Nearly Complete Genome Sequences of Eight Rabies Virus Strains Obtained from Domestic Carnivores in the Democratic Republic of the Congo

Patient Pati Pyana,[a,b] Céline Mbilo,[c,d] Julien Lannoy,[e] Simon Bonas,[e] Bobo Luntadila,[a] Jean Baptiste Kabongo,[b] Ipos Ngayi Lukusa,[a] Leonard Ntunuanga,[f] Jakob Zinsstag,[c,d] Stephanie Mauti,[e] ⓘ Laurent Dacheux[e]

[a]Institut National de Recherche Biomédicale (INRB), Laboratoire de la Rage, Gombe, Kinshasa, Democratic Republic of the Congo

[b]Université Pédagogique Nationale de Kinshasa, Kinshasa, Democratic Republic of the Congo

[c]Swiss Tropical and Public Health Institute, Basel, Switzerland

[d]University of Basel, Basel, Switzerland

[e]Institut Pasteur, Université de Paris, Unit Lyssavirus Epidemiology and Neuropathology, National Reference Center for Rabies, WHO Collaborating Center for Reference and Research on Rabies, Paris, France

[f]Division Provinciale de Pêche et Elevage, Matadi, Province du Kongo Central, Democratic Republic of the Congo

ABSTRACT In this report, we describe eight nearly complete genome sequences of rabies virus strains collected in the Democratic Republic of the Congo from domestic carnivores in 2017 and 2018. All of them clustered into a specific phylogroup among the Africa 1b lineage in the Cosmopolitan clade.

Rabies is the prototype of a neglected and tropical zoonotic disease, affecting poor and rural areas in Asia and Africa. To date, nearly 59,000 human cases of rabies are estimated worldwide each year, mainly due to transmission from dogs (1). Rabies virus (RABV) is the principal etiological agent of rabies, an acute and almost always fatal form of encephalomyelitis which can affect potentially all mammalian species. This virus belongs to the prototype species *Rabies lyssavirus* within the genus *Lyssavirus*, family *Rhabdoviridae* (order *Mononegavirales*) (2). Dog rabies is endemic in the Democratic Republic of the Congo (DRC), similarly to other sub-Saharan countries (3, 4). However, data available about the genetic diversity of RABV strains circulating in this country still remain extremely limited (5).

In this study, brain samples were collected from seven dogs and one cat suspected of being infected with rabies that originated from different health sanitary zones of Kongo Central Province in the DRC in 2017 and 2018 (Table 1). All these samples were confirmed positive by direct fluorescence antibody test (FAT) (6) at the Institut National de Recherche Biomédicale (INRB) in Kinshasa. Total RNA was extracted from one brain biopsy sample (approximatively 0.5 cm³) from each animal using the Direct-zol RNA miniprep kit (Zymo Research), following the manufacturer's instructions and performed in the rabies laboratory of INRB in Kinshasa. The RNA was then purified using Agencourt RNAClean XP beads (Beckman Coulter) at a ratio of 1:1.8, following the manufacturer's instructions, without the last resuspension step in nuclease-free water for half of these samples (Table 1). For the other half, RNA was eluted in 30 to 50 μL of nuclease-free water, and 20 μL was deposited in a 96-well plate (RNAstable; Biomatrica), before overnight air-drying in a laminar flow hood, following the manufacturer's instructions (Table 1). Dried RNAs in an RNAstable 96-well plate or on beads were shipped to Institut Pasteur, Paris, France, at ambient and cold temperature with ice packs, respectively, and resuspended in 30 μL nuclease-free water. The eight RNA samples were processed for next-generation sequencing (NGS) as previously described (7–9). Briefly, an rRNA depletion step was first carried out using Terminator 5'-phosphate-dependent exonuclease (Epicentre Biotechnologies), following the instructions of the manufacturer. After purification, the depleted RNA was reverse transcribed into cDNA

TABLE 1 Description of the genome sequences of the eight rabies viruses obtained from domestic carnivores in Kongo Central Province of the Democratic Republic of the Congo

