Virology

*e-*X*tra**

Natural Defect of a Plant Rhabdovirus Glycoprotein Gene: A Case Study of Virus–Plant Coevolution

Song Zhang,[1,2] Aijun Huang,[3] Xin Zhou,[4] Zhenghe Li,[4] Ralf G. Dietzgen,[5] Changyong Zhou,[1,2,†] 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, Beibei, Chongqing 400715, China
[3] National Navel Orange Research Center, College of Life Science, Gannan Normal University, Ganzhou, China
[4] State Key Laboratory of Rice Biology, Institute of Biotechnology, Zhejiang University, Hangzhou, Zhejiang, China
[5] Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St. Lucia, Queensland 4072, Australia
Accepted for publication 8 July 2020.

ABSTRACT

Seven isolates of a putative cytorhabdovirus (family *Rhabdoviridae*, order *Mononegavirales*) designated as citrus-associated rhabdovirus (CiaRV) were identified in citrus, passion fruit, and paper bush from the same geographical area in China. CiaRV, bean-associated cytorhabdovirus (Brazil), and papaya virus E (Ecuador) should be taxonomically classified in the species *Papaya cytorhabdovirus*. Due to natural mutations, the glycoprotein (G) and P4 genes were impaired in citrus-infecting isolates of CiaRV, resulting in an atypical rhabdovirus genome organization of 3′ leader-N-P-P3-M-L-5′ trailer. The P3 protein of CiaRV shared a common origin with begomoviral movement proteins (family *Geminiviridae*). Secondary structure analysis and *trans*-complementation of movement-deficient tomato mosaic virus and potato virus X mutants

by CiaRV P3 supported its function in viral cell-to-cell trafficking. The wide geographical dispersal of CiaRV and related viruses suggests an efficient transmission mechanism, as well as an underlying risk to global agriculture. Both the natural phenomenon and experimental analyses demonstrated presence of the "degraded" type of CiaRV in citrus, in parallel to "undegraded" types in other host plant species. This case study shows a plant virus losing the function of an important but nonessential gene, likely due to host shift and adaption, which deepened our understanding of course of natural viral diversification.

Keywords: citrus, evolution, plant rhabdovirus, gene defect, genomics, movement protein complementation, viral quasi-species, virology

The family *Rhabdoviridae* in the order *Mononegavirales* comprises ecologically important negative-strand (ns) RNA viruses that infect a wide range of hosts among vertebrates, invertebrates, and plants (Dietzgen et al. 2012; Walker et al. 2018). With an enveloped virion of approximately 100 to 430 nm in length and 45 to 100 nm in diameter, most rhabdoviruses share a rod-like (rhabdo-) ultrastructure (bullet, cone-shaped, or bacilliform), providing a morphological basis for differentiation from other families (Wagner, 1987). Many rhabdoviruses are vectored by arthropods, in which they also replicate, leading to the "arthropod origin hypothesis" (Kuzmin et al. 2009), while others are mechanically transmissible (a mammal bite or from brushing against a plant) or waterborne (Armstrong et al. 1993; Dorson et al. 1984; Jackson et al. 2009; Kuzmin and Tordo, 2012). The predominant genome type is unsegmented and encodes five prototypic structural proteins with a typical gene arrangement of 5′-N-P-M-G-L-3′, where N is the nucleoprotein gene, P

phosphoprotein gene, M matrix protein gene, G glycoprotein gene and L large RNA-dependent RNA polymerase gene. These contiguous genes are separated by conserved intergenic sequences, which are involved in regulation of transcription by capping and polyadenylation of mRNAs (Walker et al. 2015). In the assembled virion, a tight association between genomic RNA and N protein constitutes the helical ribonucleoprotein complex (RNP) bonded with L, P, and M, and enveloped by a lipid membrane with G spikes (Assenberg et al. 2010). In addition, the structural proteins are multifunctional during virus infection, replication, and spread (Assenberg et al. 2010). Furthermore, a variable number of accessory genes are interposed between adjacent genes, may overlap with or are contained within the core genes, and play various roles in the viral life cycle (Walker et al. 2011).

Plant rhabdoviruses with unsegmented genomes have been taxonomically assigned to the genera *Cytorhabdovirus* and *Nucleorhabdovirus*, and those with bisegmented genomes to the genera *Dichorhavirus* and *Varicosavirus* (Walker et al. 2018). These viruses are persistently transmitted by hemipteran insects (aphids, leafhoppers, and planthoppers), *Brevipalpus* mites, and chytrid fungi, respectively, and this dual-host characteristic (plant and vector) has been demonstrated for several of them (Dietzgen et al. 2014; Jackson et al. 2005; Sasaya et al. 2002). Arthropods and plants are the more likely primordial hosts, due to a possible single origin of the rhabdoviruses (Longdon et al. 2015). Being similar in genome architecture, cytorhabdoviruses and nucleorhabdoviruses are distinguished based on viral replication in the cytoplasm or in the nucleus, respectively of infected cells (Walker et al. 2018). The genomes of plant rhabdoviruses encode the five core structural proteins, but also contain an additional gene for viral cell-to-cell movement that would have been acquired during the adaptation of rhabdoviruses to their plant hosts (Jackson et al. 2009; Kormelink et al. 2011; Lucas 2006; Taliansky et al. 2008). This evolutionary

The author(s) declare no conflict of interest.

strategy, likely driven by horizontal gene transfer (Koonin et al. 2015; Roossinck 2008), endows plant rhabdoviruses with the ability to infect a variety of herbaceous and woody plants.

In nature RNA viruses exist as quasi-species composed of a heterogenous population of genome sequences (Holland et al. 1992). Population elasticity and plasticity act in tandem to avoid Muller's ratchet (Cervera and Elena 2016). The accumulation of sequence changes throughout virus evolution has resulted in different lineages of viruses which may share homologous genes that allow for computer-based tracing of their phylogenies, but they possess particular genes that can distinguish them (Koonin et al. 2020). The short-term fitness of viruses in their hosts has been extensively evaluated (Cervera and Elena 2016), and the study of long-term evolution has been greatly promoted by powerful sequencing technologies (Navarro et al. 2018a). However, due to occasionality, the historical nodes of diversification are poorly understood. This study reports the identification of rhabdovirus populations in citrus (*Citrus* spp.; Rutaceae), passion fruit (*Passiflora edulis*; Passifloraceae), and paper bush (*Edgeworthia chrysantha*; Thymelaeaceae) plants by RNA-sequencing (RNA-seq). A putative cytorhabdovirus was coincidentally found in all of these plant species. Comparison of isolates from citrus and the other plant species indicated that there is (i) a natural defect of two genes in only the citrus isolates and (ii) an extremely limited distribution of the citrus isolates in nature. Based on our analyses, we propose a virus–plant coevolution model. The speculation is that for viruses, genome elasticity ensures existence, but transformations occur due to plasticity under relaxed conditions.

MATERIALS AND METHODS

Sample collection. Plant materials used for RNA-seq were collected from Beibei district, Chongqing province, China in 2016 to 2018 (Fig. 1 and Supplementary Fig. S1). At the beginning of the study, petiole and mesophyll tissues from 50 citrus trees that were randomly selected in an orchard in Xiema town were combined into one sample and sequenced (data not shown), allowing the initial identification of citrus-associated rhabdovirus (CiaRV) in *Citrus* spp. that was defective in the P4 and G genes. Among 200 tested citrus trees, only three (NV157, UN41, and UN43) were finally found to be positive for CiaRV and were thus sampled individually. To explore whether CiaRV is present in plants commonly found in the vicinity of the orchard, leaf tissues of 12 other plant species (*Broussonetia papyrifera*, *Camellia japonica*, *Edgeworthia chrysantha*, *Eriobotrya japonica*, *Erythrina variegate*, *Ficus concinna*, *Koelreuteria paniculata*, *Melia azedarach*, *Morus alba*, *Passiflora edulis*, *Solanum nigrum*, and *Wisteria sinensis*) showing virus-like foliar symptoms were tested. Five symptomatic plants of each species were combined per sample and subjected to reverse transcription PCR (RT-PCR) using three conserved primer pairs for rhabdovirus nucleoprotein genes (data not shown). All primers were designed by DNAMAN 9 (Lynnon, Canada). Subsequently, CiaRV infection was confirmed in the symptomatic *P. edulis* and *E. chrysantha* plants using one of the primer pairs, NF660/NR940 (Supplementary Table S1). The leaves of three passion fruit plants (PF-1, PFYD-5, and PFRJ-10) and one paper bush plant (JX) were sampled in an orchard in Xiema town and in a garden in Ziyuntai community, respectively, and then were sequenced individually.

RNA-seq, sequence annotation, and virus sequence recovery. Total RNA was extracted from leaf tissues using the EASYspin Plus Complex Plant RNA Kit (Aidlab, China) and assessed by Nanodrop (Thermo Fisher Scientific, U.S.A.), Qubit 3.0 (Invitrogen, U.S.A.), and Agilent 2100 (Agilent, U.S.A.) for purity, concentration, and integrity, respectively. A rRNA-depleted RNA-seq library was constructed using a TruSeq RNA Sample Prep Kit (Illumina, USA) and sequenced on an Illumina HiSeq X-ten platform (Mega genomics, Beijing, China) with 150-bp paired-end reads layout. The output raw reads were processed to remove adaptors and poor-quality reads. The remaining clean reads were subjected to de novo assembly in the Trinity program (Haas et al. 2013). Contigs were annotated based on a homology-dependent method (Wu et al. 2015) using a local BLASTx search restricted to the virus taxid (10239) retrieved from the National Center for Biotechnology Information (NCBI) database.

