**Supplementary File 1: Rabies diagnosis and sequencing**

Post-mortem brain samples were collected by pathologists at the hospitals where the deceased were admitted or attended to prior to death. The brainstem and medulla samples from case 1 were collected two days after death whereas samples from all other cases were collected the same day.  Testing and sequencing was conducted at the Kilimanjaro Clinical Research Institute (KCRI), Tanzania, and at the University of Nairobi Institute of Tropical and Infectious Diseases (UNITID), Kenya. All samples were tested with rapid diagnostic tests (Antigen Rapid Rabies Ag Test Kit, BioNote) following the manufacturer’s instructions [1]. Sample storage prior to testing and sequencing differed. Samples from case 1 and 2 were stored for over a year at UNITID at -800C and samples from case 3 were stored for one month at KCRI at -800C before testing and sequencing. Samples from cases 4 and 5 were both tested on the same day as collection, then stored for three months at UNITID at -800C and for two weeks at KCRI at -800C respectively, prior to sequencing.

Additionally, a brain specimen from case 3 was fixed in formalin for immunohistochemistry [2] at the Department of Viroscience, Rotterdam, the Netherlands. The sample was first embedded in paraffin, cut in 3µM slices, stained with antibody against the nucleoprotein (RABV-N, antibody 5DF12) then with peroxidase-labelled goat-anti-mouse IgG1. Peroxidase was revealed with aminoethyl carbazole, resulting in a red precipitate. Slides were counterstained with hematoxylin. Isotype controls (Streptavidin-biotin complex) were included to assess specific staining [3].

Amplicon-based sequencing was carried out following a previously established protocol [4]. Brain tissue samples were inactivated and prepared for RNA extraction by homogenisation of a small cube of tissue (~ 30 mg) in RNA shield (Zymo Research, USA) in 2 ml reinforced tubes containing 1.4mm ceramic beads, using a TerralyzerTM (Zymo Research, USA). Total RNA was extracted using the Quick Viral RNA Mini-prep Kit (Zymo Research, USA) and cDNA synthesised by reaction of LunaScript RT Supermix (New England Biolabs, UK), nuclease free water and viral RNA. Thermocycling comprised primer annealing at 25°C for 2 minutes, cDNA synthesis at 55°C for 10 minutes and heat inactivation at 95°C for 1 minute before holding at 4°C.

To enrich the RABV DNA, multiplex PCR was conducted following an established method applied to rabies virus [5,6]. Primal scheme was used to design primers specific for East Africa [7]. Sets of previously designed multiplex primer sets used in Tanzania from 2019-2020 [8] and updated to include the greater RABV diversity in Kenya (used from 2020-2023,

https://github.com/kirstyn/artic-rabv/tree/master/rampart/tza\_protocol/rabv\_ea), were used in PCR reactions. A characteristic of the multiplex PCR approach is that primers are combined into two pools (1 and 2), enabling tens of PCR reactions in 2 individual tubes per sample. Each multiplex PCR reaction mix consisted of Q5 Hot start High Fidelity 2x Master Mix (New England Biolabs), primer pool 1/2, water and cDNA. The PCR was performed under the following conditions: 30 seconds heat activation at 98°C; 35 cycles of denaturation (15 seconds at 98°C); annealing and extension (5 minutes at 65°C); and finally, a hold temperature of 4°C. Before sequencing the DNA concentration in the samples and negative control were checked by Qubit fluorometer (version 3, Thermo Fisher Scientific, USA) using a dsDNA Quantitation high sensitivity kit (Thermo Fisher Scientific, USA). Sequencing was undertaken using a Nanopore MinION MK1B, with a ligation-based library preparation kit (SQK-LSK109) and the final library loaded on an R9.4.1 flow cell.  Native barcoding was used to differentiate samples and a negative control of nuclease-free water was included.  In total, four runs were done for the five cases; the first run included cases 1 and 2, and the remaining three runs were for cases 3, 4, and 5. Other samples from different hosts (domestic dogs and various livestock) were included in each run, generating 99 new sequences (supplementary Table 1).

A RABV-tailored version of the ARTIC network’s bioinformatics pipeline [6] was applied to process raw MinION reads, comprising basecalling (when MinKNOW live basecalling was not applied during the run) and barcode demultiplexing using Guppy software (Version 6.3.2). Reads were mapped to a reference sequence from East Africa (GenBank accession number KF155002) using Minimap with alignment post-processing to normalise reads and trim primer sequences, followed by variant calling and generation of a draft consensus sequence using Medaka [9]. ARTIC pipeline scripts were used to filter variants based on quality, and a minimum depth threshold of 20 reads. Genome coverage was calculated using 20 depth coverage (supplementary Table 2). Sequences were checked for sufficient quality (i.e. less than 50% ambiguous bases) and metadata compiled into a TSV file. SQN files were prepared for email submission to GenBank using an as-yet unreleased GLUE-based submission tool, which uses GLUE’s multiple sequence alignment capabilities [10] to annotate sequences and tbl2asn [11] to generate SQN files within a Docker container. The resulting SQN files were checked for internal stop codons before submission. Consensus FASTA files are available on NCBI GenBank (accession numbers OR045927 to OR920351).

