Complete Genome Analysis of a Novel Picorna-Like Virus From a Ladybird Beetle, *Cheilomenes Sexmaculata*

Yu-Juan He

Ningbo University

Zhuang-Xin Ye

Ningbo University

Jian-Ping Chen

Ningbo University

Chuan-Xi Zhang

Ningbo University

Gang Lu (✉ lugang@nbu.edu.cn)

Institute of Plant Virology <https://orcid.org/0000-0001-8017-2192>

Jun-Min Li

Ningbo University

Research Article

Keywords: picorna-like virus, *Cheilomenes sexmaculata*, RNA virus discovery

Posted Date: November 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1092064/v1>

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

Version of Record: A version of this preprint was published at Archives of Virology on March 7th, 2022. See the published version at <https://doi.org/10.1007/s00705-022-05400-2>.

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, temporarily named “*Cheilomenes sexmaculata* picorna-like virus 1” (CSPLV1), was identified from *C. sexmaculata*. The full-length sequence of CSPLV1 was 11,384 nucleotide (nt) in length (excluding the polyA tail) with one predicted open reading frame (ORF) encoding 3727 amino acids, a 13 nt 5' untranslated region (UTR) and a 187 nt 3' UTR. The ORF of CSPLV1 consisted of four distinct domains including an RNA virus helicase domain (3029-3319 nt), a peptidase domain (5555-6121 nt), an RNA-dependent RNA polymerase domain (7154-8101 nt) and a picorna-like coat protein domain (8606-9283nt). 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, formed as an unclassified group which is closely related to the clade *Solinviviridae*. To the best of our knowledge, CSPLV1 is the first picorna-like virus revealed in *C. sexmaculata*.

Introduction

Picornaviruses are a group of positive-sense single-stranded RNA viruses, and the order picornavirales comprises *Caliciviridae*, *Dicistroviridae*, *Iflaviridae*, *Marnaviridae*, *Picornaviridae*, *Polycipiviridae*, *Secoviridae*, and *Solinviviridae*. Picornaviruses possess a genome with a large protein attached to the 5' end but no overlapping open reading frames (ORFs), and all the RNAs are translated into a polyprotein before processing [2]. Furthermore, using the RNA-dependent RNA polymerase (RdRP) of these viruses by phylogenetic analysis keeps pace with the assignments for the established picorna-like virus family [2–5]. Hence, RdRP sequence is suitable for analyzing the diversity of picorna-like viruses. The 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]. For *Coleoptera*, several picorna-like viruses have been reported from different host insect families, such as *Scarabaeidae* (*Oryctes rhinoceros*) [9], *Chrysomelidae* (*Aulacophora lewisii*) [10] and *Coccinellidae* (*Harmonia axyridis*) [11]. *Cheilomenes sexmaculata*, another member of the family *Coccinellidae*, is an important aphidophagous predator for the biocontrol of agricultural pests, which is widely distributed in Asia [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 RNAs were extracted from a single ladybird beetle using TRIzol reagent (Invitrogen, Waltham, MA, USA). The Nanodrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA) was later employed to determine the concentrations of RNA samples for subsequent NGS analysis. The RNA library (paired-end sequencing, 150 bp) was performed on the Illumina HiSeq 4000 platform (Illumina, San Diego, CA, USA). Afterwards, the Trinity software (Version 2.8.5) with default parameters was used for the de novo assembly with the preliminary treated clean reads. Thereafter, the assembled contigs were searched in the Barcode of Life Data (BOLD) system (<http://www.boldsystems.org/>) to acquire the accurate cytochrome oxidase I (COI) sequence of the ladybird beetle species. Then, a BlastN search against the nucleotide (nt) database in National Center for Biotechnology Information (NCBI) was conducted to confirm the species of the beetles. Our result indicated that COI sequence of the collected beetles was exactly same to that of *C. sexmaculata* (GenBank Number: KM207085.1).

To identify the potential viral-like sequences in the transcriptome of *C. sexmaculata*, the assembled contigs were searched against the entire viral reference database in NCBI (<https://www.ncbi.nlm.nih.gov/genome/viruses>) by using BlastX. As the 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*. To avoid false positive match, this sequence was also aligned against the NCBI nt and non-redundant protein databases. Subsequently, the candidate sequence was further confirmed by reverse transcription PCR (RT-PCR) and Sanger sequencing. The rapid amplification of cDNA ends (RACE) assay was performed to determine the extreme 5' and 3' terminal sequences of the identified virus using the SMARTer® RACE 5'/3' kit (Takara, Beijing, China) and further confirmed by Sanger sequencing. The primers used in this study are listed in Supplementary Table S1. This picorna-like virus was tentatively named “*Cheilomenes sexmaculata* picorna-like virus 1” (CSPLV1), and its complete genome sequence was submitted to GenBank with the accession number of OK632508 (Supplementary File S1).