To reconstruct complete viral genome sequences, specific primers (Supplementary Table S1) were designed to generate overlapping fragments. RT-PCR assays were done using a one-step RT-PCR kit (Takara, Japan). The 5′ and 3′ terminal sequences were determined using a commercial rapid amplification of cDNA ends (RACE) kit (Invitrogen, Thermo Fisher Scientific). Ultimately, the virus genome was generated by the assembly of sequences obtained from cloning (Promega, China) and sequencing of the PCR amplicons. Five clones per amplicon were fully sequenced by primer walking in both directions (Tsingke, China).

Sequence and phylogenetic analysis. Search of open reading frame (ORF), prediction of the molecular weights of putative proteins, and sequence comparisons were done using the CLC Genomics Workbench 9.5 (Qiagen, U.S.A.). ANNIE (Ooi et al. 2009) was utilized for comprehensive annotation of putative proteins. RNABindR v2.0 (Terribilini et al. 2007) was used to predict RNA binding sites. PROMALS3D (Pei et al. 2008) was used to predict protein secondary structures. Phylogenetic trees were generated in Mega 7.0 (Kumar et al. 2016) using the neighbor-joining method, p-distance substitution model, pairwise deletion treatment, and 1,000 bootstrap replications.

Movement complementation experiments. To test the movement protein (MP) function of the CiaRV P3 gene, *trans*-complementation assays were done using movement-deficient and
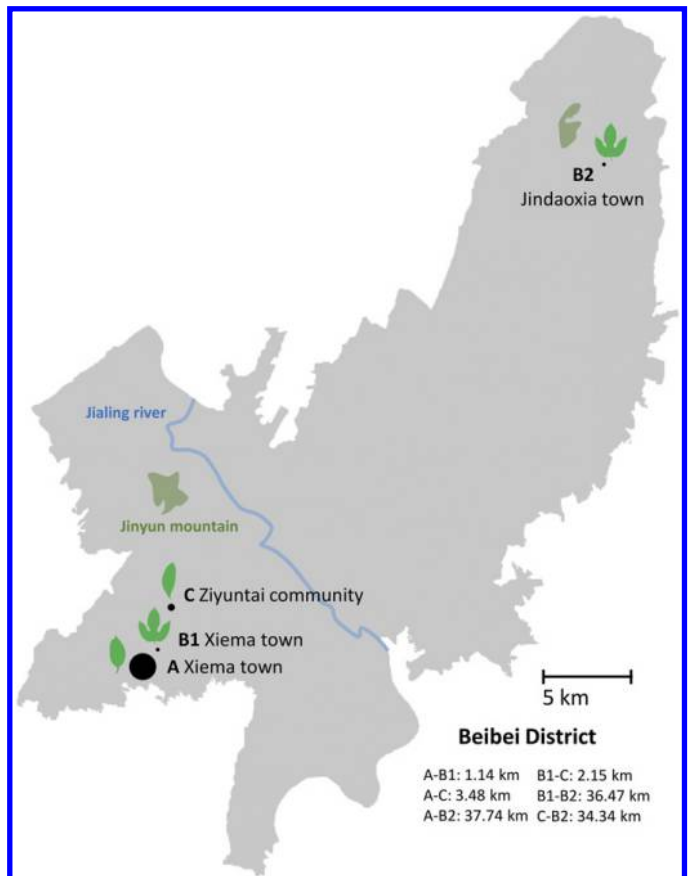

Fig. 1. Map of the location and size of sampled area in Beibei district. Citrus, passion fruit, and paper bush samples were collected in Xiema town, Xiema/Jindaoxia towns, and Ziyuntai community, respectively. The size of the black circles represents the size of the surveyed regions.

green fluorescent protein (GFP)-encoding tomato mosaic virus (ToMV-GFPΔMP) and potato virus X (PVX-GFPΔp25) recombinants, and beta-glucuronidase (GUS) expression vectors (Bayne et al. 2005; Hiraguri et al. 2011; Zhou et al. 2019). The PCR amplicons of CiaRV P3, ToMV-30K, and PVX-p25 ORFs were individually inserted into the pGD vector (Goodin et al. 2002) for transient expression using an In-Fusion HD PCR Cloning Kit (Clontech, Japan). Wild-type *Nicotiana benthamiana* plants to be used for agroinfiltration were grown in a growth chamber at 25°C under a 16-h light/8-h dark cycle. One-thousand-fold-diluted suspensions of *Agrobacterium tumefaciens* EHA 105 (initial A_{600} = 1.0) containing PVX-GFPΔp25 were mixed (1:1:1) with agrobacterial cultures carrying the TBSV p19 VSR plasmid, and pGD-P3 or pGD-PVX-p25. GFP fluorescence in the leaves was observed by an epifluorescence microscope first at 60 h postinfiltration (hpi) and then at 8 days postinfiltration (dpi). Suspensions of *A. tumefaciens* GV3101 carrying ToMV-GFPΔMP (A_{600} = 1.0) were diluted 2,000-fold and mixed 1:1 with *A. tumefaciens* cultures harboring pGD-P3 or pGD-ToMV-30K. Epifluorescence microscope observation of GFP fluorescence in the leaves was done at 50 and 60 hpi. A total of 60 GFP foci for each treatment were examined during two repeated experiments, and GFP-expressing clusters with four or more cells were considered to represent authentic movable foci.

Analysis of sequence variation in viral coding region. Viral nucleotide sequence variation types taken into account included single nucleotide polymorphisms (SNPs), insertions and deletions, which were counted using the variant detectors in the CLC Genomics Workbench 11.0, with the parameters of A (minimum coverage = 3, minimum count = 2, minimum frequency = 10%), B (minimum coverage = 3, min. count = 2, minimum frequency = 1%), and C (minimum coverage = 3, minimum count = 3, minimum frequency = 0.1%) for datasets of isolates C3 (average coverage [AC] = 9), C1 with C2 and PF2 (AC = 214 to 359), and PF1 with PF3 and PB1 (AC = 1,161 to 1,460), respectively. The degree of variation for each ORF of each isolate was measured and standardized according to the number of mutant sites (per 1,000 nucleotides) in the ORF. OriginPro 2017 software (OriginLab, U.S.A.) was used to visualize the variation, and GraphPad Prism 8.0 (GraphPad Software Inc., U.S.A.) was used for correlation analysis of sequence variation and nucleotide sequence identity.

Graft transmission of CiaRV. Virus-free seedlings of Symons sweet orange (*Citrus sinensis*) and red tangerine (*C. reticulata*), with four seedlings per variety were used as stocks to accept the bark of NV157 trees infected with CiaRV-C1. The bark of a healthy citrus tree was used as negative control. Each seedling was grafted with four pieces of bark. The treated seedlings were placed in an insect-proof glasshouse at the Citrus Research Institute, Chongqing, China, under natural light, at a controlled temperature of $25 \pm 1°C$ and relative humidity of 75%. After three months, new leaves of the stocks were tested by the one-step RT-PCR kit (Takara, Japan) with a conserved primer pair LF117/LR120 (Supplementary Table S1) designed from the L genes of all CiaRV isolates. Ten-microliter reactions contained 1 μl of total RNA from leaf samples, 5 μl of premix, 2.9 μl of RNase-free water, 0.4 μl each primer (R/F), 0.2 μl of RNase inhibitor, and 0.1 μl of enzyme mix as per the manufacturer's protocol. Thermal cycling conditions were 30 min reverse transcription at 50°C and 3 min denaturation at 95°C, followed by 35 cycles of 30 s at 95°C, 30 s at 54°C, and 40 s at 72°C, and a final 5 min extension at 72°C.

