Arch Virol (2017) 162:1373–1379
DOI 10.1007/s00705-017-3236-1

BRIEF REPORT

Complete nucleotide sequence of clematis chlorotic mottle virus, a new member of the family *Tombusviridae*

Margaret McLaughlin[1] · Ben Lockhart[2] · Ramon Jordan[3] · Geoff Denton[4] · Dimitre Mollov[1]

Received: 21 September 2016 / Accepted: 3 January 2017 / Published online: 30 January 2017

Abstract Clematis chlorotic mottle virus (ClCMV) is a previously undescribed virus associated with symptoms of yellow mottling and veining, chlorotic ring spots, line pattern mosaics, and flower distortion and discoloration on ornamental Clematis. The ClCMV genome is 3,880 nt in length with five open reading frames (ORFs) encoding a 27-kDa protein (ORF 1), an 87-kDa replicase protein (ORF 2), two centrally located movement proteins (ORF 3 and 4), and a 37-kDa capsid protein (ORF 5). Based on morphological, genomic, and phylogenetic analysis, ClCMV is predicted to be a member of the genus *Pelarspovirus* in the family *Tombusviridae*.

Clematis is a genus of woody climbing vines in the family *Ranunculaceae* widely used as an ornamental [16]. The plant is popularly called by its genus name. Because it is a perennial crop, systemic diseases, especially viruses, can have a considerable economic impact in landscape settings by reducing the vigor and aesthetic value of the plant.

Electronic supplementary material The online version of this article (doi:10.1007/s00705-017-3236-1) contains supplementary material, which is available to authorized users.

✉ Dimitre Mollov
dimitre.mollov@ars.usda.gov

[1] USDA ARS, National Germplasm Resources Laboratory, Beltsville, MD 20705, USA

[2] Department of Plant Pathology, University of Minnesota, St. Paul, MN 55108, USA

[3] USDA ARS, U.S. National Arboretum, Floral and Nursery Plant Research Unit, Beltsville, MD 20705, USA

[4] Royal Horticultural Society, Woking, Surrey GU23 6QB, UK

Viruses reported to infect clematis include tobacco streak virus (TSV) and cucumber mosaic virus (CMV) in the family *Bromoviridae* [6], tomato ringspot virus (ToRSV) in the family *Secoviridae* [3], and tomato spotted wilt virus (TSWV) in the family *Bunyaviridae* [3]. To date, there have been no viruses of the family *Tombusviridae* reported to infect clematis.

We previously reported a spherical virus (28-30 nm) associated with disease in clematis [9]. Plants obtained from the United States and the United Kingdom displayed yellow mottling and veining, chlorotic ring spots, line pattern mosaics, and flower distortion and discoloration suggestive of viral infection (Fig. 1A-E). From 2007 to 2014, more than 80 plants originating from at least four states in the Midwest and 38 samples from the U.K. tested positive for this spherical virus by transmission electron microscopy (TEM) and/or RT-PCR, with additional confirmation by sequencing. This previously undescribed virus was provisionally named clematis chlorotic mottle virus (ClCMV) and shows structural and genomic similarity to members of the family *Tombusviridae*. Preliminary analysis suggests that ClCMV may belong in the newly formed genus *Pelarspovirus* [14].

Tombusviridae is a family of plant viruses with a single-stranded, positive-sense, RNA genome. The International Committee on Taxonomy of Viruses (ICTV) currently recognizes 16 genera and 72 species of *Tombusviridae*. The sixteen genera of this family are *Tombus-, Alphacarmo-, Betacarmo-, Gammacarmo-, Aureus-, Avena-, Machlomo-, Alphanecro-, Betanecro-, Panico-, Diantho-, Gallanti-, Macana-, Zea-, Pelarspo-,* and *Umbravirus* (ICTV 2015). Viruses of this family are classified into their genera based on several structural and genomic criteria. Structurally, the virions are spherical in shape with $T = 3$ icosahedral symmetry and may have either a smooth or bumpy

