

# Viromics Unveils Extraordinary Genetic Diversity of the Family *Closteroviridae* in Wild Citrus

**Qiyan Liu**

Chinese Academy of Agricultural Sciences Citrus Research Institute

**Song Zhang**

Chinese Academy of Agricultural Sciences Citrus Research Institute

**Shiqiang Mei**

Chinese Academy of Agricultural Sciences Citrus Research Institute

**Yan Zhou**

Chinese Academy of Agricultural Sciences Citrus Research Institute

**Changyong Zhou**

Chinese Academy of Agricultural Sciences Citrus Research Institute

**Mengji Cao** (✉ [caomengji@cric.cn](mailto:caomengji@cric.cn))

Chinese Academy of Agricultural Sciences Citrus Research Institute <https://orcid.org/0000-0003-0396-262X>

---

## Research

**Keywords:** Wild citrus, Closteroviridae, Citrus tristeza virus, Modular evolution, Recombination, Genomic asymmetrically divergent

**Posted Date:** January 11th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-142743/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

## Abstract

## Background

Our knowledge of citrus viruses is largely skewed toward virus pathology in cultivated orchards; comparatively little is known about the virus diversity and ecology in natural ecosystems. The metatranscriptomics approach were used to analyze wild citrus from the Ailao Mountain region in China, a region embracing huge biodiversity in an area where citrus species originated. We investigated the evolutionary history of citrus viruses to identify the factors facilitating outbreaks of emerging viruses.

## Results

An independent wild citrus habitat were identified where citrus tristeza virus (CTV) exists along with three novel monopartite, graft-transmissible members of the family *Closteroviridae* differing from the managed agricultural systems that harbor other diverse citrus viruses. The viruses analyzed in 16 sequencing libraries of 32 citrus samples comprised five CTV genotypes, two inferred new ampelovirus species, and a monopartite ancestor type of the multipartite crinivirus that may represent a new taxon. Acquisition of heat shock protein genes (all viruses) and coat protein duplication (most), exchange of other exogenous genes (some), and alteration of expression strategy (rare) may reflect the significant historical nodes through which a *Closteroviridae* ancestor as the relic species has flourished. During viral expansion, genetic mutation coupled with recombination may be the essential impetus to update best-adapted sequences, as in the cases of our CTV and one of the ampeloviruses.

## Conclusions

This research represents the first time of virus diversity in wild citrus. The findings deepen our knowledge of *Closteroviridae* modular evolution and diversification. Given the potential emergence of similar novel viruses as pathogens, the need for surveillance of their pathogenic and epidemiological characteristics and the presence of other novel citrus viruses among wild citrus is of utmost priority for global citrus production.

## Introduction

Characterization of plant viruses has generally focused on those producing symptoms on cultivated plants. Through the use of viral metagenomics (viromics) techniques, many new viruses have been discovered, greatly enlarging our understanding of the ecology and evolution of plant viruses in nature [1–3].

Members of the family *Closteroviridae* have long, helical, filamentous particles with single-stranded, positive-sense RNA genomes. They infect a wide range of agriculturally important crops with severe

economic damage [4]. The family includes four genera: *Closterovirus*, with aphid-transmitted species; *Ampelovirus*, with mealybug/soft-scale-transmitted members; *Crinivirus*, whose members are whitefly-transmitted; and *Velarivirus*, without a known vector. Viruses of the genus *Crinivirus* have bi- or tripartite genomes that are separately encapsidated, whereas viruses belonging to *Closterovirus*, *Ampelovirus*, and *Velarivirus* have monopartite genomes [4].

The viral genomes of closterovirids contain two characteristic genomic blocks. The first is the replication-related gene block, which includes the methyltransferase (Mtr) and helicase (Hel) domains encoded by open reading frame (ORF) 1a and the RNA-dependent RNA polymerase (RdRP) domain encoded by ORF1b. The second gene block encodes proteins that function in virion assembly and transport, namely a ~ 6 kDa small hydrophobic protein (p6), a heat shock protein homolog (HSP70h), a ~ 60 kDa protein (p60), the major capsid protein (CP), and a duplicated version of the latter (the minor capsid protein, CPm). Apart from the two conserved gene blocks, additional nonconserved genes encode proteins that vary in number, arrangement, function, and origin within and between genera, with most having no detectable similarity with other viruses [4, 5]. Some of those nonconserved proteins have important functions, including suppression of plant RNA silencing, RNA repair, and host range broadening [6], whereas the functions of others are still unknown.

Recombination is one of the primary evolutionary forces affecting the *Closteroviridae*, acting by the capture of foreign genes (e.g., those encoding protease and HSP70h domains) or by duplication of intragenomic sequences (e.g., CPm). These events have shaped the extraordinary molecular and biological diversity of the family [7]. A scheme has been proposed for the genome evolution of the family *Closteroviridae* whereby a common ancestor with the RdRP, a p6-like movement protein and a single CP acquired other modules by recombination [8]. An aphid-borne monopartite virus, mint vein banding-associated virus (MVBaV), shares sequence identity with members of all genera of the family, making it a candidate ancestor [9]. However, there is still a lack of understanding of how multipartite closterovirids evolved.

The genus *Citrus* L. is considered to have originated from the southeast piedmonts of the Himalayas, a region that includes eastern Assam, northern Myanmar and western Yunnan [10]. The aphid-borne virus citrus tristeza virus (CTV), the only known citrus-infecting closterovirid, causes quick decline, seedling yellow, stem pitting, and slow deterioration that have been associated with long-term chronic losses in commercial citrus production [11]. This highly destructive viral pathogen has resulted in numerous disease outbreaks in most citrus-growing regions of the world [12].

As a part of the Himalayan biodiversity hotspot, the mountainous Yunnan Province is one of the botanically most diverse regions in the world [13]. Its peaks and deep valleys are potential barriers to species spread, and its climatic, geologic and topographic diversity provides an ideal environment for the formation and development of the flora [14]. The Ailao Mountain located in central and southern Yunnan Province and ranging in elevation from 100 to 3,000 m above sea level, preserves the largest, most continuous natural subtropical evergreen-broadleaf forest in China [15].

Due to the limits of technology, early research in wild citrus plant pathology only analyzed molecular characteristics of incomplete genomic sequences of CTV [16, 17]. Since metagenomics has been applied to the discovery and identification of novel viruses, our knowledge of the genome and epidemiology of citrus-infecting viruses has expanded tremendously. However, research has mainly focused on viral mechanisms and host pathology. While more than 30 citrus virus and viroid diseases have been identified and characterized across the world [18], little is known about virus diversity and ecology in wild ecosystems. In particular, because we know that grapevine, cherry, and blackberry can be infected by more than one closterovirus [4], there is a need to better understand the diversity of their citrus counterparts. Here we investigated the diversity of closterovirids in wild citrus trees in the Ailao Mountain region using rRNA-depleted transcriptomics. Three novel citrus-infecting closterovirids and different CTV isolates were identified and characterized, providing new insights into the virus ecology in wild citrus plants and the evolution and genetic diversity of the family *Closteroviridae*.

## Results

### Discovery of closterovirids in wild citrus

A total of 32 wild citrus samples were collected during 2018 and 2019 in Yunnan province, China (Fig. 1), and 16 libraries were prepared (Additional file 1: Table S1) and subsequently sequenced. *De novo* assembled contigs were classified as belonging to CTV and three novel closterovirids. The presence of CTV in samples was verified using contigs-specific primers (Additional file 2: Table S2), and the genomes of the three novel closterovirids were then verified and sequenced completely using RT-PCR and RACE. One of the viral contigs was identified as originating from a novel closterovirid that shared similarities with different members of the family, but it did not correspond to any established genus. The novel closterovirid was tentatively named citrus virus B (CiVB). The two other novel viruses, inferred to belong to the genus *Ampelovirus*, were tentatively named citrus-associated ampelovirus 1 (CaAV-1) and citrus-associated ampelovirus 2 (CaAV-2). The number of viral contig reads and average coverage for each sequencing library can be found in Additional file 3: Table S3. In all sequenced samples, the average coverage for CaAV-1 was much higher than for all other citrus closterovirids.

The grafted citrus trees were screened by RT-PCR with specific primers for CiVB, CaAV-1, and CaAV-2 five months after grafting. Samples YN1-2 and YNL7-2 were positive for CiVB (2/17). Most grafted plants were positive for CaAV-1 (9/17), and sample YN12-1 was positive for CaAV-2 (1/17). The results showed that all three novel viruses had the capacity to be graft-transmitted to citrus trees.

The single contigs from *de novo* assembly corresponded to nearly the complete viral genomes, and RT-PCR and Sanger sequencing using overlapping primers covering the entire genomes did not reveal any breakpoints. Based on these results, we inferred that CiVB, CaAV-1, and CaAV-2 are monopartite rather than multipartite viruses. Isolates ZLV7-2 of CiVB, YN1-62 of CaAV-1, and YN12-1 of CaAV-2 were examined for genomic features of three novel citrus closterovirids. Isolate YN1-6 of CTV was used for comparative purposes with these new citrus closterovirids.