Field survey. A total of 197 citrus trees near the NV157, UN41, and UN43 trees were surveyed for CiaRV within a circle plot (radius, 750 m), regardless of symptoms. In addition to the three passion fruit and one paper bush plants used for next-generation sequencing (NGS), eight symptomatic paper bush (Ziyuntai community) and 12 symptomatic passion fruit plants (Xiema and
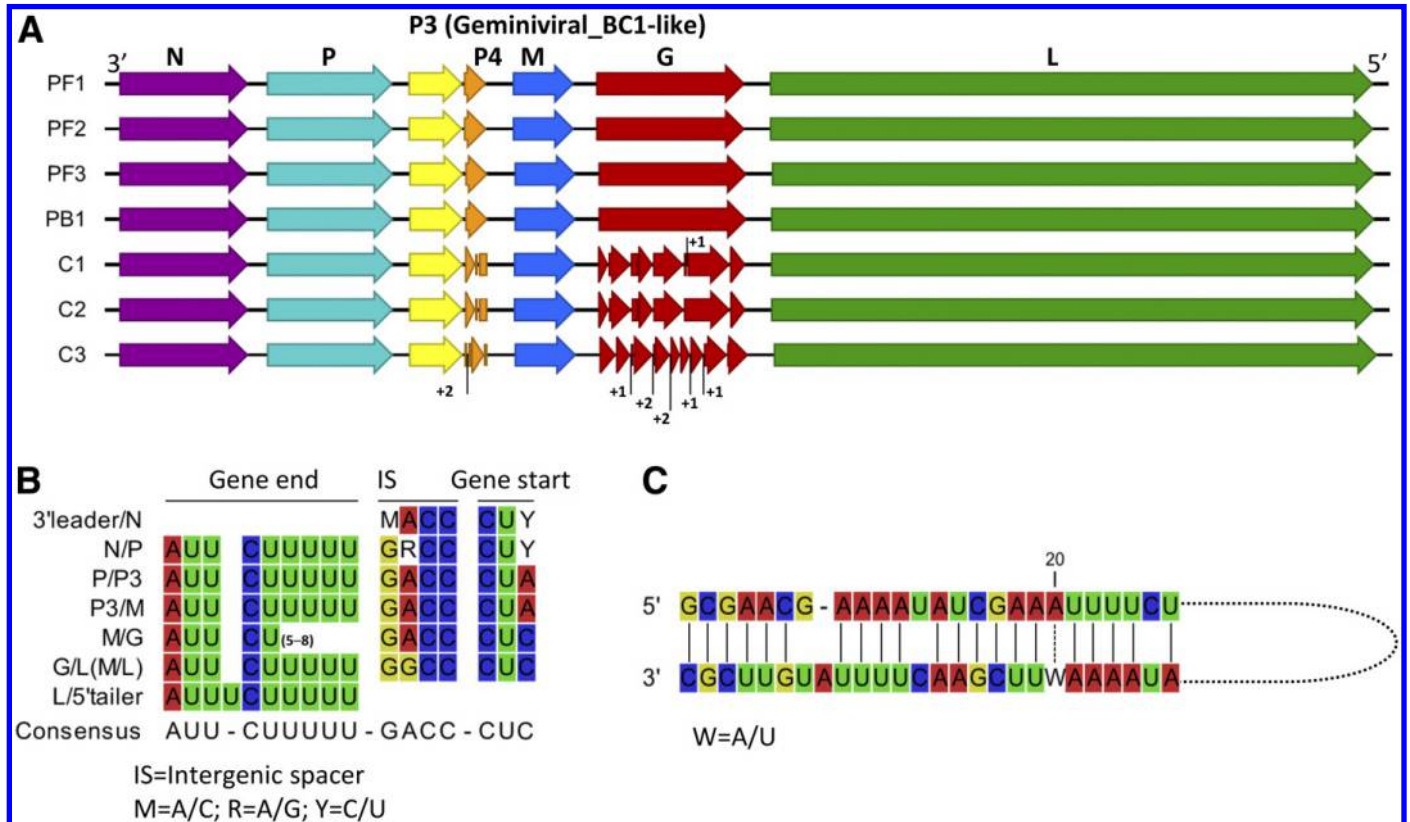

Fig. 2. A, Genomic organization of citrus-associated rhabdovirus (CiaRV) isolates. N (nucleoprotein), P (phosphoprotein), P3 (movement protein), P4 (hypothetical protein), M (matrix protein), G (glycoprotein), and L (RNA-dependent RNA polymerase. **B,** Intergenic sequences associated with mRNAs expression. **C,** Sequence complementation of genomic termini.

Jindaoxia towns) were randomly chosen for sampling in a circle plot with a radius of 200 and 100 m, respectively. The universal primers LF117 and LR120 were used for virus detection of the plant samples by RT-PCR.

RESULTS

Symptoms and location of samples. Among the samples used for HTS, while no obvious symptoms were observed on the citrus (NV157, UN41, and UN43) collected in Xiema town (location A), the passion fruit (Xiema town, location B1) and paper bush (Ziyuntai community, location C) exhibited foliar symptoms of yellow spots and chlorotic spots with yellowing, respectively (Supplementary Fig. S1). In location A, NV157, UN41, and UN43 had been planted and grown next to each another for more than a decade, and 197 citrus trees in all the regions around them were tested for CiaRV spread (Fig. 1). The distance between locations A, B1, and C is within 1 to 4 km (Fig. 1). An additional location B2 (Jindaoxia town) further away from these locations (34 to 38 km) was also tested for the presence of CiaRV in passion fruit with similar symptoms (Fig. 1).

Analysis of NGS data and virus identification. NGS of the three citrus, three passionfruit and one paper bush libraries resulted in a total of 7.85 to 12.4 G clean reads (Supplementary Table S2). Seven virus-like contigs ranging from 13,225 to 13,490 bp, one for each library, were identified and exhibited 81.4 to 97.2% amino acid (aa) sequence identity to bean-associated cytorhabdovirus (BaCV, MK202584) and papaya virus E (PVE, MH282832). These viral contigs shared >84.7% aa sequence identity with one another, suggesting the presence of a cytorhabdovirus, provisionally named citrus-associated rhabdovirus (CiaRV), in NV157, UN41, UN43, PF-1, PFYD-5, PFRJ-10, and JX plants. These cytorhabdovirus isolates are referred to as C1 to C3, PF1 to PF3, and PB1, to indicate their origins from citrus, passion fruit and paper bush respectively (MT302541 to MT302547).

Sequence characterization. The genomes of the CiaRV isolates were 13,485 to 13,526 nucleotides in size and shared 76.8 to 78.3% nucleotide and 84.7 to 99.9% amino acid sequence identity with the genomes of BaCV and PVE (Supplementary Table S3 and Supplementary Fig. S2). Seven AUG-initiated ORFs potentially encoding proteins of 189 to 2113 aa, an estimated molecular mass of 9.5 to 241.2 kDa, and an isoelectric point (IP) of 4.0 to 9.4 were predicted in the CiaRV genomes (Supplementary Fig. S3). Analysis of the proteins encoded from 3′ to 5′ by ORF1 to ORF7 suggested that they represent the nucleoprotein (N), phosphoprotein (P), movement protein (P3), hypothetical protein (P4), matrix protein (M), glycoprotein (G), and large RNA-dependent RNA polymerase (L), respectively (Supplementary Table S4). The amino acid sequences of CiaRV-encoded proteins are more than 80% identical to the analogous proteins of BaCV and PVE, with more than 90% sequence identity between the respective N and L proteins (Supplementary Fig. S2). Based on predictive algorithms, the N protein may possess several RNA binding sites concentrated at the carboxy terminus (amino acids 328 to 335, 347 to 355, 370 to 378, and 443 to 451). One phosphorylation site was found in P protein at

amino acids 152 to 155, at the end of the N-terminal disordered regions. The core region of P3 is typical of an unfolded 30K-like MP domain with a series of structural α-helix and β-strand components (Supplementary Fig. S4) (Melcher 2000; Mushegian and Elena 2015; Yu et al. 2013). The M protein showed a similar secondary structure as those of lyssaviruses (Graham et al. 2008), suggesting similar functional roles (Supplementary Fig. S5). The G protein contains five predicted glycosylation sites at amino acids 61 to 64, 229 to 232, 324 to 327, 399 to 402, and 408 to 411, a signal-peptide, and a putative transmembrane domain. These sites may aid the participation of rhabdoviral G in virion assembly and virus entry into host cells (Gaudin and Whitt, 2015; Roche et al. 2008; Sun et al. 2018). The L protein contains three predicted CDD domains at amino acids 265 to 1,105 (pfam00946), amino acids 1,136 to 1,360 (pfam14318), and amino acids 1,229 to 1,967 (TIGR04198). The former domain has been shown to be involved in replication, and the latter two are essential for mRNA capping processes; a complex consisting of N-P-L or L-(P)$_3$ is deemed to be required for full functional performance (Ivanov et al. 2011). Functions of the hypothetical accessory protein P4 are unknown. The coding region of CiaRV is flanked by untranslated 3′ leader (l) and 5′ trailer (t) sequences as for all rhabdoviruses. The termini of the genomes are highly complementary, which may contribute to the formation of a panhandle structure thought to be involved in replication. The conserved intergenic sequences separating the genes have a consensus of AUUCUUUUU/GACC/CUC and were recognized by analogy with similar sequences in other plant rhabdoviruses (Supplementary Fig. S6). In summary, CiaRV has the canonical genome arrangement 3′l-N-P-P3-(P4)-M-(G)-L-5′t of viruses in the genus *Cytorhabdovirus* (Supplementary Fig. S7), despite the P4 and G genes not being intact in all members (Fig. 2).

Defects in the P4 and G genes of citrus isolates of CiaRV. The regions of the P4 and G genes in the genomes of the three citrus isolates (C1, C2, and C3) were cloned and sequenced using specific primers (Supplementary Table S1). These genes displayed irregular patterns of various short coding regions rather than a single ORF. Missense, frameshift (+1, +2), and nonsense mutations are responsible for making the genes defective (Table 1). There were two distinct types of defects in the P4 and G genes of C1, C2, and C3. Nevertheless, citrus, paper bush, and passion fruit isolates shared a high overall nucleotide sequence identity in the P4 (78.5 to 87.1%) and G (81.9 to 87.1%) genes. It should be emphasized that when analyzing the NGS data of citrus, no sequence variant showed the presence of an intact G gene. These results suggest that the defects that occurred in the G gene were gradual, stochastic, but stable genetic transitions.

