e-**X**tra*

Identification and Characterization of Two Novel Geminiviruses Associated with Paper Mulberry (*Broussonetia papyrifera*) Leaf Curl Disease

Yuanjian Qiu,[1,2] Song Zhang,[1,2] Haodong Yu,[1,2] Zhiyou Xuan,[1,2] Liu Yang,[1,2] Binhui Zhan,[3] F. Murilo Zerbini,[4] and Mengji Cao[1,2,†]

[1] National Citrus Engineering and Technology Research Center, Citrus Research Institute, Southwest University, Beibei, Chongqing 400712, China
[2] Academy of Agricultural Sciences, Southwest University, Chongqing 400715, China
[3] State Key Laboratory for Biology of Plant Disease and Insect Pest, Institute of Plant Protection, Academy of Agricultural Sciences, Beijing 100193, China
[4] Departamento de Fitopatologia/BIOAGRO, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

Abstract

Paper mulberry (*Broussonetia papyrifera*) is a perennial woody plant used as source material for Cai Lun paper making, in traditional Chinese medicine, and as livestock feed. To identify the presence of viruses in paper mulberry plants affected by a disease with leaf curl symptoms, high-throughput sequencing of total RNA was performed. Analysis of transcriptome libraries allowed the reconstruction of two geminivirus-like genomes. Rolling-circle amplification and PCR with back-to-back primers confirmed the presence of two geminiviruses with monopartite genomes in these plants, with the names paper mulberry leaf curl virus 1 and 2 (PMLCV-1 and PMLCV-2) proposed. The genomes of PMLCV-1 (3,056 nt) and PMLCV-2 (3,757 to 3,763 nt) encode six proteins, with the V4 protein of PMLCV-1 and the V3 proteins of both viruses having low similarities to any known protein in databases.

Alternative splicing of an intron, akin to that of mastre-, becurto-, capula-, and grabloviruses, was identified by small RNA (sRNA)-seq and RNA-seq reads mapping to PMLCV-1 and PMLCV-2 antisense transcripts. Phylogenetic analyses and pairwise comparisons showed that PMLCV-1 and PMLCV-2 are most closely related to, but distinct from, two unassigned geminiviruses, citrus chlorotic dwarf associated virus and mulberry mosaic dwarf associated virus, suggesting that they are two new members of the family *Geminiviridae*. Field investigation confirmed the close association of the two viruses with leaf curl symptoms in paper mulberry plants and that coinfection can aggravate the symptoms.

Keywords: geminivirus, synergy, high-throughput sequencing, phylogeny

Paper mulberry (*Broussonetia papyrifera* [L.] Vent.), a broad-leaf, deciduous woody plant in the family Moraceae, originates from China and is now widely distributed in temperate and tropical regions of the world (Chang et al. 2015; Morgan and Overholt 2004; Peng et al. 2015). Paper mulberry is cultivated in most provinces of China, except those located in the northeastern and northwestern regions (Shen and Peng 2017). The plant, historically known as the source material of Cai Lun paper-making (Peng et al. 2019), is listed in the Chinese *Compendium of Materia Medica* and other pharmacopoeias (Li and Jia 1997). Moreover, it also can be used in livestock feed, barkcloth making, and afforestation, among other uses (Bell 1988). Nevertheless, pathogens of paper mulberry have not yet been well studied.

The geminiviruses (family *Geminiviridae*) comprise a group of plant DNA viruses with a circular, single-strand DNA genome encapsidated in twinned (geminate) icosahedral particles, and they infect a wide range of plant species in most tropical and subtropical regions of the world. Based on the type of insect vector (whiteflies, leafhoppers, treehoppers, or aphids), genome organization (mono- or bipartite), and host range (mono- or dicot plant species), geminiviruses are classified into nine genera: *Becurtovirus*, *Begomovirus*, *Capulavirus*, *Curtovirus*, *Eragrovirus*, *Grablovirus*, *Mastrevirus*, *Topocuvirus*, and *Turncurtovirus* (Zerbini et al. 2017). In addition, several divergent geminiviruses discovered with the application of rolling-circle amplification (RCA) and high-throughput sequencing (HTS) technologies (Haible et al. 2006; Roossinck et al. 2015; Wu et al. 2010), mostly from woody plants, have not yet been assigned to a genus, including citrus chlorotic dwarf associated virus (CCDaV; Loconsole et al. 2012), mulberry mosaic dwarf associated virus (MMDaV; Ma et al. 2015), apple geminivirus (Liang et al. 2015), grapevine geminivirus A (Al Rwahnih et al. 2017), camellia chlorotic dwarf associated virus (CaCDaV; Zhang et al. 2018), and passion fruit chlorotic mottle virus (PCMoV; Fontenele et al. 2018).

Most geminiviruses belong to the genus *Begomovirus*, the members of which are transmitted by whiteflies and have mono- or bipartite genomes. Viruses in the well-characterized genera *Curtovirus* and *Mastrevirus* are transmitted by leafhoppers and have monopartite genomes. The coat protein (CP, or V1/AV1 in mono- and bipartite geminiviruses, respectively) is the only structural protein, and it also functions as vector transmission mediator (Briddon et al. 1990; Poornima Priyadarshini et al. 2011). The V2/AV2 protein functions as a viral suppressor of RNA silencing, the same as the role of C4/AC4 in some viruses (Amin et al. 2011; Glick et al. 2008). Besides the V2/AV2 (New World begomoviruses lack V2/AV2), other additional open reading frames (ORFs) (V3 and V4) are found upstream of the CP gene of some geminiviruses. The replication-associated protein (Rep, or C1/AC1), which is a rolling-circle replication initiator

protein, initiates virion-sense DNA replication (Hanley-Bowdoin et al. 2013). The transcriptional-activator protein (TrAP, or C2/AC2) interferes with RNA silencing and other host defense responses (Raja et al. 2008; Trinks et al. 2005). The replication-enhancer protein (REn, or C3/AC3) is involved in viral replication (Settlage et al. 2005). Bipartite begomoviruses encode their movement proteins, the nuclear shuttle protein (NSP or BV1) and cell-to-cell movement protein (MP or BC1), on the DNA-B component (Sanderfoot and Lazarowitz 1995). Mastreviruses, grabloviruses, capulaviruses, and becurtoviruses express Rep from a spliced mRNA (C1:C2) and RepA (C1) from the complementary-sense strand (Wright et al. 1997), whereas begomoviruses, curtoviruses, turncurtoviruses, topocuviruses, and eragroviruses encode Rep in a single ORF and do not encode RepA.

In this study, we have identified two new, highly divergent geminiviruses by HTS that coinfect paper mulberry plants in China. Through field investigation of paper mulberry plants, we confirmed that the two geminiviruses are associated with leaf curl symptoms. Based on their host and symptoms, the two new geminiviruses are tentatively named paper mulberry leaf curl virus 1 and 2 (PMLCV-1 and PMLCV-2).

Materials and Methods

Plant materials. In May 2018, leaf curl symptoms on paper mulberry plants (goushu in Chinese, GS) were noted in three different locations of Chongqing province of China, from Huaye (HY) (29.82N, 106.42E), Southwest University (SWU) (29.80N, 106.40E), and Tongliang (TL) (29.84N, 106.05E). Leaf tissues were collected and were stored at –80℃. Three samples displaying either severe (GS-HY, GS-SWU) or mild (GS-TL) leaf curl (Fig. 1A, B, and D, respectively) were separately prepared for HTS analysis. Nonsymptomatic plants (Fig. 1E) surrounded by symptomatic plants were also collected at each location.

Construction of cDNA libraries, HTS, and sequence analyses. Total RNA extracts obtained from each sample using the EASYspin Plus Complex Plant RNA Kit (Aidlab, China) were used for RNA sequencing (RNA-seq), and total small RNAs (sRNAs) extracted from the GS-HY sample using the EASYspin Plant microRNA Extract kit (Aidlab) were used for sRNA sequencing (sRNA-seq). The TruSeq RNA Sample Prep kit version 2 (Illumina, U.S.A.) was used for the construction of RNA-seq cDNA libraries after depleting ribosomal RNAs. These libraries were sequenced with the Illumina HiSeq X-ten platform (paired-end reads of 150 bp) (Mega Genomics, China). The sRNA libraries were constructed using the TruSeq Small RNA Sample Prep kit (Illumina), and sequencing was carried out with the Illumina HiSeq2500 platform (1 × 50 bp reads) (Cao et al. 2019). The raw RNA-seq data are available at the Sequence Read Archive (SRA) database in NCBI (accession no. PRJNA608084). The raw sequences (reads) from RNA-seq and sRNA-seq libraries were analyzed with CLC Genomics Workbench 9.5 (Qiagen, U.S.A.) to remove sequence adaptors, low-quality reads, and especially in the case of RNA-seq data processing, RNA reads with more than 60% similarity to mulberry reference genomes (https://morus.swu.edu.cn/morusdb/). De novo assembly of the remaining clean reads into larger sequences (contigs) was carried out using the Trinity (RNA) or Velvet (sRNA) algorithms (Grabherr et al. 2011; Zerbino and Birney 2008). The resulting contigs were subjected to BLASTx and BLASTn searches (https://blast.ncbi.nlm.nih.gov/Blast.cgi) against NCBI viral (taxid: 10239) and viroidal (taxid: 2559587) databases for sequence annotation.