# Intraspecies of CTV and CaAV-1 in wild citrus plants

To better understand the intraspecific lineage classification in such a complex context, we performed a phylogenetic neighbor-net reconstruction using the complete genome sequences of different CaAV-1 and CTV isolates. The results showed that 18 CaAV-1 isolates were segregated into two different genotype classes (Fig. 2A). Twelve CTV isolates were assigned to the VT, T36, T3, T68, or RB genotype classes that covered almost all ancestors of the CTV genotypes detected in commercial citrus crops (Fig. 2B). Furthermore, the mapping reads of CTV isolates were preferentially distributed at the 3'-terminal region of the genomic RNA (Fig. 3A), similar to the observation of those viral small RNAs of 21–24 nt of CTV [19].

## Genomic organization of newly-identified closterovirids

The near-complete RNA genome of CiVB (GenBank MW365399) was determined using six RNA sequencing libraries, with some of its potentially encoded proteins sharing the highest amino acid sequence identity with those of different closterovirids. Whole-genome nucleotide sequence identity among the six CiVB isolates examined in this study was from 98.2–99.7%. The CiVB genome comprises 16,957 nucleotides (nt) with 13 ORFs (Fig. 3B). The ORF1a putatively encodes polyprotein with three conserved domains: a Mtr domain (I03298), a major facilitator superfamily domain (MFS, cl28910), and a Hel domain (pfam01443). Two TMHs are connected by hydrophilic loops and cover the MFS domain [20]. ORF1b putatively encodes RdRP with RdRP-2 (pfam00978) and RT-like superfamily (cl02808) domains. ORF2 encodes p8a with a predicted signal peptide that has a TMH. ORF3, ORF5 and ORF6 encode HSP70h, HSP90h, and CP respectively. Small ORF4 and six ORFs downstream of the CP (ORFs 7–12) encode the putative proteins p8b, p14, p9a, p9b, p34a, p11, and p34b. BLAST and CD-based searches did not reveal any statistically significant identity for these proteins in the web databases.

From 12 samples, we identified 18 isolates of a novel monopartite ampelovirus, CaAV-1 (GenBank MW365401), with 86.7–99.3% genomic nucleotide identity. The complete genomic sequence of the CaAV-1-YN1-62 consists of 16,975 nt with 13 putative ORFs (Fig. 3C). ORF1a encodes a putative polyprotein containing the signatures of Mtr (cl03298), a Hel (cl26263), and DNA polymerase III subunits gamma and tau (cl35530). ORF1b putatively encodes RdRP with the same conserved domains as CiVB. ORF2 encodes p12 with two TMHs. ORF3 encodes p24 with a thaumatin-like proteins domain (cd09218). ORF4 encodes p6 that contains a TMH. ORF5, ORF6, ORF7 and ORF8 encode HSP70h, HSP90h, CP, and CPm, respectively. ORF9 and ORF12 encode p31 and p27b that specific hit for divalent metal cations (Fe/Co/Zn/Cd) transport protein (FieF superfamily, cl30791) and DEDDy 3'-5' exonuclease of WRN class (WRN\_exo, cd06141) domains, respectively. The same domains were also identified in pistachio ampelovirus A (PAVA, MF198462) [21]. ORF10 and ORF11 encode p27a and p13, neither of which share statistically significant identity with other viral proteins in the database.

From sample YN12-1, one large contig of 13,488 nt was recognized as originating from another novel monopartite ampelovirus, CaAV-2 (GenBank MW365402). The complete sequence of the CaAV-2 genome consists of 13,407 nt and potentially encompassed 10 ORFs (Fig. 3D). ORF1 encodes polyprotein with Mtr (cl03298), basic leucine zipper (bZIP, cl21462), Hel (cl26263) and RdRP (pfam00978) domains.

Intriguingly, CaAV-2 expresses RdRP within ORF1 together with other replication-associated genes; this manner differs from all known closterovirids, which express RdRP gene in an ORF1b via a + 1 ribosomal frameshift strategy [4]. ORF2 encodes p6a that has a TMH domain. ORF3, ORF4, ORF5 and ORF encode a series of conserved proteins, namely HSP70h, HSP90h, CP, and CPm, respectively. The remaining three ORFs downstream of the CPm (ORFs 7–9), encode the putative proteins p19, p6b, and p21, respectively, with no shared identity in the databases. The p6b has a TMH domain.

## Phylogenetic analysis of CiVB, CaAV-1 and CaAV-2

The maximum likelihood (ML) phylogenetic trees for the RdRP, HSP70h, and CP genes, inferred using the corresponding closterovirid homologs, are presented in Fig. 4 (as a tanglegram) and Additional file 4: Figure S1. The major clades in the tanglegram were consistent with the ICTV classified genera in the family, while topological incongruities were observed in the phylogenies of CP and RdRP proteins for some of the clades. In particular, CiVB showed distinct evolutionary histories for CP and RdRP. More specifically, CiVB formed a separate group in all phylogenetic trees, clustering with the genus *Ampelovirus* in the RdRP phylogeny and with the genus *Velarivirus* and multipartite *Crinivirus* in the CP (Fig. 4) and the HSP70h phylogenies (Additional file 4: Figure S1). The amino acid sequences of the polyprotein, RdRP, HSP70h, and CP differed by more than 50% between CiVB and other closterovirids (Additional file 5: Figure S2). CiVB was thus inferred as a novel member of the family, with its situation resembling that of the monopartite MVBaV. The latter could not be assigned to any genus of the family, but it is transmitted by aphids, the typical vectors of members of the genus *Closterovirus* [9]. The phylogenetic trees here inferred for the RdRP, HSP70h, and CP proteins indicate that MVBaV forms a mono branch that is more closely related to ampeloviruses than to other genera. This result is distinct from that of the initial report [9] but consistent with those of more recent reports [4, 22].

The tanglegram of RdRP and CP (Fig. 4) and the phylogenetic tree of HSP70h (Additional file 6: Figure S3) showed that CaAV-1 and CaAV-2 cluster within the genus *Ampelovirus*. More specifically, CaAV-1 forms a subcluster with PAVA for both genes. However, CaAV-2 clusters with CaAV-1 and PAVA in the RdRP tree, with grapevine leafroll-associated virus 13 (GLRaV-13) in the CP tree, and with grapevine leafroll-associated virus 1 (GLRaV-1) in the Hsp70h tree. In addition, the amino acid sequence identities of the RdRP, HSP70h and CP proteins shared between CaAV-1, CaAV-2 and the ampeloviruses differ by more than 25%. Thus, these two viruses might represent novel species in the genus *Ampelovirus*.

## Frequent lateral gene transfer in newly identified closterovirids

The MFS domain, which might be related to secondary transport and appears between the Mtr and He domains, is the only identified lateral gene transfer in CiVB. CiVB ORF4 and the six ORFs downstream of the CP (ORFs 7–12) did not show any statistically significant identity among databases. A search did not identify the *CPm* gene in CiVB, perhaps because it was absent in the ancestral closterovirids from which CiVB evolved or, if present, it lacks enough similarity with the CPm of extant closterovirids. In either case, further analyses are needed to resolve this issue.

Using local blast alignment and network analysis, the external genes encoding thaumatin-like proteins were identified at different genomic regions of six other closterovirids that belong to different taxa, thus highlighting the transferability of this kind of protein in *Closteroviridae* evolution (Additional file 6: Figure S3). Besides thaumatin-like proteins, CaAV-1 and PAVA shared two cellular origin genes appearing in the same locations of the CaAV-1 and PAVA genomes. The cellular genes coding for a divalent metal cation transporter and a 3'-5' exonuclease domain correspond to p31 and p27b in CaAV-1, and p31, and p25 in PAVA, respectively. Furthermore, both CaAV-1 and PAVA lacked the leader protease domain (L-Pro) that led to them a relatively small polyprotein.

The bZIP domain has only been found in the polyprotein of CaAV-2 among members of the family *Closteroviridae*. This domain has been widely reported in both animals and plants as an enhancer-type transcription factor that plays an important role in the expression and regulation of genes [23].

## Unique mode of RdRP expression in CaAV-2

In the genome of all closterovirids sequenced to date, the first conserved gene block encodes replication-associated genes that are expressed within ORFs 1a and 1b via proteolytic processing and translational frameshifting. Specifically, ORF1 uses the start codon of ORF1a to initiate translation of the N terminal moiety of the polyprotein, and subsequently uses a + 1 frameshifting strategy to recognize ORF1b to eventually express a new polyprotein containing the RdRP domain. This polyprotein is further processed by proteolytic activity. However, the RdRP gene of CaAV-2 is expressed together with other replication-associated genes within the ORF1 and has no internal start codon, so CaAV-2 cannot initiate + 1 ribosomal frameshifting. While clearly within the clade of ampeloviruses based on sequence similarity and phylogenetic analysis, CaAV-2 seems to employ a unique strategy to express ORF1 that distinguishes it from previously studied ampeloviruses and other closterovirids.

