



Analysis of an RNA-seq Strand-Specific Library from an East Timorese Cucumber Sample Reveals a Complete *Cucurbit aphid-borne yellows virus* Genome

Solomon Maina,^{a,b,c} Owain R. Edwards,^{c,d} Luis de Almeida,^e Abel Ximenes,^f Roger A. C. Jones^{b,c,g}

School of Agriculture and Environment, Faculty of Science, University of Western Australia, Crawley, Western Australia, Australia^a; Institute of Agriculture, Faculty of Science, University of Western Australia, Crawley, Western Australia, Australia^b; Cooperative Research Centre for Plant Biosecurity, Canberra, Australian Capital Territory, Australia^c; CSIRO Land and Water, Floreat Park, Western Australia, Australia^d; Seeds of Life Project, Ministry of Agriculture, Forestry and Fisheries, Dili, East Timor^e; DNQB-Plant Quarantine, International Airport Nicolau Lobato Comoro, Dili, East Timor^f; Department of Agriculture and Food Western Australia, South Perth, Western Australia, Australia^g

ABSTRACT Analysis of an RNA-seq library from cucumber leaf RNA extracted from a fast technology for analysis of nucleic acids (FTA) card revealed the first complete genome of *Cucurbit aphid-borne yellows virus* (CABYV) from East Timor. We compare it with 35 complete CABYV genomes from other world regions. It most resembled the genome of the South Korean isolate HD118.

As part of a biosecurity project to examine possible genetic connectivity between viruses infecting crops in northern Australia and nearby Southeast Asian countries, virus genomes from East Timorese and Australian plant samples were compared (1–8). During July 2015, 15 and 22 leaf samples were collected from cucurbit plants with virus-like symptoms in East Timor and northwest Australia, respectively, and subjected to next-generation sequencing. The 15 East Timorese samples were blotted onto fast technology for analysis of nucleic acids (FTA) cards before dispatch to Australia. A complete genome of *Cucurbit aphid-borne yellows virus* (CABYV) was obtained from cucumber (*Cucumis sativus*) sample TM50 from Aileu in East Timor. CABYV was described in 1992 as causing yellow leaf symptoms in cucurbit plants in France (9) and belongs to the genus *Polevirus* within the family *Luteoviridae* (10). CABYV is a phloem-limited virus transmitted persistently by aphid vectors, including *Aphis gossypii* and *Myzus persicae* (9). It consists of a single-stranded positive-sense RNA molecule with a length of 5.7 kb (10). In sample TM50, analysis of nonpolyadenylated transcripts derived from RNA-seq strand-specific libraries (1–8, 11–14) prepared from RNA extracted from FTA cards (1–8, 15, 16) detected CABYV (designated isolate CABYV AL50).

Total RNA was extracted from FTA card disks using a ZR Plant RNA MiniPrep™ kit (Zymo Research) and treated with RNase-free DNase (Invitrogen). Extracts were subjected to library preparation using a Ribo-Zero plant kit (catalog no. RS-122-2401, Illumina) with the amount of Agencourt AMPure XP beads (Beckman Coulter, Inc.) increased to 20% in every clean-up and no RNA fragmentation step. Libraries were sequenced using a MiSeq platform with a 2 × 251 version 2 kit (Illumina) and a 1% PhiX version 3 spike control. Reads were assembled and genomes annotated using CLC Genomics Workbench version 6.5 (CLC bio) and Geneious version 8.1.7 (Biomatters) (1–8, 17, 18).

CABYV isolate AL50 yielded 4,670,138 reads and, after trimming, 4,058,029 remained. Although AL50's RNA integrity number (RIN) was only 2, this high number of

Received 16 March 2017 **Accepted** 20 March 2017 **Published** 11 May 2017

Citation Maina S, Edwards OR, de Almeida L, Ximenes A, Jones RAC. 2017. Analysis of an RNA-seq strand-specific library from an East Timorese cucumber sample reveals a complete *Cucurbit aphid-borne yellows virus* genome. *Genome Announc* 5:e00320-17. <https://doi.org/10.1128/genomeA.00320-17>.

Copyright © 2017 Maina 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 Roger A. C. Jones, roger.jones@uwa.edu.au.

reads was still obtained. *De novo* assembly generated 229 contigs and 130,416 reads mapped to the contig of interest with 31,129× coverage. The genome obtained had 5,677 nucleotides (nt) and 6 open reading frames (ORFs) organized into 2 clusters, as with other poleroviruses (19). A BLAST-based search (20), revealed that AL50 most resembled the South Korean isolate HD118 (KR231951) with 88.1% nt identity. When AL50's P3 gene sequence was compared with five CABYV-like P3 sequences from Tasmania, Australia available in GenBank, it most resembled isolate 18 (HQ543088) with 81.2% nt identity. The extent of P3 gene sequence divergence between AL50 and Tasmanian isolate 18 was too great to indicate any genetic connectivity. Further sequencing of CABYV in Tasmania and mainland Australia is required to obtain complete genomes and compare them with genomes from neighboring Southeast Asian countries.

Accession number(s). The GenBank accession number is [KY617826](https://www.ncbi.nlm.nih.gov/nucl/1000000000).

ACKNOWLEDGMENTS

Martin J. Barbetti and Mingpei You of the School of Agriculture and Environment, University of Western Australia (UWA), provided administrative support. The UWA ARC, Centre of Excellence in Plant Energy Biology and School of Molecular Sciences, and Laura Boykin also provided initial administrative support at the beginning of this project.

The Cooperative Research Centre for Plant Biosecurity and UWA provided scholarship and operating funds to Solomon Maina. The Commonwealth Scientific and Industrial Research Organization provided additional operating funds.

