Whole-Genome Sequences of Enteroviruses D94 and D111 Isolated from Stool Specimens in Angola

Shur-Wern W. Chern,[a] Nicksy Gumede,[b] Christina J. Castro,[c] W. Allan Nix,[a] Terry Fei Fan Ng[a]

aDivision of Viral Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, USA
bWorld Health Organization Regional Office for Africa, Brazzaville, Republic of Congo
cOak Ridge Institute for Science and Education, Oak Ridge, Tennessee, USA

ABSTRACT We report the whole-genome sequences of new enterovirus D94 and D111 strains, isolated from cultures from stool specimens collected from acute flaccid paralysis (AFP) cases for poliovirus surveillance in Angola during 2010.

The genus *Enterovirus* contains 15 species within *Picornaviridae*, a large family of small nonenveloped positive-sense, single-stranded RNA (~7.5-kb) viruses. Five enterovirus (EV) species D types have been described, including EV-D68, EV-D70, EV-D90, EV-D111, and EV-D120 (1). EV-D94 and EV-D111 were originally identified in the framework of poliovirus-focused acute flaccid paralysis (AFP) surveillance in Africa. Isolation and characterization of EV-D94 have been described for human stool samples collected in the Democratic Republic of the Congo (2) and from sewage specimens collected in Egypt (3). EV-D111 was isolated from human stool samples collected in Central African Republic and in Cameroon (4).

We report the whole-genome sequences of new EV-D94 and EV-D111 strains, isolated in rhabdomyosarcoma (RD) cells from stool samples collected during AFP surveillance in Angola in 2010. Stool suspensions were prepared by adding 5 ml minimal essential medium (MEM), 0.5 g glass beads, and 0.5 ml chloroform to 1 g stool sample. The mixture was mixed vigorously for 30 min in a shaker and centrifuged at 3,000 rpm for 30 min at 4°C (5). RD cells were inoculated with 50 μl of the stool suspension and incubated in 5% CO_2 at 37°C. Cell cultures were observed and harvested when they produced a cytopathic effect (5). Viral RNA was extracted from cell culture supernatants, using the QIAamp viral RNA minikit with on-column DNase digestion (Qiagen). Double-stranded cDNA was generated using the qScript Flex cDNA synthesis kit (Quantabio), using random hexamers and oligo(dT). Illumina libraries were prepared using the Nextera XT library preparation kit and sequenced on an Illumina MiSeq instrument using a 500-cycle paired-end run with multiplexing as previously described (6). Over 350,000 reads were generated per sample.

The raw data were processed through an in-house bioinformatics pipeline (6). Briefly, human reads were removed from raw FASTQ data using Bowtie 2 v2.3.3.1 (7), followed by primer trimming, adapter trimming, and Phred quality score filtering (removing reads with a score of <20) using Cutadapt v1.8.3 (8). Read deduplication was performed using the Python script Dedup.py (6), and deduplicated reads were *de novo* assembled into contigs using SPAdes v3.7.0 (9) with default k-mers. Read mapping, alignment, and sequence identity comparison with prototype reference, as well as contig annotation, were performed using Geneious vR11 (Biomatters). All tools were run with default parameters unless otherwise specified. The EV-D94 genome of 7,402 bases has a GC content of 43%. The EV-D111 of 7,403 bases has a GC content of 43%. The genome is complete as inferred by alignment and 5′ and 3′ sequence conservation of the enterovirus D species. The average read coverage for both viruses was >500-fold.

The EV-D94 genome (ANG/2010-23293) shared 85% nucleotide identity (NI) over the genome and 97% amino acid identity (AI) over the coding region compared to the prototype strain E210 (Egyptian sewage [5]). The sequences had 85%, 85%, 84%, and 87% NI and 98%, 98%, 95%, and 98% amino acid identity over the VP1, P1, P2, and P3 regions, respectively.

The EV-D111 genome (ANG/2010-23294) shared 87% NI over the genome and 97% AI over the coding region compared to strain CMR-TOK-230 (GenBank accession number MK032898) (prototype KK2640 only contained fragmented sequences). The sequences shared 88%, 87%, 87%, and 89% NI, as well as 97%, 97%, 97%, and 98% AI over the VP1, P1, P2, and P3 regions, respectively. The study strain represents the only complete genome sequence for EV-D111 in GenBank.

Enterovirus D genomes other than EV-D68 are rarely reported; these isolates and sequences will facilitate future virologic studies on enterovirus D.

Data availability. The EV-D94 and EV-D111 genome sequences have been deposited in GenBank under the accession numbers MT081370 to MT081371. The postprocessed fastq reads have been deposited in the Sequence Read Archive under the accession numbers SRR13403396 and SRR13403770.

REFERENCES

1. Walker PJ, Siddell SG, Lefkowitz EJ, Mushegian AR, Dempsey DM, Dutilh BE, Harrach B, Harrison RL, Hendrickson RC, Junglen S, Knowles NJ, Kropinski AM, Krupovic M, Kuhn JH, Nibert M, Rubino L, Sabanadzovic S, Simmonds P, Varsani A, Zerbini FM, Davison AJ. 2019. Changes to virus taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2019). Arch Virol 164:2417–2429. https://doi.org/10.1007/s00705-019-04306-w.
2. Junttila N, Lévêque N, Kabue JP, Cartet G, Mushiya F, Muyembe-Tamfum J-J, Trompette A, Lina B, Magnius LO, Chomel J-J, Norder H. 2007. New enteroviruses, EV-93 and EV-94, associated with acute flaccid paralysis in the Democratic Republic of the Congo. J Med Virol 79:393–400. https://doi.org/10.1002/jmv.20825.
3. Smura TP, Junttila N, Blomqvist S, Norder H, Kaijalainen S, Paananen A, Magnius LO, Hovi T, Roivainen M. 2007. Enterovirus 94, a proposed new serotype in human enterovirus species D. J Gen Virol 88:849–858. https://doi.org/10.1099/vir.0.82510-0.
4. Sadeuh-Mba SA, Joffret M-L, Mazitchi A, Endegue-Zanga M-C, Njouom R, Delpeyroux F, Gouandjika-Vasilache I, Bessaud M. 2019. Genetic and phenotypic characterization of recently discovered enterovirus D type 111.
PLoS Negl Trop Dis 13:e0007797. https://doi.org/10.1371/journal.pntd.0007797.
5. WHO. 2015. The 21st Informal Consultation on the Global Polio Laboratory Network, final report of recommendations.
6. Montmayeur AM, Ng TFF, Schmidt A, Zhao K, Magaña L, Iber J, Castro CJ, Chen Q, Henderson E, Ramos E, Shaw J, Tatusov RL, Dybdahl-Sissoko N, Endegue-Zanga MC, Adeniji JA, Oberste MS, Burns CC. 2017. High-throughput next-generation sequencing of polioviruses. J Clin Microbiol 55:606–615. https://doi.org/10.1128/JCM.02121-16.
7. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.
8. Kechin A, Boyarskikh U, Kel A, Filipenko M. 2017. cutPrimers: a new tool for accurate cutting of primers from reads of targeted next generation sequencing. J Comput Biol 24:1138–1143. https://doi.org/10.1089/cmb.2017.0096.
9. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.