## How recombination and variation affect asymmetrically divergent genomes

While the phylogenies indicate discordant relationships for CiVB and CaAV-2, we failed to identify any reliable recombination event for either virus. The full-length genomes of six CiVB isolates also did not reveal any recombination signals (data not shown).

The reticulation of the neighbor-net analysis indicated likely recombination both in CTV and CaAV-1, and the phylogenetic discordant topology further supported the recombination signals (Additional file 7 and 8: Figure S4 and S5). The replication-related module tree was in accord with the neighbor-net analysis, in which the same genotypes clustered together, while the HSP70h and CP trees are largely different for CTV and CaAV-1.

Recombination in CaAV-1 seemed to display a bias, with the highest frequency occurring downstream of the replication-associated module, and the recombinant minor parents mainly coming from genotype A. Of 12 recombinants detected among 17 CaAV-1 isolates, eight had large recombined fragments (> 5000 nt) in the 3'-half portion of the genome (Fig. 5A, D). Specifically, we identified two special examples in

samples ZLV8-2 and ZLV8-7, each coinfecting with two CaAV-1 genotypes. In these two samples, two CaAV-1 isolates had diverse 5'-half genome sequences but shared similar 3'-half genome sequences. As shown in Additional file 9: Figure S6, different CaAV-1 isolates from the same samples (ZLV8-2 and ZLV8-7) shared the same reads (in yellow) at the 3'-half genome portion. One difference is that they had different 5' recombination breakpoints: in HSP90h of ZLV8-2 and in the intergenic region before p24 of ZLV8-7. These two recombination breakpoints are also the most frequent 5'-breakpoints for CaAV-1 recombination.

Because recombination was by far more frequent in the 3'-half portion of the genome, we assessed the sequence identity between two CaAV-1 genotypes and found an inverse association with the recombination frequency (Fig. 5). Two CaAV-1 genotypes diverged asymmetrically throughout the genome: they shared high sequence identity (> 94%) throughout the 3'-half of the genome (downstream of the p24), diverging more in the 5'-half where the sequence identity was 76%. This asymmetrical divergence has been reported previously for CTV [24]. The CTV isolates identified here share more than 88% sequence identity in the 3'-half portion of the genome (downstream of p33) but only 72% sequence identity in the 5'-half counterpart (Fig. 5). Because "VT" is the most frequent CTV genotype in the wild citrus plants examined (Fig. 2B), it is also the most frequent recombinant material as a minor parent (Fig. 5). The intraspecific sequence identity of CTV tends to be relatively lower when the recombination frequency is higher, but this trend is not as obvious in CTV as in CaAV-1.

Regarding whether single nucleotide polymorphisms (SNPs) contribute to the asymmetrical divergence in the genomes of CTV and CaAV-1, the number of SNPs per 1000 nt in each ORF was not consistent with the trends in sequence variation (Fig. 6). The Spearman correlation analysis of the association between the lowest nucleotide sequence identity and the number of SNPs detected no correlation for either CTV ( $r = 0.198$ ) or CaAV-1 ( $r = 0.143$ ). Interestingly, we found that the average number of SNPs in the 3'-half portion of the genome was higher than in the 5'-half counterpart: 13.57 for 5'-half and 19.57 for 3'-half moieties in CTV, and 4.36 for 5'-half and 7.67 for 3'-half moieties in CaAV-1.

## Discussion

To investigate the diversity and prevalence of viruses in natural ecosystems, we sampled 32 wild citrus trees in the Ailao Mountains, an area that harbors a very diverse flora [15] and is located within the region where citrus plants originated [10]. Some viruses may have coevolved with these wild hosts over a long period of time, making the latter natural reservoirs for viruses that could cause new or re-emerging diseases.

The four closterovirids that we identified and characterized from these wild hosts comprise the well-known CTV and three novel members of the family *Closteroviridae*: CiVB, CaAV-1, and CaAV-2. We found five genotypes of CTV that had possibly dispersed to cultivated citrus areas. Alternatively, a CTV prototype may have spread from its geographical origin to other areas via appropriate vectors, then

diversified in various environments into multiple subtypes, some of which returned to the area of origin through vector(s).

Although the prevalence and potential transmission of the novel closterovirids in other wild citrus plants is largely unknown, the present data suggest that CiVB, CaAV-1, and CaAV-2 are not as widespread as CTV. The significantly higher SNP levels in CTV than in CaAV-1 may also have led to higher evolutionary rates in CTV, a view further supported by the higher CTV intraspecific diversity found in wild citrus. Higher evolutionary rates and intraspecific diversity would result in quicker adaptation, contributing to the global distribution of CTV in citrus-growing areas.

Generally speaking, the viruses newly identified in this study show that the diversity of the family *Closteroviridae* exceeds what was previously thought, since the viruses reported here: i) are highly divergent compared to known members of the family, ii) have an extraordinary ability for recombination with different sources and, iii) vary widely in their genome organization and expression strategy.

Recombination should affect the inferred phylogeny of CiVB, leading to the inconsistencies seen here. Our effort to determine putative inter- and intraspecific recombination events of CiVB with the RDP4 program failed to identify any reliable recombination signal. Given that other members of the family *Closteroviridae* may remain to be discovered, recombination events should not have been undetectable, or CiVB might be an ancestral-type virus like MVBaV [9]. In contrast to the phylogeny of MVBaV showing a closer relationship to ampeloviruses in the HSP70h and CP trees [4, 22], the HSP70h and CP phylogenies for CiVB revealed closer proximity to the bipartite criniviruses. Since CiVB is monopartite, it represents an evolutionary bridge between mono- and multipartite closterovirids. Furthermore, the characteristics of CiVB show the modular mode through which evolution operates in the family *Closteroviridae* (and perhaps in all RNA viruses): different genes evolving independently but within limits imposing conserved gene combinations and orders.

Notably, the thaumatin-like protein, which exists widely in fungi, plants, and animals and is related to host defense and developmental processes, has now been identified in seven closterovirids and is closest to the plant-derived thaumatin-like protein. The protein was only reported in these closterovirids, and it has a random genome position, unlike HSP70h that exists in all closterovirids and has a fixed genome position. However, both are encoded only by those closterovirids and not by any other viruses. Except for the thaumatin-like protein, CaAV-1 shares another two external genes with PAVA encoding two proteins containing a domain of the FieF superfamily, and a DEDDy 3'-5' exonuclease domain that have not been identified in other closterovirids.

It appears that the closterovirids have had a superior capacity to integrate exogenous genes during the long virus-plant co-evolutionary history. The members of the family *Closteroviridae* might share a common ancestor in which extra genes were integrated and subsequently underwent codivergence and cross-species transmission events. Whereby some genes like HSP70h have been subsequently retained in all extant closterovirids. Then, some genes have been under negative selection and are likely to remain in few species, such as thaumatin-like protein, L-Pro domains, and other unidentified proteins. Furthermore,

some viruses may have been under similar stresses and thereby driven to convergent evolution via horizontally acquiring the same external genes, like CaAV-1 and PAVA. The other genes with no detectable similarity in closterovirids may diverge independently, facilitating infection of different hosts or transmission by different insect vectors.

In all CTV and CaAV-1 genomes sequenced to date, the 3'-half portion of the genome is well conserved, in contrast with the 5'-half counterpart. Large fragment recombinations in the 3'-half portion of the genome are prevalent in CaAV-1. The recombination frequency in CaAV-1 showed an inverse correlation with the level of asymmetric divergence in its genome, while such a correlation was not significant in CTV. However, new CTV phylogeny lineages generated by recombination, especially of the large fragment in the 3'-half moiety of the genome, have emerged in the long history of CTV evolution [24–26]. Moreover, the tree incongruities were only present in the 3'-half portions of the CaAV-1 genomes. That indicates using a phylogenetic tree generated from complete genome sequences, or HSP70h, or CP of CTV and CaAV-1 for genotyping is not accurate. Because the translation pattern is conserved in the family *Closteroviridae*, CaAV-1 is likely to express a series of 3'-coterminal genes via sgRNAs, as in CTV. Thus, CTV and CaAV-1 might have used these sgRNAs for template exchanges for recombination.

The genes mapping at the 3'-half of the genome are responsible for a variety of functions important to their survival and reproduction. Although the genus *Citrus* comprises diverse species [10], CTV can replicate in any and all, but the viral dissemination depends on the citrus hosts [27, 28]. Sequence changes in the 3'-half portion of the genome might therefore disturb these functions and others, including vector transmission in different host species. Under selection pressure, recombination and single nucleotide variation may lead viruses to retain the more efficient fragment and to convergence of the 3'-half genome moiety in different citrus hosts.

