Complete genome analysis of a novel picorna-like virus from a ladybird beetle (*Cheilomenes sexmaculata*)

Yu-Juan He1,2 · Zhuang-Xin Ye1 · Jian-Ping Chen1,2 · Chuan-Xi Zhang1 · Gang Lu1 · Jun-Min Li1

Received: 18 November 2021 / Accepted: 18 January 2022 / Published online: 7 March 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract
The ladybird beetle *Cheilomenes sexmaculata* (family Coccinellidae, order Coleoptera) is a common insect predator of agricultural pests. In this study, the full genome sequence of a novel picorna-like virus, tentatively named “Cheilomenes sexmaculata picorna-like virus 1” (CSPLV1), was identified in *C. sexmaculata*. The full-length sequence of CSPLV1 is 11,384 nucleotides (nt) in length (excluding the polyA tail), with one predicted open reading frame (ORF) encoding a polyprotein of 3727 amino acids, a 13-nt 5' untranslated region (UTR), and a 187-nt 3' UTR. The ORF of CSPLV1 consists of four distinct domains, including an RNA virus helicase domain (nt 3029-3319), a peptidase domain (nt 5555-6121), an RNA-dependent RNA polymerase domain (nt 7154-8101), and a picorna-like coat protein domain (nt 8606-9283). Phylogenetic analysis based on the conserved RdRP sequence showed that CSPLV1, together with Wuhan house centipede virus 3, *Hypera postica* associated virus 1, and *Diabrotica undecimpunctata* virus 1, forms an unclassified group that is closely related to members of the family *Solinviviridae*. To the best of our knowledge, CSPLV1 is the first picorna-like virus discovered in *C. sexmaculata*.

Picornaviruses are a group of positive-sense single-stranded RNA viruses, and the order *Picornavirales* comprises the families *Caliciviridae*, *Dicistroviridae*, *Iflaviridae*, *MarNAVIRIDAE*, *Picornaviridae*, *Polycipiviridae*, *Secoviridae*, and *Solinviviridae*. Picornavirus virions contain a genomic RNA molecule with a large protein attached to the 5' end. The genome contains no overlapping open reading frames (ORFs), and all of the RNAs are translated into a polyprotein before processing [1]. Comprehensive phylogenetic analysis of the viral RNA-dependent RNA polymerase (RdRP) provides a basis for the taxonomy of picorna-like viruses [1–5]. Picorna-like viruses have been reported in numerous insects, such as *Drosophila* (Drosophila C virus, DcV) [6], *Apis mellifera* (sacbrood virus) [7], and *Culex spp.* (Culex picorna-like virus) [8]. In the order Coleoptera, several picorna-like viruses have been reported in members of different insect families, including *Scarabaeidae* (*Oryctes rhinoceros*) [9], *Chrysomelidae* (*Aulacophora lewisi*) [10], and *Coccinellidae* (*Harmonia axyridis*) [11]. *Cheilomenes sexmaculata*, another member of the family Coccinellidae, preys on aphids, other soft-bodied insects, and even pollen and is an important aphidophagous predator for the biocontrol of agricultural pests in Asia. Based on the competitive ability for aphid prey, *C. sexmaculata* has been used as an effective biological agent to control outbreaks of aphids [12]. In this study, the complete genome sequence of a novel picorna-like virus was identified and characterized in *C. sexmaculata*.

Colonies of ladybird beetles were collected from cucumber leaves in August 2019 in Jinhua, Zhejiang, China. Total RNA was extracted from a single ladybird beetle using TRIzol Reagent (Invitrogen, Waltham, MA, USA). A NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA) was employed to determine the concentrations of the RNA samples for subsequent next-generation sequencing (NGS) analysis. The RNA library (paired-end sequencing, 150 bp) was sequenced on an Illumina HiSeq 4000 platform without an rRNA depletion or polyA enrichment step (Illumina, San Diego, CA, USA). Quality reads were cut...
The host COI sequence was used as a positive control. A panel derived sequences in host genomic DNA using the primers shown in Detection of CSPLV1-strand cDNA produced using primer 5. (D) control 2 (NC2), primers 4 and 6 were used for PCR with a positive-strand cDNA produced using primer 2, and as negative control 1 (NC1), primers 1 and 5 were used for PCR with a negative-strand cDNA produced using primer 2 and as negative control 2 (NC2), primers 4 and 6 were used for PCR with a positive-strand cDNA produced using primer 5. (D) Detection of CSPLV1-derived sequences in host genomic DNA using the primers shown in panel A. The host COI sequence was used as a positive control.

by Trimmomatic with the following parameters: (Leading: 3; Trailing: 3; Slidingwindow: 4;15; Minlen: 36). Trinity software (Version 2.8.5) with default parameters was used for de novo assembly with the preliminary treated clean reads (SRA accession no. SRR17319554). The assembled contigs were searched using the Barcode of Life Data (BOLD) system (http://www.boldsystems.org/) to identify the cytochrome c oxidase I (COI) sequence of the ladybird beetle species. Then, a BLASTn search against the nucleotide (nt) database of the National Center for Biotechnology Information (NCBI) was conducted to confirm the species of the beetle. Our result indicated that the COI sequence of the collected beetles was exactly the same as that of C. sexmaculata (GenBank number: KM207085.1).

