esona et al. 2013b). Allele D contained lineage V US G9 strains OM46 and OM67 (Esona et al. 2013b).

Analysis of G12 strains grouped them into four distinctly assigned alleles, A, B, C, and D, by a spectrum of pink colors (Fig. 2A). Allele A consisted mostly of study strains detected in the 2015–2016 RVA seasons and G12P[8] strains detected in the 2012–2013 seasons in Nashville. Allele B included only G12P[8] strains detected in 2013 in Nashville, while C consisted of two G12P[8] strains detected in 2005 in Nashville. Allele D comprised only animal and human RVA reference strains.

For VP4, individual genotypes P[4], P[6], and P[8] (Fig. 2B) were assigned to sub-genotype allele clusters. The strains with VP4 P[8] genotypes were divided into five distinct alleles, A, B, C, D, and E, denoted by a spectrum of purple colors (Fig. 2B). Alleles A, C, and E are new designations from this study. Because of a lack of bootstrap support, previously defined alleles A and C (McDonald et al. 2009b, 2012; Roy et al. 2014) were combined to form a new allele A that consists of G1P[8], G3P[8], and G12P[8] strains along with recently reported reassortant G1P[8] (Fujii et al. 2014; Yamamoto et al. 2014; Komoto et al. 2015) and equine-like G3P[8] (Cowley et al. 2016; Perkins et al. 2017) strains. Allele B remained unchanged and consisted of previously defined allele B strains (McDonald et al. 2009b, 2012; Roy et al. 2014). Allele C formed a cluster with the RV5 P[8] component and RV5 vdG6P[8], and vdG1P[8] study strains. Allele D also remain unchanged and contained only previously reported strains, while E consisted of OP354-like P[8] strains from around the world (Zeller et al. 2015).

The strains with VP4 P[4] genotypes were divided into four distinct alleles, A, B, C, and D, denoted by a spectrum of aqua blue colors (Fig. 2B). Alleles A, B, and C were previously defined and did not change (Dennis et al. 2014). Allele A contained five study strains and contemporary US G2P[4] from the 2011 RVA season (Dennis et al. 2014), while C contained two study strains and US G2P[4] strains from the 2011–2013 seasons (Dennis et al. 2014). New Allele D consisted of reference strain DS-1 and strains belonging to the previously defined P[4] lineage I (Mascarenhas et al. 2010).

For P[6] strains, no previous alleles have been defined, so in this study, we designated four new alleles, A, B, C, and D, denoted by a spectrum of chartreuse colors (Fig. 2B). Allele A included G3P[6] strains from the current study; G1P[6], G2P[6], and G6P[6] strains from sub-Saharan Africa; and strains belonging to
previously defined P[6] lineage I (Stupka et al. 2009). Allele B consisted of porcine G9P[6] strains from Japan, while C contained human G4P[6] strains detected in Hungary in 1999–2000. Finally, allele D included human and animal P[6] strains detected in Europe in 2000–2004.

3.3 Comparative analysis of antigenic epitopes in the VP7 protein of NVSN RVA strains and vaccine strains

The VP7 gene contains three structurally defined antigenic epitopes, the 7-1a, 7-1b, and 7-2 regions (Aoki et al. 2009), which comprise 29 AA residues (Fig. 3A–D). Using the AA alignments for the VP7 proteins, we identified differences in these antigenic epitopes between the WT RVA strains and the cognate genes of the RV1 and RV5 strains (Fig. 3A–D, Table 1). Five differences (W98, Q104, K291, Q201, G264) were conserved in all study strains (Fig. 3A–C). Comparison of the allele A G1 study strains to vaccine strains showed 4–5 differences (Fig. 3A, Table 1). All were conservative in nature except M217T located in a b-strand on the protein surface (Fig. 3D-II, Table 1). Allele B G1 study strains showed 0–2 conservative differences with vaccine G1 proteins (Table 1). Compared to both vaccines, allele C strains showed two radical differences (E97G, D97G) located on the surface of the protein structure within a helical region of neutralization epitope 7-1a site (Table 1, not visible in Fig. 3D). Allele D and F strains exhibited 0–3 conservative differences compared to the vaccine strains (Table 1).

A comparison of the VP7 antigenic epitopes of study G2 strains to the G2 component of RV5 and the G1 protein of RV1 (Fig. 3A, Table 1) showed 3–11 AA differences located on the surface of the protein structure compared with the G3 component of RV5 (data not shown). A comparison of the RV5 G2 protein to G2 allele A-E strains showed 0–4 conservative/neutral AA differences (Fig. 3A and D-III, Table 1). Allele F, which comprised only animal G2 strains of porcine origin, exhibited five AA differences compared to RV5 G2 including the radical M129T substitution seen in G1 strains (Table 1).

Analysis of the G3 alleles A, B, C, and D showed 3–6 AA differences located on the surface of the protein structure compared with the G3 component of RV5 (Fig. 3A and D-IV, Table 1). In addition to four conservative/neutral AA differences maintained in allele A, B, and C strains, allele B study strains RVA/Human-wt/USA/300052558/2015/G3P[6], RVA/Human-wt/USA/300054444/2015/G3P[6], and RVA/Human-wt/USA/2014738139/2014/G3P[6] displayed an additional neutral change A146V and a radical change A221D located on the protein surface, the latter a non-polar to negatively charged AA side chain (Fig. 3A and D-IV, Table 1). All G3 alleles A, B, and C study strains exhibited a K238N substitution, which creates a potential N-linked glycosylation site on the VP7 protein and may have some effect on the antigenicity of 7-1a antigenic epitopes (Zeller et al. 2012).

Analysis of the single G4 study strain RVA/Human-wt/USA/300449267/2015/G4P[8] showed three conservative/neutral differences distributed among two of the antigenic epitopes located on the surface of the protein structure, compared with the G4 component of RV5 (data not shown) (Fig. 3A).