Conventional PCR and RCA. Total nucleic acids of leaf tissues (200 mg) from symptomatic or nonsymptomatic paper mulberry plants were extracted using a modified cetyltrimethylammonium bromide (CTAB)-based method (Li et al. 2008). The pellet was dissolved in 100 μl of nuclease-free water and was used for PCR amplification with Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific, U.S.A.) and specific back-to-back primers (Supplementary Table S1) designed from the viral contigs identified by BLAST analysis, with two pairs targeting each virus. Conventional PCR was performed using 1 μl of total DNA in a reaction mixture containing 1× Phusion HF Buffer, 0.2 mM of each dNTP, 0.2 mM of each primer (PV1-F1/R1, PV2-F1/R1, Supplementary Table S1), and 0.02 U of Phusion DNA Polymerase (2 U/μl). Thermocycling consisted of one cycle at 98℃ for 30 s followed by 35 cycles at 98℃ for 8 s, 55℃ for 25 s, and 72℃ for 90 or 120 s, and one final cycle at 72℃ for 10 min. All resulting amplicons were separated by 1% agarose gel electrophoresis, gel purified, and cloned into the pEASY-Blunt plasmid vector (TransGen Biotech, China). At least five independent clones of each fragment were sequenced on both strands by primer walking. RCA (Inoue-Nagata et al. 2004) was performed using an aliquot (1 μl) of DNA preparations and the Illustra TempliPhi 100 Amplification kit (GE Healthcare, U.S.A.) according to the manufacturer's instructions. RCA amplicons were excised with *Eco*RI, *Cla*I, *Sac*I, *Nde*I, *Kpn*I, or *Bgl*II restriction enzymes, predicted to cut PMLCV-1 and PMLCV-2 DNA at a single site, and separated by agarose gel electrophoresis.

Sequence analyses. Sequences of the amplicons were assembled and analyzed with the Lasergene sequence analysis package (DNAS-TAR), and the complete genome sequences of the viruses were submitted to the GenBank database. The reconstructed genomes were subjected to the following analyses: (i) ORF prediction using ORF Finder (https://www.ncbi.nlm.nih.gov/gorf/gorf.html); (ii) prediction of putative introns using FGENESH version 2.6 (Solovyev et al. 2006); (iii) search for conserved and functional domains in the predicted proteins using SMART (http://smart.embl.de/) and Conserved Domain Database (CDD) (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi/) (Marchler-Bauer et al. 2017); (iv) identification of transmembrane domains using TMHMM (http://www.cbs.dtu.dk/services/TMHMM/); (v) comparisons of the related viral sequences using ClustalW (Thompson et al. 1994), and construction of phylogenetic trees based on whole genome nucleotide (nt) and protein amino acid (aa) sequences in MEGA7 (Kumar et al. 2016) using maximum-likelihood algorithms with 1,000 bootstrap replications; and (vi) multiple sequence alignments performed using the MUSCLE alignment program (https://www.ebi.ac.uk/Tools/msa/muscle/), and the alignment files used to generate sequence identity matrices using Sequence Demarcation Tool (SDT) version 1.2 (Muhire et al. 2014).

Detection of the new geminiviruses in field plants. To evaluate the prevalence of the two new geminiviruses in field plants, leaf samples from 91 paper mulberry plants (81 symptomatic and 10 nonsymptomatic) were randomly collected from the same three locations where the initial samples were collected, adopting a sample plot with a diameter of approximately 1 km at each location. Total nucleic acids were extracted from leaf tissues using a CTAB-based method (Li et al. 2008). PCR was performed using the 2× Taq Master Mix kit (Quick Load) (Novoprotein, China) with specific primers that lie in the conserved genomic regions, two pairs for each virus (PV1-DF1/DR1 and PV1-DF2/DR2 for PMLCV-1, and PV2-DF1/DR1 and PV2-DF2/DR2 for PMLCV-2; Supplementary Table S1). PCR was carried out in a total volume of 10 μl with 1.0 μl of template DNA, 0.15 μl (10 μM) of each R/F primer pair, 5.0 μl of Master Mix reagent, and 3.7 μl of nuclease-free water. Thermalcycling conditions consisted of an initial denaturation at 95℃ for 5 min and 35 cycles at 95℃ for 30 s, 55℃ for 30 s, and 72℃ for 30 s, with a final extension step at 72℃ for 10 min.

Results

Discovery of two new geminiviruses in paper mulberry affected by leaf curl disease. Paper mulberry plants with virus-like leaf curl symptoms (V-shaped margin deformation, and puckering and curling of the leaves) were found in Chongqing province (Fig. 1A to D). To identify viral agents in these symptomatic plants, total RNA libraries were constructed and sequenced using Illumina platforms. Total RNA-seq clean reads of the libraries GS-HY (57,327,346 reads), GS-SWU (63,100,140 reads), and GS-TL (59,993,834 reads) were obtained after adapter trimming and filtering of low-quality reads. Additional host reads were removed by mapping the retained reads to the mulberry reference genome (reads with 60% or higher similarities were eliminated). This resulted in

9,251,344 (GS-HY), 10,176,156 (GS-SWU), and 11,874,493 (GS-TL) clean reads. Finally, de novo assembly of unmapped reads produced 73,205 (GS-HY), 74,495 (GS-SWU), and 105,492 (GS-TL) contigs for the symptomatic samples (Supplementary Table S2). The initial BLASTx and BLASTn searches revealed only two kinds of viral contigs, which were both related to the family *Geminiviridae* according to sequence identities shared between them (41.2 to 42% at nt level), with their closest relatives (BLASTx, 42 to 58%; BLASTn, 65 to 86%) being the unclassified geminiviruses CCDaV, MMDaV,

CaCDaV, and PCMoV. The viral contigs appear to be circular based on the sequencing data, and their sizes (3,056/3,057 and 3,557/3,557/3,560 bp) were similar to the full-length genomes of the related viruses.

Validation of complete viral genome sequences. Based on the contigs obtained from the three libraries and the circular topology of the geminivirus genome, back-to-back primers (Supplementary Table S1) were designed to amplify the circular DNA of each of the two putative new viruses. With these primers, amplicons of
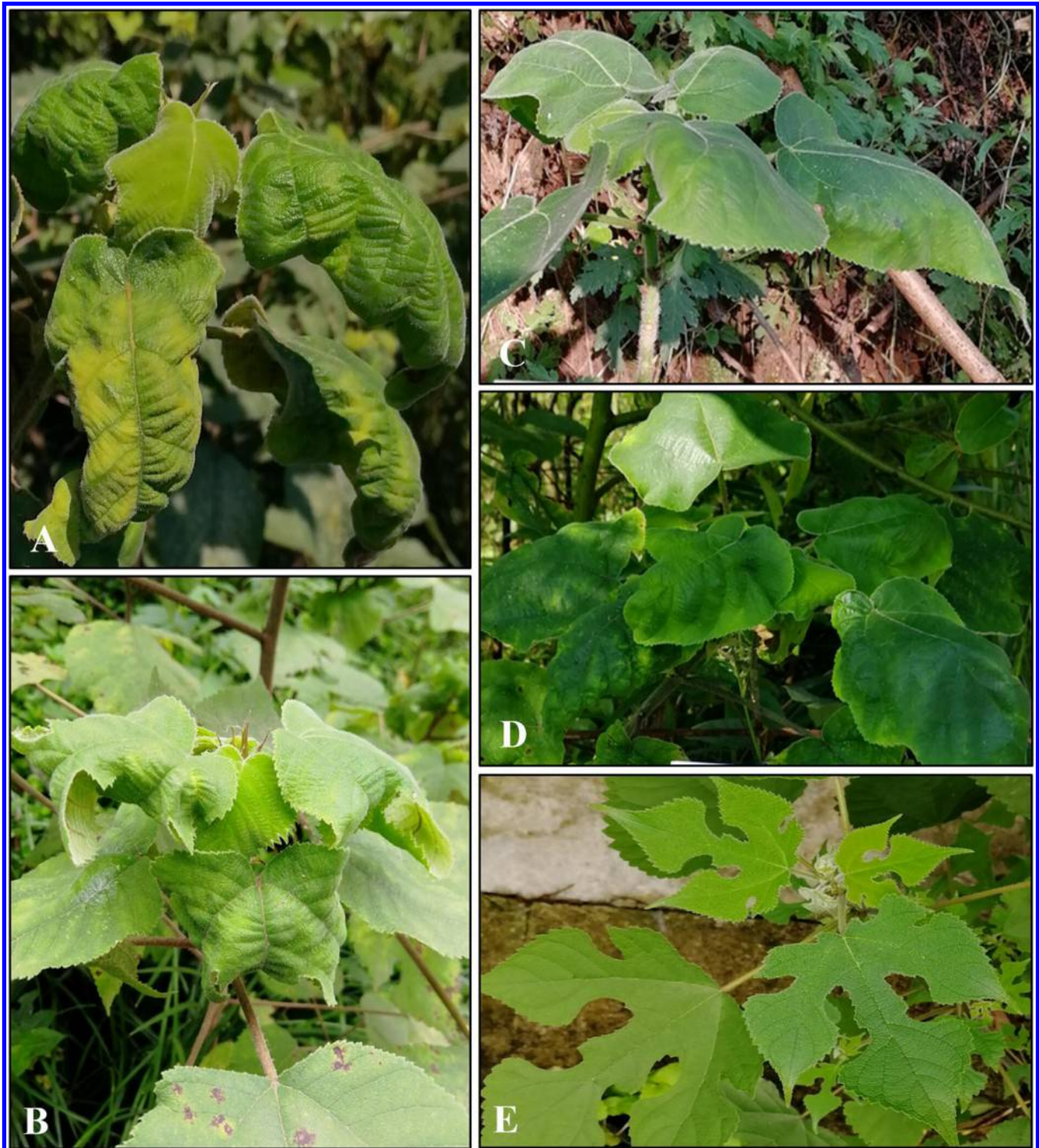

