

Genomic characterisation identifies cross-border importation of the 2026 Bundibugyo virus outbreak lineage from the Democratic Republic of the Congo into Kenya

Leonard Kingwara

leonard.kingwara@gmail.com

Kenya National Public Health Institute

Millicent Ndia

Kenya National Public Health Institute

Abdi Roba

Kenya National Public Health Institute

Lily Kirui

Kenya National Public Health Institute

Kennedy Yatich

Kenya National Public Health Institute

Max Mutuku

Kenya National Public Health Institute

Charles Rombo

Kenya National Public Health Institute

Emmanuel Okunga

Kenya National Public Health Institute

Kanana Kimonye

Kenya National Public Health Institute

Muthoka Richard

Kenya National Public Health Institute

Isabella Ayagah

Kenya National Public Health Institute

Juma Bonventure

Centers for Disease Control and Prevention

Naomi Lucchi

Centers for Disease Control and Prevention

Yao Selom Philippe Atrah

Africa Centres for Disease Control and Prevention

Collins K. Tanui

Africa Centres for Disease Control and Prevention

Yenew Kebede

Africa Centres for Disease Control and Prevention

Borus Peter

World Health Organization Regional Office for Africa

Nicksy Gumede Moeletsi

World Health Organization Regional Office for Africa

Mildred Shieshia

United States Department of State

Patrick Amoth

Ministry of Health, Kenya

Kamene Kimenye

Kenya National Public Health Institute

Mary Muthoni Muriuki

Ministry of Health, Kenya

Case Report

Keywords:

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11322777/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Bundibugyo virus (BDBV) is one of the orthoebolaviruses that cause Ebola disease in humans and is the least characterised of them; no vaccine or specific therapeutic is licensed for Bundibugyo virus disease (BVD). In 2026 the third and largest recorded BVD outbreak emerged in Ituri Province, Democratic Republic of the Congo (DRC), spread to Uganda, and prompted the declaration of a Public Health Emergency of International Concern. Here we report the first full-length BDBV genome recovered in Kenya, from a case with recent travel to the DRC confirmed on 3 October 2026 and sample collected and tested on 5th October 2026 after presenting with febrile like illness. The 18,939-nucleotide genome falls within the contemporary 2026 outbreak lineage and differs from its closest relatives in Ituri by only four single-nucleotide polymorphisms, whereas it is about 1.3% divergent from the 2007 Ugandan and 2012 DRC BDBV lineages. These data strongly support the interpretation that the Kenyan case represents a cross-border importation from the ongoing DRC rather than a locally sustained.

Main text

Bundibugyo virus (BDBV; *Orthoebolavirus bundibugyoense*) was first identified in 2007 during an outbreak of viral haemorrhagic fever in Bundibugyo District, Uganda, in which 131 suspected cases and 42 deaths were reported[1, 2], and was responsible for a second outbreak in Isiro, DRC, in 2012 (77 suspected cases and 36 deaths)[3, 4]. Although BDBV is one of the orthoebolaviruses that cause Ebola disease in humans[5, 6], it has attracted far less scientific and countermeasure investment than Zaire ebolavirus, and no vaccine or specific therapeutic is licensed for BVD.

In April 2026 a cluster of unexplained deaths was identified in the mining-intensive Mongbwalu Health Zone of Ituri Province, eastern DRC, and whole-genome sequencing confirmed BDBV on 14 May 2026[7, 8]. The outbreak expanded rapidly across Ituri, North Kivu and South Kivu, reaching 662 confirmed cases and 124 confirmed deaths in the DRC and 19 confirmed cases and 2 deaths in Uganda by 10 June 2026[7]. The World Health Organization declared a Public Health Emergency of International Concern on 17 May 2026 and the Africa Centres for Disease Control and Prevention a Public Health Emergency of Continental Security on 18 May 2026[7, 8]. Kenya, which shares a long land border with Uganda and close air and road links to eastern DRC, maintained enhanced Ebola disease surveillance throughout this period. However, on 3 October 2026, BDBV was laboratory-confirmed at the Kenya National Public Health Institute National Virology Reference Laboratory in Kenya in a patient with a history of travel to the DRC. The full genome was recovered and is reported here.