Since RV5 also expresses the G6 VP7 protein, we compared it to the G6 of RV5 dG6P[8] reassortant study strain RVA/Human-wt/USA/2014741597/2014/G6P[8]. The VP7 of strain RVA/Human-wt/USA/2014741597/2014/G6P[8] contained eight AA differences...
| Gene | Genotype | Allele* | Reference | No of AA changes | Radical AA changes\(^b\) | Location | Epitope site |
|------|----------|---------|-----------|-----------------|--------------------------|----------|--------------|
| VP7  | G1       | A       | RV1       | 4               | M217T                    | Surface – strand | 07-Feb       |
| VP7  | G1       | A       | RV5       | 5               | M217T                    | Surface – strand | 07-Feb       |
| VP7  | G1       | B (New) | RV1       | 0               |                          |           |              |
| VP7  | G1       | B       | RV5       | 2               |                          |           |              |
| VP7  | G1       | C       | RV1       | 2               | E97G, D97G               | Surface helix   | 7-1a         |
| VP7  | G1       | C       | RV5       | 2               | E97G, D97G               | Surface helix   | 7-1a         |
| VP7  | G1       | D       | RV1       | 1               |                          |           |              |
| VP7  | G1       | D       | RV5       | 1               |                          |           |              |
| VP7  | G1       | E (New)| RV1       | 0               |                          |           |              |
| VP7  | G1       | E       | RV5       | 0               |                          |           |              |
| VP7  | G1       | F (New)| RV1       | 3               |                          |           |              |
| VP7  | G1       | F       | RV5       | 1               |                          |           |              |
| VP7  | G2       | A (New)| RV5       | 3               |                          |           |              |
| VP7  | G2       | B (New)| RV5       | 4               |                          |           |              |
| VP7  | G2       | C       | RV5       | 4               |                          |           |              |
| VP7  | G2       | D (New)| RV5       | 3               |                          |           |              |
| VP7  | G2       | E (New)| RV5       | 0               |                          |           |              |
| VP7  | G2       | F (New)| RV5       | 11              | M217T                    | Surface – strand | 07-Feb       |
| VP7  | G3       | A       | RV5       | 3               |                          |           |              |
| VP7  | G3       | B (New)| RV5       | 05-Jun          | A221D                    | Protein surface | 07-Feb       |
| VP7  | G3       | C (New)| RV5       | 4               |                          |           |              |
| VP7  | G3       | D (New)| RV5       | 4               | K238D, D242A             | Surface left loop | 7-1b         |
| VP7  | G9       | A (New)| RVA/Human-tc/USA/WI61/1983/G9P[8] | 0 | | | |
| VP7  | G9       | B (New)| RVA/Human-tc/USA/WI61/1983/G9P[8] | 03-Apr | | | |
| VP7  | G9       | C (New)| RVA/Human-tc/USA/WI61/1983/G9P[8] | 01-Mar | | | |
| VP7  | G9       | D (New)| RVA/Human-tc/USA/WI61/1983/G9P[8] | 2 | | | |
| VP7  | G12      | A (New)| RVA/Human-tc/BEL/B4633/2003/G12P[8] | 0 | | | |
| VP7  | G12      | B (New)| RVA/Human-tc/BEL/B4633/2003/G12P[8] | 0 | | | |
| VP7  | G12      | C (New)| RVA/Human-tc/BEL/B4633/2003/G12P[8] | 1 | | | |
| VP7  | G12      | D (New)| RVA/Human-tc/BEL/B4633/2003/G12P[8] | 5 | | | |
| VP4  | P[8]     | A (New)| RV1       | 6               | S131R                    | Surface – strand | 08-Mar       |
| VP4  | P[8]     | A (New)| RV5       | 6               | D195G                    | Surface helix   | 08-Mar       |
| VP4  | P[8]     | B       | RV1       | 6               | S131R                    | Surface – strand | 08-Mar       |
| VP4  | P[8]     | B       | RV5       | 6               | D195G                    | Surface helix   | 08-Mar       |
| VP4  | P[8]     | C (New)| RV1       | 4               | S131R                    | Protein surface | 08-Mar       |
| VP4  | P[8]     | C (New)| RV5       | 1               |                          |           |              |
| VP4  | P[8]     | D       | RV1       | 1               |                          |           |              |
| VP4  | P[8]     | D       | RV5       | 5               | R131S                    | Protein surface | 08-Mar       |
| VP4  | P[8]     | E (New)| RV1       | 5               |                          |           |              |
| VP4  | P[8]     | E (New)| RV5       | 6               |                          |           |              |
| VP4  | P[4]     | A       | RVA/Human-tc/USA/DS-1/1976/G2P[4] | 1 | | | |
| VP4  | P[4]     | B       |           | 1               |                          |           |              |
distributed among all three antigenic epitopes (Fig. 3A). AA substitutions S146H, P213V, and T217I located on the protein surface represent radical changes.

Comparison of the VP7 epitopes of the G9 study strains to reference G9 strain RVA/Human-tc/USA/WI61/1983/G9P[8], which was isolated in 1983 in the USA, revealed 0–4 conservative/neutral differences distributed in two of the three antigenic epitopes (Fig. 3B, Table 1). Antigenic epitope 7-2 was highly conserved amongst the older and contemporary G9 strains as well as G9 study strains (Fig. 3B).

Comparison of the VP7 epitopes of the G12 alleles A, B, C, and D of the study strains to a G12 reference strain from 2003, RVA/Human-wt/BEL/B4633/2003/G12P[8], and contemporary human G12 genotypes from GenBank revealed an absolute conservation among study strains belonging to alleles A and B (Fig. 3C) and C displayed a single conservative AA difference at N100S (data not shown). Relative to reference strain RVA/Human-wt/BEL/B4633/2003/G12P[8], allele D, which consisted of porcine strain RU172 and human strain L26, detected in 2002 and 1987, respectively, exhibited 5 conservative/neutral AA changes (data not shown).

### 3.4 Comparative analysis of antigenic epitopes in the VP4 protein of NVSN RVA strain and vaccine strains

The VP4 protein, which in this study was represented by P[4], P[6], and P[8] strains, was divided into the established VP8* and VP5* regions for comparison (Estes 2013). The VP8* region contains four (8-1 to 8-4) antigenic epitopes composed of 25 AAs, while the VP5* region has five epitopes (5-1 to 5-5) with 12 AAs (Dormitzer et al. 2002) (Fig. 4A–C). Out of these thirty-seven AA residues that span these epitopes, twenty-nine are known neutralization escape mutation sites (Kobayashi, Taniguchi, and Urasawa 1990; Zhou et al. 1994; Dormitzer et al. 2004; Monnier et al. 2006; McKinney et al. 2007).

A comparison of the P[8] alleles A, B, and C of the study strains to the P[8] components of RV1 and RV5 showed that twenty-four of thirty-seven AA residues distributed throughout the antigenic epitopes of the VP4 gene were highly conserved (Fig. 4A). Allele A consisted of mostly G1P[8], G2P[8], G3P[8], G9P[8], and G12P[8] study strains and other US strains from the 2011–2013 seasons and the differences between these strains and both vaccines were mostly contained in VP8* epitopes 8-1 and 8-3 (Fig. 4A). Relative to the RV1 P[8], the P[8] component of the Allele A study strains exhibited 6 AA differences including a radical S131R substitution, located in a b-strand on the surface of the protein in neutralization epitope site 8-3 (Fig. 4A and D, Table 1). The conservative change at N113D was only present among G9P[8] strains detected in the 2014 and 2016 RVA seasons (Fig. 4A). Compared to RV5, a radical difference, D195G, was located in a helix on the protein surface (Fig. 4A, D-I and D-II, Table 1). Furthermore, the P[8] component of G12P[8] strains in allele A showed 2–5 AA differences when compared with vaccine strain P[8] proteins (Fig. 4A). Allele C which consisted of G1P[8] and G6P[8] study strains, exhibited 4 AA differences when compared to RV1 P[8] (Fig. 4A) including a radical change at position S131R. Comparing vaccine strains to strains in allele D, 1–5 differences were seen, including a radical change at position R131S located on the VP8* region of the protein surface (Fig. 4A, D-I and D-II, Table 1). Allele E contained mostly non-study OP354-like P[8] strains which exhibited 5–6 difference with vaccine P[8] proteins (Fig. 4A, Table 1).

