



OPEN ACCESS

EDITED BY

Mohammed Rohaim,
Lancaster University, United Kingdom

REVIEWED BY

Toufic Elbeaino,
International Centre for Advanced
Mediterranean Agronomic Studies, Italy
Charith Raj Adkar-Purushothama,
Université de Sherbrooke, Canada
Humberto J. Debat,
Instituto Nacional de Tecnología Agropecuaria,
Argentina

*CORRESPONDENCE

Xi Huang
✉ xihuang@hainanu.edu.cn

[†]These authors have contributed equally to this work and share first authorship

RECEIVED 31 August 2023

ACCEPTED 22 November 2023

PUBLISHED 14 December 2023

CITATION

Zhao R, Su X, Yu F, Liu Z and Huang X (2023) Identification and characterization of two closely related virga-like viruses latently infecting rubber trees (*Hevea brasiliensis*). *Front. Microbiol.* 14:1286369. doi: 10.3389/fmicb.2023.1286369

COPYRIGHT

© 2023 Zhao, Su, Yu, Liu and Huang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Identification and characterization of two closely related virga-like viruses latently infecting rubber trees (*Hevea brasiliensis*)

Ruibai Zhao^{1,2†}, Xiaoqi Su^{1†}, Fengjuan Yu¹, Zhu Liu² and Xi Huang^{1*}

¹College of Tropical Crops, Sanya Nanfan Research Institute of Hainan University, Sanya, China, ²School of Life Sciences, Hainan University, Haikou, China

A novel virga-like virus, provisionally named Rubber tree latent virus 2 (RTL2), was identified from rubber tree (*Hevea brasiliensis*). It is a close relative of the previously reported Rubber tree latent virus 1 (RTL1). The complete genomes of RTL1 and RTL2 were sequenced and comparatively analyzed in terms of genome organization, putative gene products and phylogenetic relationship. Both RTL1 and RTL2 have positive-sense single-stranded RNA genomes that encode seven open reading frames (ORFs), forming a similar genomic layout. In phylogenetic analyses based on replicase and coat protein amino acid sequences, RTL1 and RTL2 were clustered with unclassified virga-like viruses. They are distinct from currently recognized plant virus families. RTL1 and RTL2 can be distinguished from members of *Virgaviridae* by the presence of a putative coat protein duplex and a poly(A) tail at the 3'-terminus. The authenticity of RTL1 and RTL2 as infectious viruses was confirmed through field investigations and transmissibility assays. In conclusion, RTL1 and RTL2 represent a novel plant virus group that does not readily fit into current virus families.

KEYWORDS

virus discovery, high-throughput sequencing, latent virus, rubber tree latent virus 1, rubber tree latent virus 2

Introduction

The Para rubber tree (*Hevea brasiliensis* Muell.), a perennial deciduous woody plant in the family *Euphorbiaceae*, originates from the Brazilian Amazon. Today, it is widely cultivated in tropical areas worldwide, particularly in the Southeast Asia, for the commercial production of natural rubber (*cis*-1, 4-polyisoprene) (Liu et al., 2020). Rubber trees are the predominant source of natural rubber, comprising approximately one-third of the latex, which is essentially cytoplasm of the laticifers in the phloem (Tang et al., 2016). Commercial rubber tree plantations have been facing economic losses caused by several diseases, including the South American leaf blight (SALB) caused by *Pseudocercospora ulei* (Guyot and Le Guen, 2018), anthracnose caused by *Colletotrichum* species (Cao et al., 2019; Liu et al., 2023), the powdery mildew caused by *Oidium heveae* (Limkaisang et al., 2005), and tapping panel dryness syndrome (TPD), the cause of which remains unknown (Ramachandran et al., 2000; Pellegrin et al., 2004).