We sequenced using Illumina Viral surveillance panel on Miseq 100 to obtain a BDBV genome comprising of 18,939 nucleotides, with 23 ambiguous bases (0.12%). All seven canonical open reading frames (NP, VP35, VP40, GP, VP30, VP24 and L) are present, and each translates to a full-length protein with no frameshift or disruptive insertion or deletion. To place the genome in the context of the contemporary outbreak, we assembled all 1,021 BDBV records held on Pathoplexus on 9 October 2026, of which 953 were collected in 2026 (930 from the DRC, 20 from Uganda, two from Germany and one

from France), retained 921 unique full-length genomes after de-duplication, and aligned them against SUDV, EBOV, RESTV and TAFV outgroups[9–11].

Maximum-likelihood phylogenetic inference placed the Kenyan genome unambiguously within the 2026 outbreak lineage, together with all 2026 DRC genomes, the Ugandan cluster and the European importations, and not on a branch basal to that lineage (Fig. 1). The Kenyan genome differed from its four nearest neighbours by only four single-nucleotide polymorphisms (SNPs); these four genomes were sampled in Rwampara (19th June 2026), Nia Nia (27 June and 12 July 2026) and Bunia (12 August 2026), all in Ituri Province. From the 930 DRC genomes, we removed the duplicates and remained with 873. Across the 873 contemporary DRC genomes the median distance from the Kenyan genome was 15 SNPs (range 4–98; Fig. 2), and the Kenyan sequence was recovered within nested clades composed exclusively of DRC genomes (ultrafast bootstrap support of 97% for the immediate sister clade of four taxa, and of 72% and 69% for the deeper clades comprising 37 and 39 DRC genomes). No evidence of a divergent sublineage, segmental reassortment or mosaic genome was found. By contrast, the Kenyan genome differed from the 2007 Ugandan and 2012 DRC lineages by 212–265 nucleotides (about 1.3% divergence across comparable sites) and from other orthoebolaviruses by 30.9–37.3% (Fig. 2). No other BDBV genome from Kenya was present in the public sequence record at the time of analysis[7, 11].

Relative to the 2026 outbreak consensus sequence, the Kenyan genome encodes six amino-acid substitutions, at each of which the Kenyan residue is absent from both the 2007–2008 and the 2012 lineages: N26D and I528N in the nucleoprotein (NP), P344S in the GP1 subunit of the glycoprotein (GP), E192K in VP30, and F1022I and Y1562S in the RNA-dependent RNA polymerase (L) (Table 1). Ten of the 14 nucleotide positions at which the Kenyan genome differs from every other 2026 genome lie within the first 37 nucleotides of the genome, adjacent to a six-base ambiguous stretch; we cannot exclude that these represent end-of-genome assembly rather than biological variation, and we recommend confirmatory sequencing of both termini. Because BDBV-specific countermeasures remain in development, the possible effect of these substitutions on diagnostic assay inclusivity and on candidate vaccines and antibody therapeutics warrants assessment and continued monitoring of the expanding lineage.

The 2026 lineage shows temporal structure, but its rate is not yet well constrained. A root-to-tip regression across the 127 genomes of the lineage with a full collection date gave 1.3×10^{-3} substitutions per site per year (approximately 24 nucleotides per genome per year; $R^2 = 0.14$), but this estimate was unstable across taxon subsets (an alternative sample of the same lineage gave 7.5×10^{-4}), as expected over a sampling window of only five months. Three sequences were not included as they did not have collection date. Estimates derived from the divergence between the Kenyan genome and the 2007–2008 (6.6×10^{-4} ; approximately 12.5 nucleotides per genome per year) and 2012 (8.9×10^{-4} ; approximately 17) lineages were more stable, and bracket the substitution rate typical of ebolaviruses. Because of this uncertainty we do not present a point estimate for the origin of the lineage; the data place the most recent common ancestor of the 2026 lineage within the months preceding the first recognised cluster in April 2026. The within-lineage diversity also exceeds what the four-to-five-month

sampling window alone would generate, which, taken with the rate estimates, suggests that the lineage circulated undetected for several months before first detection[12].

Together these data identify the Kenyan case as a cross-border importation of the ongoing DRC outbreak lineage rather than the recognition of a novel or locally maintained BDBV transmission chain. Three operational implications follow. First, case detection, contact tracing and community engagement should be organised around the travel and contact history linking the patient to Ituri and North Kivu, rather than around local exposure. Second, because the lineage is evolving measurably, sustained near-real-time genomic surveillance with open data deposition-as already established in the DRC and Uganda – is needed to keep cross-border importations traceable. Third, in the absence of licensed BDBV countermeasures, diagnostic assurance is the principal laboratory safeguard, and imported genomes should be shared promptly so that any assay-relevant or antigenic change is detected early[13].

