



# Genome Sequence of a Wa-Like G3P[8] Rotavirus from a 12-Month-Old Child with Diarrhea in Manipur, India

Yengkhom Damayanti Devi,<sup>a</sup> Upalabdha Dey,<sup>a</sup>  Aditya Kumar,<sup>a</sup> Chongtham Shyamsunder Singh,<sup>b</sup>  Nima D. Namsa<sup>a,c</sup>

<sup>a</sup>Department of Molecular Biology and Biotechnology, Tezpur University, Napaam, Assam, India

<sup>b</sup>Department of Paediatrics, Regional Institute of Medical Sciences, Imphal, India

<sup>c</sup>Centre for Multi-disciplinary Research, Tezpur University, Napaam, Assam, India

**ABSTRACT** Rotavirus A (RVA) was detected in the stool of a 12-month-old child with diarrhea, mild fever, and vomiting. A viral metagenomic approach identified a Wa-like genotype G3P[8] strain named RVA/Human-wt/IND/RM25112/2016.

Rotavirus A (RVA) is the major etiologic agent of acute gastroenteritis in young children worldwide. Six structural proteins (VP1 to VP4, VP6, and VP7) and six nonstructural proteins (NSPs) (NSP1 to NSP6) are encoded by the 11-double-stranded RNA (dsRNA) genome (1). The VP7-VP4-VP6-VP1-VP2-VP3-NSP1-NSP2-NSP3-NSP4-NSP5/NSP6 genes of an RVA strain are denoted by the descriptor Gx-P[x]-Ix-Rx-Cx-Mx-Ax-Nx-Tx-Ex-Hx (x represents the genotype number) in the classification system (2). The Wa, DS-1, and AU-1 genotype constellations are described as G1-P[8]-I1-R1-C1-M1-A1-N1-T1-E1-H1, G2-P[4]-I2-R2-C2-M2-A2-N2-T2-E2-H2, and G3-P[9]-I3-R3-C3-M3-A3-N3-T3-E3-H3, respectively (3–6).

The presence of RVA antigen in stool was confirmed using the Premier Rotaclone enzyme-linked immunosorbent assay (ELISA) kit (Meridian Bioscience Inc., USA) and a rapid immunochromatographic test, the Rota+Adeno+Astro+Noro enzyme immunoassay (EIA) combo card kit (CerTest Biotec S.L., Spain). The sample was homogenized by vortex-mixing at room temperature for 5 min, centrifuged at 14,500 rpm for 12 min, and filtered through a 0.22- $\mu$ m filter. Total RNA was extracted using the TRIzol method (catalog number 15596018; Invitrogen, Carlsbad, CA, USA). cDNA was prepared using SuperScript III reverse transcriptase and random hexamers (Invitrogen). The dsRNA was denatured at 95°C for 5 min. The reverse transcription reaction was carried out with thermal cycling steps at 25°C for 10 min and 55°C for 60 min. VP7 and VP4 nested multiplex PCR and sequencing confirmed the isolate as the G3P[8] genotype (7).

The TruSeq stranded total RNA library preparation kit (catalog number 20020594; Illumina) was utilized to prepare the library for whole-genome sequencing, and final libraries were quantified using a Qubit v.4.0 fluorometer and a Qubit DNA high-sensitivity (HS) assay kit (catalog number Q32851; Thermo Fisher Scientific). Paired-end (2  $\times$  150-bp) sequencing was performed using an Illumina NovaSeq 6000 system. A total of 19.6 million raw reads were generated; after trimming of the adapters and barcodes, 17.4 million retained reads were utilized for the downstream analysis. The quality control of the raw data was carried out using FastQC v.0.11.9. The processed reads were submitted to Genome Detective (8) and CCMetagen (9) servers to determine the taxonomic assignment of the sequences. The cleaned reads were aligned with rotavirus reference genome sequences obtained from NCBI (GenBank accession numbers [GCA\\_003156295.1](#), [GCA\\_003259085.1](#), [GCF\\_000864225.1](#), [GCF\\_000864245.1](#), [GCF\\_000880735.1](#), [GCF\\_000890155.1](#), [GCF\\_000907835.1](#), [GCF\\_000910335.1](#), [GCF\\_000973395.3](#), [GCF\\_001343825.1](#), [GCF\\_004117615.1](#), and [GCF\\_013086085.1](#)), and the alignment statistics were measured using Bowtie2 v.2.4.2 (10). The rotavirus-specific reads were extracted using CCMetagen and *de novo* assembled with MEGAHIT v.1.2.9 (11). All software was used with default settings.

