| Virology | Announcement

Coding-complete genome sequences of rotavirus A reference strains EDIM, Ph158, and CC425

Mary C. Casey-Moore,[1] Eric Katz,[1] Slavica Mijatovic-Rustempasic,[1] Jose Jaimes,[1] Rashi Gautam,[1] Michael D. Bowen[1]

AUTHOR AFFILIATION See affiliation list on p. 4.

ABSTRACT This study reports the coding-complete genome sequences of three rotavirus A (RVA) reference strains previously adapted in tissue culture: RVA/Mouse-tc/USA/EDIM/XXXX/G16P[16] with a G16-P[16]-I7-R7-C7-M8-A7-N7-T10-E7-H9 genotype constellation, RVA/Human-tc/USA/Ph158/1998/G9P[6] with a G9-P[6]-I2-R2-C2-M2-A2-N2-T2-E2-H2 genotype constellation, and RVA/Human-tc/USA/CC425/1998/G3P[9] with a G3-P[9]-I2-R2-C2-M2-A3-N2-T1-E2-H3 genotype constellation.

Group A rotaviruses (RVAs) cause acute gastroenteritis in young children and animals (1–3). Belonging to the *Sedoreoviridae* family, RVA genomes consist of 11 double‐stranded RNA (dsRNA) segments whose classification is based on sequencing of all genome segments (4, 5). The genotype constellation for the 11 RVA segments is Gx‐P[x]‐Ix‐Rx‐Cx‐Mx‐Ax‐Nx‐Tx‐Ex‐Hx for genes VP7‐VP4‐VP6‐VP1‐VP2‐VP3‐NSP1‐NSP2‐NSP3‐NSP4‐NSP5/6, respectively (6). Here, we report the coding-complete genome sequences for reference strains RVA/Mouse-tc/USA/EDIM/XXXX/G16P[16] (referred to as EDIM), RVA/Human-tc/USA/Ph158/1998/G9P[6] (Ph158), and RVA/Human-tc/USA/CC425/1998/G3P[9] (CC425), which were previously partially characterized as follows: EDIM as G16-P[16]-I7-Rx-Cx-Mx-A7-Nx-Tx-E7-Hx (7), Ph158 as G9-P[x]-Ix-Rx-Cx-Mx-Ax-Nx-Tx-Ex-Hx (8), and CC425 as G3-P[9]-Ix-Rx-Cx-Mx-Ax-Nx-Tx-E2-Hx (9).

EDIM was initially isolated from the stool of an infected mouse and adapted for mouse models of human rotavirus infection (10). Ph158 was cultured to help characterize human G9 strains (8), and CC425 was cultured for P[9] characterization (9). These reference strains have significantly impacted the field through their use in early studies evaluating the immunological mechanisms against RVA, phylogenetic analyses of emerging strains, and assay development for genotyping; however, the gene-coding regions have yet to be fully sequenced (10, 11). In this report, next-generation sequencing (NGS) was performed to facilitate accurate interpretations of the origin of the reference strains and assist in tracing their evolutionary patterns.

Each strain was propagated in MA104 cell monolayers from existing stocks, as previously described (12). Viral stocks were maintained in IMDM media, and undiluted supernatant was used for the dsRNA extraction. The NGS methods were performed as previously described (13). Briefly, RVA dsRNA was extracted using the MagNA Pure Compact RNA Isolation Kit (Roche). The dsRNA served as a template for the cDNA sequencing library using the NEBNext Ultra RNA Library Prep Kit and Multiplex Oligos for Illumina Kit (New England Biolabs). The library was sequenced using a MiSeq Reagent Kit v.2 (500-cycle) on a MiSeq System (Illumina).