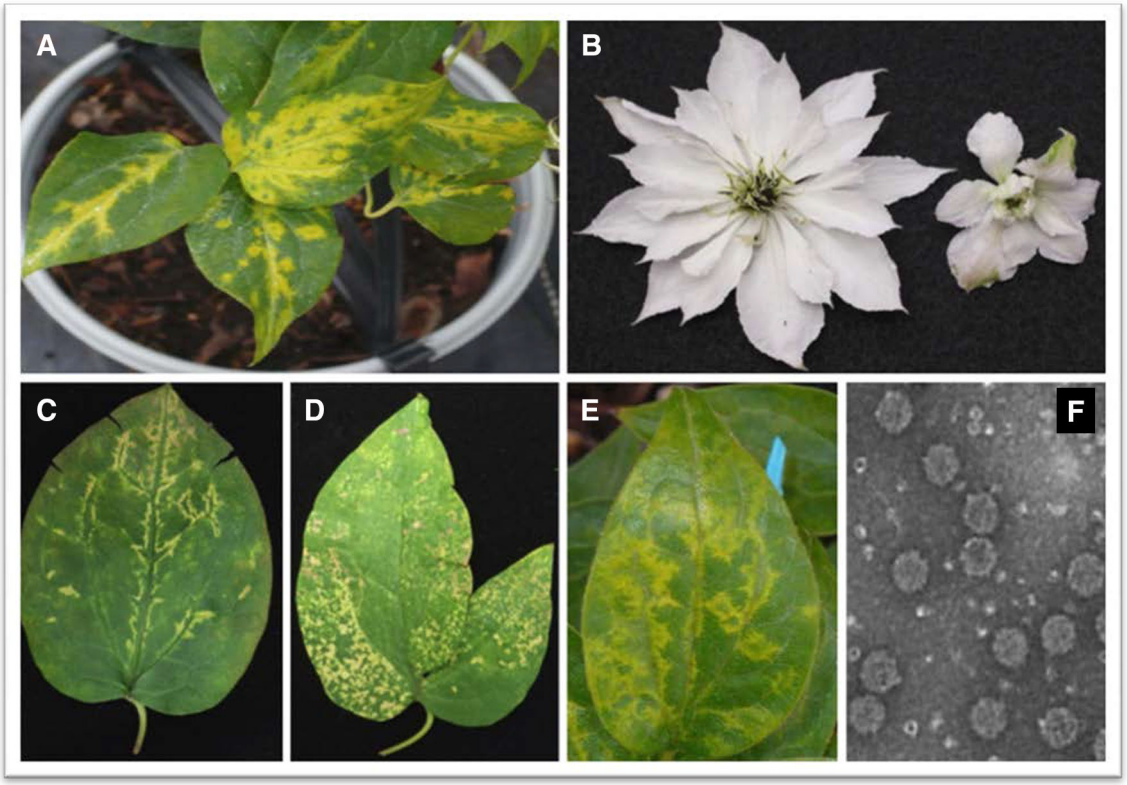

Fig. 1 Symptoms observed in clematis infected with ClCMV (virion morphology shown in F). A) chlorosis and chlorotic ring spots. B) Flower distortion and discoloration in cv. Dancing Dorien (right) and a flower from an uninfected plant (left). C-E) Yellow veining, mosaic, and yellow mottling, respectively, on infected cultivars. F) Spherical viral particles from a purified fraction observed using transmission electron microscopy

appearance [13]. Members of the family *Tombusviridae* have genomes that are non-segmented (with the exception of the genus *Dianthovirus*) and are neither 5'-capped nor 3'-polyadenylated [13]. Translation is initiated by a 3' cap-independent translational enhancer [7]. Viruses in this family encode several proteins responsible for RNA synthesis, virus assembly, and cell-to-cell viral transport [13]. A viral RNA-dependent RNA polymerase (RdRp) and an accessory protein are involved in viral synthesis [13]. These viral RdRps are a product of either a readthrough or frameshift element in the 5' ORF, depending on the genus [7]. Viral assembly is facilitated by coat proteins (CPs), and cell-to-cell viral transport is mediated by one or more proteins that enable viral transit through the plasmodesmata [7]. Another characteristic is the presence of virally encoded suppressors of RNA silencing (VSRs), which vary in structure and function among different genera [2].

Members of the genus *Pelarspovirus* have several distinguishing features. They utilize three different reading frames to translate five predicted gene products [14]. The five ORFs encode two proteins that aid in replication, two centrally located movement proteins, and the capsid protein [14]. The two movement proteins and capsid protein are encoded in a single subgenomic RNA. Members of this genus are also reported to contain a non-AUG start codon initiating the ORF containing MP2 [14]. The final distinguishing characteristic of pelarspoviruses is the lack of an AUG codon in any frame between the AUG initiation codon of MP1 and the CP open reading frame [14].

