



Complete Genome Sequence of a New Ebola Virus Strain Isolated during the 2017 Likati Outbreak in the Democratic Republic of the Congo

Tony Wawina-Bokalanga,^a Bert Vanmechelen,^a Joan Martí-Carreras,^a Valentijn Vergote,^a Kurt Vermeire,^b Jean-Jacques Muyembe-Tamfum,^c Steve Ahuka-Mundeke,^c  Piet Maes^a

^aKU Leuven, Department of Microbiology, Immunology and Transplantation, Clinical and Epidemiological Virology Division, Rega Institute, Leuven, Belgium

^bKU Leuven, Department of Microbiology, Immunology and Transplantation, Laboratory of Virology and Chemotherapy, Rega Institute, Leuven, Belgium

^cInstitut National de Recherche Biomédicale (INRB), Kinshasa, Democratic Republic of the Congo

ABSTRACT Genomic sequencing for early identification of Ebola virus remains a big challenge in low-income countries. Here, we report the complete genome sequence of an Ebola virus strain obtained during the 2017 Likati outbreak in the Democratic Republic of the Congo (DRC) by using the Oxford Nanopore Technologies (ONT) MinION sequencer.

Ebola virus is a filamentous, enveloped, nonsegmented negative-sense RNA virus that belongs to the genus *Ebolavirus*, family *Filoviridae*. This genus consists of five species, *Zaire ebolavirus* (EBOV), *Sudan ebolavirus* (SUDV), *Bundibugyo ebolavirus* (BDBV), *Tai Forest ebolavirus* (TAFV), and *Reston ebolavirus* (RESTV) (1).

Our research team was deployed to the Democratic Republic of the Congo (DRC) in May 2018 to perform real-time sequencing using the Oxford Nanopore Technologies (ONT) MinION sequencer, supporting local diagnostic teams during the Ebola outbreak. Prior to using the MinION sequencer in Equateur Province, where the outbreak was ongoing, the protocol was validated by sequencing a confirmed Ebola-positive blood sample from the May 2017 Likati outbreak at the Institut National de Recherche Biomédicale (INRB), Kinshasa, Democratic Republic of the Congo.

Of the eight reported cases in the 2017 Likati outbreak, four were fatal, but only two samples have been confirmed to be positive for EBOV (<https://www.cdc.gov/vhf/ebola/history/chronology.html>). Here, we present the complete genome sequence of the EBOV strain found in one of these two samples, obtained using the MinION sequencer. The other positive sample was deemed inadequate for sequencing due to its low viral load and suboptimal storage conditions.

Viral RNA was extracted from 50 μ l of inactivated whole blood using a viral RNA extraction kit (Qiagen) according to the manufacturer's instructions but with an extra washing step (buffer AW2). Reverse transcription-PCR (RT-PCR) was performed with each of the EBOV-specific primer pools kindly provided by the ARTIC project (2) to amplify the EBOV genome. Pooled amplicons were cleaned up with AMPure XP beads (New England Biolabs) by washing 2 times with 70% ethanol and resuspending in 50 μ l of RNase-free water. The purified DNA was quantified on a Qubit 1.0 fluorimeter (Thermo Fisher Scientific), and libraries were prepared using the 1D genomic DNA by ligation (SQK-LSK108) kit and the protocol supplied by ONT. Sequencing was performed using a R9.4.1 flow cell and MinKNOW v2.0 software.

Reads were base called with ONT Albacore v3.0.1 and subsequently quality, adapter, and primer trimmed with Porechop v0.2.3 using default parameters (3). To construct the consensus sequence, *de novo* assembly was performed using Canu v1.7.0 (4), and reference assembly was done with Minimap2 (5) using default parameters. The result-

Citation Wawina-Bokalanga T, Vanmechelen B, Martí-Carreras J, Vergote V, Vermeire K, Muyembe-Tamfum J-J, Ahuka-Mundeke S, Maes P. 2019. Complete genome sequence of a new Ebola virus strain isolated during the 2017 Likati outbreak in the Democratic Republic of the Congo. *Microbiol Resour Announc* 8:e00360-19. <https://doi.org/10.1128/MRA.00360-19>.

Editor Simon Roux, DOE Joint Genome Institute

Copyright © 2019 Wawina-Bokalanga et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Piet Maes, piet.maes@kuleuven.be.