To date, there have been four reports of natural viral infections in rubber trees, but the exact etiology of these viruses remains unconfirmed (Gama et al., 1983; Fonseca et al., 2018; Li et al., 2020; Chen et al., 2022). The first documented virus to infect rubber trees, belonging to the genus *Carlavirus* of the family *Betaflexiviridae*, was found in stunted seedlings displaying

inter-veinal chlorosis, suggesting an association with leaf disease (Gama et al., 1983). Three decades later, the second putative virus named *Hevea brasiliensis* virus (HBrV) was detected in Brazil by small RNA sequencing (sRNA-seq). Phylogenetic analysis of the 1913 amino acid (aa) viral sequence assembled from HBrV showed the closest relationship to members of the genus *Maculavirus*, family *Tymoviridae*. HBrV infection was latent, as it was detected in asymptomatic plants with uniformly green leaves (Fonseca et al., 2018).

Rubber tree plantations represent a significant sector of the tropical crops industry in Hainan, China, but they suffer serious losses due to TPD. High-throughput RNA sequencing (RNA-seq) technologies have been applied to identify pathogenic microorganisms in affected plants (Wu et al., 2015). In an effort to identify potential causal agents of TPD, a comparative transcriptome analysis of TPD-affected and healthy bark samples collected from Hainan, China, led to the discovery and identification of two viruses: Rubber tree virus 1 (RTV1) and Rubber tree latent virus 1 (RTLVL1), previously annotated as Cherry virus A (CVA) and Tobacco mosaic virus (TMV), respectively (Li et al., 2020; Chen et al., 2022). RTV1 is a member of the genus *Capillovirus* of family *Betaflexiviridae*, encoding two overlapping open reading frames (ORF) corresponding to a replicase-coat protein (CP) fusion polyprotein and a movement protein (MP), respectively. Complete genome of RTV1 consists of 6,811 nucleotides (nt), excluding the poly(A) tail (Li et al., 2020). Although both RTV1 and RTLVL1 were identified from TPD-affected barks, infection of RTLVL1 was considered to be asymptomatic, as it could be detected in 35 of 44 healthy trees. RTLVL1 was suggested to be an unassigned member of the family *Virgaviridae*, as its replicase clustered with unclassified members and was not closely related to any genus of *Virgaviridae* in phylogenetic analysis. Interestingly, RTLVL1 was most closely related to several arthropod-infecting viruses, with its genomic organization and length being distinct from those of plant-infecting *Virgaviruses* (Chen et al., 2022). RTLVL1 may potentially constitute a new virus genus while its taxonomy and evolution history still remain elusive.

In this work, we identified a novel relative of RTLVL1, provisionally named Rubber tree latent virus 2 (RTLVL2) in co-infection with RTLVL1, and confirmed their presence in commercial rubber tree plantations. The complete genome of RTLVL2 was much longer than the previously reported genome of RTLVL1, suggesting that the reported RTLVL1 genome might be incomplete, which was re-sequenced and corrected in this work. We conducted a comparative analysis of genome organization, putative gene products, and phylogenetic relationships of the two viruses. Experimental confirmation demonstrated the mechanical transmissibility and latent infection of RTLVL2.

Materials and methods

RNA-seq screening of viral agents in rubber tree bark samples

The RNA library construction and RNA-seq had been conducted previously (Li et al., 2020). Two bark samples used for RNA-seq analysis were collected from a TPD tree and a healthy tree, respectively, in a commercial plantation in Qionghai County, Hainan Province, China. The bark samples were stored at -80°C