The total number of paired-end reads was 1,267,346 for EDIM, 1,064,027 for Ph158, and 1,486,484 for CC425. The average read length ranged from 189 to 198 bp. Reference-guided assembly was performed using CLC Genomics Workbench 20.0.3

TABLE 1 Sequence details and BLASTN results of reference strains EDIM, Ph158, and CC425

Strain	Gene	Genotype	GC content (%)	Full sequence length (bp)	Coding sequence length (bp)	ORF	GenBank accession no.	Data for top BLASTN match		
								BLASTN description	Identity (%)	GenBank accession no.
EDIM	VP7	G16	39.6	1062	981	Complete	MT276814	Murine rotavirus strain ETD_822 VP7 gene, complete cds	99.8	GQ479955.1
	VP4	P[16]	45.3	2347	2325	Complete	MT276815	Murine rotavirus EW outer capsid protein VP4 gene, complete cds	100.0	U08429.1
	VP6	I7	43.7	1355	1194	Complete	MT276816	Murine rotavirus strain ETD_822 VP6 gene, complete cds	100.0	GQ479952.1
	VP1	R7	39.7	3294	3267	Complete	MT276817	Murine rotavirus strain ETD_822 VP1 gene, complete cds	99.9	GQ479947.1
	VP2	C7	40.7	2667	2637	Complete	MT276818	Murine rotavirus strain ETD_822 VP2 gene, complete cds	99.8	GQ479948.1
	VP3	M8	39.2	2583	2508	Complete	MT276819	Murine rotavirus strain ETD_822 VP3 gene, complete cds	99.9	GQ479949.1
	NSP1	A7	41.5	1599	1482	Complete	MT276820	Murine rotavirus strain ETD_822 NSP1 gene, complete cds	99.4	GQ479951.1
	NSP2	N7	39.9	1043	951	Complete	MT276821	Murine rotavirus strain ETD_822 NSP2 gene, complete cds	99.8	GQ479954.1
	NSP3	T10	41.5	1069	933	Complete	MT276822	Murine rotavirus strain ETD_822 NSP3 gene, complete cds	99.7	GQ479953.1
	NSP4	E7	43.6	740	528	Complete	MT276823	Murine rotavirus strain ETD_822 NSP4 gene, complete cds	99.6	GQ479956.1
	NSP5	H9	42.3	647	594	Complete	MT276824	Murine rotavirus strain ETD_822 NSP5/6 gene, complete cds	99.9	GQ479957.1
Ph158	VP7	G9	34.3	1027	981	Complete	MT339194	Rotavirus G9 VP7 gene for neutralizing capsid protein, strain PH158, genomic RNA	99.9	AJ491183.1
	VP4	P[6]	35.4	2347	2328	Complete	MT339195	Human rotavirus G9P[6] strain GR 10924/99 VP4 gene, complete cds	99.6	FJ183356.1
	VP6	I2	37.0	1285	1194	Complete	MT339196	Rotavirus A strain RVA/Human-wt/UGA/MSK-13–048/2013 /G9P[8] VP6 (VP6) mRNA, complete cds	99.7	KX655510.1
	VP1	R2	32.4	3273	3267	Complete	MT339197	Human rotavirus G9P[6] strain GR 10924/99 VP1 gene, complete cds	99.5	FJ183353.1
	VP2	C2	32.9	2654	2640	Complete	MT339198	Rotavirus A strain MRC-DPRU9317 segment 2, complete sequence	99.4	JN605438.1
	VP3	M2	30.8	2571	2508	Complete	MT339199	Human rotavirus G9P[6] strain GR 10924/99 VP3 gene, complete cds	99.5	FJ183355.1
	NSP1	A2	27.8	1536	1461	Complete	MT339200	Rotavirus A strain RVA/Human-wt/UGA/MSK-13–048/2013 /G9P[8] NSP1 (NSP1) mRNA, complete cds	99.4	KX655512.1
	NSP2	N2	35.8	1032	954	Complete	MT339201	Human rotavirus G9P[6] strain GR 10924/99 NSP2 gene, complete cds	99.7	FJ183361.1
	NSP3	T2	31.6	1044	933	Complete	MT339202	Rotavirus A strain MRC-DPRU9317 segment 7, complete sequence	99.6	JN605445.1
	NSP4	E2	38.5	719	528	Complete	MT339203	Human rotavirus NSP4 gene for nonstructural protein 4, strain 1071	100.0	AJ236754.1
	NSP5	H2	30.7	716	603	Complete	MT339204	Rotavirus A strain RVA/Human-wt/ZAF/MRC-DPRU9164/1999 /G9P[6] segment 11 nonstructural protein 5 (NSP5) and nonstructural protein 6 (NSP6) genes, complete cds	99.6	KJ752248.1
CC425	VP7	G3	35.8	1054	981	Complete	MT276803	Human rotavirus VP7 gene for VP7 capsid protein, genomic RNA, strain CC425, P3[9],G3	100.00	AJ311738.1
	VP4	P[9]	35.9	2355	2328	Complete	MT276804	Rotavirus A strain RVA/Human-tc/USA/Se584/1998 /G6P[9] outer capsid protein VP4 gene, complete cds	98.03	EF672605.1
	VP6	I2	38.4	1351	1194	Complete	MT276805	Equine rotavirus A strain RVA/Horse-wt/IND/ERV2/2015 /G6P[1] segment 6 VP6 gene, complete cds	97.71	OK651101.1
	VP1	R2	34.8	3292	3267	Complete	MT276806	Bovine rotavirus core protein (VP1) gene, complete cds	97.17	J04346.1
	VP2	C2	35.2	2669	2643	Complete	MT276807	Rotavirus A strain RVA/Human-tc/USA/Se584/1998 /G6P[9] core shell protein VP2 gene, complete cds	98.75	EF583042.1