Fig. 1. Symptoms on leaves of paper mulberry. **A** and **B,** Chlorosis, leaf curl, and deformed margins in a plant infected by paper mulberry leaf curl virus 1 (PMLCV-1) and paper mulberry leaf curl virus 2 (PMLCV-2); **C,** mild leaf curl in a plant infected with PMLCV-1; **D,** mild leaf curl in a plant infected with PMLCV-2; and **E,** nonsymptomatic leaves. Leaves from the plants depicted in A, B, and D were used for high-throughput sequencing (libraries GS-HY, GS-SWU, and GS-TL, respectively).

approximately 3.0 kb, greater than 3.0 kb, and below 1.0 kb were obtained using total DNA from symptomatic paper mulberry plants, whereas no amplicons were obtained from the nonsymptomatic control sample (Fig. 2A and B). However, the small amplicons (<1.0 kb) were not sequenced. Cloning and sequencing of the amplicons (>1.0 kb) produced five isolates provisionally designated as PMLCV-1-[SWU] (3,056 nt, MN595125), PMLCV-1-[HY] (3,056 nt, MN595124), PMLCV-2-[HY] (3,756 nt, MN595126), PMLCV-2-[SWU] (3,763 nt, MN595127), and PMLCV-2-[TL] (3,757 nt, MN595128). PMLCV-1 isolates share 99.7% nt sequence identity with each other, whereas those of PMLCV-2 share 89.9 to 92.5% nt sequence identities with each other.

To reconfirm the circular nature of PMLCV-1 and PMLCV-2 in the symptomatic samples and confirm the monopartite nature of each virus, RCA followed by restriction fragment length polymorphism analyses was performed using total DNA from the HY, SWU, and TL samples and a nonsymptomatic paper mulberry plant. Two fragments (3,056 and 3,757 nt) were always obtained following digestion with nearly all the selected enzymes (except *Bgl*II and the GS-TL sample, for which only the 3,757-nt large fragment was obtained), which had the same electrophoretic mobility as the PCR products of the full-length PMLCV-1 and PMLCV-2 DNA molecules used as controls (Fig. 2C). Moreover, the fragments were confirmed as PMLCV-1 and PMLCV-2 by cloning and sequencing. Extra bands were not detected, and no fragments were observed in the negative controls (Fig. 2C). Altogether, the data derived from HTS, PCR, and RCA indicate that PMLCV-1 and PMLCV-2 are two monopartite, circular DNA viruses.

Sequence analysis of PMLCV-1 and PMLCV-2. The complete genome identities between PMLCV-1 and PMLCV-2 isolates are low (39.0 to 39.6%). The viral isolates PMLCV-1-[HY] and PMLCV-2-[TL] were used for all subsequent analyses. The related sequence information of different isolates is summarized in Supplementary Table S3. PMLCV-1 is more related to MMDaV (42.2% identity; Supplementary Table S4), whereas PMLCV-2 is more closely related to CCDaV (53.7% identity; Supplementary Table S5). These data confirm that PMLCV-1 and PMLCV-2 are two distinct viruses.

The two viruses have a similar genome arrangement, with four ORFs in the virion-sense strand (V1, V2, V3, and V4) and two in the complementary-sense strand (C1 and C1:C2) (Fig. 3A and C). Similar to other geminiviruses, the putative stem-loop structure formed by GC-rich inverted repeats flanking the highly conserved nonanucleotide motif TAATATT/AC is present in the large intergenic region (LIR; Fig. 3B and D). The TATA boxes and coupled termination signals (AAT/AAA) were identified in the large and short intergenic regions, respectively, in both strands (Fig. 3A and C). Interestingly, the mapping pattern of RNA reads showed a concentration of reads in the ORF regions, suggesting that the predicted ORFs are indeed transcribed (Fig. 3A and C).

The V1 ORFs overlap with V2 and potentially encode proteins of 256 aa (PMLCV-1) and 245 aa (PMLCV-2). The V1 proteins of both PMLCV-1 and PMLCV-2 were identified as a geminivirus CP (Pfam E-values of $3.5e^{-26}$ and $4.3e^{-28}$ for PMLCV-1 and PMLCV-2, respectively). Moreover, in the PMLCV-1 CP, a nuclear localization signal was identified at aa 48 to 59 (GRRSKRTRRTRY). Pairwise aa sequence comparisons indicated a significant difference in identity of V1 of PMLCV-1 (33.1 to 35.9%) and PMLCV-2 (34.2 to 59.5%) with the CPs of MMDaV, CCDaV, CaCDaV, and PCMoV (Supplementary Tables S4 and S5). A lower level of aa sequence identity was obtained (9.6 to 24.4%) when the V1s were compared with CPs of members of the other genera in the family *Geminiviridae* (Supplementary Fig. S1).

The V2 ORFs potentially encode proteins with 148 aa for PMLCV-1 and 106 aa for PMLCV-2. The putative proteins share conserved N-terminal and central regions with those of related viruses (Supplementary Fig. S2) but only a limited aa sequence identity (17.7 to 32.7%). These proteins are related to the V2 proteins of MMDaV and CaCDaV based on BLASTp analysis.

The deduced V3 protein of PMLCV-1 has 74 aa, and that of PMLCV-2 has 73 aa. This ORF partly overlaps ORF V2. Although no significant similarity was found for related V3 genes when searching against databases, a transmembrane domain that may play a role in geminivirus trafficking was identified as a common characteristic (Supplementary Fig. S2) (Ma et al. 2015).

