



Coding-Complete Genome Sequence of a Partitivirus Isolated from Pine Processionary Moth Eggs

Franck Dorkeld,^a Réjane Streiff,^a  Carole Kerdelhué,^a  Mylène Ogliastro^b

^aCBGP, INRAE CIRAD IRD Institut Agro Université de Montpellier, Montpellier, France

^bDGIMI, INRAE Université de Montpellier, Montpellier, France

ABSTRACT Two coding-complete nucleotide sequences of a partitivirus (family *Partitiviridae*) were discovered in transcriptomic data sets obtained from eggs of the Lepidoptera *Thaumetopoea pityocampa*. Each segment encodes a single open reading frame, and these two segments are predicted to encode an RNA-dependent RNA polymerase and a coat protein, respectively.

Pine processionary moths (PPMs; *Thaumetopoea pityocampa*) are aggressive defoliators of pine trees in the Mediterranean basin (1). PPMs reproduce in summer and lay their eggs on needles at the tip of pine tree branches, where the caterpillars hatch and weave a large protective silken nest for the colony. Large transcriptome sequencing (RNA-Seq) data sets corresponding to several PPMs at all developmental stages (eggs to adults) were mined to determine the viral diversity in different populations from southern Europe. We describe here the full-length coding-complete genome sequence of a betapartitivirus (family *Partitiviridae*), associated with the eggs of a unique PPM population from Portugal that lays eggs in the spring (2).

Viruses from the family *Partitiviridae* have a small and nonenveloped capsid protecting their bisegmented double-stranded RNA (dsRNA) genomes that are believed to be separately encapsidated (3). These viruses group into five genera infecting plants, fungi, and protozoa, with host-based clades and clusters within each genus (3), including viruses infecting mildew species, fungal causal agents of diseases, and major threats to agriculture (4). Eggs from *Thaumetopoea pityocampa* were collected in the region of Leiria, Portugal, in June 2012 (2, 4). The eggs were crushed, and total RNAs were extracted with TRIzol reagent. A library was constructed using the TruSeq stranded mRNA prep kit (Illumina) with an average insert size of 450 bp and sequenced in a lane of paired-end (PE) 2 × 100-bp reads on the HiSeq 2000 platform (5). We obtained 34,478,275 paired-end reads that were trimmed (Trimmomatic v.0.33) (6) and assembled into contigs (Trinity v.2.0.2) (7). Default parameters were used for all software unless otherwise specified. Viral contigs were identified by subjecting the contigs consecutively to BLASTX (E value, 1.e-10) and BLASTN searches against the viral RefSeq and nonredundant (nr) databases (NCBI), respectively. We identified two contigs of similar sizes (2,210 and 2,164 bp) made of the assembly of 754 and 830 reads, respectively, which represent 4.5% of the total viral reads in the sample. Both genetic elements are monocistronic; each is predicted to contain a single open reading frame (ORF). The first contig (2,210 bp; GC content, 39%) encodes a putative 710-amino-acid RNA-dependent RNA polymerase (RdRP); the second contig (2,164 bp; GC content, 40%) encodes a 654-amino-acid capsid protein (CP). Taxonomic sequence assignment showed that both contigs share the same closest relative, i.e., the *Plasmopara*

Citation Dorkeld F, Streiff R, Kerdelhué C, Ogliastro M. 2021. Coding-complete genome sequence of a partitivirus isolated from pine processionary moth eggs. *Microbiol Resour Announc* 10:e00071-21. <https://doi.org/10.1128/MRA.00071-21>.

Editor Jelle Matthijnsens, KU Leuven

Copyright © 2021 Dorkeld 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 Mylène Ogliastro, marie-helene.ogliastro@inrae.fr.

Received 21 January 2021

Accepted 4 February 2021

Published 25 February 2021

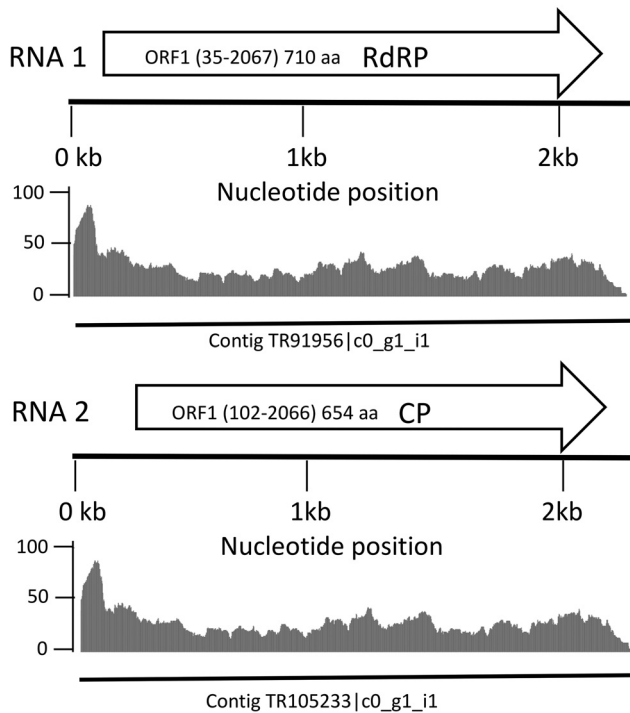
Thaumetopoea pityocampa associated Partitivirus