This report has several limitations. It rests on a single genome, so no intra-host diversity, transmission direction or outbreak-wide phylodynamic inference is possible. Most comparison sequences are nanopore consensus genomes, which carry residual error that can inflate distances and bias rate and divergence estimates, and the historical record is short and partly fragmentary. The clock signal across a five-month sampling window is correspondingly weak and unstable ($R^2 = 0.14$), and we therefore report the origin of the lineage only as a broad interval rather than a date. Finally, the clinical and epidemiological detail of the Kenyan case is not presented here.

Table 1
Amino-acid substitutions in the Kenyan genome relative to the 2026 outbreak consensus

Gene	Protein	Length (aa)	Position	2007 / 2012 lineage and 2026	Kenya
NP	Nucleoprotein	740	26	N	D
NP	Nucleoprotein	740	528	I	N
GP	Glycoprotein (GP1)	677	344	P	S
VP30	Minor nucleoprotein	290	192	E	K
L	RNA-dependent RNA polymerase	2211	1022	F	I
L	RNA-dependent RNA polymerase	2211	1562	Y	S

Methods

Sequence data retrieval. All BDBV records available on Pathoplexus (<https://pathoplexus.org/ebola-bdbv>) were retrieved on 9 October 2026 (1,021 records). Reference genomes for Sudan virus (NC_006432.1), Ebola virus (NC_002549.1), Reston virus (NC_004161.1), Taï Forest virus (NC_014372.1) and BDBV

(NC_014373.1) were obtained from NCBI GenBank. Records shorter than 15,000 nucleotides were excluded from phylogenetic analysis.

Alignment and phylogenetics. Sequences were aligned with MAFFT v7 and columns with data in fewer than 5% of BDBV sequences were trimmed, giving an alignment of 18,940 columns. Exactly identical full-length genomes were collapsed to a single representative. Maximum-likelihood trees were inferred with IQ-TREE 2 under GTR+G with constant-site correction, using 1,000 ultrafast bootstrap replicates, on a 155-taxon set spanning all three BDBV lineages and rooted on the 2007–2008 Ugandan lineage; branch lengths are shown in the source tree file, and Fig. 1 is drawn as a cladogram for legibility because the historical lineages are approximately 1.3% divergent from the 2026 lineage. Pairwise distances were computed over unambiguous, non-gap sites only.

Substitution analysis. Nucleotide and amino-acid substitutions were called against the majority-base consensus of 93 high-quality ($\geq 99\%$ complete) 2026 DRC genomes, and against consensus sequences for the 2007 and 2012 lineages; open reading frames were extracted by mapping to the annotated BDBV reference NC_014373.1.

Temporal analysis. Root-to-tip regression was performed on the 2026 lineage using collection dates, and the substitution rate was cross-checked against the divergence between the 2026 lineage and the 2007 and 2012 lineages.

Reporting. This study did not involve a pre-registered protocol or a prospective design; a single genome was characterised against publicly available sequence data.

Declarations

Data availability

The Kenyan BDBV genome has been deposited in Pathoplexus under accession PP_0080L6U.1 and will be made available on Pathoplexus (<https://pathoplexus.org/ebola-bdbv>). All comparison genomes are publicly available on Pathoplexus.

Acknowledgements

We acknowledge the Ministry of Health, Kenya, for the resources and institutional support that enabled this investigation, and the staff of the Kenya National Public Health Institute National Virology Reference Laboratory for diagnostic testing and sequencing. We thank the Africa Centres for Disease Control and Prevention and the Africa Pathogen Genomics Initiative for technical and coordination support, and the sequencing teams in the Democratic Republic of the Congo and Uganda whose timely open deposition of Bundibugyo virus genomes on Pathoplexus made this comparative analysis possible.

Funding

No explicit funding was received for the work

Author contributions

L.Kin. conceptualised the study, analysed the genomic data and drafted the manuscript. M.N., A.R., J.O., L.Kir., K.Y., M.M. and C.R. performed laboratory testing and sequencing and reviewed the manuscript. E.O., K.Kimo., I.A., J.B. and N.L. contributed to case investigation, contact tracing and writing. C.K.T. and Y.K. contributed to interpretation of the findings in the continental genomic surveillance context and to critical revision of the manuscript. P.A. and K.Kime. supervised the work. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Conflict of Interest Declaration

The authors declare that they have no affiliations with, or involvement in, any organization or entity with any financial interest in the subject matter or materials discussed in this manuscript.