Virus name	Code	Host	Animal status	Location	Collection date (day-mo-yr)	Support[a]	Total no. of reads	No. of mapped reads (%)	Avg coverage (×)	Genome length (nt)	GC content (%)	GenBank accession no.	SRA accession no.
18033RDC	1Cn3	Dog	Owned	Matadi	28-Oct-17	Plate	2,221,528	92,167 (4.1)	1,141.6	11,923	46	OK317992	SRX12433974
18065RDC	7Cn41	Dog	Owned	Boma	2-Jul-18	Beads	4,028,176	8,353 (0.2)	103.5	11,868[b]	46	OK317993	SRX12433975
18026RDC	7Cn68	Dog	Owned	Boma	16-Sep-18	Beads	3,724,300	92,361 (2.5)	1,148.4	11,899[b]	46	OK317994	SRX12433976
18093RDC	9Cn11	Dog	Unowned	Muanda	25-Nov-17	Beads	4,734,998	2,414,598 (51)	29,967.5	11,923	46	OK317995	SRX12433977
18101RDC	9Cn22	Dog	Unowned	Muanda	27-Jun-18	Plate	1,515,480	10,631 (0.7)	132.2	11,882[b]	46	OK317996	SRX12433978
18111RDC	9Ct9	Cat	Owned	Muanda	4-Nov-17	Plate	3,635,804	20,729 (0.6)	257.8	11,877[b]	46	OK317997	SRX12433979
18037RDC	11Cn1	Dog	Unowned	Mbanza-Ngungu	29-Nov-17	Plate	2,332,676	45,426 (1.9)	566	11,886[b]	46	OK317998	SRX12433980
18114RDC	KN	Dog	Owned	Kinshasa (Ngaliema)	2-Feb-18	Beads	2,561,938	85,775 (3.3)	1,066.9	11,923	46	OK317999	SRX12433981

[a] Dry RNA was stored and shipped in plates (RNAstable 96-well plate) or on beads (Agencourt RNAClean XP).
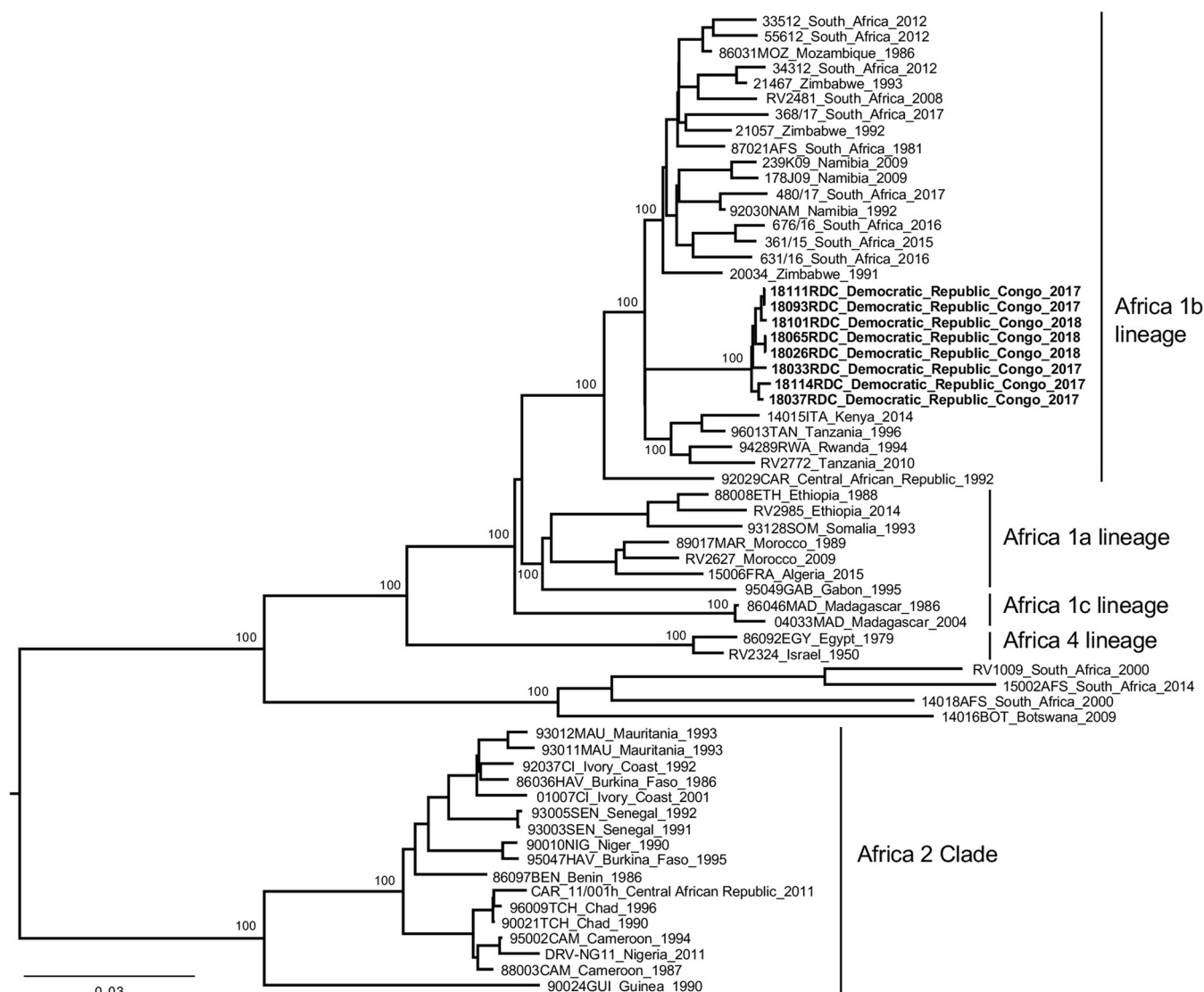[b] Genome sequence with incomplete 5′ untranslated region (UTR).