The objectives of this study were to characterize the ClCMV genome and confirm its taxonomic and phylogenic relationship to known viruses.

A ClCMV isolate was obtained from naturally infected clematis plants in Minnesota displaying the symptoms described above. Virions were purified from 40 g of symptomatic leaves, using a previously described protocol [10]. To the virion solution of 500 µl an equal volume of chloroform was added and the sample was mixed by vortexing, and centrifuged at 10,000g_{max} for 10 minutes. Aliquots of the supernatant were negatively stained separately with sodium phosphotungstate (PTA) and uranyl acetate (UA) for examination by transmission electron microscopy (TEM), and spherical virus particles 28-30 nm in diameter were observed (Fig. 1F), consistent with members of the family *Tombusviridae*. No other viral particles were observed. Virion nucleic acids were extracted from the remaining supernatant solution by the addition of SDS, phenol, and chloroform. Nucleic acids

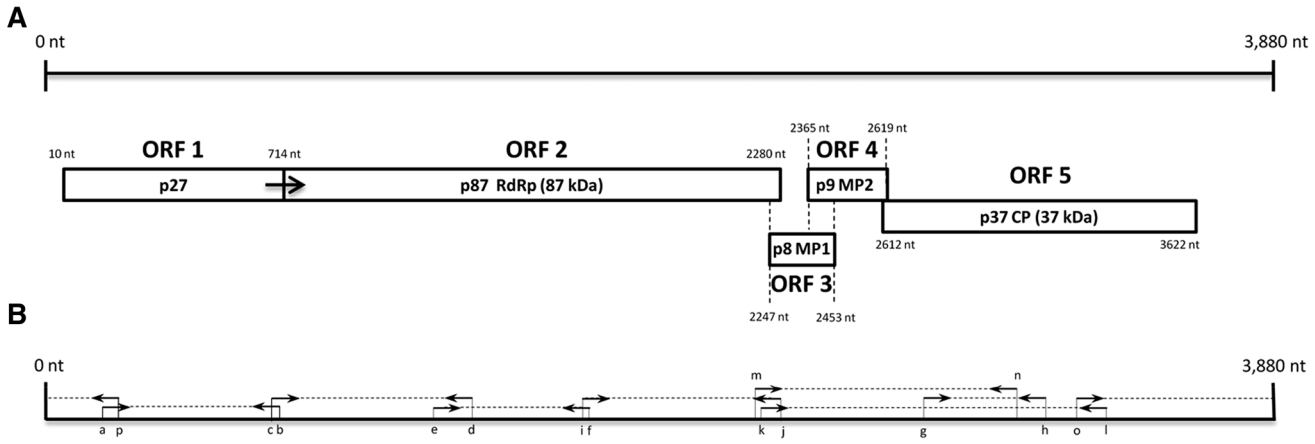

Fig. 2 A) Schematic representation of the clematis chlorotic mottle virus (ClCMV) genome organization. ORFs 1, 2, and 4 are located in frame 1, ORF 5 is located in frame 2, and ORF 3 is located in frame 3. An arrow between ORF 1 and ORF 2 represents a readthough amber stop codon. B) Positions of the primers used to amplify the ClCMV genome (see Supplementary Table 1)

were precipitated with isopropanol. The solution was centrifuged, and the resulting pellet was washed with 70% ethanol and dried under vacuum. The pellets were resuspended in RNase-free water.

Virion RNA was used to prepare a genomic library using the method described by Froussard [5]. Viral cDNA library fragments were amplified and cloned using a pGEM T-Easy^{TM} kit (Promega, Madison WI). Plasmid DNA was extracted from clones using a Zippy^{TM} Plasmid DNA extraction kit (Zymo Research) and sequenced in both directions. Sequence assembly was performed using Sequencher 4.10.1 software (Gene Codes Corporation, Ann Arbor, MI). Primer pairs designed to amplify overlapping fragments covering the entire genome were used for RT-PCR reactions. Amplicons were cloned, sequenced, and assembled as described above. The 5'-terminal sequence was determined using a second-generation RACE kit (Roche Applied Science) and a ClCMV-specific primer. The resulting sequence was cloned using a pGEM T-Easy^{TM} kit (Promega, Madison WI). Four clones were sequenced in both directions. The 5'-terminal sequence was then assembled using Geneious R9 (Biomatters, New Zealand). To determine the 3'-terminal sequence, purified viral RNA was first polyadenylated using poly(A) polymerase (Ambion) according to manufacturer's instructions. Rapid amplification of 3' cDNA ends (3' RACE) was performed using an oligo-dT_{17}V primer and 2 ClCMV-specific primers with a Clonetech SMARTer RACE 5'/3' kit. PCR amplicons were cloned into the RACE-kit-supplied In-Fusion HD cloning vector and six clones were sequenced in both directions. The complete ClCMV genome sequence was assembled using Geneious R9.