**Editor** Jelle Matthijnsens, KU Leuven

**Copyright** © 2022 Devi et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

Address correspondence to Nima D. Namsa, [namsa@tezu.ernet.in](mailto:namsa@tezu.ernet.in).

The authors declare a conflict of interest.

**Received** 19 January 2022

**Accepted** 10 June 2022

**Published** 25 July 2022

**TABLE 1** Assembly and genotyping details for the RVA genome described in this study

Segment no. (protein product)	Segment length (bp)	Mean coverage (×)	No. of reads	GC content (%)	ORF length (aa) (protein product) <sup>a</sup>	Genotype	GenBank accession no. for closest strain	GenBank accession no.	Similarity to closest strain (%)
1 (VP1)	3,299	871	6,270	31	955 (VP1)	R1	<a href="#">GU199492.1</a>	<a href="#">OL906393</a>	98.34
2 (VP2)	2,915	1,179	7,358	31	894 (VP2)	C1	<a href="#">DQ146661.1</a>	<a href="#">OL906392</a>	98.35
3 (VP3)	2,737	1,068	6,369	29	835 (VP3)	M1	<a href="#">DQ146651.1</a>	<a href="#">OL906391</a>	98.05
4 (VP4)	2,443	1,103	6,666	31	775 (VP4)	P[8]	<a href="#">DQ146652.1</a>	<a href="#">OL906390</a>	97.12
5 (NSP1)	1,673	927	3,533	29	497 (NSP1)	A1	<a href="#">EF560708.1</a>	<a href="#">OL906389</a>	98.08
6 (VP6)	1,438	1,799	6,405	36	397 (VP6)	I1	<a href="#">DQ146642.1</a>	<a href="#">OL906388</a>	98.83
7 (NSP3)	1,205	1,055	2,223	32	310 (NSP3)	T1	<a href="#">DQ146646.1</a>	<a href="#">OL906387</a>	97.32
8 (NSP2)	1,201	650	1,763	33	317 (NSP2)	N1	<a href="#">DQ146656.1</a>	<a href="#">OL906386</a>	98.32
9 (VP7)	1,189	971	2,247	34	326 (VP7)	G3	<a href="#">EF672602.1</a>	<a href="#">OL906385</a>	96.93
10 (NSP4)	922	951	2,017	38	175 (NSP4)	E1	<a href="#">DQ146658.1</a>	<a href="#">OL906384</a>	98.86
11 (NSP5/NSP6)	742	838	1,375	39	197 (NSP5)/90 (NSP6)	H1	<a href="#">DQ146681.1</a>	<a href="#">OL906383</a>	99.49

<sup>a</sup> ORF, open reading frame; aa, amino acids.

The CCMetagen protocol yielded 8.7 million reads after quality trimming. Eleven nearly complete genome segments of an RVA strain were recovered based on the length of each best hit, as shown in Table 1. This strain was typed as G3P[8] with a human Wa-like genotype constellation (I1-R1-C1-M1-A1-N1-T1-E1-H1) and a common ancestor with porcine RVA strains (5) using the ViPR (12) and RotaC (13) tools. All 11 segments showed a range of 96.93% to 99.49% similarity with known contemporary RVA strains (Table 1). This genome will serve as a reference for the region and form part of the baseline data that might be needed to address the impact of rotavirus vaccination in India (14).

This study was approved by the institutional ethical review board of Tezpur University [approval number DoRD/TUEC/10-14/2017/4(b)], and written informed consent was obtained from the parents/guardians of study participants and control subjects.

**Data availability.** The assembled genomes have been deposited in GenBank under the accession numbers [OL906383](#) to [OL906393](#). The raw reads have been deposited under SRA accession number [SRX14381993](#).

## ACKNOWLEDGMENTS

Y.D.D. is a recipient of a UGC-CSIR Senior Research Fellowship for Ph.D. This work was supported by the Research Fund of the Department of Biotechnology, Government of India (grant BT/338/NE/TBP/2012 [14 August 2014]).