Comparison of epitopes of P[4] alleles A, B, C, and D of the study strains to reference and contemporary P[4] strains from GenBank showed 0–2 AA differences (Fig. 4B, Table 1). Allele E contained mostly non-study OP354-like P[8] strains which exhibited 5–6 difference with vaccine P[8] proteins (Fig. 4B, Table 1).
3.5 Sub-genotype allele designation for the VP6, VP1-VP3, VP6, and NSP1-NSP5 genes of the study strains

Phylogenetic analysis of the VP6 gene from the study and GenBank strains grouped them into Wa-like (genotype I1) and DS-1-like (genotype I2) strains (Fig. 2C). The Wa-like VP6 strains were further divided into ten alleles designated A, B, C, D, E, F, G, H, I, and J. Alleles A–E were defined previously (McDonald et al. 2009b, 2012; Roy et al. 2014) and comprised G1P[8], G3P[8], and G12P[8] study strains and contemporary G1P[8], G3P[8], and G12P[8] strains detected in the USA, Belgium and Bangladesh. Of the four new alleles (F, G, H, I) and I comprised mostly human and animal RVA strains, and G consisted of US G1P[8] strains detected in the 2007 and 2013 RVA seasons. Allele H consisted of a single RV1 strain, 3000378227, detected during the 2016 season, and G1P[8] reference strain Wa. For the DS-1-like (I2) strains, six distinct alleles were identified: A, B, C, D, E, and F (Fig. 2C). Allele A was previously defined and comprised G2P[4] US strains detected in the 2011 season (Dennis et al. 2014) and five new alleles, B–F, were defined. Due to a lack of bootstrap support, previously defined alleles B and C (Dennis et al. 2014) were combined to form a new allele B, which consisted of US G2P[4] strains detected in 2011, DS-1-like G1P[8] strains from Japan and Vietnam, DS-1-like G8P[8] strains from the USA, Japan, and Germany, and the reassortant G2P[8] strain from the current study. Allele C consisted of G8P[8], G8P[6], G9P[6], and G12P[6] strains from Africa and Bangladesh, while D was composed of study strains DS-1-like study strains RVA/Human-wt/USA/2014737554/2014/G1P[8] and RVA/Human-wt/USA/3000347195/2015/G2P[4]. Allele E consisted of G2P[4] reference strains DS-1 and TB-Chen and G12P[4] strain L26. Allele F contained two study strains, RVA/Human-wt/USA/3000558291/2016/G2P[4] and RV5 vdG6P[8] reassortant strain RVA/Human-wt/USA/2014741579/2016/G6P[8] and the RV5 strains.

Phylogenetic analysis of the VP1 Wa-like (R1 genotype) strains divided them into three distinct alleles, A, B, and C (Fig. 2D) and all were newly defined for this genotype. Allele A consisted exclusively of G1P[8] study strains from the 2014 RVA season, while C included only contemporary US G12P[8] strains detected during the 2012–2013 seasons. Because of a lack of bootstrap support, previously defined alleles A, B, C, and D strains (McDonald et al. 2009b, 2012; Roy et al. 2014) were combined to form a new allele B which consisted of G1P[8], G3P[8], G9P[8], and G12P[8] strains. For the DS-1-like (R2 genotype) strains, eleven alleles were identified: A, B, C, D, E, G, H, I, J, and K (Fig. 2D), and eight (D, E, F, G, H, I, J, and K) were newly defined. Alleles A and B contained only US G2P[4] strains detected in the 2011 season, while C consisted of two G2P[4] study strains, RVA/Human-wt/USA/2014737624/2014/G2P[4] and RVA/Human-wt/USA/2014737459/2014/G2P[4] detected in 2014 and reassortant G2P[8] study strain RVA/Human-wt/USA/3000558215/2016/G2P[8] detected in 2016. Also included in allele C were US G2P[4] strains detected in the 2011–2013 seasons. Allele D consisted of DS-1-like G1P[8] and G3P[8] strains, while E included the reassortant G1P[8] study strains with a DS-1.
genetic backbone. Allele F consisted of G2P[4] study strains exclusively from the 2014–2015 seasons and G included G2P[4] strain TB-Chen and reference strain DS-1. Allele H consisted of only G3P[6] study strains, while I included a G3P[6] study strain and a single G6P[6] strain from Belgium. RV5-like and RV5 strains were grouped in allele K.

The VP2 gene segment of Wa-like (C1 genotype) strains were divided into ten distinct alleles designated A, B, C, D, E, F, G, H, I, and J (Fig. 2E) with six new alleles (D, E, F, G, H, I, and J) defined in this study. Previously defined alleles D and E (McDonald et al. 2012; Roy et al. 2014) were combined to form a new allele D which consisted of mostly US G1P[8] strains detected during the 2005–2008 RVA seasons. G3P[8] and G9P[8] strains that were defined as allele C previously (McDonald et al. 2009b, 2012; Roy et al. 2014) were divided into two groups: one group with G9P[8] strains from the current study and designated as allele E and a second group consisting of G1P[8], G3P[8], and G9P[8] from previous studies (McDonald et al. 2009b, 2012; Roy et al. 2014) and G1P[8] and G9P[8] study strains detected in the 2014 and 2015 RVA seasons that remained in allele C. Alleles F, G, and H consisted mostly of G1P[2] study strains detected during 2014–2016 seasons and US G12P[8] strains detected during the 2012–2013 seasons (Ogden et al. 2018). Allele I consisted of mostly G1P[8] study strains detected during the 2014 season, while J included the RV1 vaccine strain and G1P[8] reference strain Wa. Alleles A and B, defined previously, did not change. For the DS-1-like VP2 gene segment (C2 genotype), the strains were grouped into seven distinct alleles, designated A, B, C, D, E, F, and G, and six of these (B, C, D, E, F, and G) were newly defined (Fig. 2D). G2P[4] strains that were previously grouped as alleles B and C (Dennis et al. 2014) were combined to form a new allele B. Allele C consisted of DS-1-like G3P[8] strains from Germany, Japan and the USA, while D consisted of exclusively G2P[4] study strains detected in Cincinnati during the 2014–2015 RVA seasons. Allele E consisted entirely of G3P[6] study strains. Allele F included the two G1P[8] study strains with a DS-1-like backbone, and G consisted of RV5-like and RV5 strains.

