

Nucleotide sequence and genome organization of a new proposed crinivirus, tetterwort vein chlorosis virus

Fumei Zhao^{1,2} · Ran Hee Yoo^{1,2} · Seungmo Lim^{1,2} · Davaajargal Igori^{1,2} · Su-Heon Lee³ · Jae Sun Moon^{1,2} 

Received: 29 April 2015 / Accepted: 21 July 2015 / Published online: 12 August 2015
© Springer-Verlag Wien 2015

Abstract The genome of tetterwort vein chlorosis virus (TVCV) from South Korea has been completely sequenced. Its genomic organization resembles those of other criniviruses, with several new features, indicating that TVCV is a member of a new species in the genus *Crinivirus*, family *Closteroviridae*. RNA1 contains 8467 nucleotides, with at least four opening reading frames (ORFs). ORF1a encodes a protein with predicted papain-like protease, methyltransferase, and helicase activities. ORF1b encodes a putative RNA-dependent RNA polymerase that is apparently expressed through a +1 ribosomal frameshift. RNA2 contains 8113 nucleotides encoding at least nine proteins, similar to most crinivirus RNA2s. The 3′ untranslated regions of the bipartite RNA genome share 82.1 % nucleotide sequence identity.

Introduction

Crinivirus is a genus in the family *Closteroviridae*, which also contains the genera *Closterovirus*, *Ampelovirus* and *Velarivirus*. Members of the family *Closteroviridae* have positive-sense, single-stranded genomes encapsidated in long flexuous virions [1–4]. Criniviruses have bipartite genomes and are transmitted by whiteflies, whereas viruses belonging to the genera *Closterovirus* and *Ampelovirus*, which have monopartite genomes, are transmitted by aphids and mealybugs, respectively [5]. Unlike the other three genera in the family *Closteroviridae*, no known insect vector has been identified for viruses in the genus *Velarivirus* [6]. Tetterwort, otherwise known as *Chelidonium majus*, in the family *Papaveraceae*, is a valuable herbaceous plant used for the treatment of various diseases, including gastroenteritis, and as an anodyne. Tetterwort produces small, bright yellow flowers and is a common flowering plant cultivated in gardens in South Korea. Here, we present the complete genomic sequence of TVCV, identified by RNA sequencing, and its relationships to other *Crinivirus* members.

F. Zhao and R. H. Yoo contributed equally to this study.

Electronic supplementary material The online version of this article (doi:10.1007/s00705-015-2552-6) contains supplementary material, which is available to authorized users.

✉ Su-Heon Lee
suheon@knu.ac.kr

✉ Jae Sun Moon
jsmoon@kribb.re.kr

¹ Biosystems and Bioengineering Program, University of Science and Technology (UST), Daejeon 305-350, South Korea

² Plant Systems Engineering Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-806, South Korea

³ School of Applied Biosciences, Kyungpook National University, Daegu 702-701, South Korea

Materials and methods

Tetterwort plants showing typical symptoms of vein chlorosis and leaf distortion were collected in Yesan, South Korea. Total RNA was extracted using TRI Reagent (MRC, Cincinnati, OH, USA) and the ribosomal RNA was removed using a Ribo-Zero rRNA Removal Kit (Plant Leaf) (Epicentre, Madison, WI, USA), according to the manufacturer's instructions. The remaining RNA was then used for construction of an RNA sequencing library using a TruSeq RNA Sample Prep Kit (Illumina, San Diego, CA, USA). The prepared sample was sequenced on an Illumina HiSeq 2500