To identify potential viral sequences in the transcriptome of C. sexmaculata, the assembled contigs were used to search the entire viral reference database at NCBI (https://www.ncbi.nlm.nih.gov/genome/viruses) using BLASTx. As a result, a candidate picorna-like viral sequence (11,369 nt) with a relatively high abundance (53.39X coverage) was discovered from the assembled transcripts of C. sexmaculata (GenBank number: OK632508) using SMARTer® RACE kit (Takara, Japan), and this result was confirmed by Sanger sequencing. The primers used in this study are listed in Supplementary Table S1. This picorna-like virus was tentatively named “Cheiromenes sexmaculata picorna-like virus 1” (CSPLV1), and its complete genome sequence was submitted to the GenBank database with the accession number OK632508 (Supplementary File S1).

The full-length genome sequence of CSPLV1 is 11,384 nt in length (excluding the polyA tail), with a single predicted ORF (nt 14-11,197) encoding a polyprotein of 3727 amino acids, a 13-nt 5’ untranslated region (UTR), and a 187-nt 3’ UTR. In addition, based on InterProScan conserved domain prediction (http://www.ebi.ac.uk/interpro), the ORF of CSPLV1 contains four distinct domains, including an RNA virus helicase domain (Hel) (nt 3029-3319), a peptidase domain (Pep) (nt 5555-6121), an RNA-dependent RNA polymerase domain (RdRP) (nt 7154-8101) and a picornalike coat protein domain (nt 8606-9283) (Fig. 1A). The abundance and distribution of reads over the viral genome is shown in Fig. 1B. To evaluate the potential replication of CSPLV1 in C. sexmaculata, the negative strand of the CSPLV1 genome was detected by strand specific RT-PCR [13–15]. Using the primer pairs 1/5 and 2/5, PCR products of the expected sizes were obtained using viral negative-strand cDNA as a template. Sanger sequencing of these PCR products confirmed that the sequences of PCR products corresponded to those in the CSPLV1 genome (Fig. 1C).

To rule out the possibility that CSPLV1-derived sequences were integrated in the host genome, we amplified PCR fragments with four pairs of virus-specific primers (primer pairs 1/5, 2/5, 4/6, 3/5) using host DNA as a template, and no obvious bands were detected in a 1% agarose gel (Fig. 1D).

Next, phylogenetic analysis was performed using the conserved RdRP amino acid sequences of CSPLV1 and previously reported members of the order Picornavirales. Sequences were obtained from the NCBI GenBank database and aligned using MAFFT [16], and the maximum-likelihood (ML) algorithm and the Jones-Taylor-Thornton (JTT) substitution model were employed to construct a phylogenetic tree in RAxML-NG with 1000 bootstrap replications [17]. Chinese wheat mosaic virus (NP_059513.1) and barley stripe mosaic virus (AHY22372.1) were used as outgroups in the tree. The results clearly indicated that CSPLV clustered with Wuhan house centipede virus 3 (WHCV3) (YP_009342325.1) (Fig. 2). CSPLV1 and WHCV3, together with Hypera postica associated virus 1 (HPAV1) (QUS52866.1) and Diabrotica undecimpunctata virus 1 (DUV1) (QIT20099.1), formed an unclassified group and were found to be closely related to members of the family Solinviviridae (Fig. 2). In addition, pairwise comparisons of the amino acid and nucleotide sequences of RdRp were performed for CSPLV1 and the other picornaviruses. Similar to the result of the phylogenetic analysis, CSPLV1 was found to be closely related to the three viruses clustered in the unclassified group, sharing 30.1-47.5% aa and 43.8-54.9% nt sequence identity (Table 1). The host insects and geographical distribution of these four unclassified viruses are shown in Supplementary Table S3.