(Continued on next page)

TABLE 1 Sequence details and BLASTN results of reference strains EDIM, Ph158, and CC425 (*Continued*)

Strain	Gene	Genotype	GC content (%)	Full sequence length (bp)	Coding sequence length (bp)	ORF	GenBank accession no.	Data for top BLASTN match		
								BLASTN description	Identity (%)	GenBank accession no.
CC425	VP3	M2	32.4	2588	2508	Complete	MT276808	Rotavirus A strain RVA/Cow-tc/USA/NCDV/1967 /G6P6[1] VP3 gene, complete cds	97.33	DQ870495.1
	NSP1	A3	35.7	1570	1476	Complete	MT276809	Rotavirus A strain RVA/Human-tc/USA/Se584/1998 /G6P[9] nonstructural protein one gene, complete cds	98.58	EF672606.1
	NSP2	N2	35.0	1049	954	Complete	MT276810	Rotavirus A strain RVA/Human-tc/USA/Se584/1998 /G6P[9] nonstructural protein two gene, complete cds	99.27	EF672608.1
	NSP3	T1	33.7	1074	933	Complete	MT276811	Rotavirus RVA/Human-wt/USA/12US1134/2012 /G3P[9] NSP3 (NSP3) gene, complete cds	98.7	KF500516.1
	NSP4	E2	41.4	740	528	Complete	MT276812	Human Rotavirus NSP4 gene for nonstructural-protein 4, genomic RNA, strain CC425, P3[9],G3	100.0	AJ311728.1
	NSP5	H3	38.8	641	597	Complete	MT276813	Rotavirus A strain RVA/Human-wt/ITA/PAI58/1996 /G3P[9] NSP5 gene, complete cds	98.4	GU296419.1

software (14), resulting in coding-complete sequences for all segments (Table 1). To determine the open reading frames (ORFs), study sequences were cross-referenced using additional sequences with well-established ORFs from GenBank and further validated using established literature that elucidates the conventionally recognized gene-coding assignments (15, 16). Partial non-coding sequences were reported depending on coverage. Genotypes were determined according to the guidelines of the Rotavirus Classification Working Group (7) using NCBI's BLASTN tool (17).

NGS and genotyping results are presented in Table 1. The genome constellations were determined to be: EDIM as G16-P[16]-I7-R7-C7-M8-A7-N7-T10-E7-H9, Ph158 as G9-P[6]-I2-R2-C2-M2-A2-N2-T2-E2-H2, and CC425 as G3-P[9]-I2-R2-C2-M2-A3-N2-T1-E2-H3. The constellation for EDIM has been observed in RVA strains of murine origin and was most closely related to murine strain ETD_822 (6). Ph158 was found to possess a complete genogroup 2 constellation. CC425 is closely related to equine, bovine, and human-origin strains, including G6P[9] strain Se584 (18). While further analysis is needed, this report presents the complete gene sequences of common reference strains used throughout RVA research.