The VP3 genes of Wa-like (M1 genotype) strains were divided into nine alleles, designated A, B, C, D, E, F, G, H, and I (Fig. 2F). Alleles A, B, C, D, E, and F were previously defined and did not change (McDonald et al. 2009b, 2012; Roy et al. 2014) but three new alleles, G, H, and I, were defined in this study. Allele G consisted of mostly contemporary G12P[8] strains detected in Nashville in the 2013 RVA season, while H consisted of entirely US G9P[8] strains detected in 2009. Allele I consisted of exclusively G12P[8] strains from the 2014 (1 strain), 2015 (14 strains), and 2016 (2 strains) seasons and 40 G12P[8] strains detected in the 2012–2013 RVA season in Nashville (Ogden et al. 2018). For the DS-1-like VP3 gene segment (M2 genotype), the strains were grouped into seven distinct alleles, designated A, B, C, D, E, F, and G (Fig. 2F) and four of these (D, E, F, and G) were newly defined. Allele D consisted of DS-1-like G1P[8], G3P[8], and G8P[8] strains from Germany, Japan and the USA, while E consisted exclusively of G3P[6] study strains. Allele F consisted entirely of G2P[4] study strains detected during the 2014 and 2015 RVA seasons, while G was composed of RV5 strains, two G6P[8] RV5 derived strains and a single G3P[6] study strain. Among alleles which have been defined previously, A, B, and C (Dennis et al. 2014), only C contained study strains: three G2P[4] strains detected in the 2014 and 2015 RVA seasons, the reassortant G2P[8] study strain detected in 2016, and G2P[4] strains detected in the 2011 and 2013 RVA seasons.

The NSP1 strains belonging to Wa-like genotype A1 were divided into eight distinct alleles, designated A, B, C, D, E, F, G, and H and five alleles (D, E, F, G, and H) were newly defined (Fig. 2G). Newly defined alleles D and E consisted exclusively of G1P[8] study strains detected during the 2014 RVA season. Allele F consisted of mostly G12P[8] study strains from the 2014 (1 strain), 2015 (13 strains), and 2016 (2 strains) seasons and G12P[8] strains detected in the USA during the 2013 season (Ogden et al. 2018). Allele G was entirely composed of US G12P[8] strains detected in the 2013 season. Allele H consisted of G1P[6], G4P[6], and G1P[8] reference strains. Because of the lack of bootstrap support, a large number of G1P[8], G9P[8], and G12P[8] strains were designated as orphan strains. Analysis of the NSP1 gene of DS-1-like (A2) strains grouped them into six distinct alleles, designated A, B, C, D, E, and F and four of these (C–F) were newly defined (Fig. 2G). Previously defined allele A comprised G2P[4] strains from the 2011–2013 RVA seasons, seven G2P[4] study strains from the 2014–2015 seasons, and a single G3P[6] study strain detected in the 2016 season, while B consisted of exclusively G2P[4] strains from the 2011 season (Dennis et al. 2014). Allele C consisted entirely of G3P[6] study strains, while D contained reassortant G1P[8] and G3P[8] strains with a DS-1-like genetic backbone. Allele E included reassortant G1P[8] study strains with a DS-1-like genetic backbone, RVA/Human-wt/USA/2014739567/2014/G1P[8], RVA/Human-wt/USA/2014737554/2014/G1P[8], and RVA/Human-wt/USA/2014746684/2014/G1P[8] and G2P[4] strains from the current study. Allele F included two reassortant G1P[8] strains with a DS-1-like backbone from Malawi (Jere et al. 2018).

The NSP2 gene of Wa-like (N1) strains was divided into ten distinct alleles, designated A, B, C, D, E, F, G, H, I, and J (Fig. 2H), with seven new alleles (A, D, F, H, I, J) defined. Previously defined allele A was subdivided into four new alleles: A, D, E, and F. Allele A consisted of previously described US G1P[8], G9P[8], and G12P[8] strains from the 2005–2008 seasons (Roy et al. 2014). Previously defined alleles B and C (McDonald et al. 2012; Roy et al. 2014), which consisted entirely of G1P[8] strains, were confirmed. Allele D consisted of US G1P[8] strains detected during the 2005 season (McDonald et al. 2012). Allele E included only US G1P[8] detected in the 2005 and 2006 seasons (Roy et al. 2014) and G1P[8] strains detected in the 2013 season (Ogden et al. 2018), and F consisted of US G1P[8] (detected in the 2007–2009 seasons) and G12P[8] study strains, as well as some contemporary G9P[8] (detected during the 2009 season) and G12P[8] (detected in the 2012 and 2013 seasons) strains (Roy et al. 2014; Ogden et al. 2018). All G1P[8] study strains detected in the USA during the 2014 season clustered in a new allele G, and G1P[8] strains detected in the 2007–2008 seasons grouped in I. Allele H included a G1P[8], G9P[8], G12P[8] study strains and the RV1 strain, while consisted of reference strains ST-3 and Wa. Analysis of the NSP2 gene of DS-1-like (N2) strains grouped them into eight distinct alleles, designated A, B, C, D, E, F, G, and H; six of these (C–H) were newly defined (Fig. 2H). Allele D consisted of reassortant G1P[8] and G3P[8] strains with a DS-1-like genetic backbone. Allele E included G3P[6] strains from the current study and F consisted of entirely G8 strains from Africa. Allele G included reference G2P[4] strains DS-1 and TB-Chen, while H contained RV5 vdG6P[8] strain RVA/Human-wt/USA/2014741597/2014/G6P[8] detected in the 2014 season and RV5 strains.

Phylogenetic analysis of the NSP3 gene from study and GenBank strains grouped them into Wa-like (genotype T1) and DS-1-like (genotype T2) strains (Fig. 2I). The Wa-like NSP3 genes were further divided into six alleles designated A, B, C, D, E, and F. Three new alleles D, E, and F were defined for T1 strains. The majority of study strains included in this analysis clustered in previously defined alleles A, B, and C (McDonald et al. 2009b, 2012; Roy et al. 2014). Ten G1P[8] strains that were previously assigned to allele A were placed in new allele D. New allele E consisted of two US G1P[8] strains detected in the 2007 and 2013
seasons. Allele F included a single RV1 study strain, RVA/Human-wt/USA/2014741572/2014/G1P[8], and the RV1 strain. For the DS-1-like (T2) strains, eight alleles designated A, B, C, D, E, F, G, and H were identified with six new alleles (C, D, E, F, G, H; Fig. 2I). Alleles A and B consisted of previously described US G2P[4] strains detected in the 2011 season (Dennis et al. 2014). Allele C consisted of reassortant G1P[8] and G3P[8] strains with a DS-1 genetic backbone and D consisted exclusively of three G3P[6] strains detected in the current study. Allele E consisted of G2P[4] study strains and as well as US G2P[4] strains detected during the 2013 season, while F included G12P[6] and G2P[4] strains from Bangladesh. Allele G consisted of G3P[6] and G2P[4] study strains as well as G2P[4] strains from Australia, Brazil, and Canada, while H consisted of G8 strains from Africa. NSP3 genotype T6 strains which included the RV5 strains, a single RV5 vG6P[8] study strain RVA/Human-wt/USA/2014741579/2014/ G6P[8], clustered together with other T6 strains.