Evolutionary relationships. The phylogenetic trees (Fig. 3 and Supplementary Fig. S8) of the various viral proteins had a similar topology, including a distinct cluster comprising CiaRV, BaCV, PVE (Alves-Freitas et al. 2019; Bejerman and Dietzgen 2019; Medina-Salguero et al. 2019), and *Bemisia tabaci* virus-like sequences (BTCRV) separated from the other cytorhabdoviruses. High amino acid sequence identities of all encoded proteins (Supplementary Fig. S2) between CiaRV and BTCRV (whitefly) may be indicative of a new biological vector type for these

TABLE 1. List of genetic changes causing defects in the P4 and G genes of citrus-associated rhabdovirus in citrus plants[a]

Virus	P4				G			
	Start	Nonsense mutation	Insertion	Stop	Start	Nonsense mutation	Insertion	Stop
C1	AUA_{3785}	UGA_{3893}, UAG_{3902}, UAG_{3926}	−	UGA_{4012}	AUG_{5182}	UGA_{5290}, UAG_{5527}, UAG_{5602}, UGA_{5758}, UGA_{6079}, UAG_{6566}	A_{6119}	CUU_{6734}
C2	AUA_{3785}	UGA_{3893}, UAG_{3902}, UAG_{3926}	−	UGA_{4012}	AUG_{5185}	UGA_{5293}, UAG_{5530}, UAG_{5605}, UGA_{5761}, UGA_{6082}, UAG_{6568}	−	CUU_{6736}
C3	AUG_{3785}	UAA_{3847}, UAG_{3988}, UGA_{4000}	GC_{3814}	UUA_{4024}	AUG_{5198}	UAG_{5369}, UAA_{5544}, UGA_{5556}, UGA_{5777}, UAG_{6043}, UAG_{6537}	C_{5523}, AA_{5759}, AU_{5941}, U_{6149}, G_{6291}	CUU_{6762}

[a] Codons in italic: missense mutation.

cytorhabdoviruses (Bejerman and Dietzgen 2019). In fact, whiteflies are responsible for the dispersal of many important plant viruses (Navas-Castillo et al. 2011). The insertion of exogenous sequences derived from rhabdoviruses and plant RNA viruses into insect genomes has been previously demonstrated (Cui and Holmes 2012; Geisler and Jarvis 2016). Interestingly, the phylogenetic tree based on MP amino acid sequences (Fig. 3B) shows that the CiaRV virus cluster (whitefly-borne?) was more closely related to whitefly-borne geminiviruses than to other cytorhabdoviruses, suggesting an ancient recombination event between the ancestors of CiaRV and begomoviruses.

MP function of CiaRV P3. To demonstrate the cell-to-cell movement function of P3, two movement-defective heterologous RNA virus systems were tested for movement *trans*-complementation in agroinfiltrated *N. benthamiana* leaves. When ToMV-GFPΔMP (8 dpi) or PVX-GFPΔp25 (60 hpi) was coexpressed with GUS, GFP foci were confined to single cells (Fig. 4). In contrast, when coexpressed with P3, 55% of ToMV-GFP MP foci (33 out of 60) and 42% of PVX-GFP p25 foci (25 out of 60) contained GFP clusters of four or more cells (Fig. 4). For these fluorescent foci, the average numbers of GFP-expressing cells were 6.7 for ToMV-GFP MP and 8.1 for PVX-GFP p25. However, when ToMV-GFP MP or PVX-GFP p25 was coexpressed with their own MPs, i.e., ToMV MP or PVX p25, the majority of GFP foci spread across more than 10 cell boundaries. Thus, the intercellular movement of both GFP-tagged movement-defective viruses was rescued by transient expression of P3 in trans, although the complementation efficiencies by CiaRV P3 were lower than those by their autologous MPs, 30K or p25 (Fig. 4).

Viral genetic variation in HTS data. A total of 135 (1%), 289 (2.1%), 110 (0.8%), 134 (1%), 155 (1.1%), 254 (1.9%), and 147 (1.1%) variable sites were detected in ORFs based on verified genome sequences of C1, C2, C3, PF1, PF2, PF3, and PB1 isolates

(13,485 to 13,526 nt), respectively. In ORFs for the G protein (1,558 to 1,570 nt), variable sites were numbered 20 (1.3%), 28 (1.7%), 26 (1.7%), 14 (0.9%), 18 (1.1%), 38 (2.4%), and 18 (1.7%), respectively (Supplementary Table S5). These data indicate that the degree of variation in some adjacent ORFs (M and G) of all isolates are more similar and higher than in other ORFs, suggesting the two ORFs may be evolutionary hotspots (Fig. 5A). Interestingly, nucleotide sequence identities between G ORFs (81.8 to 99.7%) of all isolates tended to be lower than those of other ORFs (N, 88.6 to 99.8%; P, 87.2 to 99.9%; P3 and P4, 85 to 99.9%; M, 85 to 99.8%; and L, 87.9 to 99.8%). Pearson correlation analysis of each ORF using the lowest identity on the x-axis and the average number of variation sites per 1,000 nucleotides on the y-axis, with one unreliable point removed (M: x = 0.85, y = 21.6), showed a strong correlation coefficient of –0.7963, which suggests that viral short-term sequence variation may positively affect long-term sequence difference (Fig. 5B).

CiaRV systematic infection and prevalence. CiaRV-C1 was detected in the new leaves of all the CiaRV-infected graft-inoculated citrus stocks, but not in any of the trees grafted with negative control bark, indicating CiaRV transmission by grafting in citrus. Furthermore, systemic infection of CiaRV in a single plant indicated that the virus can overcome the vascular barriers. However, with the exception of NV157, UN41, and UN43, no other CiaRV-infected citrus trees were found. Neither the citrus field samples nor the grafted plants showed any virus-like symptoms. In contrast, 63% (5/8) of the symptomatic passion fruit trees and 75% (9/12) of the symptomatic paper bush trees tested by RT-PCR were positive for CiaRV, which may suggest an association between CiaRV infections and the observed symptoms (Supplementary Fig. S9), but this will need to be confirmed by testing a larger number of symptomatic and asymptomatic samples. Based on the differences in sequence, phylogeny, host, and possibly
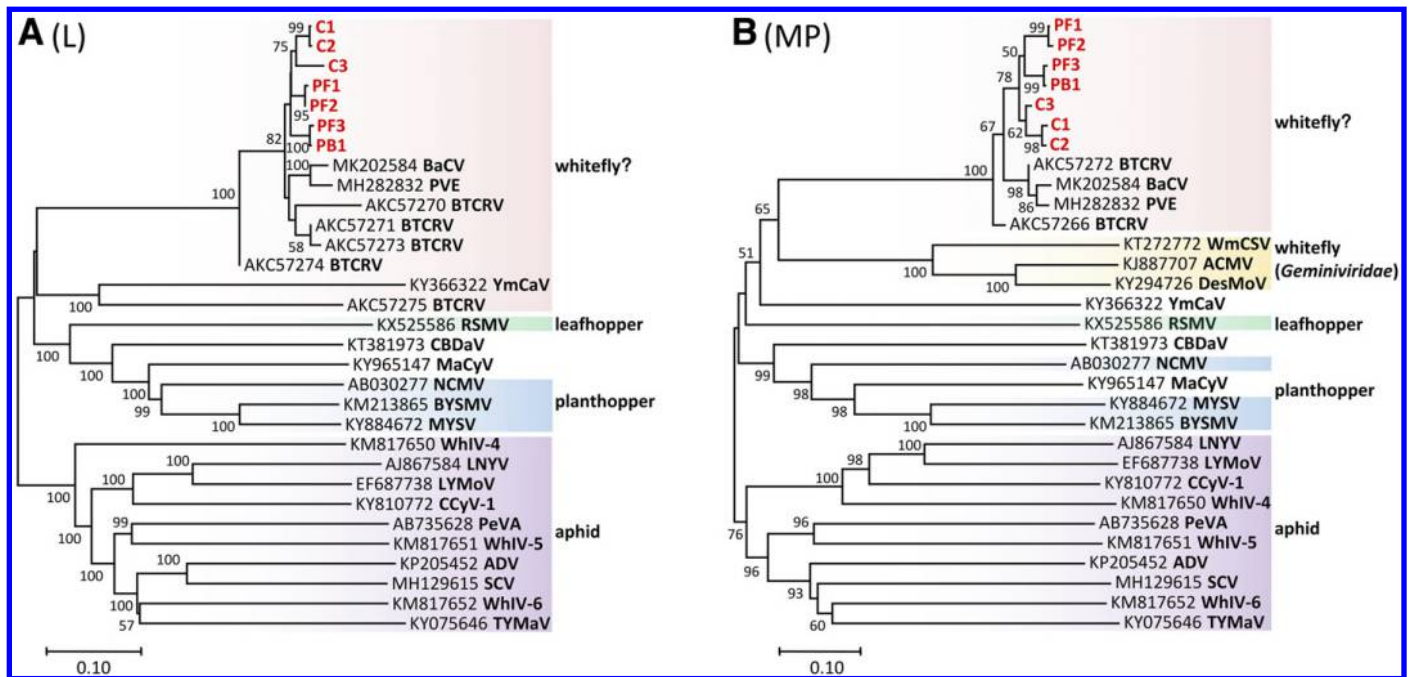

Fig. 3. Phylogenetic trees inferred from the amino acid sequence comparisons of **A,** RNA-dependent RNA polymerase (L) and **B,** movement protein (MP) of citrus-associated rhabdovirus (CiaRV) isolates (in red or gray font), bean-associated cytorhabdovirus (BaCV), papaya virus E (PVE), *Bemisia tabaci* virus-like sequences (BTCRV), and representative cytorhabdoviruses. In the MP tree, several related begomoviruses were also included. One-thousand bootstrap replications were applied, with percentage values that exceeded 50 shown at the nodes. The bar represents the number of amino acid replacements per site. The viruses used for tree construction included African cassava mosaic virus (ACMV), alfalfa dwarf virus (ADV), barley yellow striate mosaic virus (BYSMV), cabbage cytorhabdovirus-1 (CCyV-1), colocasia bobone-disease associated virus (CBDaV), desmodium mottle virus (DesMoV), lettuce necrotic yellow virus (LNYV), lettuce yellow mottle virus (LYMoV), maize-associated cytorhabdovirus (MaCyV), maize yellow striate mosaic virus (MYSV), northern cereal mosaic virus (NCMV), persimmon virus A (PeVA), rice stripe mosaic virus (RSMV), strawberry crinkle virus (SCV), tomato yellow mottle-associated virus (TYMaV), watermelon chlorotic stunt virus (WmCSV), Wuhan insect virus 4 (WhIV-4), Wuhan insect virus 5 (WhIV-5), Wuhan insect virus 6 (WhIV-6), and yerba mate chlorosis-associated virus (YmCaV).

associated symptoms, CiaRV isolates were classified into three strains viz. citrus (C), passion fruit (PF), and paper bush-passion fruit (PBF). Furthermore, the investigation of CiaRV in Beibei district showed that distribution of the virus in citrus was limited (3/200 trees = 1.5%), but was widespread in other hosts, i.e., 80% (12/15) of passion fruit trees and 67% (6/9) of paper bush, including trees sequenced by NGS.

