

Molecular characterization of a novel chlorotic leaf curl-associated foveavirus infecting prune trees

Lianshun Zheng

Zhejiang University College of Agriculture and Biotechnology

Yang Han

Beijing Institute of Genomics Chinese Academy of Sciences

Pan Gong

Chinese Academy of Agricultural Sciences Institute of Plant Protection

Xueping Zhou

Zhejiang University College of Agriculture and Biotechnology

Jianxiang Wu

Zhejiang University College of Agriculture and Biotechnology

Shuai Fu (✉ fushuai@zju.edu.cn)

Zhejiang University College of Agriculture and Biotechnology <https://orcid.org/0000-0002-8338-0074>

Research Article

Keywords:

Posted Date: August 16th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1937671/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

The *Prune* tree (*Prunus mume*) is widely distributed in East Asian countries for fruit, medicinal herbs, or ornamental purposes for more than 2000 years. High-throughput sequencing in *prune* tree leaves showing chlorosis and leaf curl symptoms sampled from the China National Botanical Garden revealed the presence of a putative novel virus belonging to the genus *Foveavirus*, the family *Betaflexiviridae*. We cloned the completion genome of this new virus and named this virus as prune tree chlorotic leaf curl-associated virus (PTCLCaV). The genome of PTCLCaV is a 9 034 nucleotide (nt) positive-sense, single strand RNA, excluding its 3' poly (A) tail, and contains five open reading frames (ORFs). ORF1 encodes a putative viral replication-associated protein (Rep), and ORFs 2, 3, and 4 encode the triple gene block (TGB) proteins. ORF5 encodes the capsid protein (CP). Our phylogenetic analysis demonstrated that PTCLCaV is evolutionarily related to members from the genus *Foveavirus* of the family *Betaflexiviridae*. Based on sequence identity and phylogenetic analysis, PTCLCaV represents a new species of the genus *Foveavirus* in the family *Betaflexiviridae*.

Main

Prune tree (*Prunus mume*), belonging to the genus *Prunus* in the family *Rosaceae*, is originated from China and widely distributed in East Asian countries for fruit production, medicinal herbs, or ornamental purposes for more than 2000 years [2]. *Prunus* is susceptible to viroid or viral infection, causing symptomatic or symptomless disease, and many of viral species infecting *Prunus* belong to the family of *Betaflexiviridae* [5]. *Betaflexiviridae* is the largest family in the order *Tymovirales* and is divided into two subfamilies (*Quinvirinae* and *Trivirinae*) according to the types of their viral movement proteins: triple gene block (TGB) movement proteins or 30K-like movement proteins [1]. The subfamily *Quinvirinae* contains genera of *Foveavirus*, *Carlavirus*, and *Robigovirus*, and three unclassified species of *Banana mild mosaic virus*, *Banana virus X*, and *Sugarcane striate mosaic-associated virus* (Kondo, 2021 #401). Foveaviruses are positive-sense, single strand RNA viruses, whose virions are flexuous filaments with about 800 to over 1000 nm in length and 12–15 nm in diameter. Their host range is mostly restricted to woody plants [8, 9].

In 2021, we collected chlorotic and leaf curl leaves from prune trees in the National Botanical Garden, Beijing City, China (Fig. 1a). To identify the potential viruses in collected leaves, RNA-seq was performed as described previously [1, 6, 7]. Briefly, the EASYspin Plus Complex Plant RNA Kit (Aidlab Biotech, Beijing, China) was used to extract total RNA from the sampled leaf tissues. After removing ribosomal RNAs using the TruSeq RNA Sample Prep Kit (Illumina, San Diego, CA, USA), the RNA library was synthesized and then sequenced on the Illumina HiSeq X-ten platform (Biomarker Technologies, Beijing, China). The CLC Genomics Workbench 9.5 software (QIAGEN, Germantown, DEU) was used to trim adapters of the paired-end raw reads and filter the low-quality reads. The Trinity program was used to *de novo* assemble 15 628 839 clean reads into large contigs. The contigs were then subjected to a BLASTx and BLASTp search against nucleotide (nt) and amino acid (aa) sequences of the NCBI databases. The 5' end and 3' end sequences of the novel viral genome were determined through the 5' end and 3' end RACE (Rapid Amplification of cDNA Ends) PCR as instructed (Roche Diagnostic, Milan, Italy). The amplified products were inserted individually into the pLB cloning vector (Tiangen, Beijing, China), and then transformed into *Escherichia coli* DH5 cells. Then, at least five positive transformed clones were DNA-sequenced. The resulting nucleotide and the predicted protein sequences were analyzed using the Snap program V4.1.9 software (GSL Biotech, San Diego, CA). Multiple alignments using nucleotide or amino acid sequences were done using the MEGA X software as described with default settings [3].