until used. Total RNA was isolated using TRIzol reagent (Invitrogen, China) according to the manufacturer's instructions. The quality and purity of the extracted RNA were examined using agarose electrophoresis, a NanoPhotometer[®] spectrophotometer (Implen, CA, USA), and the RNA 6000 Nano Kit assay with the Agilent Bioanalyzer 2,100 system (Agilent Technologies, CA, USA). Importantly, the Ribo-Zero[™] rRNA Removal Kit (Plant Leaf) was used to remove rRNA and first-strand cDNA was synthesized through reverse transcription using random hexamers so that bacterial and viral transcripts without poly(A) tails could be detected. After ligation of distinct adaptors, cDNA was purified with the Agencourt AMPure XP system (Beckman Coulter, MA, USA) to preferentially select 250 to 300 base pair (bp) fragments prior to generation of the two final libraries through PCR. Sequencing was conducted and paired-end reads were generated on the Illumina HiSeq 2,500 platform at Novogene Co. (Beijing, China). Raw reads in FASTQ format were filtered to obtain high-quality clean reads. Transcriptome assembly was accomplished using Trinity through combined short reads with a certain overlap length ($k\text{-mer} = 25$) (Grabherr et al., 2011). The assembled unigenes were clustered using default parameters (Davidson and Oshlack, 2014) and then subjected to similarity searches against the NCBI non-redundant nucleotide and protein databases (NR) and the Swissprot database for annotation and screening of viral transcripts (Altschul et al., 1997; Buchfink et al., 2015).

Sequencing and analysis of the viral genome

To sequence the complete viral genomes of RTLVL1 and RTLVL2, back-to-back primer sets producing overlapping fragments which covered nearly the complete genome (Supplementary Table S1) were designed based on the assembled unigenes from RNA-seq analysis. Rapid Amplification of cDNA Ends (RACE) was carried out to confirm the 5' and 3' terminal sequences using the 5'/3'RACE Kit, 2nd Generation (Roche, Indianapolis, IN) according to the manufacturer's instructions. In the initial 3' RACE trial, total RNA was polyadenylated prior to reverse transcription using *Escherichia coli* Poly(A) Polymerase (New England Biolabs, Ipswich, MA) in case the viruses do not possess a poly(A) tail. PCR was carried out using PrimeSTAR[®] GXL DNA Polymerase (TaKaRa, Dalian), and the resulting products were incubated with Taq-polymerase at 72°C for 10 min then ligated into the pMD-19 T vector (TaKaRa, Dalian). At least three independent positive clones of each fragment were subjected to Sanger sequencing (Sangon Biotech, China). Overlapping sequences were analyzed and assembled to attain the consensus full-length genome with SeqMan Pro 7.1.0 (Lasergene, GATC Biotech).

ORFs and the conserved functional domains were predicted using ORF Finder¹ and the Conserved Domain Database (CDD²) (Lu et al., 2020), respectively. Predicted ORFs were annotated by blast against

1 <https://www.ncbi.nlm.nih.gov/orffinder/>

2 <https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>

the NCBI NR database and the Swissprot database. Potential transmembrane domains were predicted using TMHMM 2.0 (see footnote 2). Amino acid (aa) sequence identity between homologous ORFs of RTL1 and RTL2 as well as the remote relatives from other viruses was calculated by DNAMAN software (Lynnon BioSoft).

Phylogenetic analysis

Representative species of all seven genera of the family *Virgaviridae* and the related taxa in orders *Martellivirales* and *Hepelivirales* were chosen to demonstrate the basic phylogenetic relationship of this virus group with alpha-like replicases. Top-scored relatives identified in blast searches, most of which are unassigned species with low homology to ratified members, were also included to identify probable virus group. Sequences were aligned using the ClustalW program implemented in the MEGA7 software with default parameters. Phylogenetic trees were then constructed in MEGA7 by using the Maximum Likelihood method based on the JTT matrix-based model with 1,000 bootstrap replications (Jones et al., 1992; Kumar et al., 2016). The calculated trees were further edited with Adobe Photoshop.³

Incidence survey

Leaf samples of 44 asymptomatic rubber trees were collected from rubber tree plantations in Dingan County (Hainan Province, China) in 2020. A further collection of leaves of 30 healthy rubber trees with the same sampling standards was carried out in rubber tree plantations in Qionghai County. Total RNA was extracted with 100 mg sample using the RNAPrep Pure Plant Plus Kit (TIANGEN BIOTECH, China) according to the manufacturer's instruction. cDNA was synthesized from total RNA with random hexamer primers using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific) following the manufacturer's instruction. Two individual primer sets were used for detection of each virus (Supplementary Table S1). PCR conditions were 30 cycles of denaturation at 95°C for 15 s, primer annealing at 60°C for 20 s, and extension at 72°C for 2 min (2 kb product), followed by final extension at 72°C for 5 min, using the SanTaq Plus PCR Mix (Sangon Biotech, China).