FIG 1 Genomic organization of the PPM-associated partitivirus. The complete major open reading frames (ORFs) predicted are represented by the arrows labeled RdRP (RNA-dependent RNA polymerase) and CP (capsid protein) and corresponding to RNA1 and RNA2, respectively. Matching reads mapped to the virus genomes are represented by a gray area with a scale that represents the coverage. The solid line at the bottom represents the contigs corresponding to the segments RNA1 and RNA2.

viticola lesion-associated Partitivirus 7 (GenBank accession numbers [MN556979.1](#) and [MN556980.1](#)), displaying 46% and 30% identity with the RdRP and CP, respectively. The contigs were named RNA1 and RNA2, according to the bisegmented genomic organization of this virus family (Fig. 1). According to the ICTV, taxonomic assignment in the family *Partitiviridae* is based on the RdRP homology, with a genus threshold of <24%. With 46% identity with the RdRP of the closest relative, two monocistronic contigs, and predicted proteins within the size range of this virus family, this PPM-associated partitivirus is likely a novel species in the genus *Betapartitivirus* (2). This virus was tentatively named *Thaumetopoea pityocampa*-associated partitivirus.

Members of the genus *Betapartitivirus* can infect plants and fungi, and the *Plasmopara viticola* lesion-associated Partitivirus 7 is an unclassified virus that infects the fungal causal agent of grape downy mildew (GenBank accession numbers [MN556979.1](#) and [MN556980.1](#) for RNA1 and RNA2, respectively). Although no fungal host can be associated with the novel partitivirus described here, mildew species are potential candidates, as they are commonly found in conifer forests, such as the powdery mildew that makes a white powdery substance on tree leaf surfaces. The most likely hypothesis is that infected fungi associated with pine trees have contaminated the surface of *T. pityocampa* eggs laid on the branches.

Data availability. The GenBank accession numbers are [MT799182](#) and [MT799183](#). The reads were deposited in the NCBI Sequence Read Archive (SRA) database under BioProject accession number [PRJNA663237](#). The reads from the specific metagenome used here for virus reconstruction are registered with the SRA under accession number [SRX9125299](#).

ACKNOWLEDGMENTS

This work was supported by INRAE, departments ECODIV (project EVIL) and SPE. M.O. is supported by the European Union's Horizon 2020 research and innovation program under grant agreement number 773567.

REFERENCES

1. Kerdelhué C, Zane L, Simonato M, Salvato P, Rousselet J, Roques A, Battisti A. 2009. Quaternary history and contemporary patterns in a currently expanding species. *BMC Evol Biol* 9:220. <https://doi.org/10.1186/1471-2148-9-220>.
2. Santos H, Burban C, Rousselet J, Rossi J-P, Branco M, Kerdelhué C. 2011. Incipient allochronic speciation in the pine processionary moth (*Thaumetopoea pityocampa*, Lepidoptera, Notodontidae). *J Evol Biol* 24:146–158. <https://doi.org/10.1111/j.1420-9101.2010.02147.x>.
3. Vainio EJ, Chiba S, Ghabrial SA, Maiss E, Roossinck M, Sabanadzovic S, Suzuki N, Xie J, Nibert M, ICTV Report Consortium. 2018. ICTV Virus Taxonomy Profile: Partitiviridae. *J Gen Virol* 99:17–18. <https://doi.org/10.1099/jgv.0.000985>.
4. Burban C, Rocha S, Leblois R, Rossi J-P, Sauné L, Branco M, Kerdelhué C. 2020. From sympatry to parapatry: a rapid change in the spatial context of incipient allochronic speciation. *Evol Ecol* 34:101–121. <https://doi.org/10.1007/s10682-019-10021-4>.
5. Gschloessl B, Dorkeld F, Berges H, Beydon G, Bouchez O, Branco M, Bretaudeau A, Burban C, Dubois E, Gauthier P, Lhuillier E, Nichols J, Nidelet S, Rocha S, Sauné L, Streiff R, Gautier M, Kerdelhué C. 2018. Draft genome and reference transcriptomic resources for the urticating pine defoliator *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae). *Mol Ecol Resour* 18:602–619. <https://doi.org/10.1111/1755-0998.12756>.
6. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
7. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29:644–652. <https://doi.org/10.1038/nbt.1883>.