REFERENCES

- Maina S, Edwards OR, de Almeida L, Ximenes A, Jones RAC. 2016. Complete genome sequences of the *Carlavirus Sweet potato chlorotic fleck virus* from East Timor and Australia. *Genome Announc* 4(3):e00414-16. <https://doi.org/10.1128/genomeA.00414-16>.
- Maina S, Edwards OR, Jones RAC. 2016. First complete genome sequence of *Pepper vein yellows virus* from Australia. *Genome Announc* 4(3):e00450-16. <https://doi.org/10.1128/genomeA.00450-16>.
- Maina S, Edwards OR, de Almeida L, Ximenes A, Jones RAC. 2016. Complete genome sequences of the *Potyvirus Sweet potato virus 2* from East Timor and Australia. *Genome Announc* 4(3):e00504-16. <https://doi.org/10.1128/genomeA.00504-16>.
- Maina S, Edwards OR, de Almeida L, Ximenes A, Jones RAC. 2016. First complete genome sequence of *Suakwa aphid-borne yellows virus* from East Timor. *Genome Announc* 4(4):e00718-16. <https://doi.org/10.1128/genomeA.00718-16>.
- Maina S, Edwards OR, Barbetti MJ, de Almeida L, Ximenes A, Jones RAC. 2016. Deep sequencing reveals complete genome of *Sweet potato virus G* from East Timor. *Genome Announc* 4(5):e00957-16. <https://doi.org/10.1128/genomeA.00957-16>.
- Maina S, Edwards OR, de Almeida L, Ximenes A, Jones RAC. 2016. First complete genome sequence of *Bean common mosaic necrosis virus* from East Timor. *Genome Announc* 4(5):e01049-16. <https://doi.org/10.1128/genomeA.01049-16>.
- Maina S, Edwards OR, de Almeida L, Ximenes A, Jones RAC. 2017. Metagenomic analysis of cucumber RNA from East Timor reveals an *Aphid lethal paralysis virus* genome. *Genome Announc* 5(2):e01445-16. <https://doi.org/10.1128/genomeA.01445-16>.
- Maina S, Coutts BA, Edwards OR, de Almeida L, Ximenes A, Jones RAC. 2017. *Papaya ringspot virus* populations from East Timorese and Northern Australian cucurbit crops: biological and molecular properties, and absence of genetic connectivity. *Plant Disease* [Epub ahead of print.]. <https://doi.org/10.1094/PDIS-10-16-1499-RE>.
- Lecoq H, Bourdin D, Wipf-Scheibel C, Bon M, Lot H, Lemaire O, Herrbach E. 1992. A new yellowing disease of cucurbits caused by a luteovirus, *Cucurbit aphid-borne yellows virus*. *Plant Pathol* 41:749-761. <https://doi.org/10.1111/j.1365-3059.1992.tb02559.x>.
- D'Arcy CJ, Domier LL. 2005. Luteoviridae. In Mayo MA, Maniloff J, Desseberger U, Ball LA, Fauquet CM (ed). *Virus taxonomy: VIIIth report of the International Committee on Taxonomy of Viruses*. Academic Press, New York, NY.
- Morlan JD, Qu K, Sinicropi DV. 2012. Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue. *PLoS One* 7:e42882. <https://doi.org/10.1371/journal.pone.0042882>.
- Wu Q, Ding SW, Zhang Y, Zhu S. 2015. Identification of viruses and viroids by next-generation sequencing and homology-dependent and homology-independent algorithms. *Annu Rev Phytopathol* 53:425-444. <https://doi.org/10.1146/annurev-phyto-080614-120030>.
- Nagano AJ, Honjo MN, Mihara M, Sato M, Kudoh H. 2015. Detection of plant viruses in natural environments by using RNA-Seq. *Methods Mol Biol* 1236:89-98. https://doi.org/10.1007/978-1-4939-1743-3_8.
- Sigurgeirsson B, Emanuelsson O, Lundeberg J. 2014. Analysis of stranded information using an automated procedure for strand specific RNA sequencing. *BMC Genomics* 15:631. <https://doi.org/10.1186/1471-2164-15-631>.
- Chikh-Ali M, Bosque-Pérez NA, Vander Pol D, Sembel D, Karasev AV. 2016. Occurrence and molecular characterization of recombinant *Potato virus Y^{NTN}* isolates from Sulawesi, Indonesia. *Plant Dis* 100:269-275. <https://doi.org/10.1094/PDIS-07-15-0817-RE>.
- Ndunguru J, Taylor NJ, Yadav J, Aly H, Legg JP, Aveling T, Thompson G, Fauquet CM. 2005. Application of FTA technology for sampling, recovery and molecular characterization of viral pathogens and virus-derived transgenes from plant tissues. *Viro J* 2:45. <https://doi.org/10.1186/1743-422X-2-45>.
- Kehoe MA, Coutts BA, Buirchell BJ, Jones RAC. 2014. Plant virology and next generation sequencing: experiences with a *Potyvirus*. *PLoS One* 9:e104580. <https://doi.org/10.1371/journal.pone.0104580>.
- Katoh K, Misawa K, Kuma K-i, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059-3066. <https://doi.org/10.1093/nar/gk4436>.
- Mayo MA, Ziegler-Graff V. 1996. Molecular biology of luteoviruses. *Adv Virus Res* 46:413-460.
- Bao Y, Chetverin V, Tatusova T. 2014. Improvements to pairwise sequence comparison (PASC): a genome-based web tool for virus classification. *Arch Virol* 159:3293-3304. <https://doi.org/10.1007/s00705-014-2197-x>.