The full ClCMV genome sequence was annotated and submitted to the NCBI GenBank database under accession number KX712140. The open reading frames (ORFs) were identified using Geneious R9 and later confirmed using NCBI ORF Finder [17]. Each ORF sequence was analyzed using BLASTP to identify similarities to known viral sequences [1].

The complete genomic sequence of ClCMV is 3,880 nt long with five open reading frames (ORFs) (Fig. 2). The first ORF is 702 nucleotides (234 codons) long and encodes a p27 (27-kDa) protein that has 50% sequence identity to the p27 protein of pelargonium line pattern virus (PLPV). At nt position 703-705, ORF1 terminates at an amber stop codon (UAG), which is predicted to be read through to produce the 87-kDa replicase protein (ORF 2). The sequence surrounding the UAG stop codon is 5'-AAA **UAG** GGA-3' (nt 700-708), which is consistent with the optimal consensus context (5'-AAA **UAG**GG(G/A)-3') for type III readthrough elements [4]. This translational readthrough is thought to involve the formation of a pseudoknot in the RNA that serves as a platform for replicase complex assembly [12]. Members of the family *Tombusviridae* contain conserved motifs within the polymerase coding region that enable this RNA-RNA interaction [4]. Mfold RNA secondary structure prediction analysis [18] of the genome of CClMV revealed an internal replication element (IRE) within the RdRp (see Supplementary Figure S1) similar to those found in other members of the family *Tombusviridae*, including viruses of the genus *Pelarspovirus* [12, 14]. This IRE contains one distinct conserved motif (5'-CAU UCC-3'/5'-GGC AGG-3') that has been identified in other members of the genus *Pelarspovirus* [12]. This motif is present in the replicase coding region (ORF2) of ClCMV at nt 1020 and nt 1069.

The RNA-dependent RNA polymerase (RdRp) encoded by the second ORF has 58% sequence identity to the RdRp (YP 009116634) of elderberry latent virus (ELV) and a 57% identity to the RdRp (AHZ59467) of Rosa rugosa leaf