We thank the next-generation sequencing (NGS) team at Neuberger Supratech Reference Laboratories (Ahmedabad, India) for their contribution in the preparation of NGS libraries and the generation of the NGS data.

## REFERENCES

- Estes MK. 2001. Rotaviruses and their replication, p 1747–1785. *In* Knipe DM, Howley PM (ed), *Fields virology*, 4th ed. Lippincott Williams & Wilkins, Philadelphia, PA.
- Matthijnssens J, Ciarlet M, McDonald SM, Attoui H, Bányai K, Brister JR, Buesa J, Esona MD, Estes MK, Gentsch JR, Iturriza-Gómara M, Johne R, Kirkwood CD, Martella V, Mertens PPC, Nakagomi O, Parreño V, Rahman M, Ruggeri FM, Saif LJ, Santos N, Steyer A, Taniguchi K, Patton JT, Desselberger U, Van Ranst M. 2011. Uniformity of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group (RCWG). *Arch Virol* 156:1397–1413. <https://doi.org/10.1007/s00705-011-1006-z>.
- Nakagomi O, Nakagomi T, Akatani K, Ikegami N. 1989. Identification of rotavirus genogroups by RNA-RNA hybridization. *Mol Cell Probes* 3:251–261. [https://doi.org/10.1016/0890-8508\(89\)90006-6](https://doi.org/10.1016/0890-8508(89)90006-6).
- 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>.
- Matthijnssens J, Ciarlet M, Heiman E, Arijs I, Delbeke T, McDonald SM, Palombo EA, Iturriza-Gómara M, Maes P, Patton JT, Rahman M, Van Ranst M. 2008. Full genome-based classification of rotaviruses reveals a common origin between human Wa-like and porcine rotavirus strains and human DS-1-like and bovine rotavirus strains. *J Virol* 82:3204–3219. <https://doi.org/10.1128/JVI.02257-07>.
- 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>.
- World Health Organization. 2009. Manual of rotavirus detection and characterization methods. World Health Organization, Geneva, Switzerland.
- Vilsker M, Moosa Y, Nooij S, Fonseca V, Ghysens Y, Dumon K, Pauwels R, Alcantara LC, Vanden Eynden E, Vandamme A-M, Deforche K, de Oliveira T. 2019. Genome Detective: an automated system for virus identification from high-throughput sequencing data. *Bioinformatics* 35:871–873. <https://doi.org/10.1093/bioinformatics/bty695>.
- Marcelino VR, Clausen PTL, Buchmann JP, Wille M, Iredell JR, Meyer W, Lund O, Sorrell TC, Holmes EC. 2020. CCMetagen: comprehensive and accurate identification of eukaryotes and prokaryotes in metagenomic data. *Genome Biol* 21:103. <https://doi.org/10.1186/s13059-020-02014-2>.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.

11. Li D, Liu CM, Luo R, Sadakane K, Lam TW. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31:1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>.
12. Pickett BE, Sadat EL, Zhang Y, Noronha JM, Squires RB, Hunt V, Liu M, Kumar S, Zaremba S, Gu Z, Zhou L, Larson CN, Dietrich J, Klem EB, Scheuermann RH. 2012. ViPR: an open bioinformatics database and analysis resource for virology research. *Nucleic Acids Res* 40:D593–D598. <https://doi.org/10.1093/nar/gkr859>.
13. Maes P, Matthijnsens J, Rahman M, Van Ranst M. 2009. RotaC: a Web-based tool for the complete genome classification of group A rotaviruses. *BMC Microbiol* 9:238. <https://doi.org/10.1186/1471-2180-9-238>.
14. Banerjee A, Lo M, Indwar P, Deb AK, Das S, Manna B, Dutta S, Bhadra UK, Bhattacharya M, Okamoto K, Chawla-Sarkar M. 2018. Upsurge and spread of G3 rotaviruses in Eastern India (2014–2016): full genome analyses reveals heterogeneity within Wa-like genomic constellation. *Infect Genet Evol* 63:158–174. <https://doi.org/10.1016/j.meegid.2018.05.026>.