Mechanical transmission

Thirty seedlings of rubber tree and *N. benthamiana* were used for mechanical inoculation, respectively. Rubber tree seedlings (cultivar Reyan7-33-97) at the three/four true-leaf stage were purchased from the Rubber Research Institute of Chinese Academy of Tropical Agricultural Science and cultivated in greenhouse under controlled conditions of 26 ± 2°C, 16 h/8 h light/dark before used for mechanical inoculation assay. All seedlings were tested for the presence of RTL1 and RTL2 in groups of 10 by RT-PCR to ensure that they were uninfected prior to inoculation. *N. benthamiana* seedlings were

germinated and kept in greenhouse under controlled conditions of 22 ± 2°C, 16 h/8 h light/dark. Rubber tree leaves co-infected by RTL1 and RTL2 were collected from Qionghai County and severed as inoculum. Fresh leaves were ground in 0.1 M phosphate buffered saline (PBS, pH 7.0) and then manually applied with carborundum powder (600 mesh) to Rubber tree or *N. benthamiana* seedlings. Mock inoculation was performed with PBS and carborundum powder. The inoculated plants were monitored for symptoms of infection and tested for presence of virus at 2 months post inoculation (mpi) by RT-PCR as described above.

Results

Discovery of RTL2, a novel virga-like virus by RNA-seq

We had previously conducted a comparative transcriptome analysis of rubber tree bark samples from trees affected by TPD and healthy ones (Li et al., 2020). A total of 65,986 unigenes were *de novo* assembled with Trinity and Corset and then annotated by BLAST searches against NCBI NR and Swissprot databases. This analysis found multiple virus-like sequences, leading to the discovery and identification of RTL1 and RTL2 (Li et al., 2020; Chen et al., 2022). Upon further examination of the assembled contigs, an interesting phenomenon were observed: aside from the RTL1-derived unigenes that exhibited the closest similarity to the TMV replicase large subunit, several unigenes with sizes ranging from 8,211 to 8,796 nucleotides exhibited the highest similarity to the replicase large subunit of Tomato mosaic virus (ToMV). This finding strongly suggests the presence of another virga-like virus in the bark samples, as outlined in Supplementary Table S2. Reflecting the phylogenetic relationship between TMV and ToMV, the ORF1 sequences encoding the replicase from these two groups of contigs were closely related. They shared approximately 60% nt and 60% aa identity with each other. In contrast, when compared to their closest relatives in a blastx search against the GenBank NR database, they exhibited only about 35% aa identity. Given the substantial similarity in the replicase-coding sequences, we tentatively designated this newly discovered virga-like virus as RTL2 in reference to RTL1.

Sequencing of the complete genomes of RTL1 and RTL2 by RT-PCR and RACE

Based on the alignment of the RTL2-derived unigenes, four back-to-back primer sets for RT-PCR, nested reverse primers for 5'RACE and forward primers for 3'RACE were designed to amplify the complete genome of RTL2 (Supplementary Table S1A). Expected bands of three primer sets with sizes of approximately 2 kilobases (kb) were obtained, which covered the 5' majority of the unigenes with overlaps between adjacent fragments. 5'RACE with different combinations of nested primers produced unique clear bands corresponding to the expected sizes, and an identical 5' end was obtained from the 20 clones sequenced. However, the fourth primer set covering the 3' 2 kb region of the unigenes generated a large band of nearly 6 kb, which was approximately 3.5 kb longer than expected. Moreover, 3' RACE produced multiple bands forming gradient

