Coding-Complete Genome Sequences of an Iteradensovirus and an Alphapermutotetra-Like Virus Identified from the Pine Processionary Moth (Thaumetopoea pityocampa) in Portugal

Franck Dorkeld, Rejane Streiff, Carole Kerdelhué, Mylène Ogliastro

To cite this version:
Franck Dorkeld, Rejane Streiff, Carole Kerdelhué, Mylène Ogliastro. Coding-Complete Genome Sequences of an Iteradensovirus and an Alphapermutotetra-Like Virus Identified from the Pine Processionary Moth (Thaumetopoea pityocampa) in Portugal. Microbiology Resource Announcements, American Society for Microbiology, 2021, 10 (1), 10.1128/MRA.01163-20. hal-03130972

HAL Id: hal-03130972
https://hal.inrae.fr/hal-03130972
Submitted on 4 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Coding-Complete Genome Sequences of an Iteradensovirus and an Alphapermutotetra-Like Virus Identified from the Pine Processionary Moth (*Thaumetopoea pityocampa*) in Portugal

Franck Dorkeld, Réjane Streiff, Carole Kerdelhué, Mylène Ogliastro

ABSTRACT The coding-complete genome sequences of an iteradensovirus (family *Parvoviridae*) and an alphapermutotetra-like virus (family *Permutotetraviridae*) were discovered from transcriptomic data sets obtained from *Thaumetopoea pityocampa* larvae collected in Portugal. Each of the coding-complete genome sequences of these viruses contains three main open reading frames (ORFs).
The *Thaumetopoea pityocampa* virus (TpV) is made of one contig (5,621 bp; GC content, 46.3%) assembled from 20,260 reads (mean coverage, 306-fold), which represent 4.6% of the total viral reads (Fig. 1B). Three main overlapping ORFs were predicted, displaying 30% amino acid identity with the protein of an unclassified fish virus (ORF1; GenBank accession number YP_009361866), 34% amino acid identity with the putative RNA-dependent RNA polymerase (RdRP) from the Hubei leech virus 1 (ORF2; GenBank accession number NC_032925.1), and 49% amino acid identity with the capsid protein (CP) of the *Thosea asigna* virus (ORF3; GenBank accession number NC_043232), the type species of the genus *Alphapermutotetravirus* (10). Putative small ORFs (position 4758 to 4961) and RNA secondary structures (position 5300 to 5626; 50% prediction: RNAstructure Web server, http://rna.urmc.rochester.edu/RNAstructureWeb) at the 3’ end might correspond to minor CP(s) and pseudoknots described in this virus family (11).

**Data availability.** The GenBank accession numbers are MT796426 and MT796427 for *Thaumetopoea pityocampa* iteradensovirus and MT796428 for *Thaumetopoea pityocampa* virus. The reads were deposited in the Sequence Read Archive (SRA) NCBI database under BioProject accession number PRJNA663237.

**ACKNOWLEDGMENT**

This work was supported by INRAE, department ECODIV (innovative project EVIL).

**REFERENCES**