Sequencing System, generating approximately 40 gigabase pairs of DNA sequence. BLAST results indicated that sequences of the seven viral fragments produced after contig assembly were homologous to, but distinct from, those of the criniviruses lettuce chlorosis virus (LCV), cucurbit yellow stunting disorder virus (CYSDV), and cucurbit chlorotic yellows virus (CCYV). Four of the seven sequences were related to RNA1, and the other three were related to RNA2. Reverse transcription polymerase chain reaction (RT-PCR) was used to reamplify the complete sequences of those seven fragments from an RNA template isolated from tetterwort. The PCR amplicons were sequenced directly and shown to be consistent with the original sequences. This indicates that the tetterwort plant was probably infected with an uncharacterized crinivirus. To obtain the continuous sequence, specific primers originating from two adjacent fragments were designed to amplify the missing regions. All PCR procedures for joining the fragments were performed with Pfu-X DNA Polymerase (Solgent, Daejeon, South Korea). The 5'-terminal sequences of both RNAs were determined using a previously reported method [7]. The 3'-terminal regions of the bipartite genomes were generated with the addition of poly(A) tails to the purified RNA with a poly(A) polymerase tailing kit (Epicentre), which was then used as the template for RT with an oligo(dT)₁₈ primer. The nucleotide sequences of RNA1 and RNA2 of TVCV were deposited in GenBank under accession numbers KR002686 and KR002687, respectively. Open reading frames (ORFs) were identified in RNA1 and RNA2 with ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>), and transmembrane helices in the predicted proteins were identified with the TMHMM program (<http://www.cbs.dtu.dk/services/TMHMM/>) [8]. The deduced amino acid sequences of selected proteins were aligned using ClustalW [9] to construct phylogenetic trees by the neighbor-joining method using MEGA 6 (Fig. 2) [10].

Sequence properties

TVCV RNA1 is 8467 nucleotides (nt) long and contains four ORFs, designated (by analogy to other criniviruses) as ORF1a (nt 67–6045), ORF1b (nt 6044–7561), P6a (ORF2, nt 7577–7738) and P19 (ORF3, nt 7741–8232). ORF1a encodes a predicted polyprotein of 229 kDa, with papain-like protease (P-PRO), methyltransferase (MT), and helicase (HEL) domains (Fig. 1). In an alignment of the TVCV P-PRO sequence with those of other criniviruses, the presumed cleavage site of the putative protease domain is located between Gly₄₇₆ and Val₄₇₇, and the two catalytic residues identified in the protease domain are Cys₄₀₈ and His₄₅₇. Four transmembrane helices (amino acids 1240–1259, 1266–1283, 1287–1309 and 1353–1375) were predicted in the region between the MT and HEL motifs. As in most other criniviruses, ORF1a terminates with UUUGA and overlaps ORF1b, which putatively encodes an RNA-dependent RNA polymerase (RdRp) that is presumably expressed through a +1 ribosomal frameshift. P6a shows low similarity to its orthologue in bean yellow disorder virus (BnYDV) and includes a transmembrane helix. P19 encodes a putative 19-kDa protein with no significant sequence homology to any other viral protein available in the public databases. TVCV RNA2 (8113 nt) includes the hallmark gene array of the family *Closteroviridae*, encoding a 62.7-kDa heat shock protein 70 family homologue (Hsp70h, nt 1331–3004), a 22.7-kDa major coat protein (CP, nt 5055–5780) and a 54-kDa minor coat protein (CPm, nt 5780–7204) (Fig. 1). P6b (6.4 kDa), P60 (59.7 kDa), and P9b (9.7 kDa), encoded between the Hsp70h and CP ORFs, together with P27 (26.9 kDa), encoded downstream from CPm, are similar to the corresponding proteins in other criniviruses. No transmembrane helix was found in these four ORFs (Fig. 1). Two unique small ORFs, encoding two putative proteins of 9.5 kDa

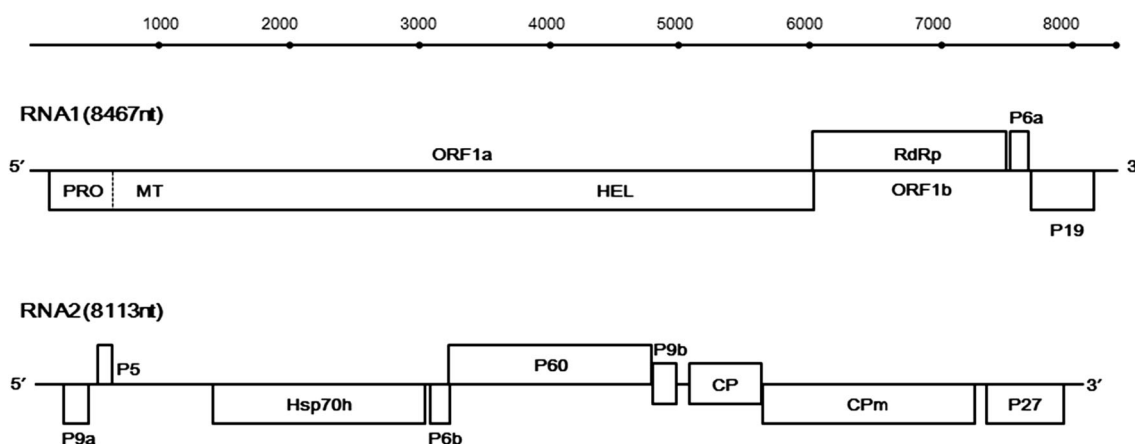


Fig. 1 Genomic structures of TVCV RNA1 and RNA2. Rectangles represent open reading frames encoding putative viral proteins. The ORFs encoding proteins P6a and P19 in RNA1 and P5, P6b, P9a/b,

and P27 in RNA2 are named according to the predicted molecular mass of each protein