³ <https://www.adobe.com>

ladders in the serial nested amplifications, suggesting the random breakage near the 3' end of viral RNA. We therefore speculated that the assembled unigenes contained a 3.5 kb missing gap near the 3' end. Based on the sequence of the 6 kb band, additional primers inside and outside of the presumed gap were designed. RT-PCR with various combinations of primers confirmed the fidelity and continuity of the 6 kb amplicon as authentic viral sequence. A second trial of 3' RACE with nested forward primers (Supplementary Table S1A) located at the 3' proximal of the 6 kb sequence produced unique clear bands corresponding to the nested primers, and an identical 3' end was obtained from the 10 clones sequenced. 3'RACE using total RNA or 3'-polyadenylated RNA gave the same results, confirming the existence of poly(A) tail at the 3' end. Sequences of all the amplicons were assembled to generate the complete genome sequence of RTL2 (GenBank accession OR500511), which turned out to be 11,866 nt in length, encoding seven ORFs.

The RTL2 genome is much longer than the previously determined the genome sequence of RTL1, which consists of 9,422 nt (GenBank accession MW774645) (Chen et al., 2022). Due to the close genetic correlation between RTL1 and RTL2, namely approximately 60% nt and 60% aa identities, it is reasonable to speculate that the previously determined RTL1 genome is incomplete. Given that the assembled RTL2-derived unigenes contained a 3.5 kb missing gap, it is possible that similar error occurred in the transcriptome analysis pipeline which generates 3'-truncated RTL1 contigs. To verify the 3' end of RTL1, 3'RACE was conducted with the 72°C extension process extended to 3 min per amplification cycle. In accordance with this assumption, serial bands of about 3 kb were obtained in serial amplifications with various combinations of nested primers which located within 1 kb upstream of the reported RTL1 3' end. Based on the consensus sequence of these 3 kb products, five forward nested primers near the 3' terminal were designed to conduct a second trial of 3'RACE (Supplementary Table S1B), still with the 72°C extension process extended to 3 min per amplification cycle for fully extension. Unique clear bands of about 0.5 kb corresponding to the nested primers were obtained, and a uniform 3' end was identified from the 10 clones sequenced. 3'RACE with total RNA excluding the polyadenylation procedure gave the same results, confirming the existence of poly(A) tail at the 3' end. The rest of RTL1 genome was fully re-sequenced by RT-PCR of overlapping fragments and 5' RACE, and sequences of all the amplicons were assembled to generate the complete genome sequence of RTL1 (GenBank accession OR500510), which turned out to be 11,923 nt in length, excluding the poly(A) tail.

Genome annotation of RTL1 and RTL2

RTL1 and RTL2 shared 58.31% identity of the genome nt sequences. Both genomes had a G/C content of 39% and were predicted to encode seven ORFs which displayed similar layout of genome organization (Figure 1). All ORFs shared marginal (30.36% for ORF7) to considerable (64.17% for ORF5) aa identity with their homologue encoded by the other virus, demonstrating the close genetic relationship between RTL1 and RTL2 (Table 1).

ORF1 in RTL1 and RTL2 shares a nucleotide identity of 60.75% and an amino acid identity of 58.87%. This ORF encodes a substantial replicase polyprotein, weighing 298.2 kDa for RTL1 and 291.3 kDa

