Next-Generation Sequencing Reveals the First Complete Genome Sequence of *Cowpea aphid-borne mosaic virus* from Uganda

E. K. Mbeyagala,[a,b] P. Tukamuhabwa,[b] J. Bisikwa,[b] T. Holton,[c] S. B. Mukasa[b]

[a]National Semi-Arid Resources Research Institute (NaSARRI)/NARO, Serere, Uganda
[b]School of Agricultural Sciences, Makerere University, Kampala, Uganda
[c]Biosciences eastern and central Africa, International Livestock Research Institute, Nairobi, Kenya

ABSTRACT We present here the first complete genome sequence of *Cowpea aphid-borne mosaic virus* (CABMV) isolated from cowpea in Uganda and compare it with five CABMV complete genome sequences from Brazil (2), India (2), and Zimbabwe (1). It most resembled the genomes of two Brazilian isolates (MG-Avr and BR1) and one Indian isolate (RR3).

Viral diseases of cowpea are prevalent in the main cowpea-growing areas of Uganda (1). *Cowpea aphid-borne mosaic virus* (CABMV) is a single-stranded positive-sense RNA virus in the genus *Potyvirus*, family *Potyviridae*. Although it was first reported in Uganda in 1981 (2), currently, there is no full or partial genome sequence for any isolate from Uganda. Presently, only five complete genome sequences of CABMV are available in GenBank (accession numbers AF348210 [Zimbabwe], HQ880243 [Brazil], HQ880242 [Brazil], KM655833 [India], and KM597165 [India]). Leaf samples from cowpea plants showing typical CABMV symptoms, such as mosaic pattern, chlorosis, mottling, leaf curling, and necrosis, were collected from farmers' fields in eastern, northern, and central Uganda. The samples were dried using silica gel. Total RNA was extracted from the samples using TRI reagent (Sigma-Aldrich, St. Louis, MO, USA). RNA quality and concentration were measured using a NanoDrop spectrophotometer (Thermo Fisher Scientific, MA, USA). RNA quality was further checked using Tris-acetate-EDTA (TAE)-formamide agarose gel electrophoresis (3). Samples that had intact 18S and 28S subunits on agarose gel and without genomic DNA contamination were selected for library preparation and sequencing. Four libraries, each comprising 10 samples from the total RNA extracts, were prepared using the TruSeq stranded mRNA sample preparation kit (catalog number RS-122-2101; Illumina, San Diego, CA, USA). The final size and concentration of each library were checked using the 2100 Bioanalyzer desktop system (Agilent, Santa Clara, CA, USA) and quantified using a Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA, USA). Paired-end sequencing (2 × 300 bp) was performed using a MiSeq Desktop sequencer platform at Biosciences eastern and central Africa (International Livestock Research Institute [BecA-ILRI], Nairobi, Kenya). Read quality checking, trimming (4), and *de novo* sequence assembly were carried out using CLC Genomics Workbench version 7.0.4 (CLC bio, Qiagen, CA, USA). The library made from eastern Ugandan samples yielded 1,761,488 reads, and after trimming, 1,329,444 reads remained. *De novo* assembly yielded 10,174 contigs.

A BLASTN search against a plant virus database of 57,417 sequences using NCBI-BLAST-2.2.29+ (5) identified a contig with the highest identity to CABMV. The BLASTN search was carried out using default settings (-word_size, 28; -reward, 1; -penalty, -2; -gapopen, 0; -gapextend, 2.5). This 9,904-nucleotide (nt) sequence, from isolate Serere1, had a single open reading frame (ORF) composed of 10 mature proteins typical of

the *Potyvirus* genus (6, 7). A pairwise nucleotide comparison of both the entire genome and polyprotein sequences revealed that the Serere1 genome most resembled those of Brazilian isolates MG-Avr (GenBank accession number HQ880243) and BR1 (HQ880242) and an Indian isolate, RR3 (KM597165), with 77.8%, 77.4%, and 77.5% nt identity, respectively.

Accession number(s). The complete genome sequence of CABMV isolate Serere1 was deposited in DDBJ/ENA/GenBank under the accession number KT726938.

ACKNOWLEDGMENTS

REFERENCES

1. Edema R, Adipala E, Florini DA. 1997. Influence of season and cropping system on occurrence of cowpea diseases in Uganda. Plant Dis 81: 465–468. https://doi.org/10.1094/PDIS.1997.81.5.465.
2. Anonymous. 1981. Viruses, virus-like diseases and mycoplasma-like diseases, p 95–110. *In* Pest control in tropical grain legumes. Centre for Overseas Pest Research, London, United Kingdom.
3. Masek T, Vopalensky V, Suchomelova P, Pospisek M. 2005. Denaturing RNA electrophoresis in TAE agarose gels. Anal Biochem 336:46–50. https://doi.org/10.1016/j.ab.2004.09.010.
4. Del Fabbro C, Scalabrin S, Morgante M, Giorgi FM. 2013. An extensive evaluation of read trimming effects on Illumina NGS data analysis. PLoS One 8:e85024. https://doi.org/10.1371/journal.pone.0085024.
5. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389.
6. Adams MJ, Antoniw JF, Fauquet CM. 2005. Molecular criteria for genus and species discrimination within the family *Potyviridae*. Arch Virol 150: 459–479. https://doi.org/10.1007/s00705-004-0440-6.
7. Adams MJ, Antoniw JF, Beaudoin F. 2005. Overview and analysis of the polyprotein cleavage sites in the family *Potyviridae*. Mol Plant Pathol 6:471–487. https://doi.org/10.1111/j.1364-3703.2005.00296.x.