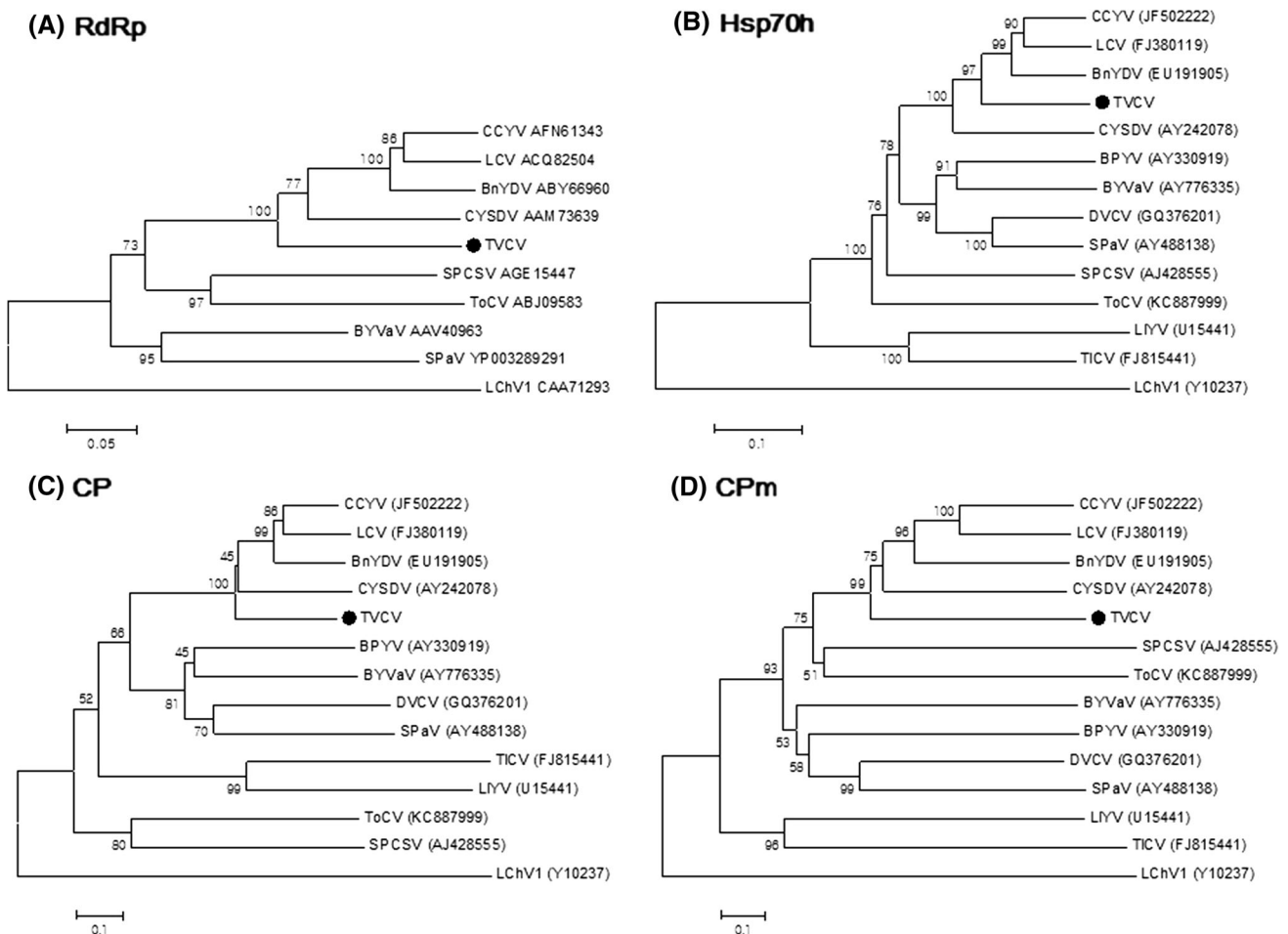


Fig. 2 Phylogenetic relationships between TVCV and other members in the family *Closteroviridae*. Phylogenetic trees were constructed based on the RdRp (A), Hsp70h (B), CP (C), and CPM (D) deduced amino acid sequences using the neighbor-joining method. The GenBank accession number of each sequence is given on the tree. Bean yellow disorder virus, BnYDV; beet pseudo-yellows virus, BPYV; blackberry yellow vein-associated virus, BYVaV; cucurbit