for RTL2. It encompasses three conserved domains predicted by the Conserved Domain Database (CDD) (see Figure 1 and Table 1): a viral methyltransferase domain (MTR, pfam01660), a viral RNA helicase domain (HEL, pfam01443) and an RNA-dependent RNA polymerase domain (RdRp, pfam00978). The predicted MTR domain is the hallmark of the 'Sindbis-like' supergroup of positive-strand RNA viruses, such as Hordei-, Tobra-, Tobamo-, Bromo- and Closteroviruses (Rozanov et al., 1992). The 800 nt helicase domain contained the conserved NTP-binding motifs GKT and DE (Koonin and Dolja, 1993), which were identified from aa 1789–1791 and 1873–1874 of RTL1 ORF1, and aa 1738–1740 and 1821–1822 of RTL2 ORF1, respectively. The well-characterized GDD motif in the RdRp domain (Koonin, 1991) was found from aa 2,452–2,454 of RTL1 ORF1 and aa 2,402–2,404 of RTL2 ORF1. In addition to the domains conservatively associated with viral replication, a 2OG-Fe(II) oxygenase superfamily (CDD accession cl21496) (Aravind and Koonin, 2001) and an ovarian tumor (OTU) superfamily (CDD accession cl45892) (Messick et al., 2008) were found in the intergenic region between MTR and HEL (Figure 1), with lower *E*-value of about 1e-10 to 1e-12. Blastx against NCBI NR database showed that ORF1 best matched the replicase polyprotein of *Oxera neriifolia* associated virus (OnV), an unclassified plant virus found in large-scale transcriptome data mining (Mifsud et al., 2022; Table 1). However, the rest of the top-scored hits were basically replicase proteins encoded by unclassified invertebrate viruses followed by replicases of Tobamoviruses (Supplementary Table S3), indicating that RTL1 and RTL2 were virga-like viruses related to the family *Virgaviridae*.

RTL1 ORF2, with a molecular weight of 24.5 kDa, and RTL2 ORF2, weighing 25.0 kDa, exhibited a shared nucleotide identity of 52.83% and an amino acid identity of 42.60%. Despite our analyses with the Conserved Domain Database (CDD), we were unable to identify any conserved domains within ORF2. Blastx against NR revealed marginal homology between RTL2 ORF2 and the putative ORF2 of OnV, which encodes a replicase polyprotein closely related to RTL1 ORF1 (Table 1). Additionally, Blastp analysis of RTL2 ORF2 against the Swissprot database revealed only one distant viral homologue (with an amino acid identity of 25.4%), which was the putative ORF2 of Hubei virga-like virus 11 (HVLV11), an unclassified invertebrate virus identified through high-throughput transcriptome data mining (Shi et al., 2016). Although we pursued similar investigations for RTL1 ORF2, we could not identify any significant hits. However, upon conducting multiple alignments (please see Supplementary Figure S1), we did identify several conserved amino acid residues within the C-terminal region of ORF2 in RTL1, RTL2, OnV, and HVLV11. This finding suggests a potential common origin for these proteins, which currently have unknown functions, and implies a distant phylogenetic relationship among these unclassified viruses.

RTL1 ORF3, with a molecular weight of 17.8 kDa, and RTL2 ORF3, weighing 18.1 kDa, shared a nucleotide identity of 59.49% and an amino acid identity of 56.69%. Standard Blastx analysis against the NCBI NR database and searches in the CDD did not yield significant hits for ORF3. However, Blastp analysis against the Swissprot database unveiled Tobamovirus coat protein (CP)-like motifs in both RTL1 and RTL2 ORF3, suggesting that ORF3 likely encodes a viral CP. Despite the ambiguous annotation of ORF3 as a CP, RTL1 ORF4, with a molecular weight of 19.6 kDa, and RTL2 ORF4, weighing 22.0 kDa, both contained a CP domain resembling that of Tobacco