FIG 1 Phylogenetic analysis of the eight RABV strains from the Democratic Republic of the Congo and different representative African strains. The tree was based on the nearly complete genome sequences (11,786 to 11,807 nt) and constructed using the maximum-likelihood approach based on the generalized time-reversible model proportion of invariable sites plus the gamma-distributed rate heterogeneity (GTR + I + Γ4), utilizing subtree pruning and regrafting (SPR) of branches, as estimated using PhyML 3.0 (12) with Smart Model Selection (http://www.atgc-montpellier.fr/phyml-sms/). The robustness of individual nodes was estimated using 100 bootstrap replicates. Only bootstrap values of ≥90 are indicated. The scale bar indicates nucleotide substitutions per site.

using random primers and Superscript III reverse transcriptase (Invitrogen), according to the manufacturer's instructions, and double-stranded DNA (dsDNA) was synthesized as already described (7–9). Finally, dsDNA libraries were constructed using the Nextera XT kit (Illumina) and sequenced using a 2 × 150-nucleotide (nt) paired-end strategy on the NextSeq 500 platform (7–9). NGS data were analyzed using *de novo* assembly and mapping (both using CLC Assembly Cell, Qiagen), with a dedicated workflow built on the Institut Pasteur Galaxy platform (7–10). Contig sequences were assembled to produce the final consensus genome using Sequencher 5.2.4 (Gene Codes Corporation). The quality and accuracy of the final genome sequences were checked after a final mapping step of the original cleaned reads and visualized using Tablet (11). Maximum likelihood (ML) phylogenetic analysis was performed on the nearly complete genome sequences (11,786 to 11,807 nt) of the eight RABV strains and different representative African strains using PhyML (12), after a multiple alignment step performed using ClustalW 2.1 (13), implemented in the Institut Pasteur Galaxy platform (10). The ML phylogenetic tree was visualized using FigTree (http://tree.bio.ed.ac.uk/) (Fig. 1). All tools were run with default parameters unless otherwise specified.

The genome sequences presented the five canonical genes encoding the nucleoprotein (N; 1,353 nt, 450 amino acids [aa]), phosphoprotein (P; 894 nt, 297 aa), matrix protein (M; 609 nt, 202 aa), glycoprotein (G; 1,575 nt, 524 aa), and RNA polymerase (L; 6,384 nt, 2,127 aa) (Table 1). The leader and trailer sequences, when complete, were 58 and 70 nucleotides long, respectively (checking done after alignment with genetically close and available complete genomes [Fig. 1]) (Table 1). The transcription initiation (TI) signal AACA and the transcription termination polyadenylation (TTP) sequences TGA_7 were observed for all the genes, except for the G gene, which presented the AGA_7 motif for TTP. The nucleotide identity between the eight genome sequences, determined using Ident and Sim software implemented in the Sequence Manipulation Suite (https://www.bioinformatics.org/sms2/ident_sim.html) (14), was high (>98.9%), and genetic analysis confirmed that they clustered together in lineage Africa 1b within the Cosmopolitan clade (Fig. 1) (15).

Data availability. The nearly complete genome sequences of the eight rabies viruses from the Democratic Republic of the Congo were deposited at GenBank under the accession numbers OK317992 to OK317999 and the BioProject accession number PRJNA767799.

REFERENCES

1. Hampson K, Coudeville L, Lembo T, Sambo M, Kieffer A, Attlan M, Barrat J, Blanton JD, Briggs DJ, Cleaveland S, Costa P, Freuling CM, Hiby E, Knopf L, Leanes F, Meslin F-X, Metlin A, Miranda ME, Müller T, Nel LH, Recuenco S, Rupprecht CE, Schumacher C, Taylor L, Vigilato MAN, Zinsstag J, Dushoff J, Global Alliance for Rabies Control Partners for Rabies Prevention. 2015. Estimating the global burden of endemic canine rabies. PLoS Negl Trop Dis 9:e0003709. https://doi.org/10.1371/journal.pntd.0003709.