The deduced V4 protein of PMLCV-2 (302 aa) is much larger than that of PMLCV-1 (100 aa). Although the PMLCV-2 V4 shares a moderate degree of sequence identity (50 to 54.7%) with related viruses (PCMoV, CCDaV, and CaCDaV), the PMLCV-1 V4 as well
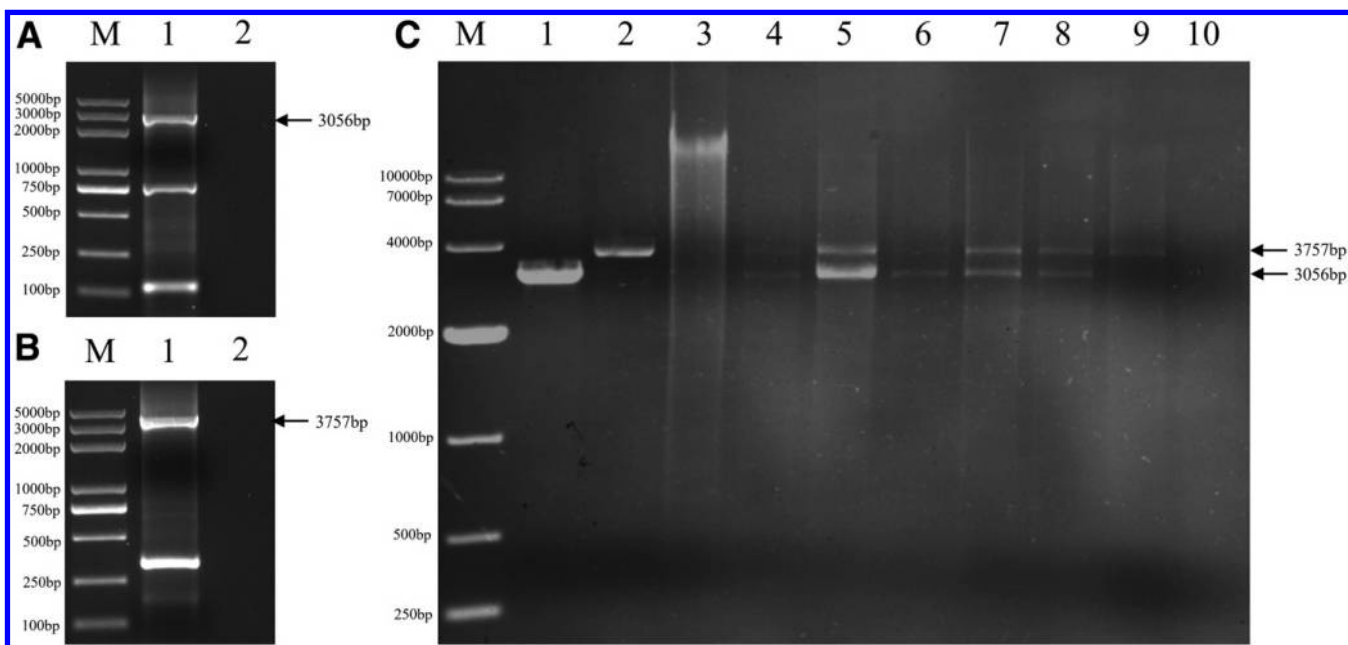

Fig. 2. Agarose gel electrophoresis of PCR products obtained using **A,** primers PV1-F1/R1 (paper mulberry leaf curl virus 1 [PMLCV-1]) and **B,** PV2-F1/R1 (paper mulberry leaf curl virus 2 [PMLCV-2]) from a symptomatic (lane 1) and a nonsymptomatic paper mulberry plant (lane 2); lane M is DL 5,000 DNA Marker (TaKaRa), with sizes of DNA fragments indicated on the left. **C,** Agarose gel electrophoresis of rolling-circle amplification products from nonsymptomatic (lane 10) and symptomatic paper mulberry plants digested with restriction enzymes *EcoR*I, *Cla*I, *Sac*I, *Nde*I, *Kpn*I, and *Bg1*II (lanes 4 to 9, respectively), or not digested (lane 3). Lanes 4 and 5 are from plant sample GS-HY, lanes 6 to 8 are from plant sample GS-SWU, and lane 9 is from plant sample GS-TL. Lane M is DL 10,000 DNA Marker (TaKaRa), with sizes of DNA fragments indicated on the left. PCR product of PMLCV-1 (lane 1) and PMLCV-2 (lane 2) from symptomatic paper mulberry plants.

as the MMDaV V4 and V5 proteins have no similarity with any protein in the databases. The secondary structures of PMLCV-2 V4 and those of related viruses (PCMoV, CCDaV, and CaCDaV) show a typical 30-kDa (30K) movement protein pattern (Ishikawa et al. 2013; Melcher 2000; Yu et al. 2013), with a series of unfolded α-helices and β-sheets (αA, β1, β2, αB, β3 to β7) from the N to the C terminus.

The two coding regions in the complementary-sense strand, C1 and C1:C2, potentially encode replication-associated proteins RepA and Rep, respectively. The aa sequence identities among PMLCV-1, PMLCV-2, and the four related viruses in C1 and C1:C2 are 34.8 to 48.9% and 43.7 to 55.8%, respectively (Supplementary Tables S4 and S5). In addition, the aa motifs reported in other geminivirus Reps were also found in the amino-terminal sequences of the Rep proteins of PMLCV-1/PMLCV-2: (i) motif I (FLTYP), required for specific dsDNA binding; (ii) rolling-circle replication (RCR) motif II (HLH/HVH), a metal-binding site that may be involved in protein conformation and DNA cleavage; (iii) RCR motif III (YIRKD/YLKKS), a catalytic site for DNA cleavage (Nash et al. 2011); and (iv) Walker A (GPSRTGKT/GPSRSGKT) and Walker B (LYN-VIDDI) domains, which are crucial components of the nt-binding site (Walker et al. 1982) (Supplementary Fig. S3).

An intron in the genomes of PMLCV-1 and PMLCV-2. Many geminiviruses in the genera *Mastrevirus*, *Becurtovirus*, *Grablovirus*, and *Capulavirus*, as well as several that have not yet been assigned to a genus, have an alternatively spliced transcript in the complementary strand (Al Rwahnih et al. 2013; Bernardo et al. 2013; Dekker et al. 1991; Fontenele et al. 2018; Ma et al. 2015; Morris et al. 1992; Mullineaux et al. 1990; Schalk et al. 1989; Varsani et al. 2014; Wright et al. 1997). An intron was found in the transcripts of both PMLCV-1 (131 nt) and PMLCV-2 (119 nt) (Fig. 4A and B). The intron contains the consensus 5′ and 3′ splice sites and branch points characteristic of U2 plant introns (Lewandowska et al. 2004) (Supplementary Fig. S4) and has an RNA U+A content (54.2% for PMLCV-1, 58.8% for PMLCV-2) close to the minimum needed for efficient splicing (Simpson and Brown 1993). In addition, a total

of 155 PMLCV-1 and 159 PMLCV-2 sRNAs spanning the exon–exon junctions, lacking the intron, were identified (Fig. 4A and B). Furthermore, nine RNA-seq reads for PMLCV-1 and 15 for PMLCV-2 map across the junctions, also indicating the presence of alternatively spliced introns (Fig. 4C and D). Together, these results provide strong evidence that unspliced and spliced transcripts capable of encoding RepA (C1) and Rep (C1:C2), respectively, are produced in vivo.

Phylogenetic analysis. Evolutionary relationships were inferred from nt sequences of the full-length viral genome and also from aa sequences of CP and Rep (Fig. 5; Supplementary Fig. S5). The CP and Rep cladograms showed a similar topology, with PMLCV-1, PMLCV-2, CCDaV, MMDaV, CaCDaV, and PCMoV clustering together. These unassigned geminiviruses formed a monophyletic cluster in the CP tree, whereas in the Rep tree, they were placed in a monophyletic cluster together with mastreviruses and capulaviruses (Supplementary Fig. S5). In the full genome tree, PMLCV-1 and PMLCV-2 were placed in two different monophyletic clusters, with PMLCV-1 sharing a closer evolutionary link to MMDaV, whereas PMLCV-2 is closer to CCDaV, CaCDaV, and PCMoV (Fig. 5). This pattern suggests two distinct evolutionary paths for these two groups of geminiviruses.

Field surveys of paper mulberry trees. Field surveys were carried out to determine the distribution of PMLCV-1 and PMLCV-2 in Chongqing province. A total of 91 paper mulberry trees, 81 of which displayed distinct leaf curl symptoms, were tested by PCR using specific primers, and the results were confirmed by Sanger sequencing. Although the nonsymptomatic paper mulberry samples were all negative, 79 out of the 81 symptomatic samples yielded fragments corresponding to the two viruses. The vast majority of the samples (70 out of 81) had both viruses, with only three single infections by PMLCV-1 and six single infections by PMLCV-2 (Supplementary Fig. S6). These results confirm the close association between the presence of PMLCV-1 and/or PMLCV-2 and paper mulberry leaf curl disease.