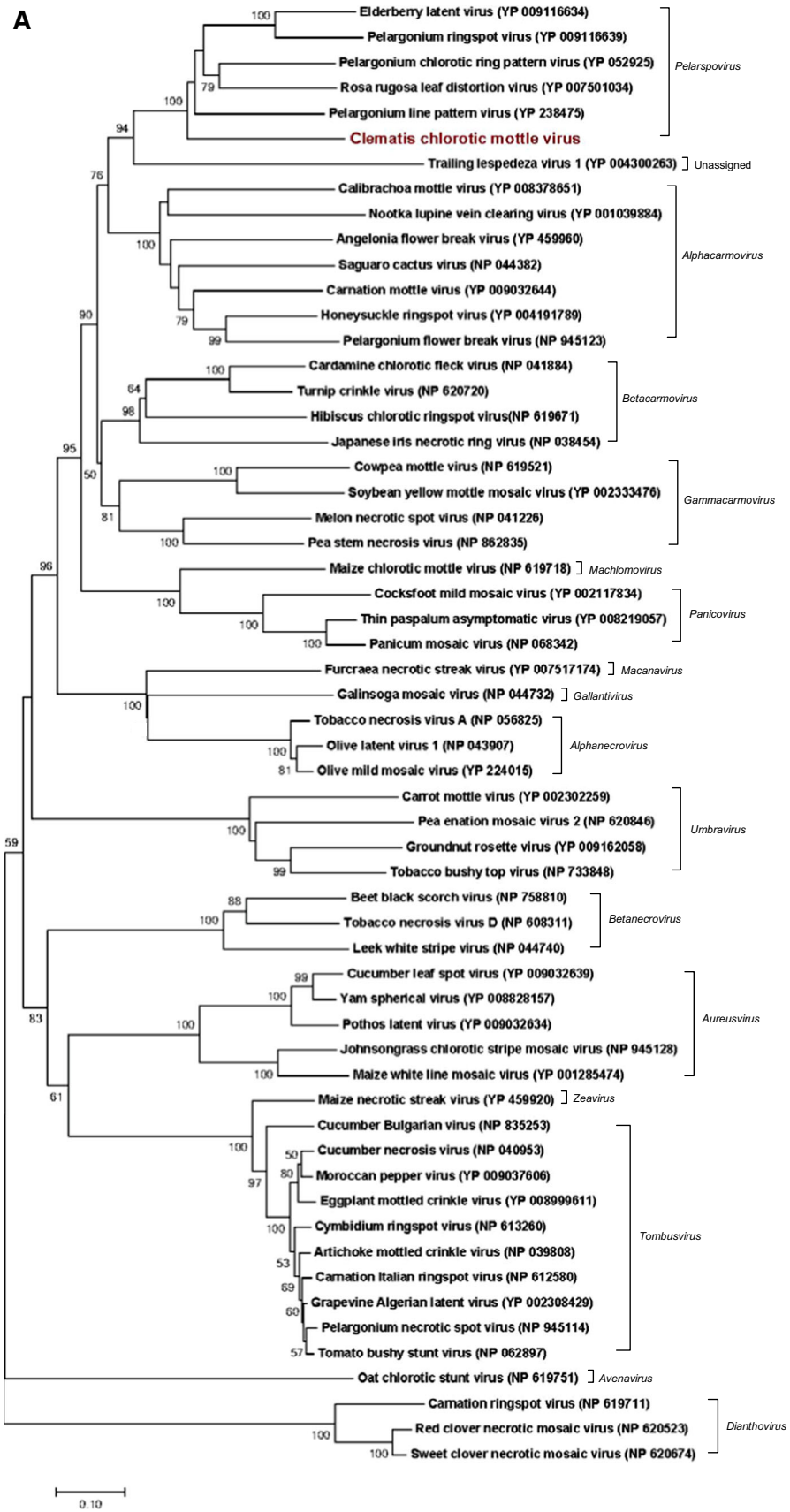

Fig. 3 Phylogenetic analysis of clematis chlorotic mottle virus (accession number KX712140) and selected members of the family *Tombusviridae* based on (A) amino acid sequences of the replicase and (B) amino acid sequences of the capsid protein. Members of the genus *Umbravirus* were omitted from the tree in panel B due to the absence of a capsid protein encoded in their genomes. Phylogenetic trees were constructed using the maximum-likelihood algorithm. The numerical values represent bootstrap (1000 replicates) values above 50. The scale bars correspond to substitutions per amino acid site. The accession numbers of the amino acid sequences are listed with each virus name