In conclusion, based on analysis of the full-length genome sequence and phylogenetic analysis, CSPLV1, a novel unclassified picorna-like virus, was identified in a ladybird beetle (C. sexmaculata). To the best of our knowledge, this is the first picorna-like virus reported in C. sexmaculata.
Fig. 2 Phylogenetic analysis of CSPLV1 and previously reported members of the order Picornavirales based on conserved amino acid sequences of RdRP. Bootstrap values are shown at each node of the tree (when > 50). Scale bars represent percentage divergence. CSPLV1, identified in this study, is shown in red font.

Table 1 Amino acid/nucleotide sequence identity based on the conserved RdRP domain

| Family          | Virus   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  | 14  | 15  | 16  | 17  | 18  |
|-----------------|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Unclassified    | CSPLV1  | ***| 47.5| 37.0| 30.1| 29.6| 30.2| 27.8| 25.7| 10.0| 18.2| 18.9| 19.1| 17.0| 15.5| 15.1| 17.5| 8.2 | 7.0 |
|                 | WHCV3   | 54.9| ***| 38.2| 30.7| 28.8| 30.1| 29.9| 30.2| 19.9| 18.4| 19.2| 17.8| 18.4| 14.0| 17.2| 18.6| 9.3 | 8.6 |
|                 | HPAV1   | 48.0| 47.5| ***| 33.2| 25.6| 28.8| 26.2| 27.6| 23.7| 21.7| 21.8| 20.8| 20.6| 17.8| 19.2| 20.7| 8.8 | 10.1|
|                 | DUV1    | 43.8| 44.9| 50.8| ***| 27.3| 26.0| 27.0| 26.2| 19.5| 18.0| 19.0| 18.3| 15.8| 17.1| 17.5| 16.3| 11.5| 9.5 |
| Soliviridae     | NFV1    | 37.4| 39.5| 37.5| 39.4| ***| 40.4| 40.4| 42.1| 23.1| 17.8| 19.1| 17.6| 17.5| 18.9| 18.3| 19.9| 11.4| 9.8 |
|                 | SINV-3  | 39.4| 42.8| 42.0| 41.3| 51.5| ***| 56.3| 43.5| 23.4| 19.7| 20.8| 18.4| 20.9| 18.3| 20.0| 21.1| 9.9 | 8.4 |
|                 | WIV12   | 36.2| 39.1| 38.8| 39.9| 50.5| 63.7| ***| 48.7| 24.4| 19.3| 22.3| 22.8| 20.4| 17.8| 18.5| 18.5| 10.8| 9.2 |
|                 | DVVV2   | 36.7| 39.3| 37.7| 37.7| 53.0| 53.0| 53.8| ***| 20.1| 16.7| 18.8| 18.8| 17.4| 16.8| 18.0| 19.3| 11.1| 9.3 |
| Flaviridae      | HCV1    | 36.9| 37.8| 37.7| 37.5| 41.9| 45.0| 43.0| 36.4| ***| 55.7| 34.4| 32.5| 26.9| 22.7| 24.5| 26.7| 9.5 | 9.8 |
|                 | TCIFV   | 36.0| 36.0| 36.6| 38.0| 37.1| 41.7| 39.0| 34.9| 63.5| ***| 36.3| 33.9| 25.6| 22.2| 24.0| 23.9| 8.9 | 9.4 |
|                 | YOIFV1  | 38.7| 34.7| 36.5| 36.9| 38.7| 43.9| 40.4| 36.3| 47.6| 46.5| ***| 72.5| 23.0| 22.4| 24.2| 24.9| 10.3| 8.8 |
|                 | ARIFV   | 38.3| 36.6| 36.3| 36.9| 41.2| 43.9| 42.0| 39.2| 47.5| 47.7| 71.1| ***| 23.3| 23.2| 24.9| 25.5| 10.1| 8.8 |
| Secoviridae     | LSV1    | 32.9| 34.0| 34.4| 33.2| 36.0| 35.5| 36.2| 33.6| 35.7| 38.5| 35.5| 37.5| ***| 29.4| 29.5| 38.6| 10.5| 10.6|
|                 | ANPMV   | 32.6| 34.6| 34.1| 33.5| 34.2| 38.2| 36.3| 35.2| 38.7| 38.0| 36.6| 36.0| 45.4| ***| 46.4| 28.8| 11.0| 8.4 |
|                 | CUMMV   | 35.5| 35.7| 35.5| 33.8| 35.9| 38.4| 38.1| 34.8| 40.6| 38.2| 36.7| 39.5| 45.0| 55.9| ***| 32.5| 11.2| 11.0|
|                 | ALSV    | 35.7| 35.4| 35.0| 34.8| 37.7| 40.5| 38.6| 34.9| 40.3| 39.0| 36.9| 38.5| 52.7| 44.2| 47.8| ***| 11.4| 10.3|
| Virgaviridae    | CWSV    | 29.8| 28.9| 30.1| 28.9| 31.0| 31.3| 29.4| 29.6| 29.9| 31.3| 32.2| 31.6| 35.2| 34.7| 33.3| 34.5| ***| 55.4|
|                 | BSMV    | 27.8| 30.4| 29.5| 26.8| 29.2| 32.5| 29.4| 28.7| 30.5| 32.4| 31.0| 28.8| 35.9| 33.5| 33.9| 35.9| 59.2| ***|