A large contig of 9 005 nt was obtained by RNA-seq and bioinformatic analysis, and the result of Blastx search against NCBI protein databases showed that it shares the highest aa sequence identity (80.42%) with the replication-associated protein (Rep) of asian prunus virus 3 (APV3), indicating the presence of a novel virus belonging to family

Betaflexiviridae in these sampled *prune* tree leaves. Based on the symptoms and the host plant, we name this novel virus as prune tree chlorotic leaf curl-associated virus (PTCLCaV). We then amplified the full-length sequence of the PTCLCaV genome by RT-PCR, followed by 5' and 3' RACE. The sequencing results showed that the genome of PTCLCaV contained 9 034 nt, excluding its poly (A) tail (GenBank accession: OP066307). Pairwise percent identity analysis showed that the genomic sequence identity between PTCLCaV and the other seven *Foveavirus* ranges from 47.3–69.9% (Table 1). ORFs in the viral genomic RNA and conserved domains of viral proteins were then predicted using ORF Finder software (<http://www.ncbi.nlm.nih.gov/orffinder>) and Conserved Domain Database (<https://www.ncbi.nlm.nih.gov/cdd>) [4]. The predicted results showed that the PTCLCaV genome composes of five ORFs, a 62-nt 5' untranslated region (UTR), and a 455-nt 3' UTR (Fig. 1b).

ORF1 starts from an AUG start codon at nt position 63 and ends at a UAG stop codon at nt position 6 161, and is predicted to encode a putative Rep protein with 2 032 aa residues. This putative Rep contains five conserved domains: a methyltransferase domain (aa position 43–353, pfam01660), a 2OG-Fe(II) oxygenase domain (aa position 808–910, cl21496), a peptidase domain (aa position 1 062 – 1 147, cl05111), a helicase domain (aa position 1 233-1 489, cl26263), and an RNA-dependent RNA polymerase domain (aa position 1 617-2 022, cl03049). Sequence alignment analysis showed that the Rep of PTCLCaV shared the highest nt sequence identity (71.1%) and aa sequence identity (80.4%) with the Rep of APV3 (Table 1). ORF2, ORF3, and ORF4 encode triple gene block (TGB1, TGB2, and TGB3) proteins, respectively. ORF2 of PTCLCaV is predicted to start from an AUG codon at nt position 6 194 and end at a UGA stop codon at nt position 6 913 (Fig. 1b), and encodes a 27.1-kDa polypeptide containing a conserved helicase domain (aa position 24–223, pfam01443) that shares the highest nt sequence identity (73.2%) with TGB1 of asian prunus virus 2 and 3 (APV2 and APV3) and the highest aa sequence identity (83.7%) with TGB1 of APV3 (Table 1). ORF3 of PTCLCaV starts at nt position 6 888 and ends at nt position 7 229 (Fig. 1b). It was predicted to encode a 12.0-kDa protein containing a plant viral movement protein domain (aa position 5-105, pfam01307), which shares the highest nt sequence identity (76.9%) with TGB2 of APV2 and the highest aa identity (87.6%) with TGB2 of asian prunus virus 1 (APV1) (Table 1). ORF4 of PTCLCaV starts at nt position 7 162 and ends at nt position 7 359, and encodes a 7.1-kDa protein containing a conserved 7k viral coat protein domain (aa position 6–61, pfam02495) (Fig. 1b), which shares the highest nt sequence identity (84.8%) with TGB3 of APV2 and the highest aa identity (83.1%) with TGB3 of APV2 (Table 1). ORF5 of PTCLCaV starts at nt position 7 416 and ends at nt position 8 579 (Fig. 1b), and encodes a 42.2-kDa protein with a conserved flexi CP domain (aa position 201–339, cl02836). This protein shared the highest nt sequence identity (64.2%) with the CP of APV1 and the highest aa sequence identity (63%) with the CP of APV2 (Table 1). The alignment analysis of the genome sequences of PTCLCaV and seven representative viruses members from the genus *Foveavirus* showed that the genome of PTCLCaV shares the highest sequence similarity (69.9%) with that of APV2 (Table 1).