Each wild host citrus tree was almost always coinfecting with different genotypes of CaAV-1. Such a pattern might be a consequence of superinfection exclusion (SIE) of CaAV-1 not functioning, since its different genotypes showed relatively large variation. The case is similar to what happens in different CTV genotypes coinfecting citrus plants, in which primary CTV infection failing to affect secondary infection of heterologous genotypes has been reported [29]. Coinfection of one host plant by different virus genotypes might result from the same mechanism, since recombination between different genotypes is advantageous for evolution and adaptation of the virus population.

Why are there so many closterovirids in wild citrus species but not other common viruses? This may be because while CTV and perhaps the ancestor of other closterovirids are native to citrus plants, other viruses may have been transferred to cultivated citrus crops from other plants, an idea that should be further explored by examining wild citrus growing in the same areas as cultivated varieties. However, extant members of the family *Closteroviridae* are known to be transmitted in cultivated areas by a few specific insect vectors, whereas in a natural environment the spectrum of such vectors may be wider. Thus, in natural citrus habitats that have had no human intervention, the insect-borne closterovirids may infect wild citrus plants over a relatively wider area compared with other viruses. In addition, the Ailao

Mountains are to some extent isolated from the rest of the world, and perhaps other citrus viruses and viroids that originated in other areas have not arrived here.

Apart from CTV, the discovery of CiVB, CaAV-1, and CaAV-2 represents the first report of new members of the family *Closteroviridae* infecting citrus plants, suggesting that these plants are potential hosts for diverse as-yet-unknown closterovirids. From one isolated region of wild citrus, we found viruses that belong to the known genus *Ampelovirus* (CaAV-1 and CaAV-2, two putative new species), *Closterovirus* (CTV), and CiVB, which may represent a new genus in the family *Closteroviridae*. These findings should serve as a warning that new citrus-infecting closterovirids may exist and may pose new problems for citrus production. Further surveys and monitoring are warranted to advance our knowledge and protect the citrus industry.

Viruses are a two-edged sword, affecting agriculture and natural biodiversity as they have coevolved with plants and have exerted selection pressure on their hosts. From a beneficial perspective, a few closterovirus-derived genes and RNA interference vectors have been developed to control plant diseases [30, 31]. In particular, CTV may serve as a tool for citrus tree protection and improvement [11, 32]. With a large genome similar to that of CTV, the novel citrus viruses may prove to be suitable for construction of new vectors.

The new insights acquired regarding virus ecology in wild citrus trees and genetic diversity of the family *Closteroviridae* support the notion that despite their genetic variation, CTV and the novel citrus closterovirids reported here evolved from the same genetic pool and that they share an ancestral-type virus, CiVB. Present-day monocultures over large areas are particularly vulnerable to a wide range of viruses, while globalization has facilitated their transport to new areas. It is in this framework, where wild and cultivated citrus species have come into contact via humans and insect vectors, that CTV became prevalent and destructive for the citrus industry. The novel viruses might disperse, as did CTV, from citrus orchards near the Ailao Mountain region. Accordingly, we need to continue monitoring and studying these novel viruses and to examine other areas harboring wild citrus in order to provide a foundation for deploying protective strategies in the future. The lessons from the recent emergence of the covid-19 pandemic apparently following a similar scenario of virus spreading from ecological niches hosting virus species to populations that were never previously exposed need to be paid attention by those concerned with citrus tree health.

## Conclusions

Collectively, our research represents the first time of virus diversity in wild citrus. We identified citrus tristeza virus (CTV) and three novel closterovirids from wild citrus plants in the Ailao Mountain region while no other common viruses, indicating that the region may be an independent geographical point of origin specific for citrus-infecting closterovirids. Each of the novel viruses had interesting features: one showed direct evidence of modular evolution, another had a distinct expression strategy, and a third had asymmetrical genomic variation similar to that of CTV, these reflect how the family *Closteroviridae*

envolved. The populations of the four closterovirids found in the wild citrus were more divergent compared with those of cultivated citrus reflecting on the extraordinary recombination ability of closterovirids resulting in highly diverse virus genomes.

## Materials And Methods

### Sample collection

This study was based on the analysis of RNA libraries from wild citrus tree samples from the Ailao Mountain region (Fig. 1A-C) in central Yunnan Province, southwestern China, collected between July 2018 and August 2019. Some samples from the same location being pooled for library construction and transcriptome sequencing. Some samples were asymptomatic, while others expressed a wide range of symptoms (e.g., boat-shaped leaf curl, leafroll, chlorosis) (Fig. 1D–G) that could be caused by viruses or by other stresses such as fungi, insect pests, nutritional deficiencies, and aging. We therefore did not attempt a correlation analysis between viruses and symptoms; this awaits further study.

Citrus branch samples were grafted onto virus-free seedlings (Daidai sour orange [*Citrus aurantium* L. var. *cyathifera* Y. Tanaka] or Morocco sour orange [*C. aurantium* L.]) using two bark patches and two buds to preserve the virus materials. All grafted plants were grown in an insect-proof greenhouse. New flush leaves of the grafted seedlings were screened by RT-PCR with viral-specific primers five months after grafting to test whether the viruses were graft-transmissible.

### RNA library construction and sequencing

Total RNA was extracted using TRIzol LS (Invitrogen), and ribosomal RNA (rRNA) was removed using a Ribo-zero rRNA Removal Kit (Epicentre). The rRNA-depleted RNA libraries were constructed using a NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB) and subsequently sequenced on the X-ten platform (Illumina), generating paired-end reads of 150 bp. The information on library construction and corresponding geographic locations of samples for each sequencing library can be found in Additional file 1: Table S1.

### Sequence assembly and virus discovery

For each library, sequencing read data were processed to remove adapter sequences and low quality reads through the CLC Genomics Workbench 9.5 (Qiagen). The resulting clean data were mapped to citrus genomes [33, 34] to filter out the host reads. *De novo* assembly was then performed using the Trinity program (Broad Institute). To identify potential viruses, we compared the assembled contigs against the GenBank database using the BLASTx program.

### Confirmation and extension of virus genomes

Each potential viral genome was further examined by one-step reverse transcription-PCR (RT-PCR) using overlapping primers designed on the assembled sequences (Additional file 2: Table S2) to confirm the results. The terminal regions of the potential viral genomes were determined using rapid amplification of

cDNA ends (RACE) kit (GeneRacer, Invitrogen, USA). The resultant cDNA products were purified with a Biospin PCR Purification Kit (BioFlux) and cloned into the pGEM-T Easy Vector (TransGen Biotech), which was later transferred into DH5 $\alpha$  competent cells (Takara). At least five clones per amplicon were randomly selected and sequenced in both directions by the BGI Company (Beijing, China).

## Virus genome annotation

For each newly-identified viral genome, open reading frames (ORFs) were predicted using ORF finder (<https://www.ncbi.nlm.nih.gov/orffinder/>). The potential functions of the encoded proteins were inferred from the NCBI protein database (<https://www.ncbi.nlm.nih.gov/protein>) and conserved domain database (<https://www.ncbi.nlm.nih.gov/cdd/>) using BLAST searches and Batch CD-Search. Relationships of diverse proteins of new and known closterovirids were assessed using local blast within TBtools [35], and were then analyzed using Cytoscape [36, 37] with evalue < 1E-5. Transmembrane domains and signal sequences were predicted with SignalP and TMHMM programs on the DTU Health Tech prediction servers (<http://www.cbs.dtu.dk/services/>). Protein hydropathy was analyzed using ProtParam (<https://web.expasy.org/protparam/>). Pairwise sequence identities of viral genomes were calculated with the CLC Genomics Workbench based on MAFFT alignments.

## Inference of virus evolutionary relationships

The amino acid sequences of RdRP, HSP70h and CP of the newly identified viruses and CTV were aligned with the corresponding homologs of representative members of the family *Closteroviridae* using the MAFFT version 7 of the E-INS-I algorithm [38]. All alignments were further trimmed using TrimAl version 1.2 in the automated mode to remove ambiguously aligned regions [39]. The corresponding ML phylogenetic trees of RdRP, HSP70h and CP were inferred using IQ-TREE, version 1.6.12, under the LG + F + R5, LG + F + R5 and LG + F + R4 substitution models [40]. Topological support was assessed with the regular bootstrap method in IQ-TREE (1000 replicates). Bootstrap values were only displayed when greater than 50%. The phylogenetic trees were visualized using FigTree version 1.4.0. The phylogenetic tree match between RdRP and CP was estimated by the tanglegram algorithm in Dendroscope [41]. The intraspecific phylogenetic lineages for CaAV-1 and CTV were assessed using the Neighbor-Net approach in SplitsTree5 [42].

## Recombination analysis

Recombination was detected using both the RDP4 package [43] and by observing phylogenetic tree structural incongruities from different regions of the viral genome. The phylogenetic tree generated from the complete nucleotide and the amino acid sequences of the polyprotein (ORF1a and 1b), HSP70h and CP genes were used to compare the tree structural incongruities. Subsequently, the aligned nucleotide sequences were imported into RDP4 and analyzed using the RDP, GENECONV, Chimaera, MaxChi, BootScan, SiScan and 3Seq methods. The potential recombination events were considered significant only when supported by at least three methods with  $P$ -value <  $10^{-6}$ . The recombination frequency of CTV and CaAV-1 was calculated based on the proportion of recombinants at a specific site among the total

isolates, and the corresponding intraspecific sequence identities were accessed using SimPlot analysis [44]. SNPs in coding regions were detected using the CLC Genomics Workbench 9.5.