Interestingly, compared with the single-infected symptomatic plants, the plants with coinfection exhibited a visibly more severe
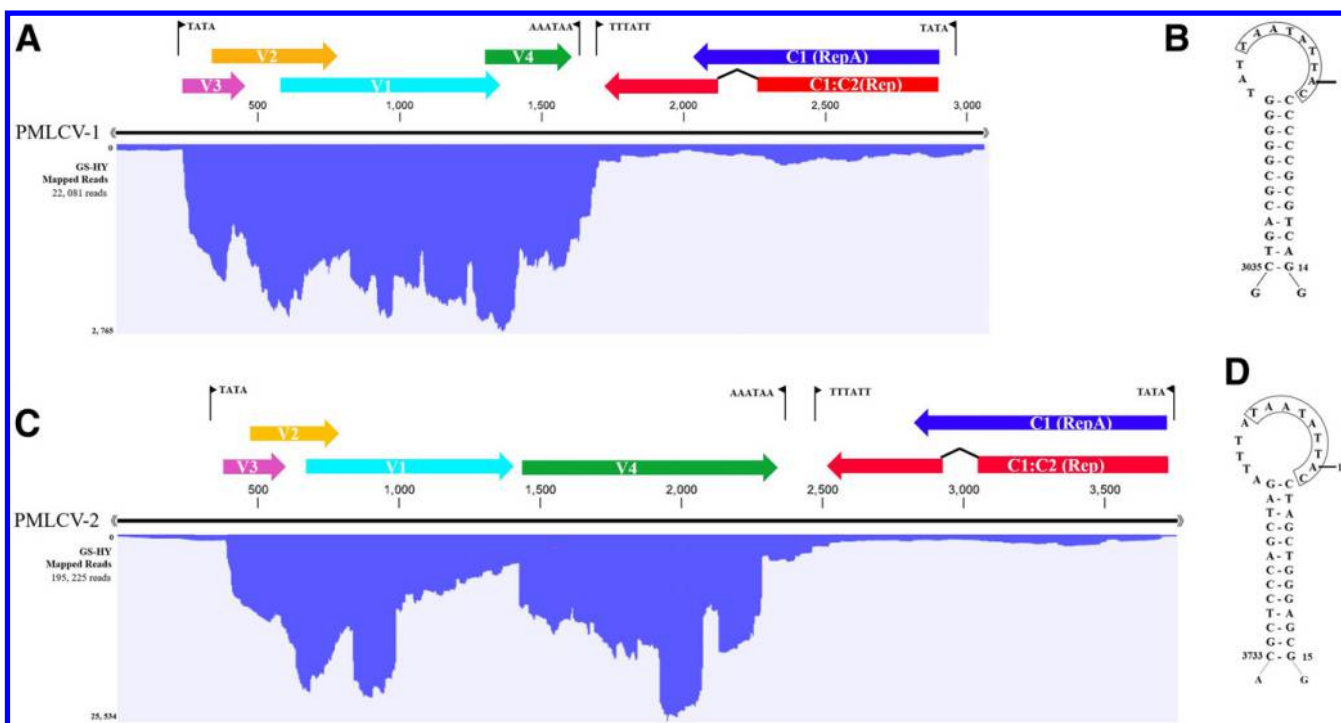

Fig. 3. Linear representations of the circular genomes of **A,** paper mulberry leaf curl virus 1 (PMLCV-1) and **C,** paper mulberry leaf curl virus 2 (PMLCV-2) with the virus-related RNA reads from the GS-HY library mapped on the viral genomes. Genome hallmarks, including TATA boxes and polyadenylation signals, are indicated. Coding sequences of the genes are denoted in the virion (V) or complementary (C)-sense strands, and the intron in C1:C2 is also indicated. Rep = replication-associated protein. **B** and **D,** Nonanucleotide sequence within the stem-loop in the large intergenic regions of PMLCV-1 and PMLCV-2, respectively. Number 1 indicates the first position in the viral genome, which, according to convention, is coincident with the origin of replication.

degree of leaf curl and vein necrosis (Fig. 1). To obtain a viral population-derived perspective on this observation, the RNA reads derived from each virus were counted in the three samples submitted to HTS (samples HY and SWU had a coinfection, whereas sample TL was infected only with PMLCV-2) (Supplementary Table S2). The proportion of PMLCV-2-derived reads was approximately 10-fold that of PMLCV-1 reads in the two coinfected samples (Supplementary Table S2). Interestingly, PMLCV-2-derived reads were much more abundant (approximately sixfold) in the coinfected samples compared with sample TL. Thus, it appears that coinfection of PMLCV-1 and PMLCV-2 enhances the viral titer of the latter. Because we did not sequence a PMLCV-1 single-infected sample, we do not know whether the titer of this virus also increases in mixed infection. Moreover, PMLCV-2 single-infected samples tended to display yellowing of the veins and a V-shaped pattern compared with the leaf curl phenotype of PMLCV-1 single-infected plants.

Discussion

Paper mulberry is widely cultivated in poverty-stricken and rocky-desert areas of China, with great scientific, ecologic, economic, and medicinal values. Its perennial nature and vegetative propagation make it prone to viral infection. However, only a few studies have been published to date on bacterial and fungal pathogens of paper mulberry (Liu et al. 2004; Wu et al. 2011; Yan et al. 2011). In this study we used HTS to identify viruses associated with leaf curl symptoms in paper mulberry plants. We identified geminivirus-like sequences from three independent symptomatic samples using BLAST searches based on sequence similarity. It must be noted that although the only viral contigs found were these geminivirus-like sequences, we should not exclude the existence of other highly divergent viruses and viroids escaping our similarity-dependent detection method.

Because DNA viruses can be integrated into the genome of the host (Bejarano et al. 1996), the existence of these viruses in episomal form was validated by other molecular approaches, including PCR with back-to-back primers and RCA. Cloning and sequencing of the obtained amplicons provided direct evidence of the episomal presence of the two viruses and of their circular nature.

In spite of their different genome sizes, PMLCV-1 and PMLCV-2 have a similar genomic organization, which is also similar to those of the unclassified geminiviruses CCDaV, MMDaV, CaCDaV, and PCMoV (Fontenele et al. 2018; Loconsole et al. 2012; Ma et al. 2015; Zhang et al. 2018). The viral genome encodes four ORFs in virion-sense (V1 to V4) and two ORFs in complementary-sense (C1 and C1:C2) strand, with the conserved nonanucleotide motif (TAATATT/AC) in the stem-loop structure located in the LIR. As well as to CCDaV, MMDaV, CaCDaV, and PCMoV, a significant level of identity at the aa level of V1 was found with the coat proteins of begomoviruses (Supplementary Fig. S1). The overlapping ORF V2s of PMLCV-1 and PMLCV-2 are homologous to the V2 of the unclassified geminiviruses MMDaV and CaCDaV as well as begomoviruses. The V2 protein of begomoviruses and MMDaV was shown to be an RNA silencing suppressor (Yang et al. 2018; Zhang et al. 2012). Despite extensive sequence differences, the V3 proteins identified in this study contain a transmembrane domain that is also present in MPs of becurto-, curto-, and mastreviruses (Al Rwahnih et al. 2013; Dekker et al. 1991; Morris et al. 1992; Mullineaux et al. 1990; Schalk et al. 1989; Wright et al. 1997). No match in the databases was found for the proteins coding for the PMLCV-1 ORF V4, whereas the V4 of PMLCV-2 has the characteristics of the 30K movement protein superfamily and shows some similarity with the BC1 (MP) protein of begomoviruses. Splicing signals were identified in the transcripts of PMLCV-1 and PMLCV-2, which are identical to the U2 plant introns (Lewandowska et al. 2004) (Supplementary Fig. S3). The splicing signal of PMLCV-1 is slightly different from that of PMLCV-2, suggesting that the expression strategy of Rep protein of the two geminiviruses is not always same. The conserved motifs of the Rep were found in the complementary-sense ORFs of PMLCV-1 and PMLCV-2, respectively (Supplementary Fig. S2), which resemble that of mastreviruses.

