



# Genome Sequence of Cauliflower Mosaic Virus Identified in Earwigs (*Doru luteipes*) through a Metagenomic Approach

Márcio Tadeu Godinho,<sup>a</sup> Débora Pires Paula,<sup>a</sup> Arvind Varsani,<sup>b,c</sup> Simone Graça Ribeiro<sup>a</sup>

Embrapa Recursos Genéticos e Biotecnologia, Brasília, Brazil<sup>a</sup>; The Structural Biology Research Unit, Department of Clinical Laboratory Sciences, University of Cape Town, Cape Town, South Africa<sup>b</sup>; The Biodesign Center for Functional Microbiomics, Institute and School of Life sciences, Center for Evolution and Medicine, Arizona State University, Tempe, Arizona, USA<sup>c</sup>

**ABSTRACT** Here we report the first complete genome sequence of a cauliflower mosaic virus from Brazil, obtained from the gut content of the predator earwig (*Doru luteipes*). This virus has a genome of 8,030 nucleotides (nt) and shares 97% genome-wide identity with an isolate from Argentina.

Vector-enabled metagenomics (VEM), which allows for the identification of viruses that are vectored by insects and those in higher insect trophic levels, has proved to be a powerful approach for identifying both novel and known plant and animal viruses circulating in various ecosystems (1–7).

As part of a study to identify arthropod trophic interactions, the gut contents from the dermapteran insect predator earwig (*Doru luteipes*) were analyzed as described in reference 8. Ten *D. luteipes* adult individuals were collected in organic farms in Central Brazil, their guts were dissected, and the total DNA was extracted and pooled. This DNA was sequenced on an Illumina MiSeq platform using a TruSeq library (8). The reads were *de novo* assembled using Abyss v1.9 (9) with default parameters and a K-mer = 64. Contigs larger than 800 nucleotides (nt) were analyzed by BLASTx against a local viral database. A contig of 8,030 nt corresponding to the complete sequence of cauliflower mosaic virus (CaMV) was identified with 97% identity with a CaMV (isolate W260 sampled in Argentina; accession number JF809616). Posterior mapping of the Illumina reads using Bowtie (10) to the CaMV *de novo* assembled genome showed a 20× depth. CaMV is a caulimovirus (family: *Caulimoviridae*) and is found worldwide, especially in temperate regions (11). CaMV infects plant species of the Brassicaceae and Solanaceae families, and although it has been reported sporadically in Brazil since the 1960s (compiled by reference 12) this is the first complete genome sequence of a CaMV from Brazil. CaMV is transmitted by *Brevicoryne brassicae*, *Myzus persicae*, and other aphid species in a noncirculative manner.

In the gut of the earwigs, diamondback moth (*Plutella xylostella*) and harlequin ladybird (*Harmonia axyridis*) were identified together with three aphid-specific bacterial symbionts (*Hamiltonella* sp., *Regiella insecticola*, *Serratia symbiotica*) and an *Aphidius* parasitoid (8). Therefore, it is highly likely that CaMV was acquired from an infected plant by aphids or diamondback moth larvae which were preyed on by either the earwigs or by harlequin ladybirds, which appear to have been preyed on by the earwigs.

At higher insect trophic levels coupled with viral metagenomics, predators are very useful for identifying new or known viruses that are present in different species associated within these trophic networks.

Received 12 January 2017 Accepted 12 January 2017 Published 16 March 2017

**Citation** Godinho MT, Paula DP, Varsani A, Ribeiro SG. 2017. Genome sequence of cauliflower mosaic virus identified in earwigs (*Doru luteipes*) through a metagenomic approach. *Genome Announc* 5:e00043-17. <https://doi.org/10.1128/genomeA.00043-17>.

**Copyright** © 2017 Godinho 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 Simone Graça Ribeiro, [simone.ribeiro@embrapa.br](mailto:simone.ribeiro@embrapa.br).