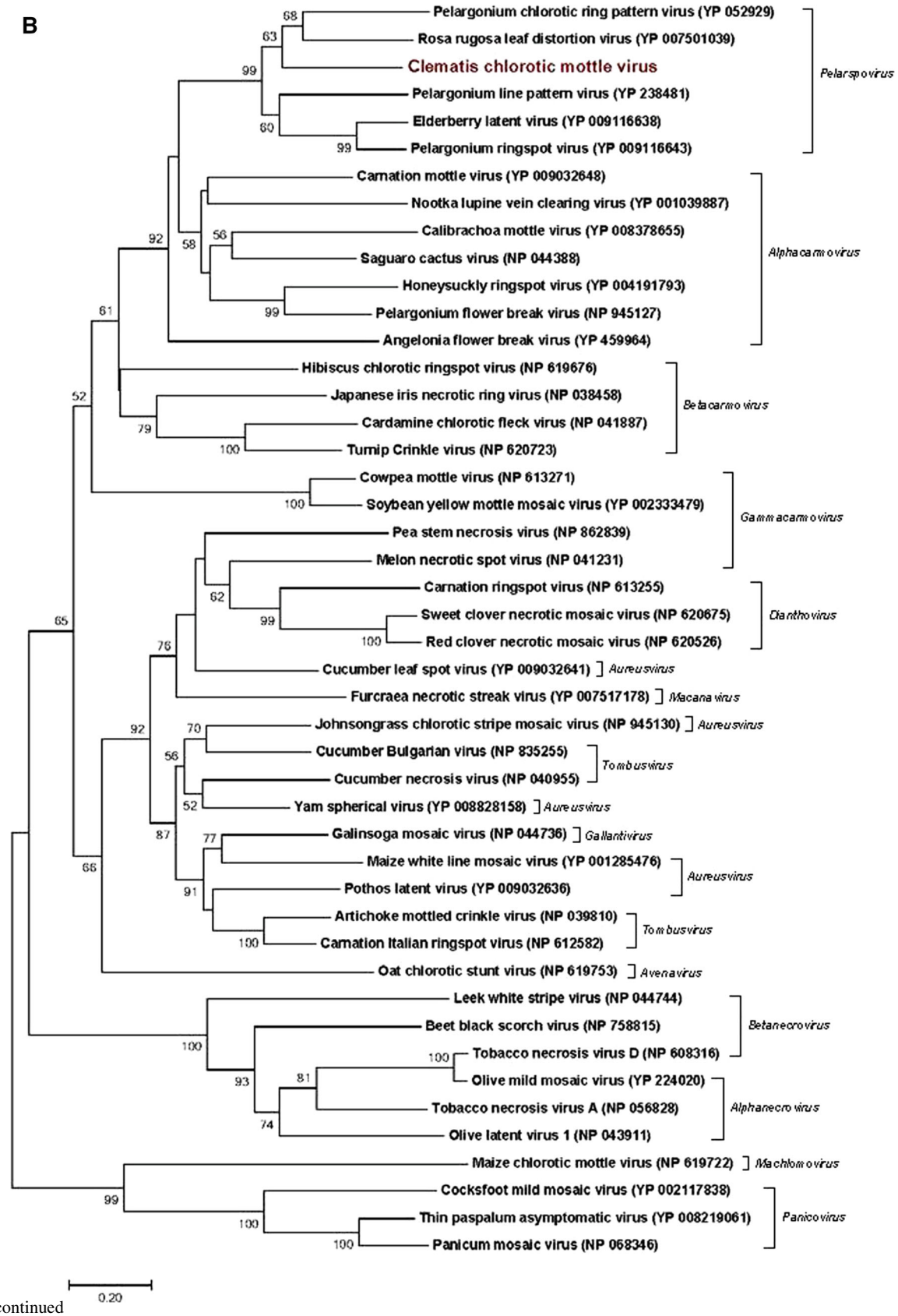

Fig. 3 continued

distortion virus (RrLDV). ORF 3 (nt 2247-2453) encodes an 8-kDa movement protein (MP1) that shows a 70% identity to the movement protein (YP 238479) of PLPV. The fourth ORF sequence is located in frame 1 and overlaps with ORF 3, which is located in frame 2. ORF 4 contains a predicted non-canonical start codon, UUG (nt 2247-2249), flanked by an optimal Kozak consensus sequence with a guanine at the +4 position and purine at the -3 position [11, 14], and it is predicted to encode a 9-kDa movement protein (MP2). This ORF has a 52% sequence identity to PLPV (AHZ59470) MP2 and 49% identity to the p12 protein (YP 007501037) of RrLDV. ORF 5 (2612-3622nt) encodes a 37-kDa capsid protein with a 47% sequence identity to the capsid protein (YP 007501039) of RrLDV. Like other members of the genus *Pelarspovirus*, there is no AUG codon located between the MP1 initiation site (ORF 3; nt 2247-2249) and the CP initiation site (nt 2612-2614).

Two primer pairs were used for RT-PCR detection of ClCMV: ClCP_F, GCGCCAAGAACAAGCTGGAC and ClCP_R, CTTAGCCCAGGCCCAGATGA, amplifying a 1043-bp fragment within the capsid protein ORF, and ClRdRp_F, CATCTTGCATTTGGCATTAG and ClCP_R2, AGGGAATCTGCGTGGATAGA, amplifying a 768-bp fragment between the RdRp and CP genes, covering the MP1 and MP2 ORFs (see Supplementary Table 1). Subsequent cloning and sequencing confirmed that the amplicons were identical to the ClCMV sequence described above.