ACKNOWLEDGMENTS

AUTHOR AFFILIATION

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

AUTHOR ORCIDs

Mary C. Casey-Moore ⓘ http://orcid.org/0000-0003-2608-4970
Michael D. Bowen ⓘ https://orcid.org/0000-0002-2564-3083

FUNDING

DATA AVAILABILITY

Project information, including data sets and the raw sequence reads, has been deposited in BioProject under accession no. PRJNA614927. NCBI GenBank accession numbers are as follows: EDIM (MT276814-MT276824), PH158 (MT339194-MT339204), and CC425 (MT276803- MT276813). Accession numbers for individual genes are listed in Table 1.

REFERENCES

1. Troeger C, Blacker BF, Khalil IA, Rao PC, Cao S, Zimsen SR, Albertson SB, Stanaway JD, Deshpande A, Abebe Z, Alvis-Guzman N, Amare AT, Asgedom SW, Anteneh ZA, Antonio CAT, Aremu O, Asfaw ET, Atey TM, Atique S, Avokpaho EFGA, Awasthi A, Ayele HT, Barac A, Barreto ML, Bassat Q, Belay SA, Bensenor IM, Bhutta ZA, Bijani A, Bizuneh H, Castañeda-Orjuela CA, Dadi AF, Dandona L, Dandona R, Do HP, Dubey M, Dubljanin E, Edessa D, Endries AY, Eshrati B, Farag T, Feyissa GT, Foreman KJ, Forouzanfar MH, Fullman N, Gething PW, Gishu MD, Godwin WW, Gugnani HC, Gupta R, Hailu GB, Hassen HY, Hibstu DT, Ilesanmi OS, Jonas JB, Kahsay A, Kang G, Kasaeian A, Khader YS, Khalil IA, Khan EA, Khan MA, Khang Y-H, Kissoon N, Kochhar S, Kotloff KL, Koyanagi A, Kumar GA, Magdy Abd El Razek H, Malekzadeh R, Malta DC, Mehata S, Mendoza W, Mengistu DT, Menota BG, Mezgebe HB, Mlashu FW, Murthy S, Naik GA, Nguyen CT, Nguyen TH, Ningrum DNA, Ogbo FA, Olagunju AT, Paudel D, Platts-Mills JA, Qorbani M, Rafay A, Rai RK, Rana SM, Ranabhat CL, Rasella D, Ray SE, Reis C, Renzaho AM, Rezai MS,

Ruhago GM, Safiri S, Salomon JA, Sanabria JR, Sartorius B, Sawhney M, Sepanlou SG, Shigematsu M, Sisay M, Somayaji R, Sreeramareddy CT, Sykes BL, Taffere GR, Topor-Madry R, Tran BX, Tuem KB, Ukwaja KN, Vollset SE, Walson JL, Weaver MR, Weldegwergs KG, Werdecker A, Workicho A, Yenesew M, Yirsaw BD, Yonemoto N, El Sayed Zaki M, Vos T, Lim SS, Naghavi M, Murray CJ, Mokdad AH, Hay SI, Reiner RC. 2018. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the global burden of disease study 2016. Lancet Infect Dis 18:1211–1228. https://doi.org/10.1016/S1473-3099(18)30362-1

2. Troeger C, Khalil IA, Rao PC, Cao S, Blacker BF, Ahmed T, Armah G, Bines JE, Brewer TG, Colombara DV, Kang G, Kirkpatrick BD, Kirkwood CD, Mwenda JM, Parashar UD, Petri WA, Riddle MS, Steele AD, Thompson RL, Walson JL, Sanders JW, Mokdad AH, Murray CJL, Hay SI, Reiner RC. 2018. Rotavirus vaccination and the global burden of rotavirus diarrhea among children younger than 5 years. JAMA Pediatr 172:958–965. https://doi.org/10.1001/jamapediatrics.2018.1960

3. Greenberg HB, Vo PT, Jones R. 1986. Cultivation and characterization of three strains of murine rotavirus. J Virol 57:585–590. https://doi.org/10.1128/JVI.57.2.585-590.1986

4. Estes MK, Kapikian AZ. 2007. Rotaviruses, p 1917–1974. In KnipeDM, B Roizman (ed), Fields Virology. Lippincott Williams & Wilkins.