The NSP4 gene of Wa-like (E1) strains was divided into five alleles, A, B, C, D, and E (Fig. 2I). Four new alleles, B, C, D, and E were defined for the Wa-like (E1) strains. Previously defined alleles C and D (mostly WT G1P[8] and RV1 strains) were combined to form a new allele B. Allele C included only G1P[8] study strains from the 2014 season. Allele D consisted of G1P[8] strains that were previously classified as allele B strains (McDonald et al. 2009b, 2012; Roy et al. 2014), while E included both older and contemporary G1P[8] strains from the years 1974 and 2007. Because of insufficient bootstrap support, several G1P[8] and G12P[8] strains that were previously assigned to allele B were designated as orphan E1 strains. Similarly, the NSP4 genes of DS-1-like (E2) strains grouped into five distinct alleles designated A, B, C, D, and E and four (B, C, D, and E) were newly defined (Fig. 2I). Because of insufficient bootstrap support for previously defined alleles B and C, G2P[4] strains were combined to form a new allele B which included reassortant G1P[8] strains on a DS-1-like genetic backbone from Thailand (Komoto et al. 2015), and Japan (Fujii et al. 2014; Yamamoto et al. 2014). G3P[8] strains from Germany (Pietsch and Liebert 2018) and the USA (Perkins et al. 2017). Allele C consisted of reassortant G1P[8] study strain RVA/Human-wt/USA/2014739067/2014/G1P[8] and two G2P[4] strains detected in Brazil during the 2005–2006 RVA seasons, while D included three G3P[6] study strains from the 2014, 2015, and 2016 RVA seasons, as well as other G3P[6] strains detected during the 2012 and 2013 RVA season in Uganda. Allele E exclusively consisted of RV5-derived and RV5 strains.

Phylogenetic analysis of the NSP5 gene from the study and GenBank strains grouped them into Wa-like (genotype H1) and DS-1-like (genotype H2) viruses (Fig. 2K). The Wa-like NSP5 strains were further divided into two alleles, A and B, which have both been defined previously (McDonald et al. 2009b, 2012; Roy et al. 2014). Allele A included G1P[8], G3P[8], G9P[8], and G12P[8] strains from the current study as well as those detected through studies conducted in the USA during the 2005–2013 RVA seasons. Allele B consisted of a RV1 strain and previously defined allele B US G1P[8] strains from 2005 to 2008. For the DS-1-like strains, bootstrap support was not sufficient to assign alleles within the NSP5 gene; hence only a single allele A was defined (Fig. 2K).

4. Discussion

This study provides a comprehensive sub-genotype allele analysis for 157 RVA study strains from 2014 to 2016 and 104 strains previously reported from the USA (McDonald et al. 2009b, 2012; Dennis et al. 2014; Roy et al. 2014). We identified 2- to 11 sub-genotypic allelic clusters across all eleven gene segments and defined 2 to 7 new alleles per gene. Allele determinations for RVA have been reported by only a few investigators previously (McDonald et al. 2009a, 2009b, 2012; Dennis et al. 2014; Roy et al. 2014) and such determinations included relatively few strains and identified only 3–5 alleles per gene (McDonald et al. 2012; Dennis et al. 2014; Roy et al. 2014). Our analysis used a comprehensive dataset with 300 whole-genome sequences representing the first comprehensive allelic analysis of US RVA strains, and these new allele designations may be extended to strains globally. Since there are currently no set criteria for defining an allele for RVA genes, we propose that RVA gene allele designation criteria should be: (1) monophyletic clustering of genes with bootstrap values of $\geq 70\%$; and (2) a within-allele NT similarity of $\geq 95\%$. Using these criteria, we found that allelic clusters were defined in most cases by both the VP7 and VP4 outer capsid genes or the consensus backbone constellation (Wa-like or DS-1-like). Roy et al. (2014) suggested that the VP7 generally predicted the allelic pattern among the other genes; however, we did not find that to be the case, most likely because a larger and more diverse dataset was used. In addition, our strict allele criteria resulted in our allelic calls not always matching previously reported allele designations (McDonald et al. 2012; Dennis et al. 2014; Roy et al. 2014). This incongruence is most likely due to the prior use of probability-based methods that are heavily dependent on the dataset and models used by other researchers (Roy et al. 2014). Hence, alleles may need to be constantly re-defined based on the nascent phylogenetic analyses of current sequences, along with strict criteria for allele designation. Also, it is worth noting that for the VP7 genotype G12 and VP4 genotype P[6], alleles were defined for the first time in this study.

A previous study suggested that the pre-vaccine or older Wa-like G1 alleles became extinct due to immune pressure (Arista et al. 2006), however, we identified G1 alleles that contained RVA strains from both the pre- and post-vaccine licensure eras, including two reassortant G1 strains with a DS-1-like genetic backbone that clustered with pre-vaccine or older Wa-like strains from allele A of the G1 genotype. It has been suggested that these older Wa-like G1 strains may provide a reservoir for selection of constellations that may lead to formation of vaccine escape strains (Roy et al. 2014).

All study strains that occupied the newly designated G2 allele D clustered together with strains detected in 2002 in Vietnam (Trinh et al. 2010), 2003 in Thailand (Khamrin et al. 2007), and 2013 in the USA (Dennis et al. 2014), indicating that these G2 study strains likely originated from either of these countries and later became established in the USA. Newly designated G3 allele B was comprised of three study strains grouped together with the RVA strains collected from Uganda in 2013 (Bwogi et al. 2017), indicating a possible African origin of these US G3 strains. Newly designated G3 allele C included a single study strain along with older G3 strains detected in the US in the 1970s (McDonald et al. 2009b), suggesting that these contemporary G3 allele C strains have been in circulation in the USA since the 1970s. As with G1 strains, continuous circulation of these older G3 strains may provide a reservoir for selection of constellations leading to vaccine escape. All study strains that occupied the newly designated G9 allele B clustered together with a single US strain, RVA/Human-wt/USA/2013755981/2013/ G9P[8], indicating these G9P[8] strains have been circulating in the USA since at least 2013. The clustering of select G12 study strains with G12 strains detected through the NVSN in the 2013 season (Ogden et al. 2018), and G12 strains from Rochester associated with an outbreak of pediatric AGE in 2007 (Payne et al.
within newly designated G12 allele A indicated that the G12 strains that occupy this allele have been circulating in the USA since as early as 2007 (Payne et al. 2009; Ogden et al. 2018).