Phylogenetic analysis based on the whole genome sequences of PTCLCaV and other representative viruses from the order *Tymovirales* indicated that PTCLCaV was most closely related to members of the genus *Foveavirus* in the family *Betaflexiviridae* (Fig. 2a and b). Phylogenetic trees constructed using aa sequences of Rep or CP of PTCLCaV and other representative viruses of the family *Betaflexiviridae* indicated that PTCLCaV is most closely similar to APV2, APV1, and APV3 (Fig. 2c and d). The PTCLCaV Rep and CP sequence identities exceed the cutoff values (i.e. 72% for the nt sequence identity and 80% for the aa sequence identity) for species discrimination in the genus *Foveavirus* that is released by the ICTV. Therefore, PTCLCaV should represent a new species in the genus *Foveavirus*.

Table 1

Pairwise percent identity values from comparisons of nucleotide (nt) and amino acid (aa) sequences of PTCLCaV and the most closely related seven member of viruses from the genus *Foveavirus*.

Virus	Accession number	Genome nt	ORF1(Rep)		ORF2(TGB1)		ORF3(TGB2)		ORF4(TGB3)		ORF5(CP)	
			nt	aa	nt	aa	nt	aa	nt	aa	nt	aa
ASPV	NC 003462.2	49.9	50.2	42.5	48.3	40.1	58.4	57.1	57.9	42.2	37.8	23.6
ApLV	NC 014821.1	49.6	50.0	42.3	48.9	40.5	56.5	55.0	57.9	40.6	37.0	24.1
GRSPaV	NC 001948.1	47.3	49.4	41.3	47.1	37.6	51.9	44.6	40.4	24.6	43.0	30.5
PCMoV	NC 009892.1	49.1	50.1	43.1	47.3	39.6	54.0	45.5	52.3	39.1	42.3	26.9
APV1	NC 025388.1	69.1	60.4	63.2	71.8	77.4	76.0	87.6	84.3	78.5	64.2	16.4
APV2	NC 028868.1	69.9	40.0	29.4	73.2	82.8	76.9	85.8	84.8	83.1	33.3	63.0
APV3	NC 028975.1	69.5	71.1	80.4	73.2	83.7	75.1	85.0	79.8	72.3	33.7	14.3
Viruses's name, abbreviations, and their GenBank accession numbers used in these analyses are shown: ASPV: apple stem pitting virus, ApLV: apricot latent virus, GRSPaV: grapevine rupestris stem pitting-associated virus, PCMoV: peach chlorotic mottle virus, APV1: asian prunus virus 1, APV2: asian prunus virus 2, and APV3: asian prunus virus 3.												

Declarations

Acknowledgements

This research was financially supported by the National Key R&D Program of China (grant no. 2019YFD1001800), China Postdoctoral Science Foundation (grant no. 2019T120521).

Author contributions

J. W. and S. F. designed the experiments; L.Z. performed the experiments; Y. H. performed the bioinformatics analysis; S. F., L. Z., and X. Z. wrote the manuscript.

