

The genome sequence of varicella-zoster virus (*Varicellovirus humanalpha3*) obtained by metagenomics from a patient presenting with an exanthem rash

Sebastian Musundi,¹ Dorcas Okanda,¹ Arnold Lambisia,¹ John Mwita Morobe,¹ Eric Muthanje,¹ Lynette Isabella Ochola-Oyier,¹ Daniel P. Carter,² Charles N. Agoti,¹ John Kiiru,^{3,4} George Githinji^{1,5}

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT Here, we report a partial genome sequence of varicella-zoster virus, recovered through metagenomic sequencing from a skin lesion sample collected from a 31-year-old male from Mombasa, Kenya, in August 2024. Phylogenetic analysis placed this isolate in varicella-zoster virus clade 5.

KEYWORDS varicella-zoster virus, metagenomics, pox viruses

Varicella-zoster virus (VZV) is a 125 kb linear double-stranded DNA alphaherpesvirus encoding 71 proteins (1). It is the causative agent of chickenpox (varicella) during primary infection and shingles (herpes zoster) upon reactivation after latency in ganglia neurons (2). VZV presents with an exanthem rash, like those observed in other poxviruses, e.g. monkeypox (Mpox) (3, 4). Here, we demonstrate that metagenomic sequence analysis can distinguish these infections.

On 23 August 2024, the KEMRI-Wellcome Trust Research Programme in Kilifi, Kenya, received a skin lesion sample stored in viral transport media (VTMs) (Revital, Kenya) (5) collected from a 31-year-old male resident of Likoni sub-county in Mombasa County for diagnostic confirmation. The patient presented with an exanthem rash, raising the suspicion of an Mpox infection.

Viral DNA was extracted from 140 µL of the VTM sample using the QIAamp viral RNA mini kit and QIAamp Fast DNA Stool Mini Kit (both Qiagen, Germany) with a modified bead beating step (6). DNA was quantified using the Qubit dsDNA High Sensitivity Assay kit on a Qubit 4 fluorometer (Thermo Fisher Scientific, USA). Library preparation employed the Native Barcoding Kit (SQK-NBD114.96) from Oxford Nanopore Technologies (ONT), and sequencing was performed on the ONT platform using an R10.4.1 flow cell. Base calling was conducted using the Dorado (v.7.3.9) software (high-accuracy model), yielding 1.27 million reads, of which 1.07 million reads remained after adapter removal using Porechop (v.0.2.4) (7). For diagnostic purposes, species identification was performed using the Chan Zuckerberg ID nanopore metagenomics bioinformatics pipeline (8). All tools used default settings unless otherwise specified.

To recover the VZV genome, human reads were removed by mapping the sequenced reads against the human reference genome hg38 (GCF_000001405.40) using Minimap2 (v.2.28) (9) and processed using SAMtools (v.1.2) (10). The remaining 193,390 reads were assembled using Flye with the *-meta* option (v.2.8.1) (11), followed by three iterative polishing rounds using Racon (v.1.5.0). The final assembly had an *N50* of 125,797, and nucleotide BLAST (12) of the 125,797 bp contig matched the human herpesvirus 3 strain isolated from the USA in 2012 (MH709330.1 and MH709332.1), 2013 (MH709360.1 and MH709349.1), and from Germany in 2007 (JN704707.1), with 99.93% identity. Our consensus genome of 124,668 bp with 46% guanine-cytosine content was derived after

Editor Kenneth M. Stedman, Portland State University, Portland, Oregon, USA

Address correspondence to George Githinji, ggithinji@kemri-wellcome.org.

The authors declare no conflict of interest.

Received 11 March 2025

Accepted 21 May 2025

Published 10 June 2025

Copyright © 2025 Musundi 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/).