**Accession number(s).** The complete genome of CaMV described in this report has been deposited in GenBank under accession number [KX434771](https://www.ncbi.nlm.nih.gov/nuccore/KX434771).

## ACKNOWLEDGMENTS

This research was sponsored by EMBRAPA, CNPq, and Fap-DF. M.T.G. was supported by a scholarship from CNPq.

## REFERENCES

- Rosario K, Seah YM, Marr C, Varsani A, Kraberger S, Stainton D, Moriones E, Polston JE, Duffy S, Breitbart M. 2015. Vector-enabled metagenomic (VEM) surveys using whiteflies (*Aleyrodidae*) reveal novel begomovirus species in the new and old worlds. *Viruses* 7:5553–5570. <https://doi.org/10.3390/v7102895>.
- Rosario K, Padilla-Rodriguez M, Kraberger S, Stainton D, Martin DP, Breitbart M, Varsani A. 2013. Discovery of a novel mastrevirus and alphasatellite-like circular DNA in dragonflies (*Epirocta*) from Puerto Rico. *Virus Res* 171:231–237. <https://doi.org/10.1016/j.virusres.2012.10.017>.
- Rosario K, Marr C, Varsani A, Kraberger S, Stainton D, Moriones E, Polston JE, Breitbart M. 2016. Begomovirus-associated satellite DNA diversity captured through vector-enabled metagenomic (VEM) surveys using whiteflies (*Aleyrodidae*). *Viruses* 8:36. <https://doi.org/10.3390/v8020036>.
- Rosario K, Dayaram A, Marinov M, Ware J, Kraberger S, Stainton D, Breitbart M, Varsani A. 2012. Diverse circular ssDNA viruses discovered in dragonflies (*Odonata: Epirocta*). *J Gen Virol* 93:2668–2681. <https://doi.org/10.1099/vir.0.045948-0>.
- Rosario K, Capobianco H, Ng TF, Breitbart M, Polston JE. 2014. RNA viral metagenome of whiteflies leads to the discovery and characterization of a whitefly-transmitted carlavirus in North America. *PLoS One* 9:e86748. <https://doi.org/10.1371/journal.pone.0086748>.
- Ng TF, Duffy S, Polston JE, Bixby E, Vallad GE, Breitbart M. 2011. Exploring the diversity of plant DNA viruses and their satellites using vector-enabled metagenomics on whiteflies. *PLoS One* 6:e19050. <https://doi.org/10.1371/journal.pone.0019050>.
- Dayaram A, Potter KA, Moline AB, Rosenstein DD, Marinov M, Thomas JE, Breitbart M, Rosario K, Argüello-Astorga GR, Varsani A. 2013. High global diversity of cycloviruses amongst dragonflies. *J Gen Virol* 94:1827–1840. <https://doi.org/10.1099/vir.0.052654-0>.
- Paula DP, Linard B, Crampton-Platt A, Srivathsan A, Timmermans MJTN, Sujii ER, Pires CSS, Souza LM, Andow DA, Vogler AP. 2016. Uncovering trophic interactions in arthropod predators through DNA shotgun-sequencing of gut contents. *PLoS One* 11:e0161841. <https://doi.org/10.1371/journal.pone.0161841>.
- Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. 2009. ABySS: A parallel assembler for short read sequence data. *Genome Res* 19:1117–1123. <https://doi.org/10.1101/gr.089532.108>.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25. <https://doi.org/10.1186/gb-2009-10-3-r25>.
- CABI. 2014. Cauliflower mosaic virus. <http://www.cabi.org/cpc>. Accessed 1 December 2016.
- Kitajima EW. 2014. Lista Comentada dos Vírus de Plantas Descritos no Brasil (2011–2013), on SBV. [http://www.sbv.org.br/site/downloads/virus/Lista\\_comentada\\_2011\\_2013.pdf](http://www.sbv.org.br/site/downloads/virus/Lista_comentada_2011_2013.pdf). Accessed 3 December 2016.