



# Complete Genome Sequence of a New *Chickpea Chlorotic Dwarf Virus* Strain Isolated from Tomato in Kenya, Obtained from Illumina Sequencing

E. K. Avedi,<sup>a,b,c</sup> C. D. Kilalo,<sup>b</sup> F. M. Olubayo,<sup>b</sup> I. Macharia,<sup>a</sup> A. O. Adediji,<sup>d</sup> E. M. Ateka,<sup>e</sup> E. M. Machuka,<sup>c</sup> J. M. Mutuku<sup>c</sup>

<sup>a</sup>Department of Phytosanitary Services and Biosafety, Kenya Plant Health Inspectorate Service, Nairobi, Kenya

<sup>b</sup>Department of Plant Science and Crop Protection, University of Nairobi, Nairobi, Kenya

<sup>c</sup>Biosciences Eastern and Central Africa, International Livestock Research Institute, Nairobi, Kenya

<sup>d</sup>Department of Crop Protection and Environmental Biology, Faculty of Agriculture, University of Ibadan, Ibadan, Nigeria

<sup>e</sup>Department of Horticulture, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

**ABSTRACT** High-throughput sequence analysis revealed the complete genome sequence of a novel, hitherto uncharacterized strain of *Chickpea chlorotic dwarf virus* (CpCDV) from tomato plants in Kenya. The sequence shared its highest nucleotide similarity (88.7%) with two CpCDV isolates from Burkina Faso.

Tomato (*Solanum lycopersicum*) is one of the world's most important vegetable crops. However, its production is constrained by viral diseases that cause yield losses (1). Many such viruses are found in the family *Geminiviridae* and the genus *Mastrevirus* with single-stranded DNA genomes (2). Seven *Mastrevirus* species infect dicotyledonous plants, including *Chickpea chlorotic dwarf virus* (CpCDV) (3–9). Nineteen CpCDV strains (strains A to S) have been described (9–11), but the presence of the virus in East Africa has not yet been reported. In this study, we identified a novel CpCDV strain, infecting tomato in Kenya, through metagenomic sequencing and phylogenetic analysis.

In 2018, five leaf samples were collected from tomato plants exhibiting stunting, yellowing, and leaf deformation in Naivasha, Nakuru County, Kenya. These samples were stored at  $-80^{\circ}\text{C}$ , and total genomic DNA was extracted according to a protocol described previously (12). Genomic DNA was pooled into one sample, Tom54, for library preparation using the Nextera DNA kit. Sequencing was performed on a MiSeq system using a v.3 kit (Illumina), with paired-end reads ( $2 \times 301$  cycles). Quality control of reads was performed using FastP software (v.0.20.0), and reads were mapped to the tomato genome (GenBank RefSeq accession number [GCA\\_000188115.3](https://www.ncbi.nlm.nih.gov/nuccore/GCA_000188115.3)) using Bowtie (v.2.3.4.3). Unmapped reads were assembled into contigs using MEGAHIT (v.1.1.3), and those representing single-stranded DNA sequences were verified using the Kaiju virus database (13). These contigs were subjected to BLASTN 2.9.0+ searches (14), while multiple sequence alignment was performed using ClustalW in BioEdit (v.7.2.3) (15). A maximum likelihood phylogenetic tree was constructed using the Jukes-Cantor model, as implemented in MEGA (v.6.06) (16). Protein prediction of open reading frames (ORFs) was performed using ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder>), and virus sequence similarities were determined using SDT (v.1.2), with pairwise gap deletions (17).

Tom54 yielded 320,207 reads (Q30, 97%), which were trimmed to 314,556 reads with an average length of 174 bp. *De novo* genome assembly yielded two contigs; one was 2,469 nucleotides long, with a G+C content of 51.73%, and the other was <138 bp long. A BLASTN-based search revealed both contigs to represent CpCDV, with the larger contig sharing the highest nucleotide similarity (88.7%) with the complete circular genome sequences of two CpCDV strains infecting tomato in Burkina Faso (GenBank

**Citation** Avedi EK, Kilalo CD, Olubayo FM, Macharia I, Adediji AO, Ateka EM, Machuka EM, Mutuku JM. 2020. Complete genome sequence of a new *Chickpea chlorotic dwarf virus* strain isolated from tomato in Kenya, obtained from Illumina sequencing. *Microbiol Resour Announc* 9:e01344-19. <https://doi.org/10.1128/MRA.01344-19>.