## Abbreviations

**Mtr:** Methyltransferase

**Hel:** Helicase

**ORF:** Open reading frame

**RdRP:** RNA-dependent RNA polymerase

**HSP70h:** Heat shock protein homolog

**CP:** Capsid protein

**CPm:** Minor capsid protein

**MVBaV:** Mint vein banding-associated virus

**CTV:** Citrus tristeza virus

**CiVB:** Citrus virus B

**CaAV-1:** Citrus-associated ampelovirus 1

**CaAV-2:** Citrus-associated ampelovirus 2

**MFS:** Facilitator superfamily domain

**PAVA:** Pistachio ampelovirus A

**GLRaV-13:** Grapevine leafroll-associated virus 13

**GLRaV-1:** Grapevine leafroll-associated virus 1

**ML:** Maximum likelihood

**L-Pro:** Leader protease domain

**SNPs:** Single nucleotide polymorphisms

## Declarations

## Ethics approval and consent to participate

Not applicable

## Consent for publication

Not applicable

## Availability of data and materials

The data that support the findings of this study are available from the corresponding author on reasonable request. Genbank accession numbers for the citrus closterovirids genomic sequences in this paper are as follows: MW365399–MW365403.

## Competing interests

The authors declare that they have no competing interests.

## Funding

This research was supported by the National Key R&D Program of China (2019YFD1001800), Fundamental Research Funds for the Central Universities (XDKJ2018AA002), National Natural Science Foundation of China (32072389), and Innovation Program for Chongqing's Overseas Returnees (cx2019013), and 111 Project (B18044).

## Authors' contributions

QL wrote the draft. SZ, MC and CZ provided critical review and editing of the paper. QL and YZ collected wild citrus samples. MC processed the viromics analysis. QL and MQ verified virus genome sequences. QL and SZ processed virus characterization. MC and CZ responsible for funding acquisition.

## Acknowledgments

We thank Lei Chen of the Industrial crop workstation of Xinping County, Yuxi, Yunnan for his help during sampling. We thank LetPub ([www.letpub.com](http://www letpub.com)) for its linguistic assistance during the preparation of this manuscript.

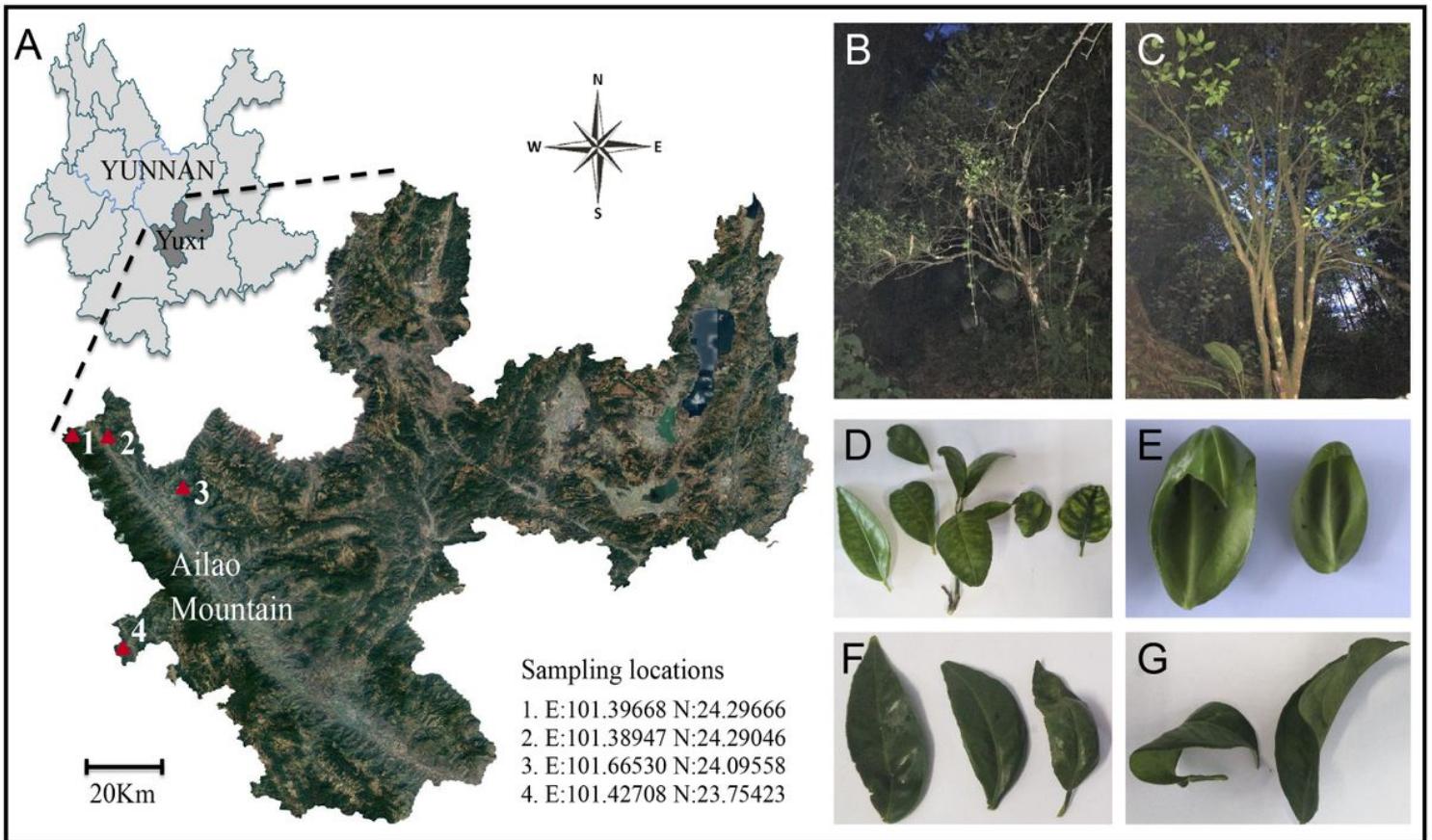
## References

1. Zhang Y-Z, Shi M, Holmes EC: Using metagenomics to characterize an expanding virosphere. *Cell* 2018, 172:1168-72.
2. Dolja VV, Krupovic M, Koonin EV: Deep roots and splendid boughs of the global plant virome. *Annual Review of Phytopathology* 2020, 58:23-53.
3. Roossinck MJ: Plant virus metagenomics: biodiversity and ecology. *Annual Review of Genetics* 2012, 46:359-69.
4. Fuchs M, Bar-Joseph M, Candresse T, Maree HJ, Martelli GP, Melzer MJ, et al: ICTV virus taxonomy profile: Closteroviridae. *Journal of General Virology* 2020, 101:364-5.
5. Dolja VV, Kreuze JF, Valkonen JPT: Comparative and functional genomics of closteroviruses. *Virus Research* 2006, 117:38-51.
6. Tatineni S, Robertson CJ, Garnsey SM, Dawson WO: A plant virus evolved by acquiring multiple nonconserved genes to extend its host range. *Proc Natl Acad Sci U S A* 2011, 108:17366-71.
7. Rubio L, Guerri J, Moreno P: Genetic variability and evolutionary dynamics of viruses of the family Closteroviridae. *Frontiers in Microbiology* 2013, 4.
8. Dolja VV, Karasev AV, Koonin EV: Molecular biology and evolution of closteroviruses: sophisticated build-up of large RNA genomes. *Annual Review of Phytopathology* 1994, 32:261-85.
9. Tzanetakis IE, Postman JD, Martin RR: A member of the Closteroviridae from mint with similarities to all three genera of the family. *Plant Disease* 2005, 89:654-8.
10. Wu GA, Terol J, Ibanez V, López-García A, Pérez-Román E, Borredá C, et al: Genomics of the origin and evolution of Citrus. *Nature* 2018, 554:311-6.
11. Dawson WO, Bar-Joseph M, Garnsey SM, Moreno P: Citrus tristeza virus: making an ally from an enemy. *Annual Review of Phytopathology* 2015, 53:137-55.
12. Moreno P, Ambrós S, Albiach-Martí MR, Guerri J, Peña L: Citrus tristeza virus: a pathogen that changed the course of the citrus industry. *Mol Plant Pathol* 2008, 9:251-68.
13. Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J: Biodiversity hotspots for conservation priorities. *Nature* 2000, 403:853-8.
14. Li XW: Floristic study of Yunnan Province. *Acta Botanica Yunnanica* 1985, 7:361-82.
15. Qiu X: Studies on the forest ecosystem in Ailao mountains, Yunnan, China. Kunming: Yunnan Sciences and Technology Press; 1998.
16. Yi L, Zhou C-y, Zhou Y, Li Z-a: Genetic evolution analysis on wild isolates of citrus tristeza virus originated in china based on coat protein genes sequences. *Agricultural Sciences in China* 2010, 9:1623-9.
17. Yi L, Zhou C-y: Phylogenetic analysis of citrus tristeza virus isolates of wild type citrus in China. *Journal of Integrative Agriculture* 2014, 13:2669-77.
18. Zhou C, da Graça JV, Freitas-Astúa J, Vidalakis G, Duran-Vila N, Lavagi I: Chapter 19 - citrus viruses and viroids. In *The Genus Citrus*. Edited by Talon M, Caruso M, Gmitter FG: Woodhead Publishing; 2020: 391-410