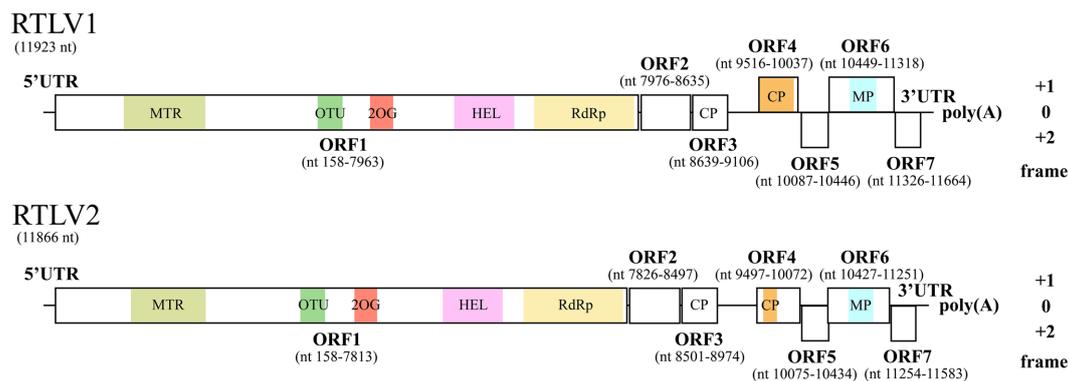


FIGURE 1

Genome organization of RTL1 and RTL2. RTL1 and RTL2 are closely related and have similar genomic architecture encoding seven pairs of ORFs. Conserved domains predicted by CDD were colored and labeled as: MTR, methyltransferase (pfam01660); OTU, ovarian tumor superfamily (cl45892); 2OG, 2OG-Fe(II) oxygenase superfamily (cl21496); HEL, RNA helicase (pfam01443); RdRp, RNA-dependent RNA polymerase (pfam00978); CP, TMV-like coat protein (pfam00721); MP, viral movement protein (pfam01107). Two uncolored putative CPs coded by ORF3 were not recognized by CDD, but shared conserved aa residues with the CDD-predicted CP coded by ORF4 and CPs of Virgaviruses (Figure 2). Genomic position of each putative ORFs were drawn to scale, and the reading frame of ORFs were determined relative to ORF1 (frame 0). UTR, untranslated region.

mosaic virus (TMV), as predicted by CDD (pfam00721). Blastx analysis against the NCBI NR database returned multiple hits of putative CP proteins encoded by various invertebrate virga-like viruses, as well as two unclassified plant viruses. Notably, RTL1 ORF4 and RTL2 ORF4 both exhibited the highest similarity to the putative CP of *Culex pipiens*-associated Tunisia virus (CpATV) (Bigot et al., 2018). Among the homologous proteins encoded by plant viruses, RTL1 ORF4 matched the putative CP of Plant-associated tobamo-like virus 1 (PaToLV1) (Rivarez et al., 2023), while RTL2 ORF4 matched the putative CP of Goji berry chlorosis virus (GBCV) (Table 1; Supplementary Table S4).

Given that both ORF3 and ORF4 contain motifs or domains resembling the TMV CP, we conducted a multiple sequence alignment involving ORF3 and ORF4 encoded by RTL1, RTL2, OnV, and several Tobamovirus-encoded CPs. This analysis identified two common motifs, RFP and FE, which are abundant in the TMV-like CP domain family pfam00721 (refer to Figure 2). However, despite the presence of these common CP-like motifs and some conserved amino acid residues in the C-terminal region, ORF3 and ORF4 exhibited extremely low amino acid identities of 18.50% for RTL1 and 17.47% for RTL2. Between the putative CPs encoded by ORF3 and ORF4, we identified a large intergenic region spanning 409 nt for RTL1 and 522 nt for RTL2 (refer to Figure 1). Within this region, ORF Finder recognized two small ORFs of 46 amino acids (nt 9,103–9,243) and 58 amino acids (nt 9,230–9,406) in RTL1, and three small ORFs of 41 amino acids (nt 8,998–9,123), 63 amino acids (nt 9,092–9,283), and 50 amino acids (nt 9,287–9,439) in RTL2. However, these ORFs appear to be hypothetical, as no significant homology resembling the pairs of ORF1–7 was detected.