Phylogenetic analysis of the VP4 gene resulted in genotypetype-specific clustering of RVA strains that further diverged to form sub-genotypic alleles. The VP4 gene of the study strains predominately clustered in newly defined P[8] alleles A and C, newly designated P[6] allele A, and previously defined P[4] alleles A and C. In addition to the common G1P[8], G3P[8], and G12P[8] strains, the novel P[8] allele A contained the recently emergent DS-1-like G1P[8] and equine-like G3P[8] strains that were not characterized until after the previous P[8] allele designations were reported. (McDonald et al. 2012; Roy et al. 2014). The new allele designations showed that symptomatic infections were not associated with any single P[8] allele.

Previous studies indicated that distinct alleles of Wa- and DS-1-like strains have been found to co-circulate within the USA (McDonald et al. 2009b, 2012; Dennis et al. 2014; Roy et al. 2014) and our phylogenetic analyses of the Wa-like study strains (G1P[8], G3P[8], G9P[8], G12P[8]) across nine RVA genes (VP1–VP3, VP6, NSP1–NSP5) predominately indicated stratification across all nine genes. Although the phylogenetic grouping of Wa-like internal genes associated with variable VP7 genotypes has been reported previously (McDonald et al. 2009b, 2012; Roy et al. 2014), this study provided a unique opportunity to examine the diversity of both older and contemporary RVA strains at the allelic level. Previous studies suggest that our observation of distinct clustering among the G1P[8] study strains collected during 2014 RVA season within alleles A (VP1), I (VP2), D and E (NSP1), and C (NSP2), may be attributed to inter-allelic reassortment events occurring between Wa-like or DS-1-like strains (McDonald et al. 2012; Roy et al. 2014), and resulting in new constellations beneficial to viral fitness. The same observation was found with the VP2 gene of G12P[8] strains from the 2014–2016 seasons within alleles F, G, and H. The comparatively increased number of alleles detected among the Wa-like strains indicated they possess extensive diversity, and our observation that no single gene segment maintains the same allele supported the existence of active and continuous reassortment events.

The DS-1-like study strains (G2P[4], G3P[6], reasortant G2P[8], G1P[8]) consistently stratified into two or three different sub-genotype alleles across the nine gene segments mentioned above, indicating high levels of diversity across multiple genes. The G2P[4] study strains consistently grouped together based on stratification of cognate alleles for the VP1–VP3 and NSP3 genes. Similarly, the G3P[6] study strains possessing a DS-1-like genetic backbone clustered together for genes VP1–VP3 and NSP1–NSP4. In addition, we observed the complete segregation of the European and Asian DS-1-like G1P[8], equine-like G3P[8], and the DS-1-like G8P[8] strains within alleles D (VP1), C (VP2), D (VP3), D (NSP1), and B (NSP4). Although previous study identified three distinct clades of G2P[4] viruses without evidence of inter-clade reassortment (Dennis et al. 2014), we observed evidence for inter-allele reassortment between the VP6 genes of G3P[6] and the DS-1-like G1P[8] study strains that clustered together in allele D, and between the NSP1 genes of G2P[4] and DS-1-like G1P[8] study strains that clustered in allele E.

We found that a subset of G1, G2, G3, G4, and P[8] RVAs contained changes in regions involved in neutralizing antibody binding and strains used in this analysis differ from those of currently available vaccines, RV1 and RV5, as noted previously (McDonald et al. 2012; Zeller et al. 2012; Roy et al. 2014; Morozova et al. 2015). However, excluding the 14 radial AA differences observed between study and vaccine-strains in the VP7 and VP4 proteins, all other AA substitutions (n = 96) were conservative or neutral. The AA substitutions observed between contemporary RVAs and vaccine strains within the neutralizing epitopes could lead to the emergence of vaccine-escape variants in response to immune pressure. Comparisons of the VP7 protein of WT G1 study strains and the G1 components of RV1 and RV5, identified two and four AA substitutions in known neutralization escape sites (Hoshino et al. 2004, 2005), respectively. The radical AA change M217T exemplified an alteration from a non-reactive and highly hydrophobic AA, methionine to a polar AA, threonine, which is common in protein functional centers and prone to phosphorylation and O-glycosylation, as such change could impact the conformation, functionality, and/or antigenicity of VP7 (Hounsell, Davies, and Renouf 1996; Betts and Russell, 2003). Similarly, radical AA changes E97G and D97G exhibited the substitution of a polar glycine in place of the larger and charged AAs, glutamic acid and aspartic acid, which could substantially alter VP7 structure, function, or sensitivity to neutralizing antibodies (Hounsell, Davies, and Renouf 1996; Betts and Russell, 2003).

Comparison of VP7 proteins between the G3 study strains and the G3 component of RV5 indicated radical change A221D, as well as neutral change K238N, which is notable because it creates a potential N-linked glycosylation site absent from the G3 protein of RV5, previously noted in Belgian G3 strains (Zeller et al. 2012). All study strains except G9 strains have an asparagine at AA position 238 of VP7, identifying a predicted glycosylation site. For the G9 strains, an aspartic acid at position 238 gives rise to a predicted integrin-binding site. Previous studies in the USA, Tunisia, Belgium, and Russia showed the presence of AA substitution K238N among G3 strains (Zeller et al. 2012; Mouna et al. 2013; Roy et al. 2014; Morozova et al. 2015). The K238N mutation appears in virulent murine RVA strains produced by serial passage of avirulent RVA strains in mice (Tsugawa, Tatsumi, and Tsutsumi 2014) and preservation of the N-glycosylation site gives the virus an advantage for propagation in cell culture (Graham et al. 2005). Mutations at position 238 have been linked to a rapid increase in viral replication efficiency in cell culture systems and likely alters antigenicity of the 7-2b neutralization epitope (Ciarel, Hoshino, and Liprandi 1997; Zeller et al. 2012). The functional impact of K238N on the integrin-binding region has not been elucidated (Graham et al. 2005). It is interesting to note that the G1 component of RV1 and RV5, and G2 and G6 components of the RV5 strains contain an asparagine at AA position 238. The use of a recently developed reverse genetics system for RVA to investigate this mutation as well as others potentially involved in neutralization-escape might be useful for future studies. (Desselberger 2017; Kanai et al. 2017).

In the P[8] protein of study strains, we identified three conservative/neutral changes in the 5-1 antigenic epitope at positions I393, A398, and I306V located in VP5 neutralization-escape sites. AA residues 382–400 in the VP5 region contained a potential membrane-interaction loop and may be essential for viral virulence, especially in facilitating viral cell attachment or penetration (Dormitzer et al. 2004; Trask et al. 2010).

Among AA substitutions described in this study, ninety per cent were conservative or neutral in nature, suggesting that purifying selection is operating on RVA genes and proteins. Radical changes in any of the neutralization epitopes may allow the protein to gain a selective advantage in the form of immune escape (Donker and Kirkwood 2012). However, as the majority of the AA changes observed in this study were conservative or neutral, it may be that RVA VP7 and VP4 proteins are under...
strong selection to maintain viral fitness (Song and Hao 2009). However, RVA immunity is thought to be polygenic and capable of conferring homotypic (same G/P type) and heterotypic (different G/P types) protection (Vesikari et al. 2007; Desselberger and Hupertz 2011). For example, RV1, which only contains the G1 and P[8] surface antigens, has proven effective at preventing moderate to severe AGE caused by non-G1P[8] strains (Staat et al. 2011; Cortese et al. 2013; Leshem et al. 2014).