19. Ruiz-Ruiz S, Navarro B, Gisel A, Peña L, Navarro L, Moreno P, et al: Citrus tristeza virus infection induces the accumulation of viral small RNAs (21-24-nt) mapping preferentially at the 3'-terminal region of the genomic RNA and affects the host small RNA profile. *Plant Mol Biol* 2011, 75:607-19.
20. Yan N: Structural advances for the major facilitator superfamily (MFS) transporters. *Trends in Biochemical Sciences* 2013, 38:151-9.
21. Al Rwahnih M, Rowhani A, Westrick N, Stevens K, Diaz-Lara A, Trouillas FP, et al: Discovery of viruses and virus-like pathogens in pistachio using high-throughput sequencing. *Plant Disease* 2018, 102:1419-25.
22. Koloniuk I, Přibylová J, Fránová J, Špak J: Genomic characterization of *Malus domestica* virus A (MdoVA), a novel velarivirus infecting apple. *Archives of Virology* 2020, 165:479-82.
23. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al: CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res* 2020, 48:D265-D8.
24. Harper S: Citrus tristeza virus: evolution of complex and varied genotypic groups. *Frontiers in Microbiology* 2013, 4.
25. Benítez-Galeano MJ, Vallet T, Carrau L, Hernández-Rodríguez L, Bertalmío A, Rivas F, et al: Complete genome sequence of a novel recombinant citrus tristeza virus, a resistance-breaking isolate from Uruguay. *Microbiology Resource Announcements* 2018, 6.
26. Wang J, Zhou T, Shen P, Zhang S, Cao M, Zhou Y, et al: Complete genome sequences of two novel genotypes of Citrus tristeza virus infecting *Poncirus trifoliata* in China. *Journal of Plant Pathology* 2020, 102:903-7.
27. Albiach-Martí MR, Grosser JW, Gowda S, Mawassi M, Satyanarayana T, Garnsey SM, et al: Citrus tristeza virus replicates and forms infectious virions in protoplasts of resistant citrus relatives. *Molecular Breeding* 2004, 14:117-28.
28. Folimonova SY, Folimonov AS, Tatineni S, Dawson WO: Citrus tristeza virus: survival at the edge of the movement continuum. *Journal of virology* 2008, 82:6546-56.
29. Bergua M, Kang S-H, Folimonova SY: Understanding superinfection exclusion by complex populations of Citrus tristeza virus. *Virology* 2016, 499:331-9.
30. Kurth EG, Peremyslov VV, Prokhnevsky AI, Kasschau KD, Miller M, Carrington JC, et al: Virus-derived gene expression and RNA interference vector for grapevine. *Journal of Virology* 2012, 86:6002.
31. Dolja V, Koonin E: The closterovirus-derived gene expression and RNA interference vectors as tools for research and plant biotechnology. *Frontiers in Microbiology* 2013, 4.
32. Folimonova SY: Citrus tristeza virus: A large RNA virus with complex biology turned into a valuable tool for crop protection. *PLOS Pathogens* 2020, 16.
33. Wu GA, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, et al: Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nature Biotechnology* 2014, 32:656-62.

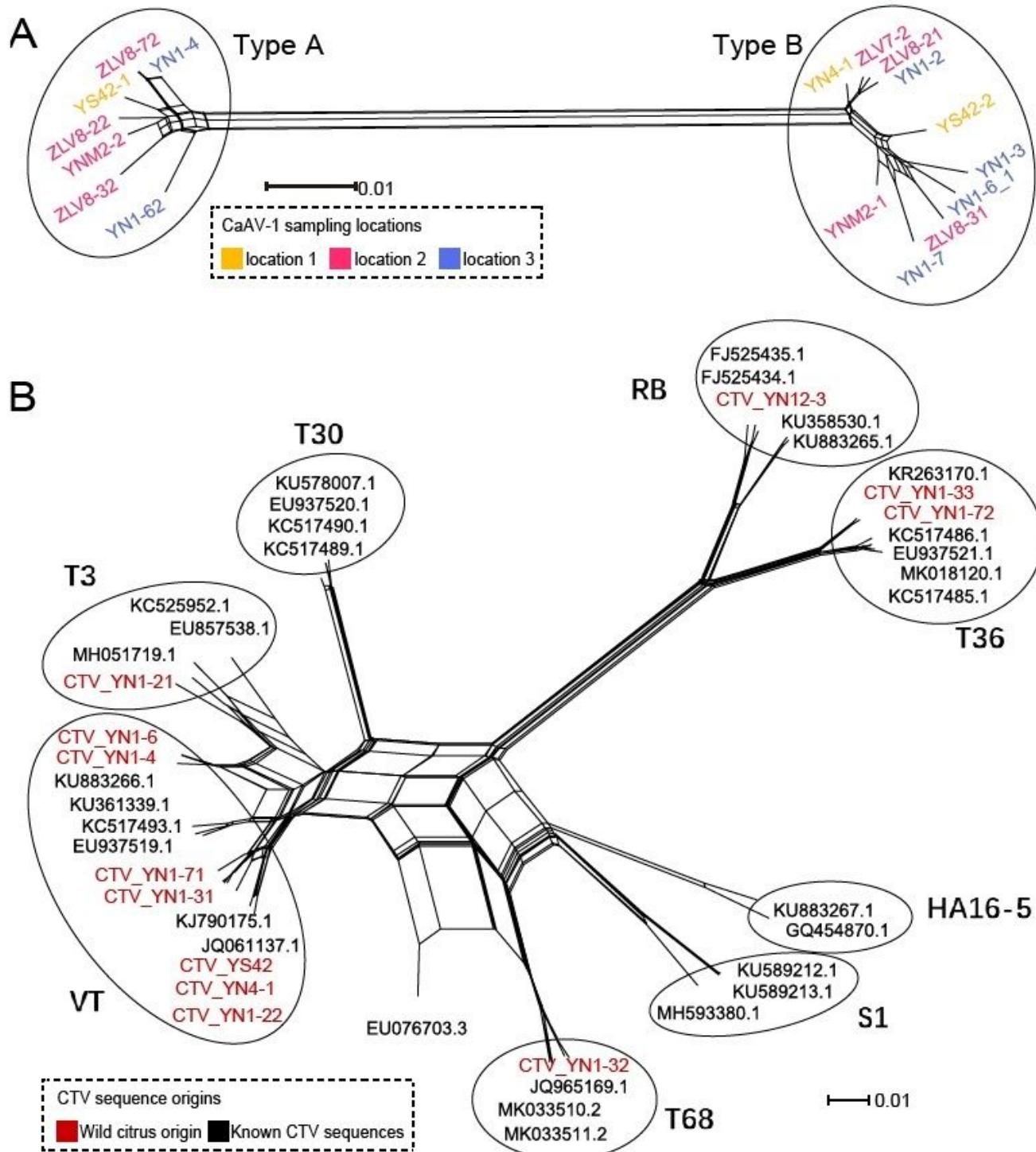
34. Xu Q, Chen L-L, Ruan X, Chen D, Zhu A, Chen C, et al: The draft genome of sweet orange (*Citrus sinensis*). *Nature Genetics* 2013, 45:59-66.
35. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al: TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant* 2020, 13:1194-202.
36. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al: Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003, 13:2498-504.
37. Atkinson HJ, Morris JH, Ferrin TE, Babbitt PC: Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLOS ONE* 2009, 4:e4345.
38. Katoh K, Standley DM: MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 2013, 30:772-80.
39. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T: trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009, 25:1972-3.
40. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al: IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* 2020, 37:1530-4.
41. Huson DH, Scornavacca C: Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Systematic Biology* 2012, 61:1061-7.
42. Huson DH: SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* 1998, 14:68-73.
43. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B: RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution* 2015, 1.
44. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, et al: Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 1999, 73:152-60.