DISCUSSION

In this study, the utilization of NGS led to the identification of CiaRV in citrus, passion fruit, and paper bush plants grown within a confined geographic region in China (Beibei district, ~750 km²). Analysis of sequence differences, phylogenetic relationships, host range, and symptomatology suggested that there may be at least three distinct CiaRV strains, namely, C (C1 to C3), PF (PF1 and PF2), and PBF (PB1 and PF3). Analysis of genome organization led to the differentiation of two genotypes, gene-defective (citrus strain) and gene-intact (other strains). CiaRV is a close relative of the cytorhabdoviruses BaCV and PVE that were recently reported in South America (Alves-Freitas et al. 2019; Bejerman and Dietzgen 2019; Medina-Salguero et al. 2019). The species demarcation criteria for the genus *Cytorhabdovirus* have recently been revised to reflect the increasing number of newly discovered cytorhabdoviruses: (i) nucleotide sequence identity less than 75% for the complete genome sequence and (ii) amino acid sequence identity in all cognate ORFs less than 80% (Freitas-Astúa et al. 2020). CiaRV genome is >75% identical to BaCV and PVE and its encoded proteins are 81 to 97% identical to the cognate proteins of these viruses that are both classified in the species *Papaya cytorhabdovirus*. Therefore, we suggest that CiaRV should be taxonomically

classified in the established species *Papaya cytorhabdovirus* in the genus *Cytorhabdovirus*.

CiaRV genome organization is similar to that of other unsegmented plant rhabdoviruses, with five structural proteins in the order 3′l-N-P-P3-(P4)-M-(G)-L-5′t. For all rhabdoviruses, the core proteins N, P, and L are required for viral replication, and their combination with genomic RNA generates a nucleocapsid (NC) complex (Leyrat et al. 2011). The P chaperones N (N-P), associates with L (L-P), and probably support the intracellular transport of viral components in host cells (Leyrat et al. 2011). In plant, the NC interacts with a cognate MP likely to form a MP-NC complex as a basic unit of viral intercellular transport (Hiraguri et al. 2012; Zhou et al. 2019).

The P3 of CiaRV, encoded by an accessory gene located between P and M genes, was identified as an MP related to geminiviruses. This is consistent with the genomic modularity of viruses in eukaryotes that manifests as recombination or reassortment (Koonin et al. 2015; Navarro et al. 2018a; Roossinck 2008). The MPs of plant rhabdoviruses and geminiviruses belong to the 30K superfamily (Huang et al. 2005; Mann et al. 2016; Mushegian and Elena 2015; Navarro et al. 2019), and CiaRV P3 is characterized by having a 30K domain (Melcher 2000). Our results with CiaRV P3 confirm previous findings that rhabdoviral MPs can rescue the cell-to-cell movement of movement-defective positive-strand RNA viruses (Hiraguri et al. 2012; Huang et al. 2005; Mann et al. 2016), but these MPs are unlikely to be interchangeable (Zhou et al. 2019). The plasmodesmata of plant cells are gateways for viral cell-to-cell and systematic infection (Benitez-Alfonso et al. 2010). The MPs of TMoV (30K) and PVX (triple gene block) are known to trigger plasmodesmata gating, which may direct viral transport between cells (Angell et al. 1996; Howard et al. 2004; Tamai and Meshi
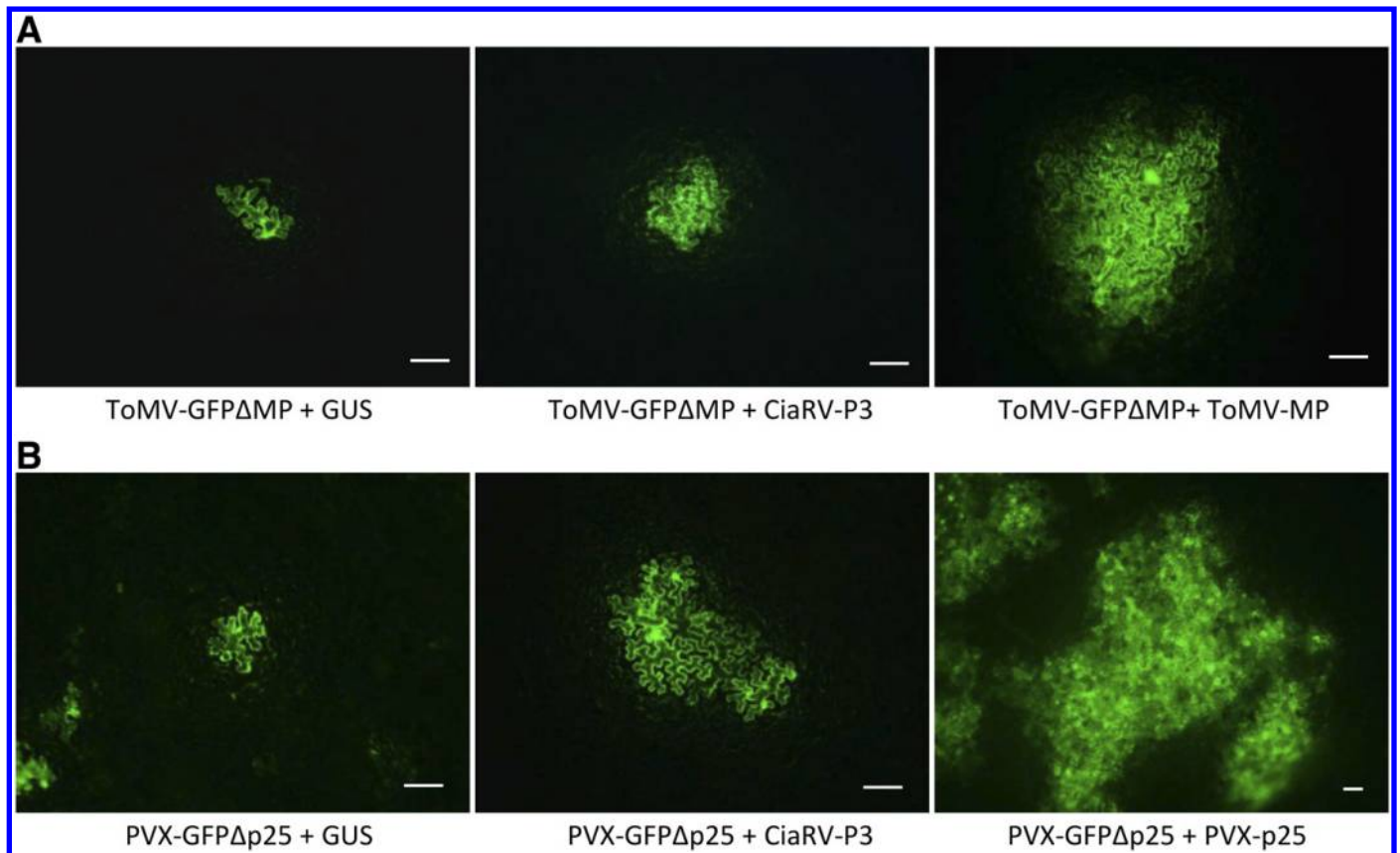

Fig. 4. *Trans*-complementation of the movement protein (MP)-mutant tomato mosaic virus (ToMV) and potato virus X (PVX), using coinfiltration of agrobacterial mixtures carrying a transcription plasmid of **A,** ToMV-GFPΔMP or **B,** PVX-GFPΔp25, and a binary plasmid (pGD-X) of citrus-associated rhabdovirus (CiaRV)-P3, ToMV MP, PVX p25, or GUS as a negative control, into the leaves of *Nicotiana benthamiana* plants. The GFP signal was observed under a fluorescence microscope at 60 hpi (A) or 8 dpi (B). Bars (regardless of length), 200 μm.

2001; Yang et al. 2000). The CiaRV P3 expressed in trans is compatible with movement-deficient TMoV and PVX, which may suggest a similar mechanism. However, how CiaRV P3 functions in planta is yet to be determined.