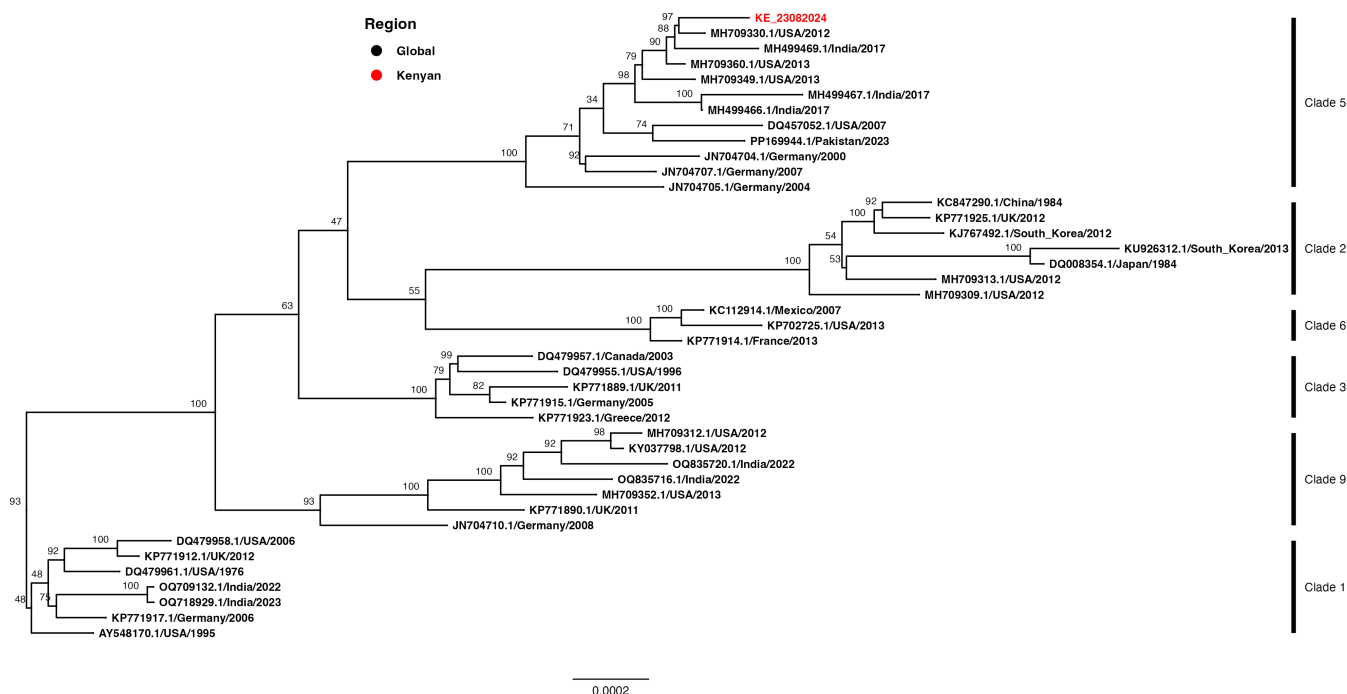


FIG 1 A maximum likelihood tree generated using complete human herpesvirus 3 nucleotide global sequences ($n = 40$) with the Kenyan isolate (KE_23082024) highlighted in red. The tips are colored by geographic origin. Bootstrap values are displayed in the nodes and indicate the level of support for each branch in the phylogenetic tree. The substitution rate is shown in the horizontal bar.

8,806 reads mapped to MH709330.1 (99.9% coverage and 34× average depth). Genome annotation was performed in Geneious Prime (v.2024) (Dotmatrix, USA) via a homology approach using the closest reference.

We retrieved VZV genomes ($n = 40$) from GenBank, aligned them using MAFFT (v.7.25) (13), and generated a maximum likelihood tree in IQ-TREE (v.2.3.3) (14) with 1,000 bootstraps using the general time reversible model as selected by ModelFinder (15). The tree was visualized in R (v.4.4.1) using the “treeio” package (16). Based on reported genotyping schemes (17, 18), the Kenyan isolate was classified under clade 5 (Fig. 1).

ACKNOWLEDGMENTS

Special thanks to the Likoni Health Facility for their support in collecting and preparing the sample.

This work was supported by funding from the UK Department of Health and Social Care via UK-PHRST research funds. The UK Public Health Rapid Support Team is funded by UK Aid from the Department of Health and Social Care and is jointly run by the UK Health Security Agency and the London School of Hygiene & Tropical Medicine. In addition, funding support was received from Wellcome. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health and Social Care.

This work was published with permission from the Director of KEMRI.

AUTHOR AFFILIATIONS

¹KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya

²UK Public Health Rapid Support Team, London, United Kingdom

³Kenya National Public Health Institute, Nairobi, Kenya

⁴Ministry of Health, Nairobi, Kenya

⁵Department of Biochemistry and Biotechnology, Pwani University, Kilifi, Kenya

AUTHOR ORCID*s*

Sebastian Musundi  <http://orcid.org/0000-0003-2940-5717>

Arnold Lambisia  <http://orcid.org/0000-0001-5312-0960>

Charles N. Agoti  <http://orcid.org/0000-0002-2160-567X>

George Githinji  <http://orcid.org/0000-0001-9640-7371>

AUTHOR CONTRIBUTIONS

Sebastian Musundi, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review and editing | Dorcas Okanda, Formal analysis, Methodology, Writing – original draft, Writing – review and editing | Arnold Lambisia, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing | John Mwita Morobe, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing | Eric Muthanje, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing | Lynette Isabella Ochola-Oyier, Formal analysis, Funding acquisition, Project administration, Supervision, Writing – review and editing | Daniel P. Carter, Formal analysis, Methodology, Resources, Writing – review and editing | Charles N. Agoti, Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Writing – original draft, Writing – review and editing | John Kiiru, Resources, Writing – review and editing | George Githinji, Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

Raw sequencing data are available in the Sequence Read Archive (BioProject [PRJNA1175889](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1175889), [SRX26453068](https://www.ncbi.nlm.nih.gov/bioproject/SRX26453068)), and the complete nucleotide sequence is under GenBank accession number [PQ505474](https://www.ncbi.nlm.nih.gov/nuclseq/PQ505474).