Ethics approval declaration

This study was reviewed and approved by the KEMRI Scientific and Ethics Review Unit (SERU), which served as the Institutional Review Board (IRB) and research ethics committee of record for the protocol. Approval was granted for the full protocol, covering all field, laboratory, and data activities described herein under a non-KERI protocol number 5199

Ethics declaration

All procedures described in this protocol are conducted in accordance with the Declaration of Helsinki and with applicable Kenyan national research ethics and regulatory requirements, including review by the KEMRI Scientific and Ethics Review Unit (SERU), a research license from the National Commission for Science, Technology and Innovation (NACOSTI), an environmental permit from the National Environment Management Authority (NEMA), and the Kenya Data Protection Act 2019.

Consent to Participate declaration

not applicable. In this study, no individual identifiers are generated, and no link is created between the result and any specific person, so individual informed consent does not apply. This is a component of implantation science being done during an outbreak investigation

Consent to Publish declaration

not applicable. This manuscript contains no individual person's data in any form. All data presented are pooled environmental measurements, aggregate counts and summary statistics that cannot be

attributed to an identifiable individual, and no individual images, case details, or personally identifiable information are included.

References

1. Wamala JF et al. Ebola hemorrhagic fever associated with novel virus strain, Uganda, 2007–2008. *Emerg Infect Dis* 16, (2010).
2. Towner JS et al. Newly discovered Ebola virus associated with hemorrhagic fever outbreak in Uganda. *PLoS Pathog* 4, (2008).
3. MacNeil A et al. Proportion of deaths and clinical features in Bundibugyo Ebola virus infection, Uganda. *Emerg Infect Dis* 16, (2010).
4. Hulseberg CE et al. Molecular analysis of the 2012 Bundibugyo virus disease outbreak. *Cell Rep Med* 2, (2021).
5. Kuhn JH et al. New filovirus disease classification and nomenclature. *Nature Reviews Microbiology* vol. 17 Preprint at <https://doi.org/10.1038/s41579-019-0187-4> (2019).
6. Harrison JE, Weber S, Jakob R, Chute CG. ICD-11: an international classification of diseases for the twenty-first century. *BMC Medical Informatics and Decision Making* vol. 21 Preprint at <https://doi.org/10.1186/s12911-021-01534-6> (2021).
7. Zhang Y, Ren J, Ma X, He X. The 2026 Bundibugyo virus disease outbreak in the Democratic Republic of the Congo: Virology, epidemiology, continental response, and research priorities. *Infectious Medicine* vol. 5 Preprint at <https://doi.org/10.1016/j.imj.2026.100270> (2026).
8. Parums DV, Editorial. The WHO Identifies Ebola Disease Due to Bundibugyo Virus as a Public Health Emergency of International Concern (PHEIC) as Vaccine Development Accelerates. *Medical Science Monitor* vol. 32 Preprint at <https://doi.org/10.12659/MSM.954627> (2026).
9. Amuri-Aziza A, et al. Emergence of a Bundibugyo virus variant in the 2026 outbreak in the Democratic Republic of the Congo and Uganda. *Nat Med*. 2026. <https://doi.org/10.1038/s41591-026-04628-8>.
10. Mooring EQ et al. Modeled Scenario Projections for the Ebola Disease Outbreak Caused by Bundibugyo Virus, 2026. *MMWR Surveillance Summaries* 75, (2026).
11. Nsawotebba A et al. Clinical profile and genomic characterization of the 2026 Bundibugyo virus index case in Uganda. *Nat Med* 32, (2026).
12. Kopp K. Bundibugyo Virus Disease: Diagnostics and Medical Countermeasures for a Neglected Ebolavirus. *Viruses* vol. 18 Preprint at <https://doi.org/10.3390/v18070775> (2026).
13. Scotch M, Magge A, Valente M, ZooPhy. A bioinformatics pipeline for virus phylogeography and surveillance. *Online J Public Health Inf*. 11, (2019).