A naturally G gene-defective cytorhabdovirus was discovered here for the first time, although viral genomes in which a potential G gene was not detected have been previously reported for some other nsRNA viruses (Li et al. 2015; Navarro et al. 2018a, b; Rott et al. 2018; Shi et al. 2016; Xin et al. 2017). One might speculate that simplification of viral genomes to adapt to plants without requiring an arthropod vector provides an evolutionary advantage, especially in fruit trees that are propagated artificially by asexual modes, such as cutting and grafting. Alternatively, the lack of a G gene might be a common attribute of different lineages of nsRNA viruses in various plants in the natural world. In a natural setting, what directly drives the diversification of nsRNA viruses is the high error rate during virus replication (Drake et al. 1998; Steinhauer et al. 1992). Experimental evidence has shown that in the absence of G and M proteins, the NC complex can still be assembled during rhabdovirus infections, but with minimal infectivity (Wang et al. 2015; Whelan et al. 2004). This provides flexibility for rhabdoviruses with an additional accessory MP for cell-to-cell movement in plants, but differs from arthropods where the G protein is required for cell recognition and entry (Deom et al. 1992; Lucas 2006; Navarro et al. 2019; Roche et al. 2008). This strongly underlines the viral intrinsic ability of losing nonessential genes in specific hosts, prior to any extrinsic environmental effects on the mutant rhabdovirus. In addition, the quasi-species nature of RNA viruses results in a flexible, mutation-tolerant population where several genomes may be dominant (Holland et al. 1992; Roossinck 2008), and the functional genomes may succumb to disabled ones. In support of this hypothesis, a significant number of tolerated CiaRV gene mutations were detected by NGS in passion fruit and paper bush. Here we highlight the use of NGS in probing the genetic structure of a quasi-species. In contrast to traditional techniques such as PCR, whereby the information obtained is partial and may lead to biases, NGS yields vast amounts of data at the whole-genome level at a particular point in time. However, intrinsic errors may still exist in the output data. Therefore, based on the above information, it is reasonable to assume that an initial variant of the genetically incorrect CiaRV-C was not naturally eliminated. Given that a virus population is elastic in order to repair deleterious mutations (Cervera and Elena 2016), genetic drift or a founder effect could be
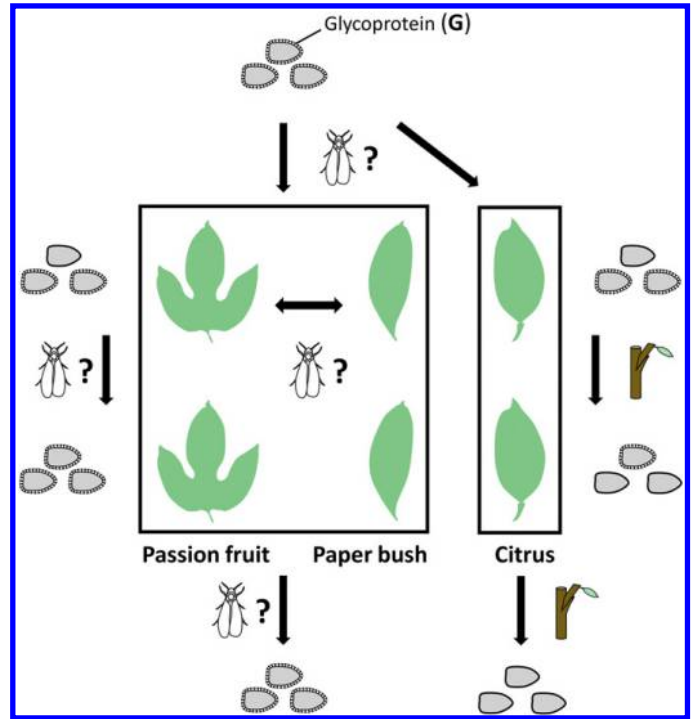

Fig. 6. Schematic illustrating the putative evolution of citrus-associated rhabdovirus (CiaRV) in citrus or passion fruit and paper bush plants. White-flies as possible insect vectors carried intact CiaRV to citrus and other plants in different epidemic seasons. As CiaRV keeps gene integrity in passion fruit and paper bush, the CiaRV of citrus are likely selected simultaneously by pressures from the host and an environment where the vectors have migrated to other places, with its P4 and glycoprotein (G) genes damaged. The grafting of citrus may contribute to fixing the genetic changes in the virus genome.
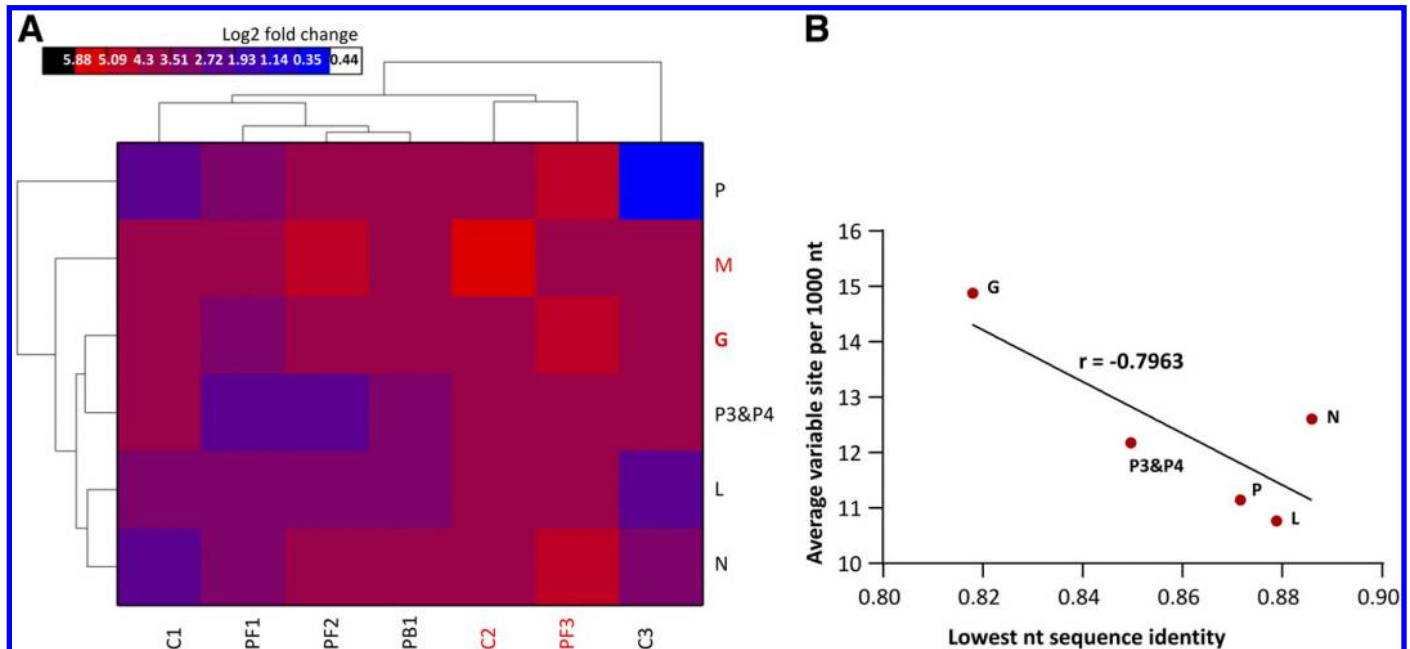

Fig. 5. A, Heatmap of sequence variation (variable sites per 1,000 nt) of open reading frames (ORFs) from isolates C1, C2, C3, PF1, PF2, PF3, and PB1 and **B,** Pearson correlation between degree of sequence variation and sequence identity. N (nucleoprotein), P (phosphoprotein), P3 and P4 (movement protein and hypothetical protein), M (matrix protein), G (glycoprotein), and L (large RNA-dependent RNA polymerase). r = Pearson correlation coefficient.

invoked to explain why the molecular transitions of defective-CiaRV in citrus were retained and stabilized, as opposed to a mutational meltdown (Delatte et al. 2007). Otherwise, natural selections may play a role in the fixation of the mutant genes. Other vital genes are also exposed to pressures from the environment, and the asymmetric effects on essential and nonessential genes may yield a defective virus with more beneficial mutations in essential genes and more fitness in citrus than the original virus. In summary, it appears that without significant selection pressure from the environment, virus evolution is stochastic among the many given possibilities based on the viral intrinsic properties such as genetic plasticity (spontaneous mutation), genome modularity (recombination or reassortment ability), and gene function specialization (nonessential function) (Roossinck 2008). For instance, when gain or loss of genes is not restricted by the environment, viruses may acquire or lose genes as long as this is nonlethal, such as G-defects of CiaRV-C in citrus. But if the outcomes are selected by the environment, viral genetic changes may induce lethality. These hypotheses could well explain that different viruses infecting the same host under the same environmental conditions vary in genome type (such as RNA or DNA and circular or linear), genome size, gene number, protein structure, and/or sequence homology, but share the essential core elements for replication and encapsidation (Koonin et al. 2020).

