



Metagenomic Analysis of Cucumber RNA from East Timor Reveals an *Aphid lethal paralysis 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 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 We present here the first complete genomic *Aphid lethal paralysis virus* (ALPV) sequence isolated from cucumber plant RNA from East Timor. We compare it with two complete ALPV genome sequences from China, and one each from Israel, South Africa, and the United States. It most closely resembled the Chinese isolate LGH genome.

As part of a project to examine possible connectivity between viruses infecting crops in northern Australia and nearby Southeast Asian countries, virus genomes from plant samples from East Timor and Australia were compared (1–6). During July and September 2015, 15- and 21-cucurbit leaf samples were collected in East Timor and Broome in northwest Australia, respectively, and subjected to next-generation sequencing. A complete genome of *Aphid lethal paralysis virus* (ALPV) was obtained from East Timorese cucumber (*Cucumis sativus*) sample TM19, but not from any other samples. ALPV belongs to the genus *Cripavirus*, family *Dicistroviridae*. It was first identified in 1988 in South Africa and contains a 9.7-kb polyadenylated ssRNA (7, 8). Five complete ALPV genomes are available in GenBank: two from China, and one each from the United States, Israel, and South Africa (9, 10). Metagenomics offers prospects for discovering novel viruses with distinct genome ontologies (11–16). RNA-Seq with rRNA-Plant depletion provides reliable metagenomic detection of polyadenylated and nonpolyadenylated RNA viruses (17). This approach detected ALPV in sample TM19.

The 15 East Timorese samples were blotted onto Fast Technology for Analysis of nucleic acids (FTA) cards (18) before dispatch to Australia. The Australian samples studied were recently collected leaves. Total RNA was extracted from both sample types using a ZR plant RNA mini prep kit (Zymo Research). The total RNA extracts were treated with RNase-free DNase (Invitrogen) and measured using Qubit (Invitrogen). RNA integrity was confirmed using RNA screen Tape (TapeStation 2200, Agilent Technologies). Libraries were prepared from total RNA using a TruSeq stranded Total RNA sample preparation Ribo-Zero Plant kit (catalogue no. RS-122-2401, Illumina) (1–6). Sequencing was by MiSeq using a V2 kit (Illumina) with 2 × 151 cycles of paired-end reads. Reads were assembled and genomes annotated using CLC Genomics Workbench version 6.5 (CLC bio) and Geneious version 8.1.7 (Biomatters) (19). Further alignment was by MAFFT (20).

FTA card sample TM19 yielded 2,287,025 reads and, after trimming, 2,248,678 remained. *De novo* assembly generated 270 contigs and 858,904 reads mapped to the

Received 26 October 2016 **Accepted** 8 November 2016 **Published** 12 January 2017

Citation 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:e01445-16. <https://doi.org/10.1128/genomeA.01445-16>.

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.

contig of interest with coverage of 30,410×. The final complete genome length was 9,789 nucleotides (nt). As with other members of the genus *Cripavirus*, it coded for a nonstructural polyprotein (encoding putative helicase, protease, and RNA-dependent RNA polymerase) and a capsid protein precursor (encoding proteins VP1 to VP4). A BLAST-based search with the pairwise sequence comparison (PASC) tool (21), revealed that the TM19 genome sequence most resembled Chinese isolate LGH (accession no. KR021407) with 93.8% nt identity. As mentioned above, the purpose of these studies was to search for plant rather than insect viruses. Possibly, ALPV resembles a related aphid virus, *Rhopalosiphum padi virus*, in spreading horizontally from the initial viruliferous aphid feeding site rapidly via the phloem, allowing nonviruliferous aphids to acquire it throughout the plant (22), which might explain its detection in a cucumber sample.

Accession number(s). This sequence was deposited in GenBank under the accession number [KX830963](https://www.ncbi.nlm.nih.gov/nuclseq/KX830963).

ACKNOWLEDGMENTS

Martin 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 Chemistry and Biochemistry, and Laura Boykin also provided initial administrative support at the beginning of this project.

The Cooperative Research Centre for Plant Biosecurity and the University of Western Australia provided scholarship and operating funds to Solomon Maina. The Commonwealth Scientific and Industrial Research Organisation 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 the complete genome sequence 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>.
- Dombrovsky A, Luria N. 2013. The *Nerium oleander* aphid *Aphis nerii* is tolerant to a local isolate of *Aphid lethal paralysis virus* (ALPV). *Virus Genes* 46:354–361. <https://doi.org/10.1007/s11262-012-0846-2>.
- Williamson C, Rybicki EP, Kasdorf GGF, Von Wechmar MB. 1988. Characterization of a new picorna-like virus isolated from aphids. *J Gen Virol* 69:787–795. <https://doi.org/10.1099/0022-1317-69-4-787>.
- Liu S, Vijayendran D, Carrillo-Tripp J, Miller WA, Bonning BC. 2014. Analysis of new *Aphid lethal paralysis virus* (ALPV) isolates suggests evolution of two ALPV species. *J Gen Virol* 95:2809–2819. <https://doi.org/10.1099/vir.0.069765-0>.
- Van Munster M, Dulleman AM, Verbeek M, Van Den Heuvel JF, Clérivet A, Van Der Wilk F. 2002. Sequence analysis and genomic organization of aphid lethal paralysis virus: a new member of the family *Dicistroviridae*. *J Gen Virol* 83:3131–3138. <https://doi.org/10.1099/0022-1317-83-12-3131>.
- Ge X, Li Y, Yang X, Zhang H, Zhou P, Zhang Y, Shi Z. 2012. Metagenomic analysis of viruses from bat fecal samples reveals many novel viruses in insectivorous bats in China. *J Virol* 86:4620–4630. <https://doi.org/10.1128/JVI.06671-11>.
- Barba M, Czosnek H, Hadidi A. 2014. Historical perspective, development and applications of next-generation sequencing in plant virology. *Viruses* 6:106–136. <https://doi.org/10.3390/v6010106>.
- Adams I, Fox A. 2016. Diagnosis of plant viruses using next-generation sequencing and Metagenomic analysis. p 323–335. *In* Wang A, Zhou X (ed), *Current research topics in plant virology*. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-319-32919-2_14.
- Ndunguru J, Sseruwagi P, Tairo F, Stomeo F, Maina S, Djikeng A, Djikeng A, Kehoe M, Boykin LM. 2015. Analyses of twelve new whole genome sequences of *Cassava brown streak viruses* and *Ugandan cassava brown streak viruses* from East Africa: diversity, supercomputing and evidence for further speciation. *PLoS One* 10:e0139321. <https://doi.org/10.1371/journal.pone.0139321>.
- 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.
- 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. *Virol 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>.

20. Katoh K, Misawa K, Kuma KI, 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/gkf436>.
21. Bao Y, Chetvernin 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>.
22. Ban L, Didon A, Jonsson LM, Glinwood R, Delp G. 2007. An improved detection method for the *Rhopalosiphum padi virus* (RhPV) allows monitoring of its presence in aphids and movement within plants. *J Virol Methods* 142:136–142. <https://doi.org/10.1016/j.jviromet.2007.01.014>.