5. Matthijnssens J, Attoui H, Bányai K, Brussaard CPD, Danthi P, del Vas M, Dermody TS, Duncan R, Fāng Q, Johne R, Mertens PPC, Mohd Jaafar F, Patton JT, Sasaya T, Suzuki N, Wei T. 2022. ICTV virus taxonomy profile: sedoreoviridae 2022. J Gen Virol 103. https://doi.org/10.1099/jgv.0.001781

6. Matthijnssens J, Miño 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, Parreño V, Bányai K, Barrandeguy M, Van Ranst M. 2012. Complete molecular genome analyses of equine rotavirus a strains from different continents reveal several novel genotypes and a largely conserved genotype constellation. J Gen Virol 93:866–875. https://doi.org/10.1099/vir.0.039255-0

7. Matthijnssens J, Ciarlet M, Rahman M, Attoui H, Bányai K, Estes MK, Gentsch JR, Iturriza-Gómara M, Kirkwood CD, Martella V, Mertens PPC, Nakagomi O, Patton JT, Ruggeri FM, Saif LJ, Santos N, Steyer A, Taniguchi K, Desselberger U, Van Ranst M. 2008. Recommendations for the classification of group a rotaviruses using all 11 genomic RNA segments. Arch Virol 153:1621–1629. https://doi.org/10.1007/s00705-008-0155-1

8. Laird AR, Gentsch JR, Nakagomi T, Nakagomi O, Glass RI. 2003. Characterization of serotype G9 rotavirus strains isolated in the United States and India from 1993 to 2001. J Clin Microbiol 41:3100–3111. https://doi.org/10.1128/JCM.41.7.3100-3111.2003

9. Griffin DD, Nakagomi T, Hoshino Y, Nakagomi O, Kirkwood CD, Parashar UD, Glass RI, Gentsch JR, National Rotavirus Surveillance System. 2002. Characterization of nontypeable rotavirus strains from the United States: identification of a new rotavirus reassortant (P2A[6],G12) and rare P3[9] strains related to bovine rotaviruses. Virology 294:256–269. https://doi.org/10.1006/viro.2001.1333

10. Ward RL, McNeal MM, Sheridan JF. 1990. Development of an adult mouse model for studies on protection against rotavirus. J Virol 64:5070–5075. https://doi.org/10.1128/JVI.64.10.5070-5075.1990

11. Gautam R, Mijatovic-Rustempasic S, Esona MD, Tam KI, Quaye O, Bowen MD. 2016. One-step multiplex real-time RT-PCR assay for detecting and genotyping wild-type group a rotavirus strains and vaccine strains (Rotarix and Rotateq) in stool samples. PeerJ 4:e1560. https://doi.org/10.7717/peerj.1560

12. Arnold M, Patton JT, McDonald SM. 2009. Culturing, storage, and quantification of rotaviruses. Curr Protoc Microbiol Chapter 15:15C. https://doi.org/10.1002/9780471729259.mc15c03s15

13. Esona MD, Gautam R, Katz E, Jaime J, Ward ML, Wikswo ME, Betrapally NS, Rustempasic SM, Selvarangan R, Harrison CJ, Boom JA, Englund J, Klein EJ, Staat MA, McNeal MM, Halasa N, Chappell J, Weinberg GA, Payne DC, Parashar UD, Bowen MD. 2021. Comparative genomic analysis of genogroup 1 and genogroup 2 rotaviruses circulating in seven US cities, 2014-2016. Virus Evol 7:veab023. https://doi.org/10.1093/ve/veab023

14. QIAGEN. QIAGEN CLC genomics workbench: NGS data analysis for any Species, any platform, any workflow, on QIAGEN website. Available from: https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-clc-genomics-workbench. Accessed Sep 2, 2022

15. Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2014. Genbank. Nucleic Acids Res 42:D32–7. https://doi.org/10.1093/nar/gkt1030

16. Estes MK, Cohen J. 1989. Rotavirus gene structure and function. Microbiol Rev 53:410–449. https://doi.org/10.1128/mr.53.4.410-449.1989

17. NCBI Resource Coordinators. 2016. Database resources of the national center for biotechnology information. Nucleic Acids Res 44:D7–19. https://doi.org/10.1093/nar/gkv1290

18. Heiman EM, McDonald SM, Barro M, Taraporewala ZF, Bar-Magen T, Patton JT. 2008. Group a human rotavirus genomics: evidence that gene constellations are influenced by viral protein interactions. J Virol 82:11106–11116. https://doi.org/10.1128/JVI.01402-08