The intact P4 and G genes are not essential for replication of CiaRV in citrus but are present in CiaRV isolates that infect other plant species. The function of the hypothetical protein P4 is unknown, while G protein is required for acquisition by arthropod vectors. The natural transmission of CiaRV and related viruses occurs at three spatial dimensions: (i) infection of a few citrus plants, likely by grafting, (ii) transspecies infection of different plants of the same geographic region, and (iii) transspecies and transcontinent infection of different plant species, considering the closely related viruses reported in French bean and papaya in South America. Spread of these related viruses in different plant species implies viral transmission through biological vectors, which are possibly whiteflies due to the proposed long-term interaction of these insects and the CiaRV-related viruses that is suggested by sequence and phylogenetic analysis. Based on the available data, we propose an evolutionary model for CiaRV (Fig. 6). At first, in one or more epidemic seasons, the viruses were carried by a potential arthropod vector, likely the whitefly. Dynamic populations of the vector fed on citrus, passion fruit, and paper bush, and transmitted viruses containing few defective variants. Genetic drift in citrus led to a virus population dominated by defective genomes. The defective viruses were occasionally disseminated to other citrus plants by vegetative propagation or contamination of secateurs, and thereafter colonized following a founder effect. As perennial plants provide a stable host environment, the coevolution of CiaRV and citrus, which may be a "specific host" for this virus (Bedhomme et al. 2012), can be considered as the main contributor to the disrepair of the P4 and G genes. Since the G gene is considered crucial for vector transmission, the lack of virus variant selection by the vector may constitute another important factor. Besides, more frequent variation in the G gene (Fig. 5), which could increase the probability of errors in the ORF, may have had a synergistic effect. The G protein may affect virion morphogenesis (Wang et al. 2015), and the surface of virions of CiaRV in citrus may lack receptor spikes. Ultimately, defective CiaRV may have established a stable host–virus pathosystem in citrus. The key checkpoint of the model is whether CiaRV is more widespread in citrus to determine if the gene defect in CiaRV is an isolated event or not. This will require further extensive surveys in citrus across the region, country, and globally.

LITERATURE CITED

Alves-Freitas, D. M., Pinheiro-Lima, B., Faria, J. C., Lacorte, C., Ribeiro, S. G., and Melo, F. L. 2019. Double-stranded RNA high-throughput sequencing reveals a new cytorhabdovirus in a bean golden mosaic virus-resistant common bean transgenic line. Viruses 11:90.

Angell, S. M., Davies, C., and Baulcombe, D. C. 1996. Cell-to-cell movement of potato virus X is associated with a change in the size-exclusion limit of plasmodesmata in trichome cells of *Nicotiana clevelandii*. Virology 216:197-201.

Armstrong, R., Robinson, J., Rymes, C., and Needham, T. 1993. British Columbia. Infectious hematopoietic necrosis in Atlantic salmon in British Columbia. Can. Vet. J. 34:312.

Assenberg, R., Delmas, O., Morin, B., Graham, S., De Lamballerie, X., Laubert, C., Coutard, B., Grimes, J. M., Neyts, J., Owens, R. J., Brandt, B. W., Gorbalenya, A., Tucker, P., Stuart, D. I., Canard, B., and Bourhyb, H. 2010. Genomics and structure/function studies of *Rhabdoviridae* proteins involved in replication and transcription. Antiviral Res. 87:149-161.

Bayne, E. H., Rakitina, D. V., Morozov, S. Y., and Baulcombe, D. C. 2005. Cell-to-cell movement of potato potexvirus X is dependent on suppression of RNA silencing. Plant J. 44:471-482.

Bedhomme, S., Lafforgue, G., and Elena, S. F. 2012. Multihost experimental evolution of a plant RNA virus reveals local adaptation and host-specific mutations. Mol. Biol. Evol. 29:1481-1492.

Bejerman, N., and Dietzgen, R. G. 2019. Bean-associated cytorhabdovirus and papaya cytorhabdovirus are strains of the same virus. Viruses 11:230.

Benitez-Alfonso, Y., Faulkner, C., Ritzenthaler, C., and Maule, A. J. 2010. Plasmodesmata: Gateways to local and systemic virus infection. Mol. Plant-Microbe Interact. 23:1403-1412.

Cervera, H., and Elena, S. F. 2016. Genetic variation in fitness within a clonal population of a plant RNA virus. Virus Evol. 2:vew006.

Cui, J., and Holmes, E. C. 2012. Endogenous RNA viruses of plants in insect genomes. Virology 427:77-79.

Delatte, H., Holota, H., Moury, B., Reynaud, B., Lett, J. M., and Peterschmitt, M. 2007. Evidence for a founder effect after introduction of tomato yellow leaf curl virus–mild in an insular environment. J. Mol. Evol. 65:112-118.

Deom, C. M., Lapidot, M., and Beachy, R. N. 1992. Plant virus movement proteins. Cell 69:221-224.

Dietzgen, R. G., Calisher, C. H., Kurath, G., Kuzman, I. V., Rodriguez, L. L., Stone, D. M., Tesh, R. B., Tordo, N., Walker, P. J., Wetzel, T., and Whitfield, A. E. 2012. Family *Rhabdoviridae*. Pages 686-713 in: Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses. A. M. Q. King, M. J. Adams, E. B. Carstens, and E. J. Lefkowitz, eds. Elsevier-Academic Press, London, U.K.

Dietzgen, R. G., Kuhn, J. H., Clawson, A. N., Freitas-Astúa, J., Goodin, M. M., Kitajima, E. W., Kondo, H., Wetzel, T., and Whitfield, A. E. 2014. *Dichorhavirus*: A proposed new genus for *Brevipalpus* mite-transmitted, nuclear, bacilliform, bipartite, negative-strand RNA plant viruses. Arch. Virol. 159:607-619.

Dorson, M., Torchy, C., Chilmonczyk, S., De Kinkelin, P., and Michel, C. 1984. A rhabdovirus pathogenic for perch, *Perca fluviatilis* L.: Isolation and preliminary study. J. Fish Dis. 7:241-245.

Drake, J. W., Charlesworth, B., Charlesworth, D., and Crow, J. F. 1998. Rates of spontaneous mutation. Genetics 148:1667-1686.

Freitas-Astúa, J., Dietzgen, R. G., Walker, P. J., Blasdell, K. R., Breyta, R., Fooks, A. R., Kondo, H., Kurath, G., Kuzmin, I. V., Stone, D. M., Tesh, R. B., Tordo, N., Vasilakis, N., Whitfield, A. E., and Ramos-González, P. L. 2020. Create twelve new species in the genus *Cytorhabdovirus*, family *Rhabdoviridae*. Approved ICTV proposal 2019.030M. https://talk.ictvonline.org/

Gaudin, Y., and Whitt, M. A. 2015. Rhabdovirus glycoproteins. Pages 49-73 in: Biology and Pathogenesis of Rhabdo- and Filoviruses. A. K. Pattnaik and M. A. Whitt, eds. World Scientific Publishing Co., Singapore.

Geisler, C., and Jarvis, D. L. 2016. Rhabdovirus-like endogenous viral elements in the genome of *Spodoptera frugiperda* insect cells are actively transcribed: Implications for adventitious virus detection. Biologicals 44: 219-225.

Goodin, M. M., Dietzgen, R. G., Schichnes, D., Ruzin, S., and Jackson, A. O. 2002. pGD vectors: Versatile tools for the expression of green and red fluorescent protein fusions in agroinfiltrated plant leaves. Plant J. 31: 375-383.

Graham, S. C., Assenberg, R., Delmas, O., Verma, A., Gholami, A., Talbi, C., Owens, R. J., Stuart, D. I., Grimes, J. M., and Bourhy, H. 2008. Rhabdovirus matrix protein structures reveal a novel mode of self-association. PLoS Pathog 4:e1000251.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., Henschel, R., LeDuc, R. D., Friedman, N., and Regev, A. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8:1494-1512.

Hiraguri, A., Hibino, H., Hayashi, T., Netsu, O., Shimizu, T., Uehara-Ichiki, T., Omura, T., Sasaki, N., Nyunoya, H., and Sasaya, T. 2012. The movement protein encoded by gene 3 of rice transitory yellowing virus is associated with virus particles. J. Gen. Virol. 93:2290-2298.

Hiraguri, A., Netsu, O., Shimizu, T., Uehara-Ichiki, T., Omura, T., Sasaki, N., Nyunoya, H., and Sasaya, T. 2011. The nonstructural protein pC6 of rice grassy stunt virus trans-complements the cell-to-cell spread of a movement-defective tomato mosaic virus. Arch. Virol. 156:911-916.

Holland, J. J., De La Torre, J. C., and Steinhauer, D. A. 1992. RNA virus populations as quasispecies. Curr. Top. Microbiol. 176:1-20.

Howard, A. R., Heppler, M. L., Ju, H. J., Krishnamurthy, K., Payton, M. E., and Verchot-Lubicz, J. 2004. Potato virus X TGBp1 induces plasmodesmata gating and moves between cells in several host species whereas CP moves only in *N. benthamiana* leaves. Virology 328:185-197.

Huang, Y. W., Geng, Y. F., Ying, X. B., Chen, X. Y., and Fang, R. X. 2005. Identification of a movement protein of rice yellow stunt rhabdovirus. J. Virol. 79:2108-2114.

Ivanov, I., Yabukarski, F., Ruigrok, R. W. H., and Jamin, M. 2011. Structural insights into the rhabdovirus transcription/replication complex. Virus Res. 162:126-137.

Jackson, A. O., Dietzgen, R. G., Fang, R. X., Goodin, M. M., Hogenhout, S. A., Deng, M., and Bragg, J. N. 2009. Plant rhabdoviruses. Pages 270-279 in: Desk Encyclopedia of Plant and Fungal Virology. M. H. V. van Regenmortel and B. W. J. Mahy, eds. Elsevier-Academic Press, Oxford, U.K.

Jackson, A. O., Dietzgen, R. G., Goodin, M. M., Bragg, J. N., and Deng, M. 2005. Biology of plant rhabdoviruses. Annu. Rev. Phytopathol. 43: 623-660.

Koonin, E. V., Dolja, V. V., and Krupovic, M. 2015. Origins and evolution of viruses of eukaryotes: The ultimate modularity. Virology 479:2-25.