**Editor** Simon Roux, DOE Joint Genome Institute

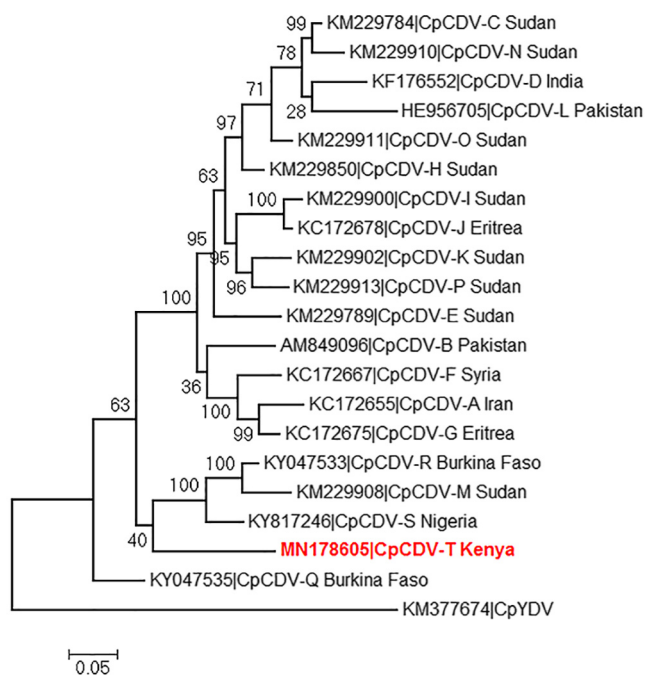
**Copyright** © 2020 Avedi et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to E. K. Avedi, [e.avedi@kephis.org](mailto:e.avedi@kephis.org).

**Received** 5 November 2019

**Accepted** 17 December 2019

**Published** 30 January 2020



**FIG 1** Maximum likelihood phylogenetic tree of the full genome of a novel *Chickpea chlorotic dwarf virus* strain from tomato in Kenya (CpCDV-T). Representative sequences of other CpCDV strains (CpCDV-A to CpCDV-S) were selected. The tree was rooted with an isolate of *Chickpea yellow dwarf virus* (CpYDV). The tree was drawn using MEGA (v.6.06).

accession numbers [KY047532](#) and [KY047533](#)) (11). Four ORFs, typical of *Mastrevirus* sp. genomes, were identified, i.e., V1, V2, C1, and C2. The V1 coat protein ORF encoded 283 amino acids, whereas V2 encoded the putative movement protein of 100 amino acids. C1 and C2 encoded the replication-associated proteins A and B via transcript splicing, with 302 and 143 amino acids, respectively. Phylogenetic analyses revealed a divergence of this isolate into a distinct and previously unreported clade, sharing a common ancestor with CpCDV strains M, R, and S (Fig. 1). Based on these properties, the larger contig was designated a complete, novel CpCDV genome and was deposited in GenBank under accession number [MN178605](#).

According to the species demarcation criteria for *Mastrevirus* spp. (18), our sequence qualifies to be considered to represent a distinct CpCDV strain; therefore, we propose the name CpCDV-T. Although Kenya shares a border with Sudan, where several CpCDV isolates have been reported (8), the high level of similarity of our isolate to one from West Africa suggests a possible introduction via trade. To our knowledge, this is the first report of CpCDV in Kenya.

**Data availability.** The sequence described here was deposited in GenBank under accession number [MN178605](#). Raw data were deposited under SRA accession number [PRJNA556271](#) with SRA identification number [SRR9737059](#), while the Tom54 sample was deposited under BioSample accession number [SAMN12346850](#).

## ACKNOWLEDGMENTS

We thank the Kenya Plant Health Inspectorate Service for supporting E.K.A. to undertake the Africa Biosciences Challenge Fund (ABCF) fellowship at the Biosciences Eastern and Central Africa-International Livestock Research Institute (BeCA-ILRI) Hub. We also appreciate SLU-Sweden for training E.K.A. and A.O.A. in genomic data analysis through funding from the Community Network for African Vector-Borne Plant Viruses (CONNECTED). CONNECTED is supported by the Global Challenges Research Fund (GCRF) from the Government of the United Kingdom. J.M.M. was partially funded by UK Biotechnological and Biological Sciences Research

Council GCRF grant number BB/P023223/1. This project was supported by the BecA-ILRI Hub through the ABCF program. The ABCF program is funded by the Australian Department for Foreign Affairs and Trade through the BecA-CSIRO partnership, the Syngenta Foundation for Sustainable Agriculture, the Bill and Melinda Gates Foundation, the UK Department for International Development, and the Swedish International Development Cooperation Agency.