Amino acid sequence identity is indicated by bold text. Nucleotide sequence identity is indicated by non-bold text. The numbers 1-18 represent the corresponding viruses in the left column. Virus names and GenBank accession numbers are listed in Supplementary Table S2.
A novel picorna-like virus from a ladybird beetle

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00705-022-05400-2.

Funding This work was funded by the National Natural Science Foundation of China (U20A2036), the Project of State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-products (ZS20190102), the Ningbo Science and Technology Innovation 2025 Major Project (2019B10004), and the Natural Science Foundation of Ningbo City (2019A610404). This work was sponsored by the K.C. Wong Magna Fund of Ningbo University.

Declarations

Competing interests The authors declare no conflict of interest.

Ethical statement No experimental work with humans was done in this study.

References

1. Culley AI, Lang AS, Suttle CA (2003) High diversity of unknown picorna-like viruses in the sea. Nature 424:1054–1057
2. Koonin EV, Dolja VV, Morris TJ (1993) Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. Crit Rev Biochem Mol 28:375–430
3. Zanotto PD, Gibbs MJ, Gould EA, Holmes EC (1996) A reevaluation of the higher taxonomy of viruses based on RNA polymerases. J Virol 70:6083–6096
4. Mari J, Poulos BT, Lightner DV, Bonami J-R (2002) Shrimp Taura syndrome virus: genomic characterization and similarity with members of the genus Cricket paralysis-like viruses. J Gen Virol 83:915–926
5. Koonin EV, Dolja VV, Krupovic M, Varsani A, Wolf YI, Yutin N, Zerbini FM, Kuhn JH (2020) Global organization and proposed metagateway of the virus world. Microbiol Mol Biol Rev 84:e00061-e119
6. Johnson KN, Christian PD (1998) The novel genome organization of the insect picorna-like virus Drosophila C virus suggests this virus belongs to a previously undescribed virus family. J Gen Virol 79:191–203
7. Ghosh R, Ball B, Wilcock M, Carter M (1999) The nucleotide sequence of sacbrood virus of the honey bee: an insect picorna-like virus. J Gen Virol 80:1541–1549
8. Cholleti H, Hayer J, Safetine J, Berg M, Blomström AL (2018) Genetic characterization of a novel picorna-like virus in Culex spp. mosquitoes from Mozambique. Virol J 15:1–10
9. Etebari K, Shelomi M, Furlong MJ (2020) Identification of a novel Picorna-like virus in coconut rhinoceros beetles (Oryctes rhinoceros). Virus Res 287:198100
10. Ye ZX, Li YH, Chen JP, Zhang CX, Li JM (2020) Complete genome analysis of a novel ilavirus from a leaf beetle, Aulacophora lewisii. Arch Virol 166:309–312
11. Liu YJ, Ren GW, Jiang LQ, Feng CC, Liu DY, Liu YJ, Xu PJ (2020) Sequencing and phylogenetic characterization of a novel RNA virus genome from Harmonia axyridis. Mol Biol Rep 47:4015–4019
12. Omkar BR (2004) Prey quality dependent growth, development and reproduction of a biocontrol agent, Cheilomenes sexmaculata (Fabricius) (Coleoptera: Coccinellidae). Biocontrol Sci Technol 14:665–673
13. Nigg JC, Kuo Y-W, Falk BW (2020) Endogenous viral element-derived piwi-interacting RNAs (piRNAs) are not required for production of ping-pong-dependent piRNAs from Diaphorina citri Densovirus. MBio 11:e02209-e2220
14. Ho EC, Donaldson ME, Saville BJ (2010) Detection of antisense RNA transcripts by strand-specific RT-PCR. Methods Mol Biol 630:125–138
15. Mancarella A, Procopio FA, Achsel T, Crignis ED, Foley BT, Corradin G, Bagni C, Pantaleo G, Graziosi C (2019) Detection of human immunodeficiency virus type 1 (HIV-1) antisense protein (ASP) RNA transcripts in patients by strand-specific RT-PCR. J Vis Exp 7:153
16. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780
17. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A (2019) RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453–4455

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.