Koonin, E. V., Dolja, V. V., Krupovic, M., Varsani, A., Wolf, Y. I., Yutin, N., Zerbini, F. M., and Kuhn, J. H. 2020. Global organization and proposed megataxonomy of the virus world. Microbiol. Mol. Biol. Rev. 84: e00061-e19.

Kormelink, R., Garcia, M. L., Goodin, M., Sasaya, T., and Haenni, A. L. 2011. Negative-strand RNA viruses: The plant-infecting counterparts. Virus Res. 162:184-202.

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.

Kuzmin, I. V., Novella, I. S., Dietzgen, R. G., Padhi, A., and Rupprecht, C. E. 2009. The rhabdoviruses: Biodiversity, phylogenetics, and evolution. Infect. Genet. Evol. 9:541-553.

Kuzmin, I. V., and Tordo, N. 2012. Genus *Lyssavirus*. Pages 37-58 in: Rhabdoviruses: Molecular Taxonomy, Evolution, Genomics, Ecology, Host–Vector Interactions, Cytopathology and Control. R. G. Dietzgen and I. M. Kuzmin, eds. Caister Academic Press, Norfolk, U.K.

Leyrat, C., Ribeiro, E. A., Géérard, F. C., Ivanov, I., Ruigrok, R. W., and Jamin, M. 2011. Structure, interactions with host cell and functions of rhabdovirus phosphoprotein. Future Virol. 6:465-481.

Li, C. X., Shi, M., Tian, J. H., Lin, X. D., Kang, Y. J., Chen, L. J., Qin, X. C., Xu, J., Holmes, E. C., and Zhang, Y. Z. 2015. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. eLife 4:e05378.

Longdon, B., Murray, G. G., Palmer, W. J., Day, J. P., Parker, D. J., Welch, J. J., Obbard, D. J., and Jiggins, F. M. 2015. The evolution, diversity, and host associations of rhabdoviruses. Virus Evol. 1:vev014.

Lucas, W. J. 2006. Plant viral movement proteins: Agents for cell-to-cell trafficking of viral genomes. Virology 344:169-184.

Mann, K. S., Bejerman, N., Johnson, K. N., and Dietzgen, R. G. 2016. Cytorhabdovirus P3 genes encode 30K-like cell-to-cell movement proteins. Virology 489:20-33.

Medina-Salguero, A. X., Cornejo-Franco, J. F., Grinstead, S., Mollov, D., Mowery, J. D., Flores, F., and Quito-Avila, D. F. 2019. Sequencing, genome analysis and prevalence of a cytorhabdovirus discovered in Carica papaya. PLoS One 14:e0215798.

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

Mushegian, A. R., and Elena, S. F. 2015. Evolution of plant virus movement proteins from the 30K superfamily and of their homologs integrated in plant genomes. Virology 476:304-315.

Navarro, B., Minutolo, M., De Stradis, A., Palmisano, F., Alioto, D., and Di Serio, F. 2018a. The first phlebo-like virus infecting plants: a case study on the adaptation of negative-stranded RNA viruses to new hosts. Mol. Plant Pathol. 19:1075-1089.

Navarro, B., Zicca, S., Minutolo, M., Saponari, M., Alioto, D., and Di Serio, F. 2018b. A negative-stranded RNA virus infecting citrus trees: The second member of a new genus within the order *Bunyavirales*. Front. Microbiol. 9: 2340.

Navarro, J. A., Sanchez-Navarro, J. A., and Pallas, V. 2019. Key checkpoints in the movement of plant viruses through the host. Adv. Virus Res. 104:1-64.

Navas-Castillo, J., Fiallo-Olivé, E., and Sánchez-Campos, S. 2011. Emerging virus diseases transmitted by whiteflies. Annu. Rev. Phytopathol. 49: 219-248.

Ooi, H. S., Kwo, C. Y., Wildpaner, M., Sirota, F. L., Eisenhaber, B., Maurer-Stroh, S., Wong, W. C., Schleiffer, A., Eisenhaber, F., and Schneider, G. 2009. ANNIE: Integrated de novo protein sequence annotation. Nucleic Acids Res. 37:W435-W440.

Pei, J., Kim, B. H., and Grishin, N. V. 2008. PROMALS3D: A tool for multiple protein sequence and structure alignments. Nucleic Acids Res. 36: 2295-2300.

Roche, S., Albertini, A., Lepault, J., Bressanelli, S., and Gaudin, Y. 2008. Structures of vesicular stomatitis virus glycoprotein: Membrane fusion revisited. Cell. Mol. Life Sci. 65:1716-1728.

Roossinck, M. J. 2008. Plant Virus Evolution. Springer-Verlag, Berlin, Heidelberg.

Rott, M. E., Kesanakurti, P., Berwarth, C., Rast, H., Boyes, I., Phelan, J., and Jelkmann, W. 2018. Discovery of negative-sense RNA viruses in trees infected with apple rubbery wood disease by next-generation sequencing. Plant Dis. 102:1254-1263.

Sasaya, T., Ishikawa, K., and Koganezawa, H. 2002. The nucleotide sequence of RNA1 of Lettuce big-vein virus, genus *Varicosavirus*, reveals its relation to nonsegmented negative-strand RNA viruses. Virology 297:289-297.

Shi, M., Lin, X. D., Tian, J. H., Chen, L. J., Chen, X., Li, C. X., Qin, X. C., Li, J., Cao, J. P., Eden, J. S., Buchmann, J., Wang, W., Xu, J., Holmes, E. C., and Zhang, Y. Z. 2016. Redefining the invertebrate RNA virosphere. Nature 540:539-543.

Steinhauer, D. A., Domingo, E., and Holland, J. J. 1992. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. Gene 122:281-288.

Sun, K., Zhou, X., Lin, W., Zhou, X., Jackson, A. O., and Li, Z. 2018. Matrix–glycoprotein interactions required for budding of a plant nucleorhabdovirus and induction of inner nuclear membrane invagination. Mol. Plant Pathol. 19:2288-2301.

Taliansky, M., Torrance, L., and Kalinina, N. O. 2008. Role of plant virus movement proteins. Pages 33-54 in: Plant Virology Protocols: From Viral Sequence to Protein Function. G. D. Foster, I. E. Johansen, Y. G. Hong, and P. D. Nagy, eds. Humana Press, Totowa, NJ.

Tamai, A., and Meshi, T. 2001. Tobamoviral movement protein transiently expressed in a single epidermal cell functions beyond multiple plasmodesmata and spreads multicellularly in an infection-coupled manner. Mol. Plant-Microbe. Interact. 14:126-134.

Terribilini, M., Sander, J. D., Lee, J. H., Zaback, P., Jernigan, R. L., Honavar, V., and Dobbs, D. 2007. RNABindR: A server for analyzing and predicting RNA-binding sites in proteins. Nucleic Acids Res. 35:W578-W584.

Wagner, R. R. 1987. Rhabdovirus biology and infection. Pages 9-74 in: The Rhabdoviruses. R. R. Wagner, ed. Springer-Verlag, Berlin, Heidelberg.

Walker, P. J., Blasdell, K. R., Calisher, C. H., Dietzgen, R. G., Kondo, H., Kurath, G., Longdon, B., Stone, D. M., Tesh, R. B., Tordo, N., Vasilakis, N., and Whitfield, A. E., and ICTV Report Consortium. 2018. ICTV virus taxonomy profile: *Rhabdoviridae*. J. Gen. Virol. 99:447-448.

Walker, P. J., Dietzgen, R. G., Joubert, D. A., and Blasdell, K. R. 2011. Rhabdovirus accessory genes. Virus Res. 162:110-125.

Walker, P. J., Firth, C., Widen, S. G., Blasdell, K. R., Guzman, H., Wood, T. G., Paradkar, P. N., Holmes, E. C., Tesh, R. B., and Vasilakis, N. 2015. Evolution of genome size and complexity in the *Rhabdoviridae*. PLoS Pathog 11:e1004664.

Wang, Q., Ma, X., Qian, S., Zhou, X., Sun, K., Chen, X., Zhou, X., Jackson, A. O., and Li, Z. 2015. Rescue of a plant negative-strand RNA virus from cloned cDNA: Insights into enveloped plant virus movement and morphogenesis. PLoS Pathog 11:e1005223.

Whelan, S., Barr, J., and Wertz, G. 2004. Transcription and replication of nonsegmented negative-strand RNA viruses. Pages 61-119 in: Biology of Negative Strand RNA Viruses: The Power of Reverse Genetics. Y. Kawaoka, ed. Springer-Verlag, Berlin, Heidelberg.

Wu, Q., Ding, S. W., Zhang, Y., and Zhu, S. 2015. Identification of viruses and viroids by next-generation sequencing and homology-dependent and homology-independent algorithms. Annu. Rev. Phytopathol. 53: 425-444.

Xin, M., Cao, M., Liu, W., Ren, Y., Zhou, X., and Wang, X. 2017. Two negative-strand RNA viruses identified in watermelon represent a novel clade in the order *Bunyavirales*. Front. Microbiol. 8:1514.

Yang, Y., Ding, B., Baulcombe, D. C., and Verchot, J. 2000. Cell-to-cell movement of the 25K protein of Potato virus X is regulated by three other viral proteins. Mol. Plant-Microbe. Interact. 13:599-605.

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.

Zhou, X., Lin, W., Sun, K., Wang, S., Zhou, X., Jackson, A. O., and Li, Z. 2019. Specificity of plant rhabdovirus cell-to-cell movement. J. Virol. 93:e00296-e19.