Recombination has played a significant role in the evolution of geminiviruses, with both nonhomologous and homologous recombination events contributing to the loss or gain of genomic regions
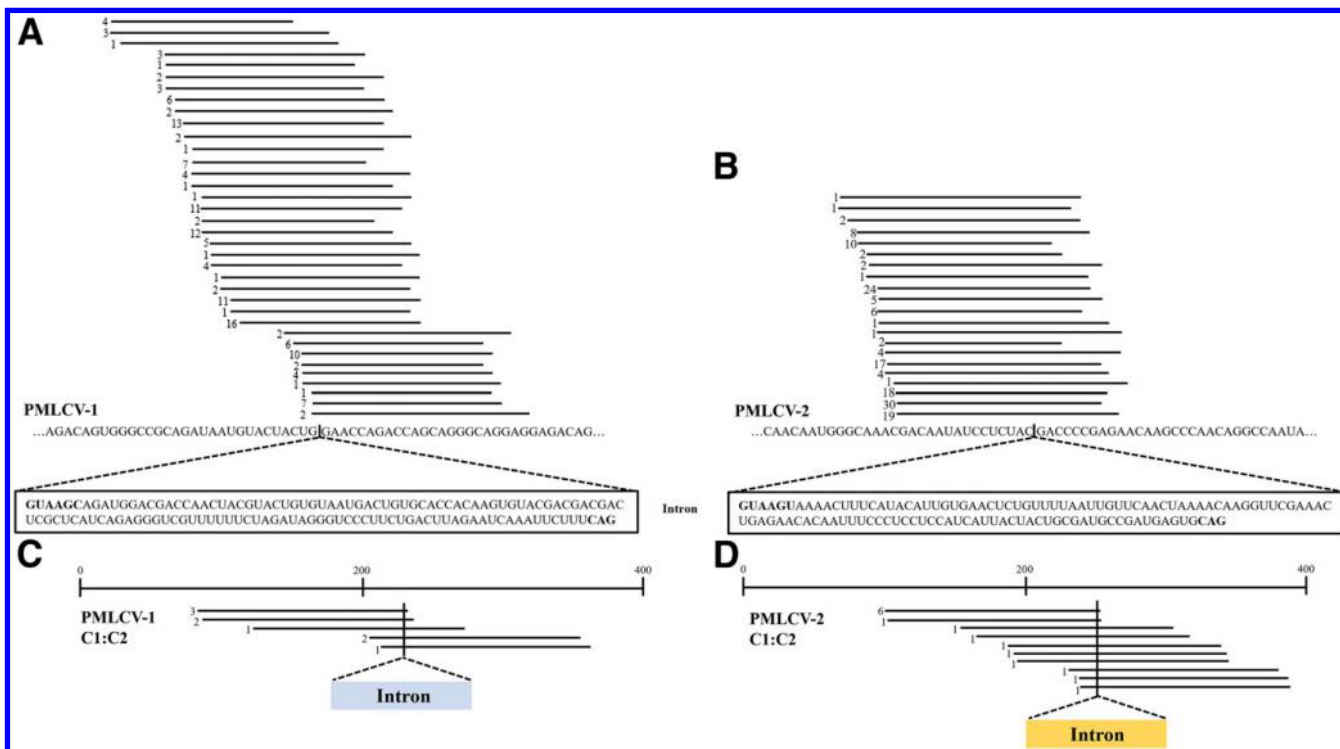

Fig. 4. Sequences of the intron (in the rectangle) and flanking exons identified in **A,** paper mulberry leaf curl virus 1 (PMLCV-1) and **B,** paper mulberry leaf curl virus 2 (PMLCV-2) mRNAs, with the number of sRNA reads (spanning the exon–exon junctions) from the GS-HY sRNA library indicated on the left. The 5′ and 3′ splicing signals characteristic of U2 plant introns are indicated in bold. The C1:C2 alternate splicing transcripts of **C,** PMLCV-1 and **D,** PMLCV-2 were also found in the GS-HY RNA-seq library.

(Briddon et al. 1996; Stanley et al. 1986). Intriguingly, likely due to recombination with different DNA sequences, the genomic regions between V1 and C2 of PMLCV-1, PMLCV-2, and the unclassified viruses (CCDaV, MMDaV, CaCDaV, and PCMoV) are highly variable, which has led to the formation of three types of V4: (i) separate from V1 and unrelated to any known viral protein (MMDaV); (ii) separate from V1 but related to begomovirus BC1 (MP) (PMLCV-2 as an example); and (iii) overlapped with V1 and classified as an unknown protein (PMLCV-1).

Fontenele et al. (2018) proposed that according to genome size, sequence similarity, and phylogenetic relationships, the 3.5-kb-type geminiviruses are evolutionary intermediates of the genetic transition between monopartite and bipartite begomoviruses. A monopartite geminivirus (~2.7-kb genome) would have gained the V4 gene, homologous to the DNA-B-encoded MP gene, by recombination, leading to the ~3.5-kb genome. The next step would have been the acquisition of the NSP gene, leading to a ~5.2-kb genome, and followed by the split that generated the two components (~2.6 kb each). For bipartite begomoviruses, the DNA-B component is not always essential for infectivity (Rochester et al. 1990; Saunders et al.

2002). Thus, it is not unreasonable to propose that the partial loss of V4 would have produced the 3.0-kb-type viruses. Due to the low identities between the V4 of PMLCV-1 and the 3.5-kb-type viruses, a recombination event involving the 3.5-kb-type viruses and an unknown viral or host gene to replace the V4, which produced the 3.0-kb-type viruses, may also have occurred. Because the C-terminal sequence variability of V1 and C2 proteins of PMLCV-1, PMLCV-2, and the unclassified viruses was found to be low (Supplementary Fig. S2), it is reasonable to propose that the recombination region is located between these two genes.

Based on their genomic features and phylogenetic analyses, we propose to classify paper mulberry leaf curl virus 1 and paper mulberry leaf curl virus 2 as two new species in the family *Geminiviridae*. Together with the other four geminiviruses that are currently not assigned to genera, they should compose a new genus of monopartite geminiviruses with >3.0-kb genomes. Furthermore, based on sequence comparisons and phylogenetic analysis, the new genus could be subdivided into two subgenera, one including PMLCV-1 and MMDaV and the other including PMLCV-2, CCDaV, CaCDaV, and PCMoV. However, it should be noted that geminate particles
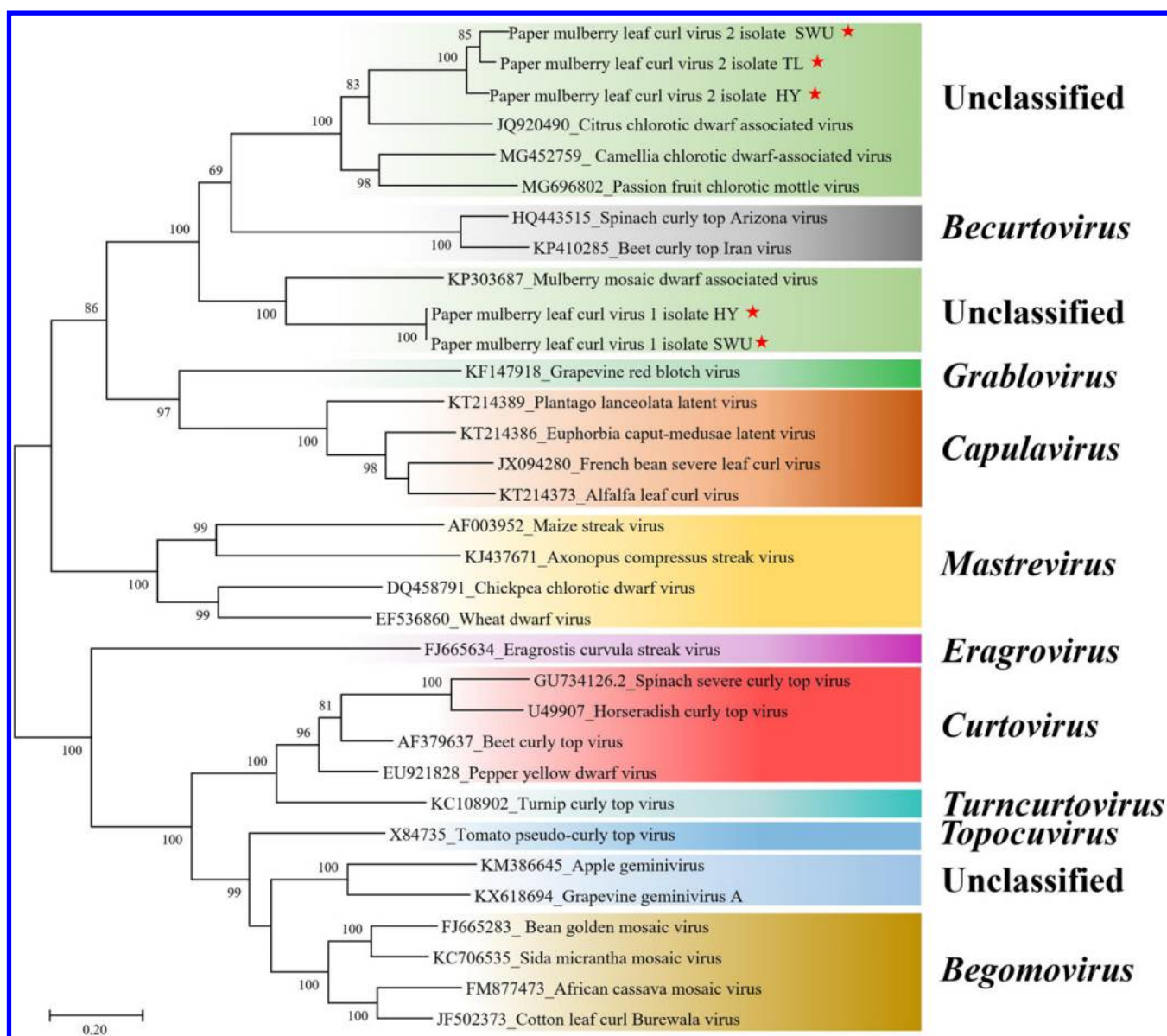

Fig. 5. Phylogenetic analysis of paper mulberry leaf curl virus 1 (PMLCV-1), paper mulberry leaf curl virus 2 (PMLCV-2), and representative members of the nine genera and unclassified viruses in the family *Geminiviridae*, based on full-genome nucleotide sequences using the maximum-likelihood method. GenBank accession numbers are placed before each virus name. Bar = 0.2 substitutions per nucleotide position. Bootstrap percentages (higher than 50%), determined from 1,000 bootstrap replicates, are shown at the branches.

have not been identified for any of these six viruses, and their vectors have not been unequivocally identified. Thus, classification at the genus/subgenus levels may be subject to changes in the future, once those properties are determined for at least a few of these viruses.