In the U.S., the virus was detected in over 80 samples of clematis cultivars, including 'Barbara', 'Bees Jubilee', 'Dancing Dorien', 'Gypsy Queen', 'Henryi', 'Jackmanii', 'Natascha', 'Niobe', and 'Still Waters'. Thirty-five of 38 symptomatic plants received from the U.K. also tested positive for this virus, including the following varieties: 'Barbara', 'Beata', 'Black Tea', 'Diana', 'Dorthy Walton', 'Gypsy Queen', 'Honoras', 'Jackmanii', 'John Huxtable', 'Kinju Atarashi', 'Lady Betty Balfor', 'Margaret Hunt', 'Marmori', 'North Star Ponjnel', Perle d'Azur', 'Rahvarinne', 'Ramona', 'Caddick's Cascade Semu', 'Tie Dye', 'Valge Daam', 'Hagley Hybrid', 'Dutch Sky Throcroft', 'Marmori', and 'Ville de Lyon'.

Phylogenetic analysis was done by comparing the 87-kDa replicase (ORF 2) and 37 kDa capsid protein (ORF 5) amino acid sequences of ClCMV to those of other members of the family *Tombusviridae*. The sequences were aligned with MEGA 7 software [8] using the ClustalW algorithm [15]. Once aligned, two phylogenetic trees were generated using the neighbor-joining method in MEGA 7 software with a maximum-likelihood algorithm of 1,000 bootstrap replicates [8]. Bootstrap values above 50% are displayed. Phylogenetic trees were generated using the RdRp sequences (Fig. 3A) and the capsid protein sequences (Fig. 3B) of select *Tombusviridae* viruses of multiple genera. In both trees, ClCMV groups most closely with other members of the genus *Pelarspovirus*.

Based on morphological, genomic, and phylogenetic criteria, we conclude that clematis chlorotic mottle virus (ClCMV) is a previously uncharacterized member of the genus *Pelarspovirus* in the family *Tombusviridae*, and the first virus of this family reported to infect clematis.

Compliance with ethical standards

All authors declare they have no conflict of interest. The research reported in this manuscript did not involve experimentation with human or animal subjects. Plant samples were moved within the U.S.A. under USDA APHIS permit P526P-13-03719 and from the U.K. under permit P526P-14-00688.

References

1. Altschul SF, Madden TL, Schäffer AA et al (1997) Gapped BLAST and PSI-BLAST:a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402
2. Burgyan J, Havelda Z (2011) Viral suppressors of RNA silencing. Trends Plant Sci 16:265–272
3. Chatzivassiliou EK, Boubourakas I, Drossos E et al (2002) Weeds in greenhouses and tobacco fields are differentially infected by tomato spotted wilt virus and infested by its vector species. Plant Dis 85:40–46
4. Firth AE, Brierley I (2012) Non-canonical translation in RNA viruses. J Gen Virol 93:1385–1409
5. Froussard P (1992) A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. Nucleic Acids Res 20:2900
6. Guy PL (2006) Effects and incidence of Cucumber mosaic virus, Watermelon mosaic virus and Zucchini yellow mosaic virus in New Zealand's only native cucurbit, Sicyos australis. Aust Plant Pathol 35:29–35
7. Jiwan SD, Wu B, White KA (2011) Subgenomic mRNA transcription in *Tombusviridae*. RNA Biology 8(2):287–294
8. Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:msw054. doi:10.1093/molbev/msw054
9. Mollov D, Lockhart B, Phibbs A, Creswell T, Ruhl G, Dorman E, Kinard G, Jordan R (2014) Clematis chlorotic mottle virus, a novel virus occurring in clematis in the USA. Phyopathology 104(S3):81
10. Mollov D, Lockhart B, Zlesak D (2013) Complete nucleotide sequence of rose yellow mosaic virus, a novel member of the family Potyviridae. Arch Virol 158:1917–1923
11. Nakagawa S, Niimura Y, Gojobori T et al (2008) Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. Nucleic Acids Res 36(3):861–871
12. Nicholson BL, Lee PKK, White KA (2012) Internal RNA replication elements are prevalent in *Tombusviridae*. Front Microbiol 3:279
13. Rochon D, Rubino L, Russo M et al (2012) Virus taxonomy. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds) Virus taxonomy, 9th edn. Elsevier, Amsterdam, pp 1111–1138
14. Scheets K, Jordan R, White KA, Hernandez C (2015) Pelarspovirus, a proposed new genus in the family Tombusviridae. Arch Virol 160:2385–2393

15. 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
16. Wang W-T (2000) Notes on the genus *Clematis* (Ranunculaceae). Acta Phytotaxon Sin 38:305–336
17. Wheeler DL, Church DM, Federhen S et al (2003) Database resources of the national center for biotechnology. Nucleic Acids Res 31:28–33
18. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research 31:3406–3415