Compliance with ethical standards

Conflict of interest

All authors declare they have no conflict of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Fu S, Zhang T, He M, Sun B, Zhou X, Wu J (2021) Molecular characterization of a novel wheat-infecting virus of the family Betaflexiviridae. *Archives of virology* 166:2875-2879
2. Gong XP, Tang Y, Song YY, Du G, Li J (2021) Comprehensive Review of Phytochemical Constituents, Pharmacological Properties, and Clinical Applications of *Prunus mume*. *Front Pharmacol* 12
3. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular biology and evolution* 35:1547-1549
4. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GH, Song JS, Thanki N, Yamashita RA, Yang M, Zhang D, Zheng C, Lanczycki CJ, Marchler-Bauer A (2020) CDD/SPARCLE: the conserved domain database in 2020. *Nucleic acids research* 48:D265-D268
5. Marais A, Faure C, Theil S, Candresse T (2018) Molecular Characterization of a Novel Species of Capillovirus from Japanese Apricot (*Prunus mume*). *Viruses* 10
6. Wang D, Fu S, Wu H, Cao M, Liu L, Zhou X, Wu J (2022) Discovery and Genomic Function of a Novel Rice Dwarf-Associated Bunya-like Virus. *Viruses* 14
7. Xie Y, Fu S, Xie L, Wang Y, Cao M, Zhou X, Wu J (2022) Identification and Characterization of Two Novel Noda-like Viruses from Rice Plants Showing the Dwarfing Symptom. *Viruses* 14
8. Youssef F, Marais A, Faure C, Barone M, Gentit P, Candresse T (2011) Characterization of *Prunus*-infecting Apricot latent virus-like Foveaviruses: Evolutionary and taxonomic implications. *Virus Research* 155:440-445
9. Zheng L, Chen M, Li R (2020) *Camellia* ringspot-associated virus 4, a proposed new foveavirus from *Camellia japonica*. *Archives of virology* 165:1707-1710

Figures

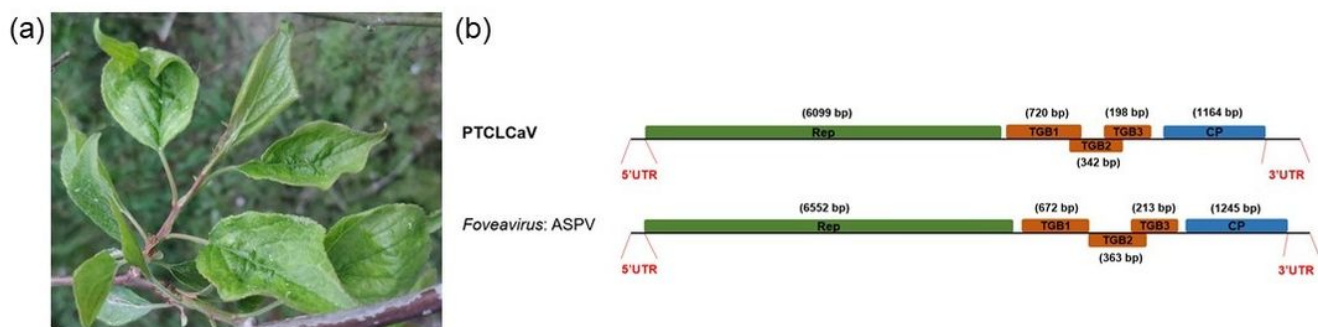


Figure 1

Symptoms and the genome organization of prune tree chlorotic leaf curl-associated virus (PTCLCaV). (a) The sampled prune tree leaves showing chlorosis and leaf curl symptoms. (b) The genome organization of PTCLCaV. Apple stem pitting virus (ASPV, NC 003462) represents a typical member of the genus *Foveavirus*. The viral replication associated protein (Rep) is shown in the green box. The triple gene block proteins (TGB) are shown in the orange boxes. The capsid protein (CP) is shown in the blue box. 5'-UTR: 5' untranslated region. 3'-UTR: 3' untranslated region.