Unusual reassortant strains characterized in this study included strain RVA/Human-wt/USA/3000558215/2016/G2P[8] with a genetic constellation of G2-P[8]-I2-R2-C2-M2-A2-N2-T2-E2-H2, which was detected during the 2016 season in Seattle and appears to be VP4 reassortant between DS-1-like and Wa-like RVAs. Two reassortant DS-1-like G1P[8] strains, RVA/Human-wt/USA/2014737554/2014/G1P[8] and RVA/Human-wt/USA/2014739067/2014/G1P[8], detected in Houston during the 2014 season had genetic constellations of G1-P[8]-I2-R2-C2-M2-A2-N2-T2-E2-H2 and appear to have originated from 3 or 4 reassortment events between a DS-1-like G2P[4] strain and Wa-like G1P[8] strain. Though rare, the reassortment of the outer capsid-encoding genes with strains from a different major genotype has been reported (Gentsch et al. 2005; Banyai et al. 2012). Atypical DS-1-like G1P[8] strains emerged in Thailand, Vietnam, the Philippines, and Japan in the 2012–2013 seasons (Fuji et al. 2014; Yamamoto et al. 2014; Komoto et al. 2015) and in Malawi in the 2013–2014 seasons (Jere et al. 2018). Phylogenetic analysis comparison of the US DS-1-like G1P[8] strains RVA/Human-wt/USA/2014737554/2014/G1P[8] and RVA/Human-wt/USA/2014739067/2014/G1P[8] to recently emerged DS-1-like G1P[8] from Asia and Malawi showed that US strains possessed similar VP7, NSP2, NSP3, and NSP5 genes, but distinct VP1-VP4, VP6, NSP1, and NSP4 genes. This suggests that VP7, NSP2, NSP3, and NSP5 gene segments most likely originated in countries where these strains first emerged. Distant phylogenetic clustering of the remaining seven genes (VP1-VP4, VP6, NSP1, and NSP4), however, suggests that they most likely emerged independently in the USA through reassortment events between currently circulating strains. Monophyletic clusters of genes other than VP4 and VP7 of US DS-1-like G1P[8] strains with other US DS-1-like strains were not detected in this study. Cross genotype reassortment events that have been reported in Asia (Jere et al. 2014; Yamamoto et al. 2014; Komoto et al. 2015), Africa (Jere et al. 2018), and now the USA are a possible result of vaccine pressure. Hence, continuous surveillance in the post-vaccine era is needed to monitor the evolution, spread and genetic stability of novel reassortant RVA strains derived from such events.

Novel vaccine reassortants were also identified and sequenced in this study. A single RV1 reassortant strain detected from a 9-month-old child in Houston, with no vaccination history, possessed the genetic constellation G1-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1 and was likely derived from reassortment of an RV1 strain, RV5 strain, and WT Wa-like strain. The VP7, VP2, NSP1, and NSP3 gene segments of this strain were identical to cognate gene sequences of RV1, while the VP4 gene was identical to the P[8] of RV5 strain RVA/Vaccine/USA/RotaTeq-WI79-4/1992/G6P[5]A[8]. The VP1, VP3, VP6, NSP2, NSP4, and NSP5 genes were closely related to Wa-like genotype 1 strains circulating in the USA during the 2005–2011 RVA seasons, and these genes possibly contributed to strain virulence. Reassortment events between RV1 strain and WT RVA strains in circulation have been previously reported (Rose et al. 2013; da Silva et al. 2015), the effects of such, including disease-causing potential, remain unclear. Reassortment events involving the RV1, RV5, and a WT strain in circulation are unprecedented and demonstrate that live-attenuated vaccine strains can reassort with one another in vaccinated populations. Strain RVA/Human-wt/USA/2014746684/2014/G1P[8] was likely derived from reassortment of RV1 and a WT DS-1-like strain, exhibiting a genetic constellation of G1-P[8]-I2-R2-C2-M2-A2-N2-T2-E2-H2. Two RV5 reassortants were detected in this study, a RV5 vdG6P[5] strain possessing a genetic constellation of G6-P[5]-I2-R2-C2-M2-A3-N2-T6-E2-H3, which was likely derived from reassortment of the vaccine components strain RVA/Vaccine/USA/RotaTeq-WI79-4/1992/G6P[1A][8] with either RVA/Vaccine/USA/RotaTeq-WI78-8/1992/G3P[7] or RVA/Vaccine/USA/RotaTeq-BrB-9/1996/G4P[7]S, and RVA/Human-wt/USA/2014741597/2014/G6P[8], which was derived from reassortment of a WT G6 strain and RV5 strain RVA/Vaccine/USA/RotaTeq-WI79-4/1992/G6P[1A][8]. Although reassortment events between RV5 strains resulting in vdG6P[1A] reassortants have been reported previously (Anderson 2008; Payne et al. 2010; Donato et al. 2012), this report of a RV5 vdG6P[5] reassortant strain and a reassortment event occurring between RV5 and WT G6 strain is novel. It is possible that these RV5 reassortant strains may possess characteristics of increased virulence when compared with the original ‘naturally attenuated’ bovine reassortant strain (Payne et al. 2010).

In conclusion, this study describes a detailed analysis of NGS data for 157 RVA strains from the NVSN for 2014–2016 RVA seasons and reports novel G1P[8], G2P[8] and vaccine strain reassortants. This study also presents a comprehensive analysis on the RVA alleles circulating in the USA along with defined criteria for RVA allele designation. Finally, the study highlights the need for continuous whole-genome determination initiatives to study in detail the dynamics of circulating RVA genotypes and the possible impact of the RVA vaccines on virus evolution in the USA.

Supplementary data
Supplementary data are available at Virus Evolution online.

Acknowledgements
We thank the staff of the Viral Gastroenteritis Branch at the Centers for Disease Control and Prevention and those in the seven NVSN sites for invaluable assistance. We also thank Ms. Mary Ann K Hall for reading and editing this manuscript. This study was supported by the Centers for Disease Control and Prevention.

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.

Conflict of interest: None declared.

References
Abe, M. et al. (2011) ‘Whole Genome Characterization of New Bovine Rotavirus G21P[29] and G24P[33] Strains Provides
Graham, K. L. et al. (2005) 'Rotaviruses Interact with alpha4beta7 and alpha4beta1 Integrins by Binding the Same Integrin Domains as Natural Ligands', The Journal of General Virology, 86: 3397–408.

Hoshino, Y. et al. (2004) '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, 78: 7795–802.

