First Complete Genome Sequence of *Cherry virus A*

Hiroaki Koinuma, Takamichi Nijo, Nozomu Iwabuchi, Tetsuya Yoshida, Takuya Keima, Yukari Okano, Kensaku Maejima, Yasuyuki Yamaji, Shigetou Namba

Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan

The 5′-terminal genomic sequence of *Cherry virus A* (CVA) has long been unknown. We determined the first complete genome sequence of an apricot isolate of CVA (7,434 nucleotides [nt]). The 5′-untranslated region was 107 nt in length, which was 53 nt longer than those of known CVA sequences.

Address correspondence to Shigetou Namba, anamba@mail.ecc.u-tokyo.ac.jp.

Cherry virus A (CVA) is a member of the genus *Capillovirus* in the family *Betaflexiviridae*. CVA possesses flexuous filamentous particles containing a monopartite, single-stranded, and positive-sense RNA genome with two open reading frames (ORFs). Although the draft genome sequence (7,383 nucleotides [nt]) of CVA was first reported in 1995 (type isolate, GenBank accession no. X82547), its 5′ end was not determined (1).

In 2015, leaves of an apricot (*Prunus armeniaca* cv. Heiwamaru) showing symptoms of vein clearing were collected in Tokyo, Japan. Total RNA was extracted from leaves using the RNeasy plant minikit (Qiagen, Germany). The cDNA library was constructed using the TruSeq RNA sample prep kit version 2 (Illumina, USA), and the library was paired-end sequenced by MiSeq (Illumina) with reagent kit version 2 (500 cycles). The reads were *de novo* assembled using the Oases program (2), and the assembled contigs were subjected to BLASTn search (3) against the GenBank database. A single contig (7,408 nt) showing 86% sequence identity with the CVA type isolate was obtained, and no other contigs showed sequence similarity with viruses/viroids. The 5′ end of the genome was amplified by rapid amplification of cDNA ends (RACE) using the Gene Racer kit (Invitrogen, USA). The 3′ end of the genome was amplified by reverse transcription-PCR (RT-PCR) using CVA-specific and oligo(dT) primers. The amplified fragments were cloned into pCR-Blunt II-TOPO vector (Invitrogen), after which 6 and 8 clones of the 5′ and 3′ ends of the genome were sequenced, respectively.

The complete genome sequence of the apricot isolate of CVA, designated CVA-J, was 7,434 nt long, excluding the poly(A) tail at its 3′ end. A comparative analysis between CVA-J and other isolates of CVA was conducted using the SDT program (4) based on pairwise alignments using the MUSCLE algorithm (5). ORF1 (nt 108 to 7136) of CVA-J shared 81.2 to 98.0% nt identity and 85.0 to 98.4% amino acid (aa) identity with five other CVA isolates. ORF2 (nt 5453 to 6844) of CVA-J shared 90.6 to 98.8% nt identity and 88.1 to 97.8% aa identity with 11 other CVA isolates. Two isolates from noncherry hosts (accession numbers LN879388 and HQ267856) showed high nucleotide sequence identity (ORF1, 98.0%; ORF2, 98.6 to 98.8%) with CVA-J, while isolates from sweet-and-sour cherries showed low nucleotide sequence identity (ORF1, 81.2 to 85.9%; ORF2, 90.6 to 92.9%) with CVA-J, in accordance with a previous report (6).

The length of the 5′ untranslated region (UTR) of CVA-J was 107 nt, which was 53 nt longer than those of known CVA sequences. This is probably because these earlier sequences were determined using forward primers designed on the uncompleted 5′-terminal sequence of the type isolate (7). The 5′-terminal nucleotide of CVA-J was G, which is observed in many viruses in the family *Betaflexiviridae* and *Alphaflexiviridae*. For flexiviruses (to which both of these families belong), G at the 5′ end is believed to serve as an efficient transcription initiation site of genomic and subgenomic RNAs (8).

In this study, we determined for the first time the complete genome sequence of CVA. Our results will be helpful for further investigations of the sequence diversity of CVA.

Nucleotide sequence accession number. The complete genome sequence of CVA-J has been deposited in the DNA Data Bank of Japan under accession no. LC125634.

REFERENCES

1. **Jelkmann W.** 1995. Cherry virus A: cDNA cloning of dsRNA, nucleotide sequence analysis and serology reveal a new plant capillovirus in sweet cherry. J Gen Virol **76:**2015–2024. http://dx.doi.org/10.1099/0022-1317-76-8-2015.

2. **Schulz MH, Zerbino DR, Vingron M, Birney E.** 2012. *Oases*: robust *de novo* RNA-seq assembly across the dynamic range of expression levels. Bioinformatics **28:**1086–1092. http://dx.doi.org/10.1093/bioinformatics/bts094.

3. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ**. 1990. Basic Local Alignment Search Tool. J Mol Biol **215:**403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.

4. **Muhire BM, Varsani A, Martin DP**. 2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One **9:**e108277. http://dx.doi.org/10.1371/journal.pone.0108277.

5. **Edgar RC**. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:**1792–1797 http://dx.doi.org/10.1093/nar/gkh340.

6. **Marais A, Svanella-Dumas L, Barone M, Gentit P, Faure C, Charlot G, Ragozzino A, Candresse T**. 2012. Development of a polyvalent RT–PCR detection assay covering the genetic diversity of *cherry capillovirus A*. Plant Pathol **61:**195–204 http://dx.doi.org/10.1111/j.1365-3059.2011.02488.x.

7. **Noorani MS, Awasthi P, Singh RM, Ram R, Sharma MP, Singh SR, Ahmed N, Hallan V, Zaidi AA**. 2010. Complete nucleotide sequence of cherry virus A (CVA) infecting sweet cherry in India. Arch Virol **155:**2079–2082. http://dx.doi.org/10.1007/s00705-010-0826-6.

8. **Komatsu K, Hirata H, Fukagawa T, Yamaji Y, Okano Y, Ishikawa K, Adachi T, Maejima K, Hashimoto M, Namba S**. 2012. Infection of capilloviruses requires subgenomic RNAs whose transcription is controlled by promoter-like sequences conserved among flexiviruses. Virus Res **167:**8–15. http://dx.doi.org/10.1016/j.virusres.2012.02.019.