ETHICS APPROVAL

Ethical approval was obtained from the KEMRI Scientific and Ethics Review Unit (protocol #5070).

REFERENCES

- Cohen JL. 2010. The varicella-zoster virus genome. *Curr Top Microbiol Immunol* 342:1–14. https://doi.org/10.1007/82_2010_10
- Gershon AA, Breuer J, Cohen JL, Cohrs RJ, Gershon MD, Gilden D, Grose C, Hambleton S, Kennedy PGE, Oxman MN, Seward JF, Yamanishi K. 2015. Varicella zoster virus infection. *Nat Rev Dis Primers* 1:15016. <https://doi.org/10.1038/nrdp.2015.16>
- McNaughton A, Karsenti N, Kwan J, Adawi A, Mansuri S, Boggild AK. 2024. Primary varicella infection in a young adult from the democratic republic of the Congo: a case report and mini-review. *Infect Dis Rep* 16:628–637. <https://doi.org/10.3390/idr16040048>
- Leung J, McCollum AM, Radford K, Hughes C, Lopez AS, Guagliardo SAJ, Nguete B, Likafi T, Kabamba J, Malekani J, Shongo Lushima R, Pukuta E, Karhemere S, Muyembe Tamfum JJ, Reynolds MG, Wemakoy Okitolonda E, Schmid DS, Marin M. 2019. Varicella in Tshuapa Province, Democratic Republic of Congo, 2009–2014. *Trop Med Int Health* 24:839–848. <https://doi.org/10.1111/tmi.13243>
- Ngetsa C, Osoti V, Okanda D, Marura F, Shah K, Karanja H, Mugo D, Gitonga J, Mutunga M, Lewa C, Orindi B, Bejon P, Ochola-Oyier LI. 2023. Validation of saline, PBS and a locally produced VTM at varying storage conditions to detect the SARS-CoV-2 virus by qRT-PCR. *PLoS One* 18:e0280685. <https://doi.org/10.1371/journal.pone.0280685>
- Knudsen BE, Bergmark L, Munk P, Lukjancenko O, Priemé A, Aarestrup FM, Pamp SJ. 2016. Impact of sample type and dna isolation procedure on genomic inference of microbiome composition. *mSystems* 1. <https://doi.org/10.1101/064394>
- Wick R, Volkening J. 2018. Github. Porechop: adapter trimmer for oxford nanopore reads. Available from: <https://github.com/rrwick/Porechop>
- Kalantar KL, Carvalho T, de Bourcy CFA, Dimitrov B, Dingle G, Egger R, Han J, Holmes OB, Juan Y-F, King R, et al. 2020. IDseq-An open source cloud-based pipeline and analysis service for metagenomic pathogen detection and monitoring. *Gigascience* 9:giaa111. <https://doi.org/10.1093/gigascience/giaa111>
- Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>
- 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)
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274. <https://doi.org/10.1093/molbev/msu300>
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 14:587–589. <https://doi.org/10.1038/nmeth.4285>
- Wang LG, Lam TTY, Xu S, Dai Z, Zhou L, Feng T, Guo P, Dunn CW, Jones BR, Bradley T, Zhu H, Guan Y, Jiang Y, Yu G. 2020. Treeio: an R package for

- phylogenetic tree input and output with richly annotated and associated data. *Mol Biol Evol* 37:599–603. <https://doi.org/10.1093/molbev/msz240>
17. Jensen NJ, Rivailler P, Tseng HF, Quinlivan ML, Radford K, Folster J, Harpaz R, LaRussa P, Jacobsen S, Scott Schmid D. 2017. Revisiting the genotyping scheme for varicella-zoster viruses based on whole-genome comparisons. *J Gen Virol* 98:1434–1438. <https://doi.org/10.1099/jgv.0.000772>
 18. Kumar A, Rajan LS, Sabarinath PS K, Shete AM, Sahay RR, Patil DY, Ingole N, K K, Padinakarai AC, Gb S, Shastri J, Padukone S, Joshi Y, Patil S, Majumdar T, Verma A, Yemul J, Shende N, Kumari V, Vedpathak P, Sathe S, Gawande P, Yadav PD. 2023. First detection of Varicella Zoster Virus clade 9 cases in India during mpox surveillance. *Ann Med* 55. <https://doi.org/10.1080/07853890.2023.2253733>