



Analysis of an RNA-seq Strand-Specific Library Sample Reveals a Complete Genome of *Hardenbergia mosaic virus* from Native Wisteria, an Indigenous Virus from Southwest Australia

Solomon Maina,^{a,b} Roger A. C. Jones^{b,c}

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; Department of Agriculture and Food Western Australia, South Perth, Western Australia, Australia^c

ABSTRACT Analysis of an RNA-seq strand-specific library revealed a complete genome of *Hardenbergia mosaic virus* (HarMV) from RNA extracted from a native wisteria (*Hardenbergia comptoniana*) plant from southwest Australia. We compared it with eight other complete HarMV genomes. It most resembled (85.8% nucleotide identity) the genome of HarMV isolate MD4-D.

In July 2016, an apical leaf sample showing obvious mosaic and deformation symptoms was collected from a plant of the indigenous southwest Australian species native wisteria (*Hardenbergia comptoniana*; family *Fabaceae*). The plant sampled also exhibited symptoms typical of phytoplasma infection (leaf chlorosis, proliferation of axillary shoots, and stunting). It was growing on a fence line bordering a playing field area in the Victoria Park suburb of Perth in southwest Australia. *Hardenbergia mosaic virus* (HarMV; genus *Potyvirus*, family *Potyviridae*) causes a conspicuous disease in native wisteria plants, is spread nonpersistently by aphids, and, like native wisteria itself, is indigenous to the region (1–5). It invades introduced lupin species (*Lupinus* spp.) at the interface of the local ancient ecosystem and recent agroecosystem (6, 7). Analysis of polyadenylated transcripts derived from RNAseq-stranded libraries (8–17) prepared from RNA extracted from the collected sample (designated VPK) detected one complete HarMV genome.

RNA was extracted from the VPK sample using a ZR Plant RNA MiniPrep™ kit (Zymo Research) and treated with RNase-free DNase (Invitrogen). The extract was subjected to library preparation using a TruSeq-stranded Ribo-Zero plant kit (Illumina, catalog no. RS-122-2401) and was subsequently subjected to quality control (8–17). The library was sent to Macrogen, Inc. (South Korea), where sequencing was done using the HiSeq 2500 platform with a TruSeq SBS version 4 kit (Illumina) with 151 cycles of paired-end reads. Reads were then assembled and genomes annotated using CLC Genomics Workbench version 6.5 (CLC bio) and Geneious version 8.1.7 (Biomatters) (18, 19).

The VPK sample yielded 13,609,056 reads and, after trimming, 12,972,958 remained. *De novo* assembly generated 25 contigs and 622,179 reads mapped to the contig of interest with a coverage of 9,107×. The complete genome obtained was named VPK-1. It consisted of 9,621 nucleotides (nt) and coded for 10 proteins, which is similar to other potyviruses (20, 21). There were eight other complete HarMV genomes already in GenBank (3, 7). A BLAST-based search (22) revealed that sequence VPK-1 most resembled the sequence of HarMV isolate MD4-D (KJ152157) with an 85.8% nt identity. In addition, the analysis revealed a partial sequence (6,002 nt in length) named VPK-2,

Received 8 May 2017 Accepted 8 June 2017 Published 27 July 2017

Citation Maina S, Jones RAC. 2017. Analysis of an RNA-seq strand-specific library sample reveals a complete genome of *Hardenbergia mosaic virus* from native wisteria, an indigenous virus from southwest Australia. Genome Announc 5:e00599-17. <https://doi.org/10.1128/genomeA.00599-17>.

Copyright © 2017 Maina and Jones. 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.

which belonged to a different HarMV strain. When sequences VPK-1 and MD4-D were truncated to match partial sequence VPK-2, the VPK-2 sequence was only distantly related to VPK-1 (83.1% nt identity) and MD4-D (82.5% nt identity). This study provides yet another example of the accuracy and reliability of high-throughput sequencing using both *de novo* and reference assembly approaches to separate different strains of the same virus present in the same sample (3, 7, 23, 24). There was no agriculture in southwest Australia until European colonization in 1829 but the region has a very diverse native flora. This diversity is reflected by the wide genetic diversity found within HarMV in this region (1–7). Such diversity is typical of indigenous viruses that have evolved over a very long period in isolation in native vegetation in remote regions (1–5, 25).

