# Genomic Characterization of an Echovirus 7 Isolate of Southeast Asian Ancestry Recovered from a Child in Nigeria with Acute Flaccid Paralysis

T. O. C. Faleye,[a,b] O. M. Adewumi,[a] J. A. Adeniji[a,c]

[a]Department of Virology, College of Medicine, Faculty of Basic Medical Sciences, University of Ibadan, Ibadan, Nigeria
[b]Center for Human Virology and Genomics, Microbiology Department, Nigerian Institute of Medical Research, Lagos, Nigeria
[c]WHO National Polio Laboratory, University of Ibadan, Ibadan, Nigeria

**ABSTRACT** Here, we describe the genome of an echovirus 7 (E7) isolate of Southeast Asian ancestry recovered from a child in Nigeria with acute flaccid paralysis (AFP). The genome has 7,295 nucleotides (nt) and an open reading frame (ORF) with 2,195 amino acids.

Echovirus 7 (E7) belongs to species B within the genus *Enterovirus*, family *Picornaviridae*, order *Picornavirales* (http://www.picornaviridae.com/enterovirus/enterovirus.htm), and it has been associated with various clinical manifestations (1). In 2014, we detected an E7 isolate from the feces of a child with acute flaccid paralysis (AFP) in Nigeria (2). The isolate, however, seemed to be of Southeast Asian ancestry (2). Here, we describe the full genome sequence of this isolate.

For this study, the isolate was recovered from storage at −20°C and passaged twice in an RD cell line culture (3). Subsequently, genomic RNA was extracted using a total RNA extraction kit (Jena Bioscience, Jena, Germany) and converted to cDNA using the Script cDNA synthesis kit (Jena Bioscience) in accordance with the manufacturer's instructions. Amplification of the genome was done in overlapping fragments as previously described (4); fragments were pooled and shipped to a commercial sequencing facility (MR DNA, TX, 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.

Quality checking and *de novo* assembly were done using FastQC v1.0.4 and the Kiki assembler v0.0.9, respectively, with default parameters. 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, with a coverage of 62,187×. It has an open reading frame with 2,195 amino acids. A BLASTn search of the GenBank database showed the genome to be most similar (83.64%) to strain 40/Longyou/ZJ (GenBank accession number KU355273), which was recovered from cerebrospinal fluid from a case of viral encephalitis in Zhejiang, China (unpublished data). The same search done with only the VP1 region showed it to be most similar (91.48%) to strain N-311 (GenBank accession number JN203720), which was recovered from the stool of a child with AFP in India (5). This confirmed our previous findings of its Southeast Asian ancestry and its association with neurological manifestations (2).

When an E7 isolate recovered from sewage-contaminated water in Nigeria in 2010 (strain_Ibadan_NGR_2010, GenBank accession number MH732737) (2) was compared using MEGA5 (6) with the strain recovered in this study, the complete open reading frame (ORF), P1, and VP1 were 79.7%, 75.9%, and 76.5% similar at the nucleotide level, respectively. At the amino acid level, they were 96.6%, 95.2%, and 94.7% similar, respectively. When the similarities between the E7 genome described here and strain_

Address correspondence to O. M. Adewumi, adewumi1@hotmail.com.

Ibadan_NGR_2010 (MH732737) were reestimated with the exclusion of the wobble positions (third nucleotide) of all codons, the values for the similarities at the nucleotide level were reduced to 96.8%, 96.1%, and 95.2% for the ORF, P1, and VP1 genomic regions, respectively. The similarities between these and the amino acid values suggest that the divergence between the similarities at the nucleotide and amino acid levels could be almost entirely accounted for by codon bias.

We describe here an E7 genome, recovered from a child in Nigeria with AFP, that is of Southeast Asian ancestry and belongs to a lineage associated with neurological manifestations. More genomes from this lineage might provide the information needed to figure out the basis of their predilection for the central nervous system (CNS).

**Data availability.** The genome described has been deposited in GenBank and SRA under the accession numbers MK159694 and PRJNA505103, respectively.

## REFERENCES

1. Tapparel C, Siegrist F, Petty TJ, Kaiser L. 2013. Picornavirus and enterovirus diversity with associated human diseases. Infect Genet Evol 14:282–293. https://doi.org/10.1016/j.meegid.2012.10.016.
2. Faleye TOC, Adewumi MO, Japhet MO, David OM, Oluyege AO, Adeniji JA, Famurewa O. 2017. Non-polio enteroviruses in faeces of children diagnosed with acute flaccid paralysis in Nigeria. Virol J 14:175. https://doi.org/10.1186/s12985-017-0846-x.
3. World Health Organization. 2004. Polio laboratory manual, 4th ed. World Health Organization, Geneva, Switzerland.
4. Faleye TOC, Adewumi OM, Adeniji JA. 2018. Reference echovirus 7 and 19 genomes from Nigeria. Microbiol Resour Announc 7:e01465-18. https://doi.org/10.1128/MRA.01465-18.
5. Rao CD, Yergolkar P, Shankarappa KS. 2012. Antigenic diversity of enteroviruses associated with nonpolio acute flaccid paralysis, India, 2007–2009. Emerg Infect Dis 18:1833–1840. https://doi.org/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. Mol Biol Evol 28:2731–2739. https://doi.org/10.1093/molbev/msr121.