Figures

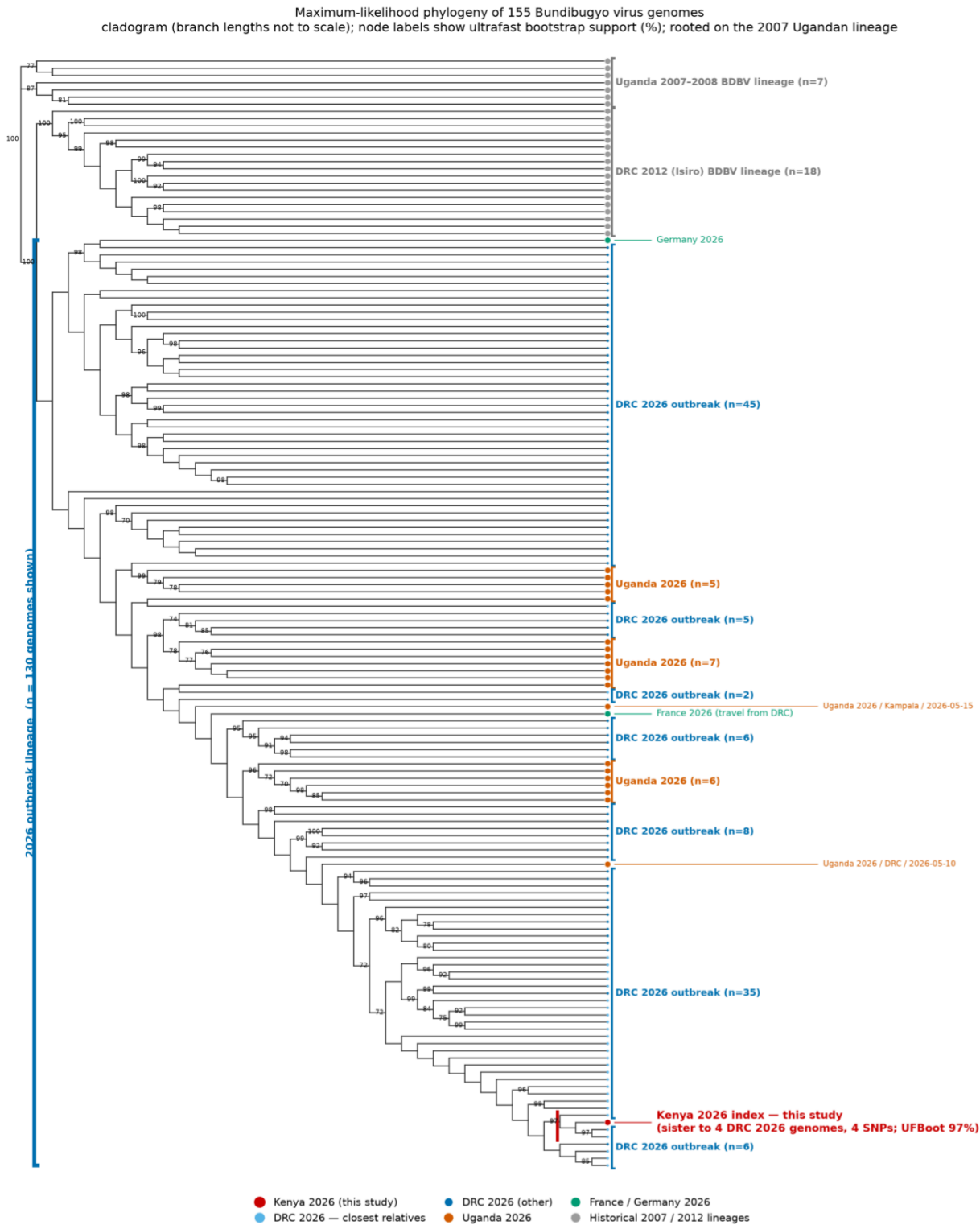


Figure 1

Maximum-likelihood phylogeny of 155 Bundibugyo virus genomes, including the 2026 Kenyan case. Cladogram (branch lengths not to scale) rooted on the 2007–2008 Ugandan BDBV lineage. Node labels show ultrafast bootstrap support (%); tips are coloured by country and year. The Kenyan genome falls within the 2026 outbreak lineage and is sister to four DRC genomes from Ituri Province (4 SNPs; ultrafast bootstrap 97%).

Cladogram of the Kenyan genome; the 30 genomes genetically closest to it; a geographically and temporally stratified sample of 2026 DRC genomes; all 2026 Ugandan, French and German genomes; and representatives of the 2007–2008 Ugandan and 2012 DRC BDBV lineages. The tree was inferred under GTR+G with constant-site correction from 646 variable sites with 1,000 ultrafast bootstrap replicates, and rooted on the 2007–2008 Ugandan lineage. Branch lengths are not to scale; genetic distances are given in **Fig. 2**. Node labels show ultrafast bootstrap support (% shown where ≥ 70). The Kenyan genome falls within the 2026 outbreak lineage and is sister to four DRC genomes sampled in Ituri Province (4 SNPs; ultrafast bootstrap 97%). The 2026 lineage comprises 130 of the genomes shown.