RABV-GLUE was used for major and minor clade assignment [10], followed by more resolved lineage designation and assignment using MADDOG [12]. Sequenced viruses were classified using the nomenclature <Major clade Minor clade\_Lineage>, e.g. Cosmopolitan AF1a\_A1.1. To build contextual datasets, publicly available sequences from the identified lineages were obtained from RABV-GLUE, and aligned with the newly generated sequences using the MAFFT FFT-NS-2 algorithm [13]. Maximum likelihood trees were built using IQTREE2 with model selection [14] and 1000 ultrafast bootstrap replicates [15]. Patristic distances were calculated using the get\_pairwise\_distances function from the R package castor [16], which calculates the phylogenetic distance between tip pairs. Trees of widespread lineages were checked to identify potential incursions relevant to human cases. If absent, these lineages were subset to only include sequences from Kenya or Tanzania up to 10 years prior to each human case (i.e. only sequences relevant to understanding transmission to the human case). Locations of cases were mapped using precise coordinates where available. Where exact coordinates were not available, the village or county centroid was used and jittered for mapping. Cases with no within-country location information were excluded from maps but included in phylogenetic analyses.

**List of abbreviations**

GLUE Gene-linked by Underlying Evolution

IgG1 Goat-anti-mouse

KCRI Kilimanjaro Clinical Research Institute

PCR Polymerase chain reaction

RABV Rabies virus

RNA Ribonucleic acid

UNITID University of Nairobi Institute of Tropical and Infectious Diseases

DNA Deoxyribonucleic acid

cDNA complementary DNA

dsDNA double-stranded deoxyribonucleic acid

Reference

1. Rabies Ag - BioNote - Canine Rapid Test - Pet Rapid Test. [cited 20 Oct 2022]. Available: https://www.qbiotech.gr/pet-rapid-test/canine-rapid-test/bionote-canine/rabies-ag.html?\_\_\_store=en&\_\_\_from\_store=el

2. Niezgoda M, Satheshkumar PS. Immunohistochemistry Test for the Lyssavirus Antigen Detection from Formalin-Fixed Tissues. JoVE J Vis Exp. 2021; e60138. doi:10.3791/60138

3. CDC - Diagnosis: Immunohistochemistry (IHC) - Rabies. 10 May 2022 [cited 20 Oct 2022]. Available: https://www.cdc.gov/rabies/diagnosis/immunohistochemistry.html

4. Bautista C, Jaswant G, French H, Campbell K, Durrant R, Gifford R, et al. Author Spotlight: A Cost-Effective Genomic Workflow for Advancing Rabies Control in Resource-Limited Settings. JoVE J Vis Exp. 2023; e65414. doi:10.3791/65414

5. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat Protoc. 2017;12: 1261–1266. doi:10.1038/nprot.2017.066

6. Brunker K, Jaswant G, Thumbi SM, Lushasi K, Lugelo A, Czupryna AM, et al. Rapid in-country sequencing of whole virus genomes to inform rabies elimination programmes. Wellcome Open Res. 2020;5: 3. doi:10.12688/wellcomeopenres.15518.2

7. PrimalScheme: primer panels for multiplex PCR. [cited 20 Oct 2022]. Available: https://primalscheme.com/

8. Artic Network. [cited 17 Mar 2022]. Available: https://artic.network/

9. Medaka. Oxford Nanopore Technologies; 2022. Available: https://github.com/nanoporetech/medaka

10. Singer JB, Thomson EC, McLauchlan J, Hughes J, Gifford RJ. GLUE: a flexible software system for virus sequence data. BMC Bioinformatics. 2018;19: 532. doi:10.1186/s12859-018-2459-9

11. What is table2asn? [cited 7 Dec 2023]. Available: https://www.ncbi.nlm.nih.gov/genbank/table2asn/

12. Campbell K, Gifford RJ, Singer J, Hill V, O’Toole A, Rambaut A, et al. Making genomic surveillance deliver: A lineage classification and nomenclature system to inform rabies elimination. PLOS Pathog. 2022;18: e1010023. doi:10.1371/journal.ppat.1010023

13. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30: 3059–3066. doi:10.1093/nar/gkf436

14. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020;37: 1530–1534. doi:10.1093/molbev/msaa015

15. Minh BQ, Nguyen MAT, von Haeseler A. Ultrafast Approximation for Phylogenetic Bootstrap. Mol Biol Evol. 2013;30: 1188–1195. doi:10.1093/molbev/mst024

16. Using the Castor API R package for data analysis. In: Castor [Internet]. [cited 2 May 2023]. Available: https://helpdesk.castoredc.com/link-castor-to-another-application/using-the-castor-api-r-package-for-data-analysis