Accession number(s). Sequence VPK-1 has been deposited in GenBank under the accession number [MF040762](https://doi.org/10.1128/genomeA.01049-16).

ACKNOWLEDGMENTS

Martin J. Barbetti and Mingpei You of the School of Agriculture and Environment, University of Western Australia (UWA), provided administrative support. UWA provided operating funds.

REFERENCES

- Webster CG, Coutts BA, Jones RAC, Jones MGK, Wylie SJ. 2007. Virus impact at the interface of an ancient ecosystem and a recent agroecosystem: studies on three legume-infecting potyviruses in the southwest Australian floristic region. *Plant Pathol* 56:729–742. <https://doi.org/10.1111/j.1365-3059.2007.01653.x>.
- Coutts BA, Kehoe MA, Webster CG, Wylie SJ, Jones RAC. 2011. Indigenous and introduced potyviruses of legumes and *Passiflora* spp. from Australia: biological properties and comparison of coat protein nucleotide sequences. *Arch Virol* 156:1757–1774. <https://doi.org/10.1007/s00705-011-1046-4>.
- Wylie SJ, Jones MGK. 2011. Characterisation and quantitation of mutant and wild-type genomes of Hardenbergia mosaic virus isolates co-infecting a wild plant of *Hardenbergia comptoniana*. *Arch Virol* 156:1251–1255. <https://doi.org/10.1007/s00705-011-0965-4>.
- Jones RAC. 2009. Plant virus emergence and evolution: origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. *Virus Res* 141:113–130. <https://doi.org/10.1016/j.virusres.2008.07.028>.
- Vincent SJ, Coutts BA, Jones RAC. 2014. Effects of introduced and indigenous viruses on native plants: exploring their disease causing potential at the agro-ecological interface. *PLoS One* 9:e91224. <https://doi.org/10.1371/journal.pone.0091224>.
- Luo H, Wylie SJ, Coutts BA, Jones RAC, Jones MGK. 2011. A virus of an isolated indigenous flora spreads naturally to an introduced crop species. *Ann Appl Biol* 159:339–347. <https://doi.org/10.1111/j.1744-7348.2011.00496.x>.
- Kehoe MA, Coutts BA, Buirchell BJ, Jones RAC. 2014. Hardenbergia mosaic virus: crossing the barrier between native and introduced plant species. *Virus Res* 184:87–92. <https://doi.org/10.1016/j.virusres.2014.02.012>.
- 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 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 Dis* 101:985–993. <https://doi.org/10.1094/PDIS-10-16-1499-RE>.
- Maina S, Coutts BA, Edwards OR, de Almeida L, Kehoe MA, Ximenes A, Jones RAC. 2017. *Zucchini yellow mosaic virus* populations from East Timorese and northern Australian cucurbit crops: molecular properties, genetic connectivity and biosecurity implications. *Plant Dis* 101:1236–1245. <https://doi.org/10.1094/PDIS-11-16-1672-RE>.
- Maina S, Edwards OR, de Almeida L, Ximenes A, Jones RAC. 2017. Analysis of an RNA-seq strand-specific library from East Timorese cucumber sample reveals a complete *Cucurbit aphid-borne yellows virus* genome. *Genome Announc* 5(19):e00320-17.
- 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/gfk436>.
- 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>.
- Revers F, Garcia JA. 2015. Molecular biology of potyviruses. *Adv Virus Res* 92:101–199. <https://doi.org/10.1016/bs.aivir.2014.11.006>.
- Bao Y, Chetvermin 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>.
- 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>.
- Kehoe MA, Jones RAC. 2016. Improving *Potato virus Y* strain nomenclature: lessons from comparing isolates obtained over a 73-year period. *Plant Pathol* 65:322–333. <https://doi.org/10.1111/ppa.12404>.
- Jones RAC, Coutts BA. 2015. Spread of introduced viruses to new plants in natural ecosystems and the threat this poses to plant biodiversity. *Mol Plant Pathol* 16:541–545. <https://doi.org/10.1111/mp.12268>.