## REFERENCES

- Hanssen IM, Lapidot M, Thomma BP. 2010. Emerging viral diseases of tomato crops. *Mol Plant Microbe Interact* 23:539–548. <https://doi.org/10.1094/MPMI-23-5-0539>.
- Zerbini FM, Briddon RW, Idris A, Martin DP, Moriones E, Navas-Castillo J, Rivera-Bustamante R, Roumagnac P, Varsani A. 2017. ICTV virus taxonomy profile: *Geminiviridae*. *J Gen Virol* 98:131–133. <https://doi.org/10.1099/jgv.0.000738>.
- Hadfield J, Thomas JE, Schwinghamer MW, Kraberger S, Stainton D, Dayaram A, Parry JN, Pande D, Martin DP, Varsani A. 2012. Molecular characterisation of dicot-infecting mastreviruses from Australia. *Virus Res* 166:13–22. <https://doi.org/10.1016/j.virusres.2012.02.024>.
- Horn NM, Reddy SV, Roberts IM, Reddy D. 1993. Chickpea chlorotic dwarf virus, a new leafhopper-transmitted geminivirus of chickpea in India. *Ann Appl Biol* 122:467–479. <https://doi.org/10.1111/j.1744-7348.1993.tb04050.x>.
- Radouane N, Ezrari S, Accotto GP, Benjelloun M, Lahlali R, Tahiri A, Vaira AM. 2019. First report of *Chickpea chlorotic dwarf virus* in watermelon (*Citrullus lanatus*) in Morocco. *New Dis Rep* 39:2. <https://doi.org/10.5197/j.2044-0588.2019.039.002>.
- Thomas JE, Parry J, Schwinghamer M, Dann E. 2010. Two novel mastreviruses from chickpea (*Cicer arietinum*) in Australia. *Arch Virol* 155:1777–1788. <https://doi.org/10.1007/s00705-010-0763-4>.
- Kraberger S, Harkins GW, Kumari SG, Thomas JE, Schwinghamer MW, Sharman M, Collings DA, Briddon RW, Martin DP, Varsani A. 2013. Evidence that dicot-infecting mastreviruses are particularly prone to inter-species recombination and have likely been circulating in Australia for longer than in Africa and the Middle East. *Virology* 444:282–291. <https://doi.org/10.1016/j.virol.2013.06.024>.
- Kraberger S, Kumari SG, Hamed AA, Gronenborn B, Thomas JE, Sharman M, Harkins GW, Muhire BM, Martin DP, Varsani A. 2015. Molecular diversity of *Chickpea chlorotic dwarf virus* in Sudan: high rates of intra-species recombination: a driving force in the emergence of new strains. *Infect Genet Evol* 29:203–215. <https://doi.org/10.1016/j.meegid.2014.11.024>.
- Kraberger S, Mumtaz H, Claverie S, Martin DP, Briddon RW, Varsani A. 2015. Identification of an Australian-like dicot-infecting mastrevirus in Pakistan. *Arch Virol* 160:825–830. <https://doi.org/10.1007/s00705-014-2299-5>.
- Zia-Ur-Rehman M, Hameed U, Herrmann HW, Iqbal MJ, Haider MS, Brown JK. 2015. First report of *Chickpea chlorotic dwarf virus* infecting tomato crops in Pakistan. *Plant Dis* 99:1287. <https://doi.org/10.1094/PDIS-02-15-0202-PDN>.
- Ouattara A, Tiendrébéogo F, Lefeuvre P, Hoareau M, Claverie S, Traoré EV, Barro N, Traoré O, Varsani A, Lett JM. 2017. New strains of *Chickpea chlorotic dwarf virus* discovered on diseased papaya and tomato plants in Burkina Faso. *Arch Virol* 162:1791–1794. <https://doi.org/10.1007/s00705-017-3262-z>.
- Pratap D, Kashikar AR, Mukherjee SK. 2011. Molecular characterization and infectivity of a Tomato leaf curl New Delhi virus variant associated with newly emerging yellow mosaic disease of eggplant in India. *Virol J* 8:305. <https://doi.org/10.1186/1743-422X-8-305>.
- Menzel P, Ng KL, Krogh A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 7:11257. <https://doi.org/10.1038/ncomms11257>.
- Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. *J Comput Biol* 7:203–214. <https://doi.org/10.1089/10665270050081478>.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30:2725–2729. <https://doi.org/10.1093/molbev/mst197>.
- Muhire BM, Varsani A, Martin DP. 2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One* 9:e108277. <https://doi.org/10.1371/journal.pone.0108277>.
- Muhire B, Martin DP, Brown JK, Navas-Castillo J, Moriones E, Zerbini FM, Rivera-Bustamante R, Malathi VG, Briddon RW, Varsani A. 2013. A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus *Mastrevirus* (family *Geminiviridae*). *Arch Virol* 158:1411–1424. <https://doi.org/10.1007/s00705-012-1601-7>.