## Figures



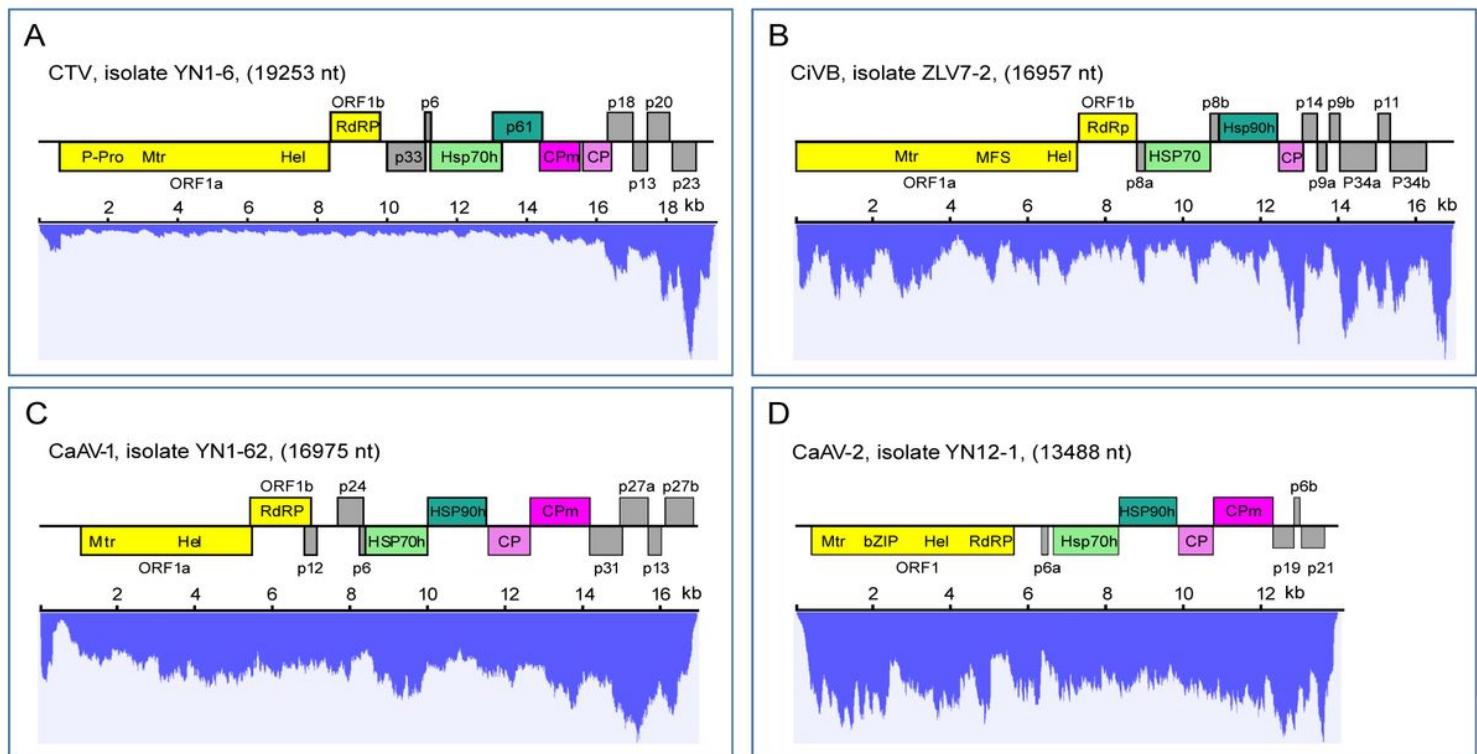
**Figure 1**

Geography of the study area and symptoms of the wild citrus trees examined. Sampling locations within the Ailao Mountain area of Yuxi City, Yunnan Province (A). Habitat of wild citrus trees (B, C). Diverse symptoms of sampled wild citrus trees (D–G). Chlorosis (D), boat-shaped leaf curl (E), leafroll (F, G). Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.



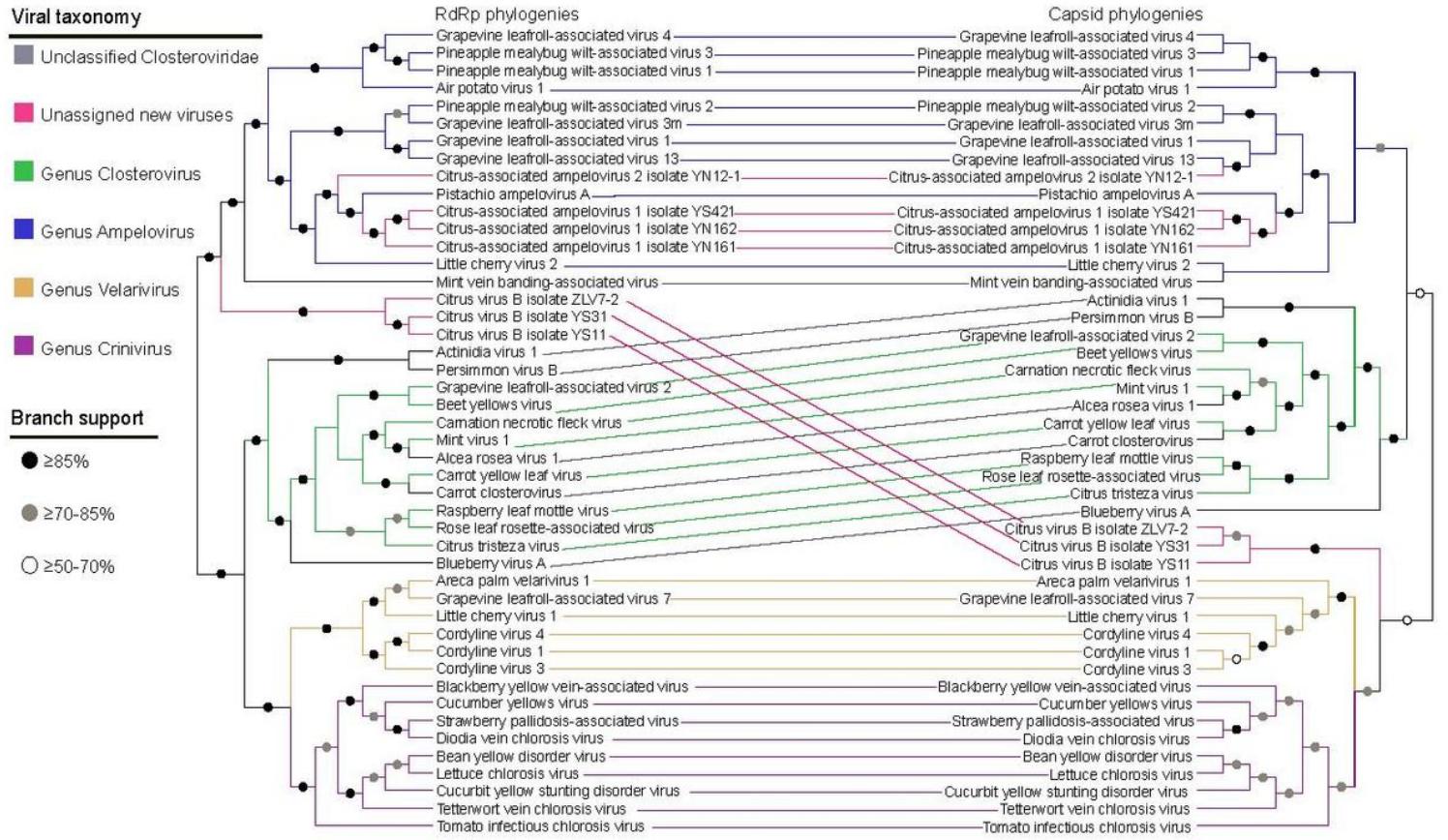
**Figure 2**

Intra-species genotyping analysis of citrus tristeza virus (CTV) and citrus-associated ampelovirus 1 (CaAV-1). Neighbor-net generated from the complete genome sequences of different CTV isolates (A) and CaAV-1 isolates (B). Parallel paths among the different viral isolates represent putative recombination events.



