

Title: Genomic characterization of an Echovirus 7 isolate of South-East Asian ancestry recovered from a child with AFP in Nigeria

Authors: ^{1,2}Faleye T.O.C., ¹Adewumi O.M. ^{1,3}Adeniji J.A.,

Affiliation:

1. Department of Virology, College of Medicine, Faculty of Basic Medical Sciences, University of Ibadan, Ibadan, Nigeria
2. Department of Microbiology, Faculty of Science, Ekiti State University, Ado-Ekiti, Nigeria
3. WHO National Polio Laboratory, University of Ibadan, Ibadan, Nigeria

FALEYE Temitope Oluwasegun Cepas:

faleyetemitope@gmail.com, temitope.faleye@eksu.edu.ng

ADEWUMI Olubusuyi Moses: adewumi1@hotmail.com

ADENIJI Johnson Adekunle: adek1808@yahoo.com

ABSTRACT

Here we describe the genome of an Echovirus 7 isolate of South-East Asian ancestry recovered from a child with acute flaccid paralysis (AFP) in Nigeria. The genome has 7,295nt and an open reading frame with 2,195 amino acids.

Keywords: Echovirus 7, AFP, Nigeria, South-East Asia, uORF

Word count: 38 words

Echovirus 7 (E7) belongs to species B within the genus Enterovirus, family *Picornaviridae*, order *Picornavirales* (1) and has been associated with various clinical manifestations (2). In 2014, we (3) detected an E7 from the faeces of a child with AFP in Nigeria. The isolate however, seemed to be of South-East Asian ancestry (3). Here we describe the full genome of the isolate.

After recovery from -20°C storage and passaged twice in RD cell line (4), genomic RNA was extracted using the Total RNA extraction kit (Jena Bioscience, Jena, Germany) and converted to cDNA using the SCRIPT cDNA synthesis kit (Jena Bioscience, Jena, Germany) in accordance with the manufacturer's instructions. Amplification of the genome was done in overlapping fragments as previously described (5), pooled and shipped to a commercial sequencing facility (MR DNA, Texas, USA) where library preparation (using the Nextera DNA sample preparation kit, Illumina) and paired end Illumina sequencing (300 cycles using the HiSeq system) were done. Subsequently, quality check was done using FastQC v1.0.4. and assembly was done using the Kiki assembler v0.0.9.

The E7 genome is 7,295 nucleotides (Nt) (G+C content of 48.1%.) long and was assembled from 3,073,179 (93.87%) of the 3,273,740 reads generated. It has an open reading frame with 2,195 amino acids (Aa). A BLASTn search of the GenBank database, showed the genome to

be most similar to KU355273 (strain 40/Longyou/ZJ) recovered from the cerebro-spinal fluid of a case of viral encephalitis in Zhejiang, China (unpublished data; GenBank). The same done with only the VP1 region showed it to be most similar to JN203720 (strain N-311) recovered from the stool of a child with AFP in India (5). This confirmed our previous finding of its South-East Asian ancestry and association with neurological manifestations (3).

When compared (using MEGA5 (6)) with MH732737 (strain_Ibadan_NGR_2010); an E7 isolate recovered from sewage contaminated water in Nigeria in 2010 (7), the complete ORF, P1 and VP1 were 79.7%, 75.9% and 76.5% similar at the Nt level. At the Aa level, they were 96.6%, 95.2% and 94.7% similar, respectively. When the similarity between the E7 genome described here and MH732737 (strain_Ibadan_NGR_2010) was re-estimated but with the exclusion of the wobble positions (third nucleotide) of all codons, the values for the similarities at the Nt level reduced to 96.8%, 96.1% and 95.2% for the ORF, P1 and VP1 genomic regions, respectively. The similarity between these and the Aa values suggests that the divergence between the similarities at the Nt and Aa level could almost entirely be accounted for by codon bias. As recently described (8), all E7 genomes compared here have a 67Aa uORF in the +1 reading frame relative to the 2,195Aa ORF and overlapping it.

The E7 genome we describe belongs to a lineage associated with neurological manifestations. More genomes from this lineage are needed to figure out their predilection for the CNS.

Accession number

The genome described and raw reads have been deposited in GenBank and SRA under the accession numbers MK159694 and PRJNA505103, respectively.

Word count: 497 words

Funding information

This study was funded by a TETFund grant to AJA.

References

1. <http://www.picornaviridae.com/enterovirus/enterovirus.htm> last accessed 12th November 2018
2. Tapparel C, Siegrist F, Petty TJ, Kaiser L. 2013. Picornavirus and enterovirus diversity with associated human diseases. *Infect Genet Evol.*14:282-93. doi: 10.1016/j.meegid.2012.10.016.
3. Faleye TOC, Adewumi MO, Japhet MO, David OM, Oluyeye AO, Adeniji JA, Famurewa O. 2017. Non-polio enteroviruses in faeces of children diagnosed with acute flaccid paralysis in Nigeria. *BMC Virology Journal*, 14:175, DOI 10.1186/s12985-017-0846-x
4. World Health Organisation, Polio laboratory Manual. Geneva, Switzerland, 4th Edition, 2004.
5. Rao CD, Yergolkar P, Shankarappa KS. 2012. Antigenic diversity of enteroviruses associated with nonpolio acute flaccid paralysis, India, 2007-2009. *Emerg Infect Dis.* 18(11):1833-40. doi: 10.3201/eid1811.111457.
6. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28(10):2731-9.s
7. Faleye T.O.C., Adewumi O.M. Adeniji J.A. 2018. Reference Echovirus 7 and 19 Genomes from Nigeria. *Microbiol Resour Announc* 7:e01465-18. <https://doi.org/10.1128/MRA.01465-18..>
8. Lulla V, Dinan AM, Hosmillo M, Chaudhry Y, Sherry L, Irigoyen N, Nayak KM, Stonehouse NJ, Zilbauer M, Goodfellow I, Firth AE. 2018. An upstream protein-coding region in enteroviruses modulates virus infection in gut epithelial cells. *Nat Microbiol.* doi: 10.1038/s41564-018-0297-1.