(P9a) and 5.4 kDa (P5), were also identified at the 5'-terminal region of RNA2. P9a is likely to be a component of the membrane, whereas P5 has a transmembrane domain between amino acids 2 and 24, but its function is uncharacterized. The 3'-untranslated region (UTR) of TVCV RNA2 is 213 nt long and shares 82.1 % identity with that of RNA1, which is 235 nt long.

Pairwise comparisons of the amino acid sequences of the highly conserved RdRp and CP of TVCV and those of other criniviruses indicated that TVCV is most closely related to BnYDV, sharing the highest amino acid identities of 75.8 % and 62.4 %, respectively, with the BnYDV orthologues. Phylogenetic analysis of the RdRp, Hsp70h, CP, and CPM amino acid sequences of criniviruses placed TVCV in a cluster including BnYDV, CYSDV, LCV, and CCYV (Fig. 2). In conclusion, the differences and similarities in the genome organization (ORF numbers, sizes,

chlorotic yellows virus, CCYV; cucurbit yellow stunting disorder virus, CYSDV; diodia vein chlorosis virus, DVCV; little cherry virus, LChV1; lettuce chlorosis virus, LCV; lettuce infectious yellows virus, LIYV; strawberry pallidosis-associated virus, SpaV; sweet potato chlorotic stunt virus, SPCSV; tomato infectious chlorosis virus, TICV; tomato chlorosis virus, ToCV

and positions) of TVCV and other criniviruses available in the database, as described above, support the proposition that TVCV is a new member of the genus *Crinivirus* [5].

Acknowledgments This work was performed with the support of the Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ01007703 and Project No. PJ01112803), Rural Development Administration, Republic of Korea.

References

- Agranovsky AA (1996) Principles of molecular organization, expression and evolution of closteroviruses: over the barriers. *Adv Virus Res* 47:119–158
- Al Rwahnih M, Dolja VV, Daubert S, Koonin EV, Rowhani A (2012) Genomic and biological analysis of Grapevine leafroll-associated virus 7 reveals a possible new genus within the family *Closteroviridae*. *Virus Res* 163:302–309

3. Karasev AV (2000) Genetic diversity and evolution of Closteroviruses. *Annu Rev Phytopathol* 38:293–324
4. Martelli GP, Abou Ghanem-Sabanadzovic N, Agranovsky AA, Al Rwahnih M, Dolja VV, Dovaş CI, Fuchs M, Gugerli P, Hu JS, Jelkmann W (2012) Taxonomic revision of the family *Closteroviridae* with special reference to the grapevine leafroll-associated members of the genus *Ampelovirus* and the putative species unassigned to the family. *J. Plant Pathol* 94:7–19
5. Martelli GP, Agranovsky AA, Bar-Joseph M, Boscia D, Candresse T, Coutts RHA, Dolja VV, Hu JS, Jelkmann W, Karasev AV, Martin RR, Minafra A, Namba S, Vetten HJ (2011) Family Closteroviridae. In: King A, Adams MJ, Carstens EB, Lefkowitz E (eds) *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*. Elsevier-Academic Press, Amsterdam, pp 987–1001
6. Melzer M, Ayin C, Sugano J, Uchida J, Kawate M, Borth W, Hu J (2013) Differentiation and distribution of Cordyline viruses 1–4 in Hawaiian ti plants (*Cordyline fruticosa* L.). *Viruses* 5:1655–1663
7. Nam M, Kim SM, Domier LL, Koh S, Moon JK, Choi HS, Kim HG, Moon JS, Lee SH (2009) Nucleotide sequence and genomic organization of a newly identified member of the genus *Carmovirus*, soybean yellow mottle mosaic virus, from soybean. *Arch Virol* 154:1679–1684
8. Krogh A, Larsson B, von Heijne B, Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567–580
9. Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
10. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30:2725–2729