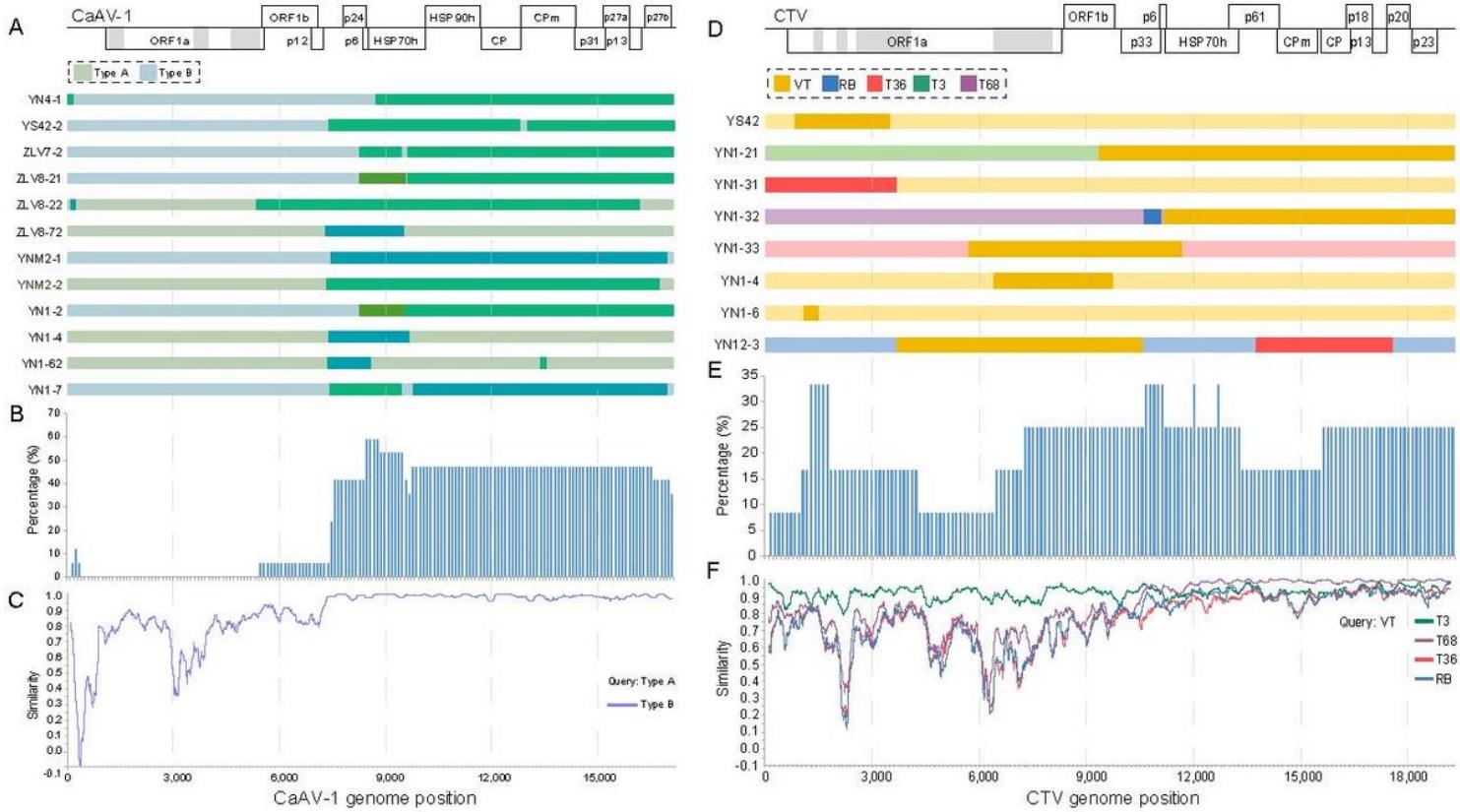
**Figure 3**

Genomic organization and transcriptome mapping analysis for citrus closterovirids. Genomic organization and transcriptome mapping analysis of citrus tristeza virus (CTV) (A), citrus virus B (CiVB) (B), citrus associated ampelovirus 1 (CaAV-1) (C) and citrus associated ampelovirus 2 (CaAV-2) (D). Mtr, methyltransferase; Hel, helicase; RdRP, RNA-directed RNA polymerase; HSP70h, heat shock protein 70 homolog; HSP90h, heat shock protein 90 homolog; CP, major coat protein; CPm, minor coat protein.



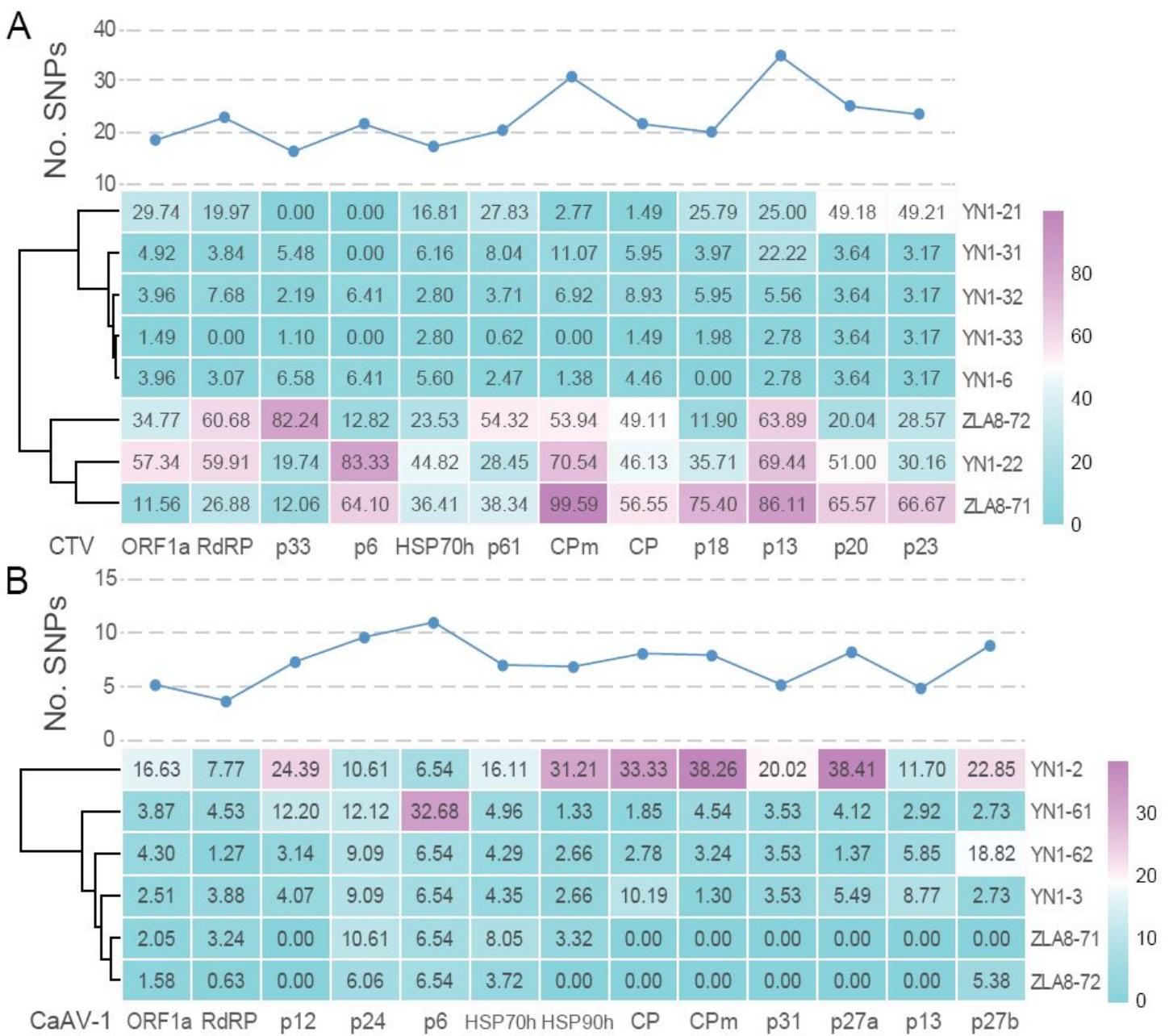
**Figure 4**

Co-evolutionary analysis for the family Closteroviridae. Tanglegram of maximum likelihood analyses of newly-identified viruses and known closterovirids based on the amino acid sequence of the RNA-directed RNA polymerase (RdRP) and the capsid protein (CP). Branch supports were inferred by bootstrapping with 1000 replicates.



**Figure 5**

Recombination analysis of citrus associated ampelovirus 1 (CaAV-1) and citrus tristeza virus (CTV). Maps of recombination patterns and parental lineages (A, D). The top part of each panel shows the genome organization for CaAV-1 YN1-62 and CTV YN1-6 isolates, and the gray shading represents identified motifs. The different color schemes depict different CaAV-1 and CTV lineages. The major parents are in a light shade, and the minor parents in a darker shade. Rate of recombinant isolates among total isolate numbers in a sliding 100 nucleotide window (B, E). The y-axis represents the proportion of recombinant sites. Intraspecific sequence identity of CaAV-1 and CTV (C, F). The map was made using SimPlot. The y-axis represents the intraspecific sequence identity. The x-axis represents CaAV-1 and CTV genome position.



**Figure 6**

Sequence variation scheme of citrus tristeza virus (CTV) and citrus associated ampelovirus 1 (CaAV-1). The levels of sequence variation per 1000 nt in coding regions of CTV (A) and CaAV-1 (B) isolates are shown in heatmaps; means are shown in the line charts (top). RdRP, RNA-dependent RNA polymerase; HSP70h, heat shock protein 70 homolog; HSP90h, heat shock protein 90 homolog; CP, major coat protein; CPm, minor coat protein.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1TableS1.xlsx
- Additionalfile2TableS2.xlsx
- Additionalfile3TableS3.xlsx
- Additionalfile4FigureS1.jpg
- Additionalfile5FigureS2.jpg
- Additionalfile6FigureS3.jpg
- Additionalfile7FigureS4.jpg
- Additionalfile8FigureS5.jpg
- Additionalfile9FigureS6.jpg
- Additionalfile10.fa