Accumulating evidence suggests that viral coinfections are the rule rather than the exception in nature (Moreno and López-Moya 2020). Coinfection is common among geminiviruses (Rentería-Canett et al. 2011; Roye et al. 1999). Field surveys of the new viruses showed that 86% of the tested symptomatic samples had a coinfection. Thus, an association between the presence of both PMLCV-1 and PMLCV-2 with paper mulberry leaf curl disease was established. However, whether one of the viruses can cause leaf curl in single infection remains to be determined. Although a small number of single-infected plants showed only mild symptoms, this observation must be extended to a larger sample size. Analysis of one PMLCV-2 single-infected plant and two coinfected samples by measuring the proportion of RNA reads indicated a higher accumulation of PMLCV-2 in the plants with coinfections. This suggests a synergistic effect when the two viruses are present in paper mulberry, which has also been observed on cassava and pepper infected with different geminiviruses (Fondong et al. 2000; Rentería-Canett et al. 2011).

Taking into account the wide geographical distribution in tropical and subtropical regions and the fast growth of paper mulberry, the two viruses described here could spread, and the geographic distribution of the viruses needs to be investigated. In summary, we have detected and characterized two new candidate geminiviruses associated with a leaf curl disease of paper mulberry. These findings are relevant for the management of these viruses and for the selection of resistant paper mulberry cultivars.

Acknowledgments

Literature Cited

Al Rwahnih, M., Alabi, O. J., Westrick, N. M., Golino, D., and Rowhani, A. 2017. Description of a novel monopartite geminivirus and its defective subviral genome in grapevine. Phytopathology 107:240-251.

Al Rwahnih, M., Dave, A., Anderson, M. M., Rowhani, A., Uyemoto, J. K., and Sudarshana, M. R. 2013. Association of a DNA virus with grapevines affected by red blotch disease in California. Phytopathology 103:1069-1076.

Amin, I., Hussain, K., Akbergenov, R., Yadav, J. S., Qazi, J., Mansoor, S., and Briddon, R. W. 2011. Suppressors of RNA silencing encoded by the components of the cotton leaf curl begomovirus-betasatellite complex. Mol. Plant-Microbe Interact. 24:973-983.

Bejarano, E. R., Khashoggi, A., Witty, M., and Lichtenstein, C. 1996. Integration of multiple repeats of geminiviral DNA into the nuclear genome of tobacco during evolution. Proc. Natl. Acad. Sci. U.S.A. 93:759-764.

Bell, L. A. 1988. Papyrus, Tapa, Amate and Rice Paper: Papermaking in Africa, the Pacific, Latin America and Southeast Asia. Liliaceae Press, McMinnville, OR.

Bernardo, P., Golden, M., Akram, M., Naimuddin, Nadarajan, N., Fernandez, E., Granier, M., Rebelo, A. G., Peterschmitt, M., Martin D. P., and Roumagnac, P. 2013. Identification and characterisation of a highly divergent geminivirus: Evolutionary and taxonomic implications. Virus Res. 177:35-45.

Briddon, R. W., Bedford, I. D., Tsai, J. H., and Markham, P. G. 1996. Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, tomato pseudo-curly top virus, suggests a recombinant origin. Virology 219:387-394.

Briddon, R. W., Pinner, M. S., Stanley, J., and Markham, P. G. 1990. Geminivirus coat protein gene replacement alters insect specificity. Virology 177:85-94.

Cao, M., Zhang, S., Li, M., Liu, Y., Dong, P., Li, S., Mi, K., Li, R. H., and Zhou, Y. 2019. Discovery of four novel viruses associated with flower yellowing disease of green Sichuan pepper (Zanthoxylum armatum) by virome analysis. Viruses 11:696.

Chang, C. S., Liu, H. L., Moncada, X., Seelenfreund, A., Seelenfreund, D., and Chung, K. F. 2015. A holistic picture of Austronesian migrations revealed by phylogeography of Pacific paper mulberry. Proc. Natl. Acad. Sci. U.S.A. 112:13537-13542.

Dekker, E. L., Woolston, C. J., Xue, Y. B., Cox, B., and Mullineaux, P. M. 1991. Transcript mapping reveals different expression strategies for the bicistronic RNAs of the geminivirus wheat dwarf virus. Nucleic Acids Res. 19:4075-4081.

Fondong, V. N., Pita, J. S., Rey, M. E. C., De Kochko, A., Beachy, R. N., and Fauquet, C. M. 2000. Evidence of synergism between African cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. J. Gen. Virol. 81:287-297.

Fontenele, R. S., Abreu, R. A., Lamas, N. S., Alves-Freitas, D. M. T., Vidal, A. H., Poppiel, R. R., Melo, F. L., Lacorte, C., Martin, D. P., Campos, M. A., Varsani, A., and Ribeiro, S. G. 2018. Passion fruit chlorotic mottle virus: Molecular characterization of a new divergent geminivirus in Brazil. Viruses 10:169.

Glick, E., Zrachya, A., Levy, Y., Mett, A., Gidoni, D., Belausov, E., Citovsky, V., and Gafni, Y. 2008. Interaction with host SGS3 is required for suppression of RNA silencing by tomato yellow leaf curl virus V2 protein. Proc. Natl. Acad. Sci. U.S.A. 105:157-161.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q. D., Chen, Z. H., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29:644-652.

Haible, D., Kober, S., and Jeske, H. 2006. Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. J. Virol. Methods 135:9-16.

Hanley-Bowdoin, L., Bejarano, E. R., Robertson, D., and Mansoor, S. 2013. Geminiviruses: Masters at redirecting and reprogramming plant processes. Nat. Rev. Microbiol. 11:777-788.

Inoue-Nagata, A. K., Albuquerque, L. C., Rocha, W. B., and Nagata, T. 2004. A simple method for cloning the complete begomovirus genome using the bacteriophage φ29 DNA polymerase. J. Virol. Methods 116:209-211.

Ishikawa, K., Maejima, K., Komatsu, K., Netsu, O., Keima, T., Shiraishi, T., Okano, Y., Hashimoto, M., Yamaji, Y., and Namba, S. 2013. Fig mosaic emaravirus p4 protein is involved in cell-to-cell movement. J. Gen. Virol. 94: 682-686.

Kumar, S., Stecher, G., and Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33: 1870-1874.

Lewandowska, D., Simpson, C. G., Clark, G. P., Jennings, N. S., Barciszewska-Pacak, M., Lin, C. F., Makalowski, W., Brown, J. W., and Jarmolowski, A. 2004. Determinants of plant U12-dependent intron splicing efficiency. Plant Cell 16:1340-1352.

Li, R., Mock, R., Huang, Q., Abad, J., Hartung, J., and Kinard, G. 2008. A reliable and inexpensive method of nucleic acid extraction for the PCR-based detection of diverse plant pathogens. J. Virol. Methods 154:48-55.

Li, Z., and Jia, M. 1997. Pharmacognostic identification of the root bark of Broussonetia papyrifera, an adulterant of cortex mori. J. Chin. Med. Mater. 20:66-68.

Liang, P., Navarro, B., Zhang, Z., Wang, H., Lu, M., Xiao, H., Wu, Q., Zhou, X., Di Serio, F., and Li, S. 2015. Identification and characterization of a novel geminivirus with monopartite genome infecting apple trees. J. Gen. Virol. 96: 2411-2420.

Liu, Q. Z., Wu, T. Q., Davis, R. E., and Zhao, Y. 2004. First report of witches' broom disease of Broussonetia papyrifera and its association with a phytoplasma of elm yellows group (16SrV). Plant Dis. 91:227.

Loconsole, G., Saldarelli, P., Doddapaneni, H., Savino, V., Martelli, G. P., and Saponari, M. 2012. Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family Geminiviridae. Virology 432:162-172.