Received 28 March 2019

Accepted 13 April 2019

Published 16 May 2019

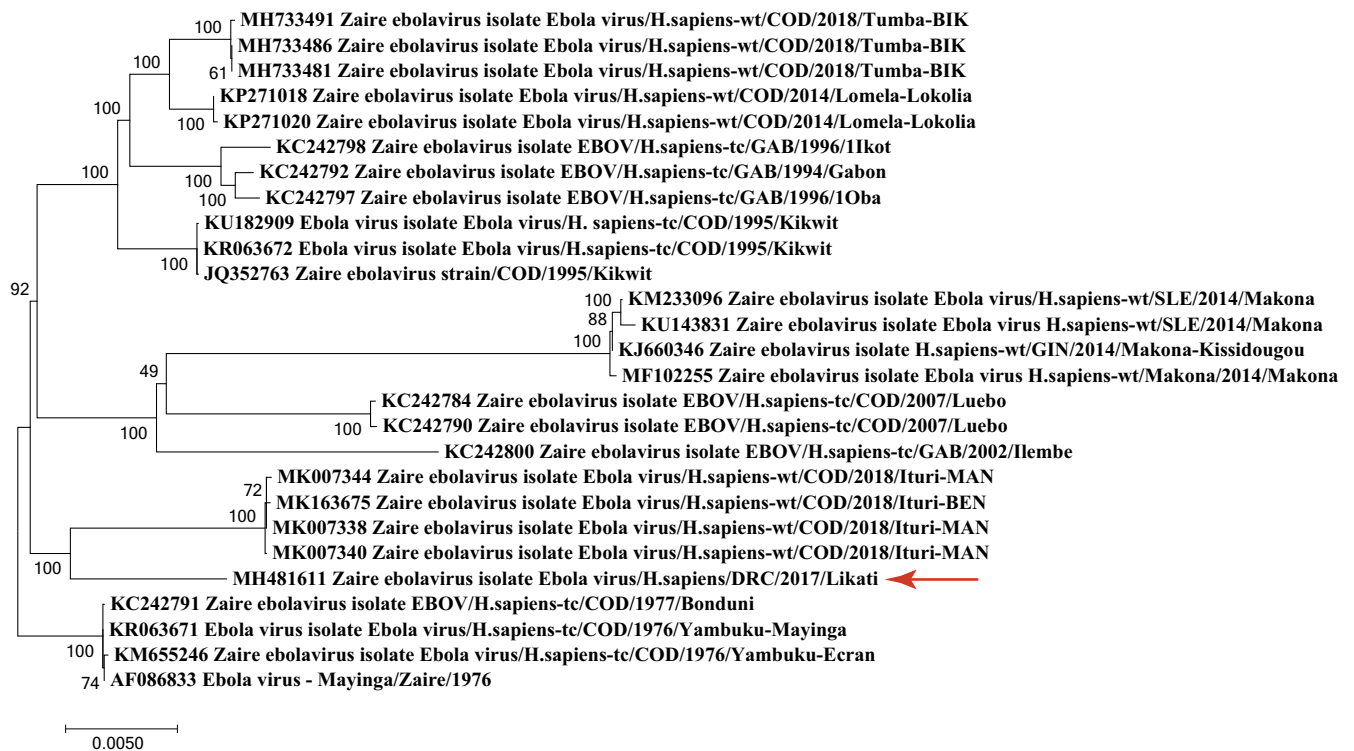


FIG 1 Phylogenetic maximum-likelihood tree based on the Hasegawa-Kishino-Yano model, using a bootstrap value of 500, of the 2017 Likati EBOV sequence (red arrow) compared to that of other full-length EBOV genomes from different Ebola virus outbreaks.

ing contigs consist of 896,252 reads with an average depth coverage of 19,604 \times and a total length of 485,957,350 bp. Reads were mapped against both draft assemblies and posteriorly polished with Nanopolish v0.9.2 (6). BLAST was used to assess the completeness of the draft consensus sequence using the previous EBOV strains (7).

This complete genome is 18,898 nucleotides in length with a GC content of 41.12%, and it consists of a 3' leader sequence, seven genes (encoding nucleoprotein, viral proteins VP24, VP30, VP35, and VP40, glycoprotein, and polymerase), and the 5' trailer sequence.

A phylogenetic tree (Fig. 1) was constructed using a multiple-sequence alignment of full-length EBOV genomes by MUSCLE in MEGA v7.0.21 (8), using default parameters. There is a nucleotide similarity of 98.69% between the newly reported sequence and the most closely related genome, EBOV/COD/1976/Yambuku-Mayinga (GenBank accession number [KR063671](#)).

Data availability. The complete genome sequence of the 2017 Likati EBOV strain has been deposited in GenBank under the accession number [MH481611](#). Raw data were submitted to the National Center for Biotechnology Information Sequence Read Archive (SRA) under the accession number [SRR8959866](#).

ACKNOWLEDGMENTS

This work was supported by grants from the Special Research Fund, KU Leuven (Bijzonder Onderzoeksfonds, KU Leuven) 3M170314 C14/17/100. B.V. is supported by an SB Ph.D. fellowship at FWO (Research Foundation–Flanders), project number 1528617N. J.M.-C. is supported by a doctoral grant from the HONOURS Marie-Sklodowska-Curie training network (721367).

We are grateful to the ARTIC project (Wellcome Trust Collaborative Award project number 206298/B/17/Z) for the provision of the primer pools used for viral genome sequencing in this project.

REFERENCES

1. Bukreyev AA, Chandran K, Dolnik O, Dye JM, Ebihara H, Leroy EM, Muhlberger E, Netesov SV, Patterson JL, Paweska JT, Saphire EO, Smither SJ, Takada A, Towner JS, Volchkov VE, Warren TK, Kuhn JH. 2014. Discussions and decisions of the 2012–2014 International Committee on Taxonomy of Viruses (ICTV) Filoviridae Study Group, January 2012–June 2013. *Arch Virol* 159:821–830. <https://doi.org/10.1007/s00705-013-1846-9>.
2. Meredith L. 2018. Ebola virus sequencing kit-list. <http://artic.network/ebov/ebov-seq-kit.html>.
3. Wick R. 2017. Porechop: a tool for finding and removing adapters from Oxford Nanopore reads. <https://github.com/rwick/Porechop>.
4. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
5. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
6. Simpson J. 2015. Nanopolish: software package for signal-level analysis of Oxford Nanopore sequencing data. <https://github.com/jts/nanopolish>.
7. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
8. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874. <https://doi.org/10.1093/molbev/msw054>.