2. Walker PJ, Blasdell KR, Calisher CH, Dietzgen RG, Kondo H, Kurath G, Longdon B, Stone DM, Tesh RB, Tordo N, Vasilakis N, Whitfield AE, ICTV Report Consortium. 2018. ICTV virus taxonomy profile: Rhabdoviridae. J Gen Virol 99:447–448. https://doi.org/10.1099/jgv.0.001020.

3. Mbilo C, Coetzer A, Bonfoh B, Angot A, Bebay C, Cassamá B, De Benedictis P, Ebou MH, Gnanvi C, Kallo V, Lokossou RH, Manjuba C, Mokondjimobe E, Mouillé B, Mounkaila M, Ndour APN, Nel L, Olugasa BO, Pato P, Pyana PP, Rerambyath GA, Roamba RC, Sadeuh-Mba SA, Suluku R, Suu-Ire RD, Tejiokem MC, Tetchi M, Tiembre I, Traoré A, Voupawoe G, Zinsstag J. 2021. Dog rabies control in West and Central Africa: a review. Acta Trop 224:105459. https://doi.org/10.1016/j.actatropica.2020.105459.

4. Mbilo C, Kabongo J-B, Pyana PP, Nlonda L, Nzita RW, Luntadila B, Badibanga B, Hattendorf J, Zinsstag J. 2019. Dog ecology, bite incidence, and disease awareness: a cross-sectional survey among a rabies-affected community in the Democratic Republic of the Congo. Vaccines 7:98. https://doi.org/10.3390/vaccines7030098.

5. Mbuyi GT, Kawaya EK, Twabela AT, Cattoli G, Walandila JS, Naletoski I, Masumu J, Dundon WG. 2020. Molecular characterization of rabies viruses from two western provinces of the Democratic Republic of the Congo (2008–2017). Virus Genes 56:651–656. https://doi.org/10.1007/s11262-020-01784-y.

6. WHO. 2018. The direct fluorescent antibody test, p 108–129. In Rupprecht C, Fooks AR, Abel-Ridder B (ed), Laboratory techniques in rabies, vol 1. World Health Organization, Geneva, Switzerland.

7. Luo D-S, Li B, Shen X-R, Jiang R-D, Zhu Y, Wu J, Fan Y, Bourhy H, Hu B, Ge X-Y, Shi Z-L, Dacheux L. 2021. Characterization of novel rhabdoviruses in Chinese bats. Viruses 13:64. https://doi.org/10.3390/v13010064.

8. Pallandre L, Luo D, Feuvrier C, Lieffrig F, Pozet F, Dacheux L, Bigarré L. 2020. Revisiting the classification of percid perhabdoviruses using new full-length genomes. Viruses 12:649. https://doi.org/10.3390/v12060649.

9. Dacheux L, Dommergues L, Chouanibou Y, Doméon L, Schuler C, Bonas S, Luo D, Maufrais C, Cetre-Sossah C, Cardinale E, Bourhy H, Métras R. 2019. Co-circulation and characterization of novel African arboviruses (genus Ephemerovirus) in cattle, Mayotte Island, Indian Ocean, 2017. Transbound Emerg Dis 66:2601–2604. https://doi.org/10.1111/tbed.13323.

10. Mareuil F, Doppelt-Azeroual O, Ménager H. 2017. A public Galaxy platform at Pasteur used as an execution engine for Web services. Galaxy Commun Conf 2017. https://f1000research.com/posters/6-1030.

11. Milne I, Stephen G, Bayer M, Cock PJ, Pritchard L, Cardle L, Shaw PD, Marshall D. 2013. Using Tablet for visual exploration of second-generation sequencing data. Brief Bioinform 14:193–202. https://doi.org/10.1093/bib/bbs012.

12. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321. https://doi.org/10.1093/sysbio/syq010.

13. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948. https://doi.org/10.1093/bioinformatics/btm404.

14. Stothard P. 2000. The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28:1102–1104. https://doi.org/10.2144/00286ir01.

15. Troupin C, Dacheux L, Tanguy M, Sabeta C, Blanc H, Bouchier C, Vignuzzi M, Duchene S, Holmes EC, Bourhy H. 2016. Large-scale phylogenomic analysis reveals the complex evolutionary history of rabies virus in multiple carnivore hosts. PLoS Pathog 12:e1006041. https://doi.org/10.1371/journal.ppat.1006041.