



Near-Complete Genome Sequence of Ndumu Virus from Garissa, Kenya, 1997

 Naazneen Moolla,^{a,b} Natalie Viljoen,^{a,d} Venessa Patharoo,^a Antoinette Grobbelaar,^a  Arshad Ismail,^e Jacqueline Weyer^{a,b,c}

^aCentre for Emerging Zoonotic and Parasitic Diseases, National Institute for Communicable Diseases, National Health Laboratory Service, Johannesburg, South Africa

^bDepartment of Medical Virology, Faculty of Health Sciences, University of Pretoria, Tshwane, South Africa

^cDepartment of Microbiology and Infectious Diseases, Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa

^dDivision of Virology, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa

^eSequencing Core Facility, National Institute for Communicable Diseases, National Health Laboratory Service, Johannesburg, South Africa

ABSTRACT We report a nearly complete genome sequence of Ndumu virus (NDUV) identified using a metagenomics approach. The sequence was derived from a viral isolate obtained from a bovine calf following a diagnostic investigation of the 1997 to 1998 Rift Valley fever (RVF) outbreak in the Garissa District of northeastern Kenya.

Isolated from *Mansonia uniformis* mosquitoes in South Africa in 1959, Ndumu virus (NDUV) belongs to the family *Togaviridae*, genus *Alphavirus* (1, 2). NDUV has also been detected in mosquitoes in Kenya (3, 4) and Senegal (5). NDUV was identified in domestic pigs and the blood meals of mosquitoes that fed on goats and sheep in a study conducted in Uganda (6, 7). The pathogenicity of the virus in humans remains unclear, although serological evidence of exposure in humans exists (1).

A blood sample collected from a calf presenting with Rift Valley fever (RVF)-like symptoms during the RVF outbreak in Kenya was submitted for diagnostic investigation to the National Institute for Communicable Diseases (NICD) (then the National Institute for Virology), South Africa (8). The sample tested negative for RVF (8), but an unidentified virus isolate was obtained during the investigation by passage in suckling mouse brains and amplified by serial passage in Vero cells (ATCC CRL-1586). The cleared lysates from infected Vero cells were stored frozen. The identity of the latter remained unknown until the investigation reported here. The isolate was retrieved from the archival collection of the Centre for Emerging Zoonotic and Parasitic Diseases, NICD, and was passaged once on Vero C1008 cells (ATCC CRL-1586) using standard cell culture techniques. Viral genomic RNA was extracted from the culture supernatant using the QIAamp viral RNA kit (Qiagen, Germany), and sequence-independent single-primer amplification (SISPA) as detailed by Djikeng et al. was performed after DNase treatment (9, 10). Briefly, cDNA was retrotranscribed from the extracted RNA using the SuperScript III kit (Invitrogen, USA) and the primer FR-26-RV-N (9). Second-strand DNA was synthesized using Klenow exo-DNA polymerase. Purified cDNA was amplified using SISPA with MyTaq Red mix (Meridian Bioscience USA), using the primer FR-20-RV (9). The DNA was quantified using a Qubit fluorimeter (Thermo Fisher Scientific, USA); paired-end libraries were prepared using the Nextera DNA Flex library preparation kit, followed by sequencing (2 × 150 bp) on a NextSeq 550 platform (Illumina, Inc., USA). A total of 4.99 million sequencing reads were obtained with an average read length of 149 bp. The FASTQ files were uploaded to the Galaxy Web platform, and the data were analyzed using the public server at <http://usegalaxy.org/> (11). All tools were run for paired-end reads using default parameters unless otherwise specified. The reads were quality trimmed (qualified quality Phred score, 20; length required, 40) using fastp V0.20.1 (12) and classified using Kraken V1.1.1 with the

Citation Moolla N, Viljoen N, Patharoo V, Grobbelaar A, Ismail A, Weyer J. 2021. Near-complete genome sequence of Ndumu virus from Garissa, Kenya, 1997. *Microbiol Resour Announc* 10:e00551-21. <https://doi.org/10.1128/MRA.00551-21>.

Editor John J. Dennehy, Queens College CUNY

Copyright © 2021 Moolla 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 Naazneen Moolla, naazneenm@nicd.ac.za.

Received 5 July 2021

Accepted 30 July 2021

Published 26 August 2021

viruses database (13); 1.72 million reads were mapped to NDUV. The paired-end reads were mapped against the NDUV (GenBank accession number [NC_016959.1](#)) reference sequence using Bowtie2 V2.3.4.1 with the very sensitive end-to-end preset (14). The mapped reads were curated into a pileup file using SAMtools V1.9 (15), the consensus sequence determined using iVar V1.2.2 (16), and annotation performed using Vgas (17) at <http://guolab.whu.edu.cn/vgas>. The sequence was 11,648 nucleotides long with 50.1% GC content and an average sequencing coverage of 7,612×. The sequence obtained had 98% nucleotide and 99% amino acid similarity to the reference sequence as determined using Sequence Identity and Similarity (SIAS), available at <http://imed.med.ucm.es/Tools/sias.html>. Preliminary analysis suggested that three nucleotide insertions ([NC_016959.1:g.7427_7428insGC](#); [NC_016959.1:g.7440_7441insG](#)) may affect the coding sequence of the structural polypeptide [[YP_005351233.1:p.\(Pro13_Pro17delinsArgGlnProTyrArgArg\) ins\(1\)](#)]. The identification of NDUV from bovines has not been previously reported, to our knowledge. These data also contribute to the limited genomic sequence data available for NDUV.