Hounsell, E. F., Davies, M. J., and Renouf, D. V. (1996) 'O-Linked Protein Glycosylation Structure and Function', Glycoconjugate Journal, 13: 19–26.

Jere, K. C. et al. (2014) 'Novel NSP1 Genotype Characterised in an African Camel G8P[11] Rotavirus Strain', Infection, Genetics and Evolution, 21: 58–66.

Kanai, Y. et al. (2017) 'Entirely Plasmid-Based Reverse Genetics System for Rotaviruses', Proceedings of the National Academy of Sciences of the United States of America, 114: 2349–54.

Khamrin, P. et al. (2007) 'Changing Pattern of Rotavirus G Genotype Distribution in Chiang Mai, Thailand from 2002 to 2004: Decline of G9 and Reemergence of G1 and G2', Journal of Medical Virology, 79: 775–82.

Kobayashi, N., Taniguchi, K., and Urasawa, S. (1990) 'Identification of Operationally Overlapping and Independent Cross-Reactive Neutralization Regions on Human Rotavirus VP4', J Gen Virol, 71 (Pt 11): 2615–23.

Kohyama, T., and Urasawa, S. (1990) 'Full Genome-Based Classification of a Sialic Acid-Independent Human Rotavirus Strain', Acta Virologica, 34: 19–22.

Leshem, E. et al. (2014) 'Distribution of Rotavirus Strains and Strain-Specific Effectiveness of the Rotavirus Vaccine after Its Introduction: A Systematic Review and Meta-Analysis', The Lancet. Infectious Diseases, 14: 847–56.

Liu, L. et al. (2016) 'Global, Regional, and National Causes of under-5 Mortality in 2000-15: An Updated Systematic Analysis with Implications for the Sustainable Development Goals', The Lancet (London, England), 388: 3027–35.

Maes, P. et al. (2009) 'RotaC: A Web-Based Tool for the Complete Genome Classification of a Rotavirus', BMC Microbiology, 9: 238.

Mascarenhas, J. D. et al. (2010) 'Identification of Two Sublineages of Genotype G2 Rotavirus among Diarrheic Children in ParaSupebas, Southern Para State, Brazil', Journal of Medical Virology, 82: 712–9.

Matthijnssens, J., et al. (2008a) '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, 82: 3204–19.

Pietsch, C., and Liebert, U. G. (2018) 'Molecular Characterization of Different Equine-like G3 Rotavirus Strains from Germany', Infection, Genetics and Evolution, 57: 46–50.

Pottgieter, A. C. et al. (2009) 'Improving Sibling Transmission of Vaccine-Derived Rotavirus (RotaTeq) Associated with Rotavirus Gastroenteritis', Pediatrics, 125: e438–441.

Perkins, C. et al. (2017) 'Genomic Characterization of the First Equine-Like G3P[8] Rotavirus Strain Detected in the United States', Genome Announcements, 5: e01341–17.

Petersen, E. F. et al. (2004) 'UCSF Chimera—a Visualization System for Exploratory Research and Analysis', Journal of Computational Chemistry, 25: 1605–12.

Ragim, R. F. (1997) 'Genetics of the Rotaviruses', Annual Review of Microbiology, 51: 225–55.

Roy, S. et al.; New Vaccine Surveillance Network. (2014) 'Comparative Genomic Analysis of Genogroup 1 (Wa-like) Rotaviruses Circulating in the USA, 2006-2009', Infection, Genetics and Evolution, 28: 513–23.

Song, X. F., and Hao, Y. (2009) 'Adaptive Evolution of Rotavirus VP7 and NSP4 Genes in Different Species', Computational Biology and Chemistry, 33: 344–9.
Staat, M. A. et al.; the New Vaccine Surveillance Network (NVSN). (2011) ‘Effectiveness of Pentavalent Rotavirus Vaccine against Severe Disease’, PEDIATRICS, 128: e267-275-e275.

Stupka, J. A. et al.; Argentinean National Surveillance Network for Diarrheas. (2009) ‘National Rotavirus Surveillance in Argentina: High Incidence of G9P[8] Strains and Detection of G4P[6] Strains with Porcine Characteristics’, Infection, Genetics and Evolution, 9: 1225–31.

Tamura, K. et al. (2013) ‘MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0’, Molecular Biology and Evolution, 30: 2725–9.

Tate, J. E. et al. (2009) ‘Decline and Change in Seasonality of US Rotavirus Activity after the Introduction of Rotavirus Vaccine’, Pediatrics, 124: 465–71.

Trask, S. D. et al. (2010) ‘A Rotavirus Spike Protein Conformational Intermediate Binds Lipid Bilayers’, Journal of Virology, 84: 1764–70.

Trinh, Q. D. et al. (2010) ‘Sequence Analysis of the VP7 Gene of Human Rotaviruses G2 and G4 Isolated in Japan, China, Thailand, and Vietnam during 2001-2003’, Journal of Medical Virology, 82: 878–85.

Troeger, C. et al. (2018) ‘Rotavirus Vaccination and the Global Burden of Rotavirus Diarrhea among Children Younger than 5 Years’, JAMA Pediatrics, 172: 958–65.

Trojnar, E. et al. (2013) ‘Identification of an Avian Group a Rotavirus Containing a Novel VP4 Gene with a Close Relationship to Those of Mammalian Rotaviruses’, Journal of General Virology, 94: 136–42.

Tsugawa, T., Tatsumi, M., and Tsutsumi, H. (2014) ‘Virulence-Associated Genome Mutations of Murine Rotavirus Identified by Alternating Serial Passages in Mice and Cell Cultures’, Journal of Virology, 88: 5543–58.

Vesikari, T. et al. (2007) ‘Efficacy of Human Rotavirus Vaccine against Rotavirus Gastroenteritis during the First 2 Years of Life in European Infants: Randomised, Double-Blind Controlled Study’, The Lancet, 370: 1757–63.

Ward, M. L. et al. (2016) ‘Molecular Characterization of the First G24P[14] Rotavirus Strain Detected in Humans’, Infection, Genetics and Evolution, 43: 338–42.

Waterhouse, A. et al. (2018) ‘SWISS-MODEL: Homology Modelling of Protein Structures and Complexes’, Nucleic Acids Research, 46: W296–W303.

Yamamoto, S. P. et al. (2014) ‘Gastroenteritis Outbreaks Caused by a DS-1-like G1P[8] Rotavirus Strain, Japan, 2012–2013’, Emerging Infectious Diseases, 20: 1030–3.

Yinda, C. K. et al. (2016) ‘Novel Highly Divergent Reassortant Bat Rotaviruses in Cameroon, without Evidence of Zoonosis’, Scientific Reports, 6: 34209.

Zeller, M. et al. (2012) ‘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, 50: 966–76.

Zhou, Y. J. et al. (1994) ‘Localization of Rotavirus VP4 Neutralization Epitopes Involved in Antibody-Induced Conformational Changes of Virus Structure’, Journal of Virology, 68: 3955–64.