The full-length genome sequence of CSPLV1 was 11,384 nt in length (excluding the polyA tail) with only one predicted ORF (14-11,197 nt) encoding 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 contained four distinct domains, including an RNA virus helicase domain (Hel) (3029-3319 nt), a Peptidase domain (Pep) (5555-6121 nt), an RNA-dependent RNA polymerase domain (RdRP) (7154-8101 nt) and a picorna-like coat protein domain (8606-9283 nt) (Fig. 1A). The relatively abundant reads across CSPLV1 genome was indicated in Fig. 1B. Moreover, to evaluate the potential replication of CSPLV1 in *C. sexmaculata*, reverse strand of CSPLV1 was detected by strand specific RT-PCR. Primers 1/5 and primers 2/5 PCR products using viral antisense cDNA as a template were obtained the anticipated size. Sanger sequencing of these PCR products confirmed that the sequences of PCR products were consistent with those indicated in the CSPLV1 genome (Fig. 1C). Next, phylogenetic analysis was performed using the conserved RdRp amino acid sequences of CSPLV1 and the previously reported viruses in the Picornavirales. Sequences were obtained from NCBI GenBank and aligned using MAFFT [13]. Afterwards, 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 [14]. The Chinese wheat mosaic virus (NP_059513.1) and Barley stripe mosaic virus (AHY22372.1) were used as outgroups of the tree. The results of phylogenetic tree analysis clearly indicated that CSPLV1 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, which is closely related to the clade *Solinviviridae* (Fig. 2). Additionally, pairwise analysis for the identities of viral amino acid/nucleotide (RdRP) was performed for CSPLV1 and the other picornaviruses. Similar to the result of the phylogenetic analysis, CSPLV1 is also closely related to these three viruses clustered in the unclassified group (Fig. 2), which shared 30.1-47.5% (aa)/43.8-54.9% (nt) identities, respectively (Table 1).

In conclusion, based on analysis of the full-length genome sequence and phylogenetic analysis, CSPLV1, a novel unclassified picorna-like virus, is identified from the ladybird beetle, *C. sexmaculata*. To the best of our knowledge, it is the first picorna-like virus reported in *C. sexmaculata*.

Declarations

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), the Natural Science Foundation of Ningbo City (2019A610404). This work was sponsored by K.C.Wong Magna Fund in Ningbo University.

Compliance with ethical standards

Conflict of interest

The authors declare no conflict of interest.

Ethical statement

No experimental work with humans was done in this study.

References

1. Bolling BG, Weaver SC, Tesh RB, Vasilakis N (2015) Insect-specific virus discovery: significance for the arbovirus community. *Viruses* 7:4911–4928

2. Culley AJ, Lang AS, Suttle CA (2003) High diversity of unknown picorna-like viruses in the sea. *Nature* 424:1054–1057

3. 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

4. Zannotto 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

5. 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

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, Willcocks 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, Fafetine 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 iflavirus 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. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780

14. 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

Tables

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

	Identity	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
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
	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
	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
	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.
<i>Soliniviridae</i>	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.
	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
	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.
	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.
<i>Iflaviridae</i>	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
	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
	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.
	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.
<i>Secoviridae</i>	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.
	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.
	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.
	ALSV	35.7	35.4	35.0	34.8	37.7	40.5	38.6	34.3	40.9	39.0	36.9	38.5	52.7	44.2	47.8	***	11.
<i>Virgaviridae</i>	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	***
	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.

Protein identity is indicated by bold text. Nucleotide identity is indicated by non-bold text. The numbers 1-18 represent the corresponding viruses in the left column. Virus names and GeneBank accession numbers are listed in Supplementary Table S2.

Figures

Figure 1

Genetic structure (A) and transcriptome raw read coverage (B) of *Cheilomenes sexmaculata* picorna-like virus 1 (CSPLV1). Hel, RNA virus helicase core domain; Pep, Peptidase C3 superfamily domain; RdRP, RNA-dependent RNA polymerase domain; CP, Coat protein domain. (C). Detection of the minus strand of CSPLV1 in *C. sexmaculata* by reverse transcription PCR using primers shown in (A). Primer 1 or primer 6 was used to generate cDNA product from viral antisense or sense strand, respectively. Primers 1/5 and primers 2/5 PCR products were amplified using antisense cDNA as a template. Primers 4/6 and primers 3/5 PCR products were amplified using sense cDNA as a template. Primers 1/5 PCR product was produced using an antisense cDNA produced by primer 2 as a negative control (NC).

Figure 2

Phylogenetic analysis of CSPLV1 and previously reported viruses in the Picornavirales based on the conserved RdRP amino acid sequences. Bootstrap values are placed over each node of the tree (when>50). Scale bars represent percentage divergence. The CSPLV1 identified in this study is indicated with red font.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFileS1.fasta](#)
- [SupplementaryTableS1.docx](#)
- [SupplementaryTableS2.docx](#)