Data availability. This sequence has been deposited in GenBank under the accession number [MZ091531](#). The version described here is the first version. The raw reads were deposited in the NCBI Sequence Read Archive (SRA) under the accession numbers [PRJNA728540](#) (BioProject) and [SAMN19091165](#) (BioSample).

REFERENCES

- Kokernot RH, McIntosh BM, Worth CB. 1961. Ndumu virus, a hitherto unknown agent, isolated from culicine mosquitoes collected in northern Natal, Union of South Africa. *Am J Trop Med Hyg* 10:383–386. <https://doi.org/10.4269/ajtmh.1961.10.383>.
- Walker PJ, Siddell SG, Lefkowitz EJ, Mushegian AR, Adriaenssens EM, Dempsey DM, Dutilh BE, Harrach B, Harrison RL, Hendrickson RC, Junglen S, Knowles NJ, Kropinski AM, Krupovic M, Kuhn JH, Nibert M, Orton RJ, Rubino L, Sabanadzovic S, Simmonds P, Smith DB, Varsani A, Zerbini FM, Davison AJ. 2020. Changes to virus taxonomy and the statutes ratified by the International Committee on Taxonomy of Viruses (2020). *Arch Virol* 165:2737–2748. <https://doi.org/10.1007/s00705-020-04752-x>.
- Crabtree M, Sang R, Lutomiah J, Richardson J, Miller B. 2009. Arbovirus surveillance of mosquitoes collected at sites of active Rift Valley fever virus transmission: Kenya, 2006–2007. *J Med Entomol* 46:961–964. <https://doi.org/10.1603/033.046.0431>.
- Ochieng C, Lutomiah J, Makio A, Koka H, Chepkorir E, Yalwala S, Mutisya J, Musila L, Khamadi S, Richardson J, Bast J, Schnabel D, Wurapa E, Sang R. 2013. Mosquito-borne arbovirus surveillance at selected sites in diverse ecological zones of Kenya; 2007–2012. *Virology* 45:109–114. <https://doi.org/10.1016/j.virol.2013.04.010>.
- Ndiaye EH, Diallo D, Fall G, Ba Y, Faye O, Dia I, Diallo M. 2018. Arboviruses isolated from the Barkedji mosquito-based surveillance system, 2012–2013. *BMC Infect Dis* 18:642. <https://doi.org/10.1186/s12879-018-3538-2>.
- Masembe C, Michuki G, Onyango M, Rumberia C, Norling M, Bishop RP, Djikeng A, Kemp SJ, Orth A, Skilton RA, Ståhl K, Fischer A. 2012. Viral metagenomics demonstrates that domestic pigs are a potential reservoir for Ndumu virus. *Virology* 433:218–224. <https://doi.org/10.1016/j.virol.2012.05.018>.
- Lutomiah J, Omondi D, Masiga D, Mutai C, Mireji PO, Ongus J, Linthicum KJ, Sang R. 2014. Blood meal analysis and virus detection in blood-fed mosquitoes collected during the 2006–2007 Rift Valley fever outbreak in Kenya. *Vector Borne Zoonotic Dis* 14:656–664. <https://doi.org/10.1089/vbz.2013.1564>.
- Woods CW, Karpati AM, Grein T, McCarthy N, Gaturuku P, Muchiri E, Dunster L, Henderson A, Khan AS, Swanepoel R, Bonmarin I, Martin L, Mann P, Smoak BL, Ryan M, Ksiazek TG, Arthur RR, Ndikuyezu A, Agata NN, Peters CJ, World Health Organization Hemorrhagic Fever Task Force. 2002. An outbreak of Rift Valley fever in Northeastern Kenya, 1997–98. *Emerg Infect Dis* 8:138–144. <https://doi.org/10.3201/eid0802.010023>.
- Djikeng A, Halpin R, Kuzmickas R, DePasse J, Feldblyum J, Sengamalay N, Afonso C, Zhang X, Anderson NG, Ghedin E, Spiro DJ. 2008. Viral genome sequencing by random priming methods. *BMC Genomics* 9:5. <https://doi.org/10.1186/1471-2164-9-5>.
- Reyes GR, Kim JP. 1991. Sequence-independent, single-primer amplification (SISPA) of complex DNA populations. *Mol Cell Probes* 5:473–481. [https://doi.org/10.1016/s0890-8508\(05\)80020-9](https://doi.org/10.1016/s0890-8508(05)80020-9).
- Jalili V, Afgan E, Gu Q, Clements D, Blankenberg D, Goecks J, Taylor J, Nekrutenko A. 2020. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. *Nucleic Acids Res* 48:W395–W402. <https://doi.org/10.1093/nar/gkaa434>.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ pre-processor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 15:R46. <https://doi.org/10.1186/gb-2014-15-3-r46>.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* 10:giab008. <https://doi.org/10.1093/gigascience/giab008>.
- Grubaugh ND, Gangavarapu K, Quick J, Matteson NL, De Jesus JG, Main BJ, Tan AL, Paul LM, Brackney DE, Grewal S, Gurfield N, Van Rompay KKA, Isern S, Michael SF, Coffey LL, Loman NJ, Andersen KG. 2019. An amplification-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol* 20:8. <https://doi.org/10.1186/s13059-018-1618-7>.
- Zhang K-Y, Gao Y-Z, Du M-Z, Liu S, Dong C, Guo F-B. 2019. Vgas: a viral genome annotation system. *Front Microbiol* 10:184. <https://doi.org/10.3389/fmicb.2019.00184>.