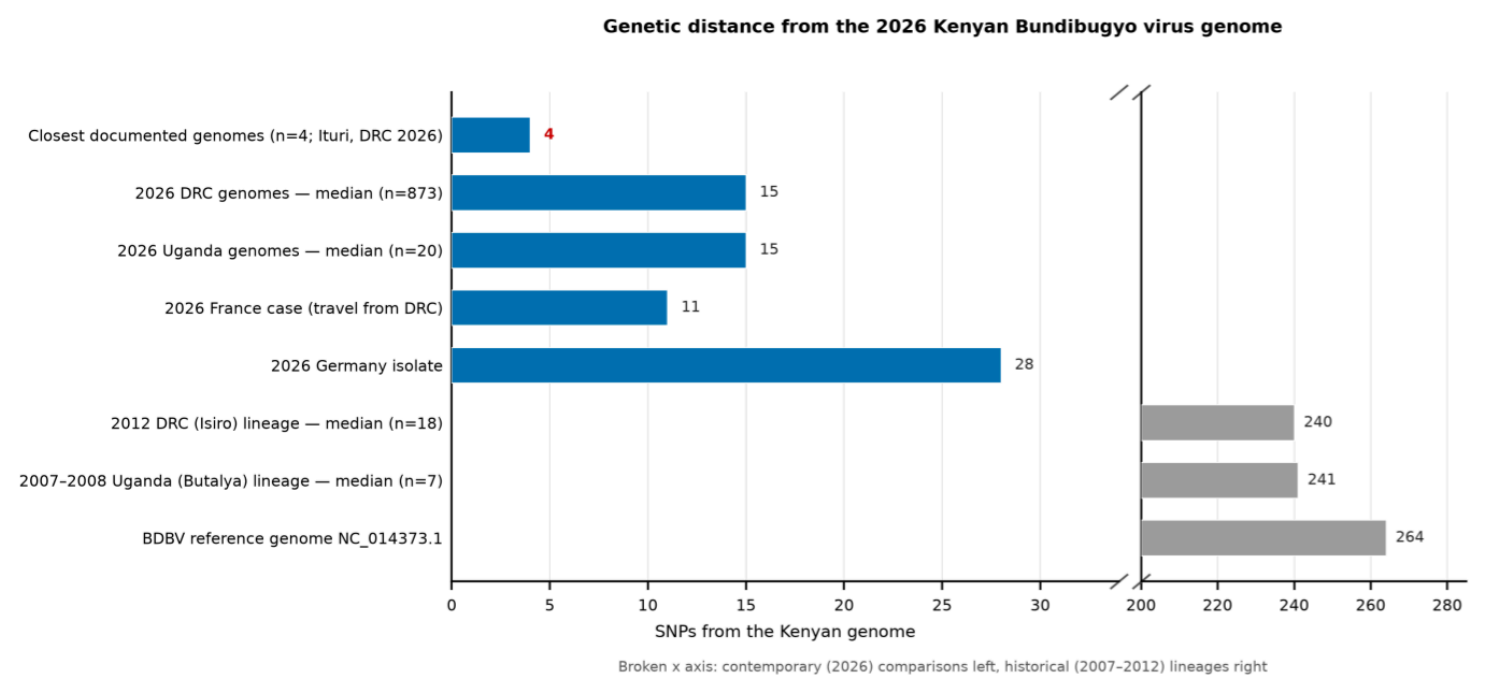


Figure 2

Genetic distance from the 2026 Kenyan Bundibugyo virus genome to other BDBV genomes. Minimum (closest group) or median SNP distance between the Kenyan genome and each group; the x axis is broken between contemporary (2026) and historical comparisons. SNP counts exclude ambiguous and gap sites.

Bars show the minimum (closest group) or median number of single-nucleotide polymorphisms (SNPs) between the Kenyan genome and each group of genomes. SNPs were counted over unambiguous, non-gap sites only, excluding positions at which either sequence carried an ambiguous base. The x axis is broken between the contemporary (2026) comparisons (4–28 SNPs) and the historical 2007–2012 lineages (240–264 SNPs). The closest documented genomes differ from the Kenyan genome by four SNPs (n = 4, all Ituri Province). n indicates the number of genomes in each group.