



Deep Sequencing Reveals a Divergent *Ugandan cassava brown streak virus* Isolate from Malawi

Willard Mbewe,^{a,b*} Stephan Winter,^c Settumba Mukasa,^b Fred Tairo,^d
Peter Sseruwagi,^d Joseph Ndunguru,^d Siobain Duffy^a

Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, New Jersey, USA^a; School of Agricultural Science, Makerere University, Kampala, Uganda^b; DSMZ Plant Virus Department, Leibniz Institute, Braunschweig, Germany^c; Mikochei Agricultural Research Institute, Dar es Salaam, Tanzania^d

ABSTRACT Illumina sequencing of RNA from a cassava cutting from northern Malawi produced a genome of *Ugandan cassava brown streak virus* (UCBSV-MW-NB7_2013). Sequence comparisons revealed stronger similarity to an isolate from nearby Tanzania (93.4% pairwise nucleotide identity) than to those previously reported from Malawi (86.9 to 87.0%).

Ugandan cassava brown streak virus (UCBSV) is one of two distinct ipomoviruses (family: *Potyviridae*) that cause the devastating cassava brown streak disease (CBSD) (1, 2) of cassava (*Manihot esculenta*) in eastern and central Africa. In Malawi in 2013, cassava cuttings with conspicuous symptoms of CBSD were collected, transferred to the DSMZ Plant Virus Department, and rooted and grown under glasshouse conditions optimized for cassava and viruses. Virus assays were conducted to identify and discriminate between CBSV and UCBSV essentially following the protocol of Mohamed et al. (3). Total RNA of a cassava cutting from Nkhata Bay District was extracted using an RNeasy Plant minikit (Qiagen, Germany), and rRNA was removed using a RiboMinus plant kit (Invitrogen). Following random cDNA synthesis and second-strand synthesis with random octamer primers, an Illumina library was prepared using a Nextera XT library kit and subsequently run on a MiSeq instrument as paired-end reads (2 × 301). In total, 1,241,000 reads were generated. *De novo* assembly was conducted in both CLC Workbench and Geneious version 10.0.5 (4). A well-supported sequence contig was obtained (average 300× coverage) that showed highest nucleotide identity with UCBSV. This 9,070-nucleotide sequence, UCBSV-MW-NB7_2013, was compared to all complete UCBSV genome sequences available in GenBank (Tanzania, FJ039520; Kenya, FN433931, FN433930, KR911721, KR911722, KR911723, KR911724, KR911725, KR911726, and KR911727; Uganda, HM181930, FJ185044, HG96522, and FN434109; and Malawi, FN433932 and FN433933) using MEGA version 7 (5).

Pairwise nucleotide sequence similarity of UCBSV-MW-NB7_2013 with UCBSV genome sequences showed 87.9 to 93.4% nucleotide identities, with the highest identity to the sequence of the UCBSV MLB3 isolate (FJ039520) from Tanzania. The amino acid sequence of both polyproteins was 95.8% identical. UCBSV-MW-NB7_2013 was more distantly related to previously characterized Malawian isolates, with a maximum percentage identity of 87.0% with isolate Ma 43 (FN433933). The 230 nucleotides of the 3' untranslated region (UTR) had high identity to extant UCBSV genomes (97% identity to Ma 43, with the terminal 52 nucleotides identical). The terminal 8 nucleotides of the 5' UTR of our UCBSV isolate (TCACATAC) were very divergent compared to all previously sequenced UCBSVs, which is most probably artifactual due to the low coverage at the ends of the contig. However, newly sequenced full UCBSV genomes from Kenya have

Received 3 July 2017 Accepted 7 July 2017 Published 17 August 2017

Citation Mbewe W, Winter S, Mukasa S, Tairo F, Sseruwagi P, Ndunguru J, Duffy S. 2017. Deep sequencing reveals a divergent *Ugandan cassava brown streak virus* isolate from Malawi. *Genome Announc* 5:e00818-17. <https://doi.org/10.1128/genomeA.00818-17>.

Copyright © 2017 Mbewe 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 Willard Mbewe, mbewewillard@yahoo.co.uk.

* Present address: Willard Mbewe, Department of Agriculture Research Services, Bvumbwe Agriculture Research Station, Limbe, Malawi.

revealed unexpected diversity in the 5' UTR (6), and further sequencing of UCBSV isolates could reveal that the 5' terminus is more variable than previously thought.

A BLAST search for the UCBSV-MW-NB7_2013 sequence against the GenBank non-redundant database showed the top two hits to be the Tanzanian full genome and the only other UCBSV sequence from Tanzania, the partial genome sequence KR108839. The similarities between this isolate from northern Malawi and UCBSV from Tanzania suggest that it may have been introduced from Tanzania, since is not as closely related to CBSV-causing viruses previously detected in Malawi.

Accession number(s). The whole-genome sequence of UCBSV-MW-NB7_2013 has been deposited in DDBJ/ENA/GenBank under the accession no. [MF379362](https://doi.org/10.1093/bioinformatics/bts199).

ACKNOWLEDGMENTS

We acknowledge full financial support from the Bill & Melinda Gates Foundation (grants 51466 and 1052391) and the UK Department for International Development, under the auspices of the Mikocheni Agricultural Research Institute through the Disease Diagnostics for Sustainable Cassava Production in Africa project. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

We also acknowledge Ehud Zelzion of the Rutgers University Genome Cooperative for assembling the genome in CLC Workbench.

REFERENCES

- Mbanzibwa DR, Tian YP, Tugume AK, Mukasa SB, Tairo F, Kyamanywa S, Kullaya A, Valkonen JPT. 2009. Genetically distinct strains of *Cassava brown streak virus* in the Lake Victoria basin and the Indian Ocean coastal area of East Africa. *Arch Virol* 154:353–359. <https://doi.org/10.1007/s00705-008-0301-9>.
- Winter S, Koerber M, Stein B, Pietruszka A, Paape M, Butgereit A. 2010. Analysis of cassava brown streak viruses reveals the presence of distinct virus species causing cassava brown streak disease in East Africa. *J Gen Virol* 91:1365–1372. <https://doi.org/10.1099/vir.0.014688-0>.
- Mohammed IU, Abarshi MM, Muli B, Hillocks RJ, Maruthi MN. 2012. The symptom and genetic diversity of cassava brown streak viruses infecting cassava in East Africa. *Adv Virol* 2012:795697. <https://doi.org/10.1155/2012/795697>.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>.
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874. <https://doi.org/10.1093/molbev/msw054>.
- Kathurima T, Nyende A, Kiarie S, Ateka E. 2016. Genetic diversity and distribution of cassava brown streak virus and ugandan cassava brown streak virus in major cassava-growing regions in Kenya. *Annu Res Rev Biol* 10:1–9. <https://doi.org/10.9734/ARRB/2016/26879>.