Ma, Y., Navarro, B., Zhang, Z., Lu, M., Zhou, X., Chi, S., Di Serio, F., and Li, S. 2015. Identification and molecular characterization of a novel monopartite geminivirus associated with mulberry mosaic dwarf disease. J. Gen. Virol. 96:2421-2434.

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., Geer, L. Y., and Bryant, S. H. 2017. CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. Nucleic Acids Res. 45:D200-D203.

Melcher, U. 2000. The '30K' superfamily of viral movement proteins. J. Gen. Virol. 81:257-266.

Moreno, A. B., and López-Moya, J. J. 2020. When viruses play team sports: Mixed infections in plants. Phytopathology 110:29-48.

Morgan, E. C., and Overholt, W. A. 2004. Wildland Weeds: Paper Mulberry, Broussonetia papyrifera. IFAS Extension, no. 2. University of Florida, Gainesville, FL.

Morris, B. A., Richardson, K. A., Haley, A., Zhan, X., and Thomas, J. E. 1992. The nucleotide sequence of the infectious cloned DNA component of tobacco yellow dwarf virus reveals features of geminiviruses infecting monocotyledonous plants. Virology 187:633-642.

Muhire, B., Varsani, A., and Martin, D. P. 2014. SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One 9: e108277.

Mullineaux, P. M., Guerineau, F., and Accotto, G. P. 1990. Processing of complementary sense RNAs of Digitaria streak virus in its host and in transgenic tobacco. Nucleic Acids Res. 18:7259-7265.

Nash, T. E., Dallas, M. B., Reyes, M. I., Buhrman, G. K., Ascencio-Ibañez, J. T., and Hanley-Bowdoin, L. 2011. Functional analysis of a novel motif conserved across geminivirus Rep proteins. J. Virol. 85:1182-1192.

Peng, X., Liu, H., Chen, P., Tang, F., Hu, Y., Wang, F., Pi, Z., Zhao, M., Chen, N., Chen, H., Zhang, X., Yan, X., Liu, M., Fu, X., Zhao, G., Yao, P., Wang, L., Dai,

H., Li, X., Xiong, W., Xu, W., Zheng, H., Yu, H., and Shen, S. 2019. A chromosome-scale genome assembly of paper mulberry (*Broussonetia papyrifera*) provides new insights into its forage and papermaking usage. Mol. Plant 12:661-677.

Peng, X., Teng, L., Yan, X., Zhao, M., and Shen, S. 2015. The cold responsive mechanism of the paper mulberry: Decreased photosynthesis capacity and increased starch accumulation. BMC Genet. 16:898.

Poornima Priyadarshini, C. G., Ambika, M. V., Tippeswamy, R., and Savithri, H. S. 2011. Functional characterization of coat protein and V2 involved in cell to cell movement of cotton leaf curl Kokhran virus-Dabawali. PLoS One 6:e26929.

Raja, P., Sanville, B. C., Buchmann, R. C., and Bisaro, D. M. 2008. Viral genome methylation as an epigenetic defense against geminiviruses. J. Virol. 82:8997-9007.

Rentería-Canett, I., Xoconostle-Cázares, B., Ruiz-Medrano, R., and Rivera-Bustamante, R. F. 2011. Geminivirus mixed infection on pepper plants: Synergistic interaction between PHYVV and PepGMV. Virol. J. 8:104.

Rochester, D. E., Kositratana, W., and Beachy, R. N. 1990. Systemic movement and symptom production following agroinoculation with a single DNA of tomato yellow leaf curl geminivirus (Thailand). Virology 178:520-526.

Roossinck, M. J., Martin, D. P., and Roumagnac, P. 2015. Plant virus metagenomics: Advances in virus discovery. Phytopathology 105:716-727.

Roye, M. E., Wernecke, M. E., McLaughlin, W. A., Nakhla, M. K., and Maxwell, D. P. 1999. Tomato dwarf leaf curl virus, a new bipartite geminivirus associated with tomatoes and peppers in Jamaica and mixed infection with tomato yellow leaf curl virus. Plant Pathol. 48:370-378.

Sanderfoot, A. A., and Lazarowitz, S. G. 1995. Cooperation in viral movement: The geminivirus BL1 movement protein interacts with BR1 and redirects it from the nucleus to the cell periphery. Plant Cell 7:1185-1194.

Saunders, K., Salim, N., Mali, V. R., Malathi, V. G., Briddon, R., Markham, P. G., and Stanley, J. 2002. Characterisation of Sri Lankan cassava mosaic virus and Indian cassava mosaic virus: Evidence for acquisition of a DNA B component by a monopartite begomovirus. Virology 293:63-74.

Schalk, H. J., Matzeit, V., Schiller, B., Schell, J., and Gronenborn, B. 1989. Wheat dwarf virus, a geminivirus of graminaceous plants needs splicing for replication. EMBO J. 8:359-364.

Settlage, S. B., See, R. G., and Hanley-Bowdoin, L. 2005. Geminivirus C3 protein: Replication enhancement and protein interactions. J. Virol. 79:9885-9895.

Shen, S., and Peng, X. 2017. The black tech in "Paper Mulberry Poverty Alleviation" fills the lack of crude protein feed material of Chinese animal husbandry. Sci. Technol. Dev. 13:435-442.

Simpson, C. G., and Brown, J. W. 1993. Efficient splicing of an AU-rich antisense intron sequence. Plant Mol. Biol. 21:205-211.

Solovyev, V., Kosarev, P., Seledsov, I., and Vorobyev, D. 2006. Automatic annotation of eukaryotic genes, pseudogenes and promoters. Genome Biol. 7:S10.

Stanley, J., Markham, P. G., Callis, R. J., and Pinner, M. S. 1986. The nucleotide sequence of an infectious clone of the geminivirus beet curly top virus. EMBO J. 5:1761-1767.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. 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.

Trinks, D., Rajeswaran, R., Shivaprasad, P. V., Akbergenov, R., Oakeley, E. J., Veluthambi, K., and Pooggin, M. M. 2005. Suppression of RNA silencing by a geminivirus nuclear protein, AC2, correlates with transactivation of host genes. J. Virol. 79:2517-2527.

Varsani, A., Martin, D. P., Navas-Castillo, J., Moriones, E., Hernández-Zepeda, C., Idris, A., Murilo Zerbini, F. M., and Brown, J. K. 2014. Revisiting the classification of curtoviruses based on genome-wide pairwise identity. Arch. Virol. 159:1873-1882.

Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. 1:945-951.

Wright, E. A., Heckel, T., Groenendijk, J., Davies, J. W., and Boulton, M. I. 1997. Splicing features in maize streak virus virion- and complementary-sense gene expression. Plant J. 12:1285-1297.

Wu, P. S., Chen, K., Du, H. Z., Yan, J., and Zhang, Q. E. 2011. First report of *Cochliobolus sativus* causing leaf spot on paper mulberry in China. Plant Dis. 95:1586.

Wu, Q., Luo, Y., Lu, R., Lau, N., Lai, E. C., Li, W. X., and Ding, S. W. 2010. Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. Proc. Natl. Acad. Sci. U.S.A. 107:1606-1611.

Yan, J., Wu, P. S., Du, H. Z., and Zhang, Q. E. 2011. First report of black spot caused by *Colletotrichum gloeosporioides* on paper mulberry in China. Plant Dis. 95:880.

Yang, X., Ren, Y., Sun, S., Wang, D., Zhang, F., Li, D., Li, S., and Zhou, X. 2018. Identification of the potential virulence factors and RNA silencing suppressors of mulberry mosaic dwarf-associated geminivirus. Viruses 10:472.

Yu, C., Karlin, D. G., Lu, Y., Wright, K., Chen, J., and MacFarlane, S. 2013. Experimental and bioinformatic evidence that raspberry leaf blotch emaravirus P4 is a movement protein of the 30K superfamily. J. Gen. Virol. 94:2117-2128.

Zerbini, F. M., Briddon, R. W., Idris, A., Martin, D. P., Moriones, E., Navas-Castillo, J., Rivera-Bustamante, R., Roumagnac, R., and Varsani, A., and ICTV Report Consortium. 2017. ICTV virus taxonomy profile: *Geminiviridae*. J. Gen. Virol. 98:131-133.

Zerbino, D. R., and Birney, E. 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18:821-829.

Zhang, J., Dong, J., Xu, Y., and Wu, J. 2012. V2 protein encoded by tomato yellow leaf curl China virus is an RNA silencing suppressor. Virus Res. 163:51-58.

Zhang, S., Shen, P., Li, M., Tian, X., Zhou, C., and Cao, M. 2018. Discovery of a novel geminivirus associated with camellia chlorotic dwarf disease. Arch. Virol. 163:1709-1712.