Among the seven pairs of ORFs, ORF5 encoded the most conserved protein between RTL1 and RTL2, sharing 65.29% nt and 64.17% aa identities (Table 1). When Blastx analysis was conducted against the Swissprot database, limited similarity with a few proteins containing kinase domains from plants and transmembrane proteins from bacteria was observed. However, this similarity only covered about 15 to 30% of the sequence. In contrast, Blastx analysis against the NCBI NR database failed to yield any significant hits. It's

noteworthy that ORF5 was identified as the sole probable transmembrane protein according to TMHMM predictions, with relatively low probabilities (approximately 0.6, 0.7, and 0.2) assigned to three potential transmembrane regions. Similar to ORF2, which exhibited no apparent homology with known proteins and may be specific to viruses within their clade, our alignment of the amino acid sequences of ORF5 from RTL1, RTL2, and OnV revealed several conserved residues, indicating their phylogenetic relationship (please refer to Supplementary Figure S2).

Moving on, RTL1 ORF6 (32.5 kDa) and RTL2 ORF6 (30.4 kDa) encoded the largest protein excluding the replicase polyprotein. CDD analyses identified a viral movement protein domain (pfam01107) in the center region of ORF6 (Figure 1). Blastx against NR revealed the most related viral homologue of ORF6 was the putative MP of *Pinus nigra* virus 1 (Rastrojo et al., 2018) (Caulimoviridae, Table 1). Blastx against Swissprot returned hits of viral MPs from families Caulimoviridae, Betaflexiviridae, and Rhabdoviridae. pfam01107 also includes MPs from the family Virgaviridae, e.g., the “30 K” MP of TMV and Tobacco rattle virus (TRV) (Melcher, 2000), which have similar molecular weight as ORF6. However, significant hits of members from Virgaviridae were not found in blast searches, suggesting that their homology with ORF6 was limited.

Finally, RTL1 ORF7 (12.4 kDa) and RTL2 ORF7 (12.2 kDa) encoded the smallest protein in the viral genome. Blast searches found no apparent homology with known proteins and no conserved domains were identified. Interestingly, ORF7 was the most divergent in the seven pairs of ORF, and the 30.36% aa identity was much lower than the 47.49% nt identity shared between ORF7 of RTL1 and RTL2, suggesting a species-specific differentiation.

Phylogenetic analysis

Preliminary Blastx search of ORF1 (the replicase polyprotein) against NR found homologues encoded by unclassified invertebrate- or plant-infecting virga-like viruses and assigned members from genus *Tobamovirus* of family Virgaviridae

TABLE 1 ORFs and their putative gene products encoded by RTL1 and RTL2.

ORF	Amino acids (aa)	MW (kDa)	Shared nt identity (%)	Shared aa identity (%)	Protein annotation	Similarity search against the NCBI NR database						
						Best matched plant virus	Classification	Host/origin	Protein accession	Query cover (%)	aa identity (%)	E-value
RTL1 ORF1	2,601	298.2	60.75	58.87	Replicase polyprotein	Oxera neriifolia associated virus	Unclassified	<i>Oxera neriifolia</i>	CAI5383846	71	37.89	0.0
RTL2 ORF1	2,551	291.3				Oxera neriifolia associated virus	Unclassified	<i>Oxera neriifolia</i>	CAI5383846	73	38.66	0.0
RTL1 ORF2	219	24.5	52.83	42.60	Unknown	Oxera neriifolia associated virus*	Unclassified	<i>Oxera neriifolia</i>	CAI5383847	NA	NA	NA
RTL2 ORF2	223	25.0				Oxera neriifolia associated virus	Unclassified	<i>Oxera neriifolia</i>	CAI5383847	61	29.71	2E-03
RTL1 ORF3	155	17.8	59.49	56.69	Coat protein	Oxera neriifolia associated virus*	Unclassified	<i>Oxera neriifolia</i>	CAI5383848	NA	NA	NA
RTL2 ORF3	157	18.1				Oxera neriifolia associated virus*	Unclassified	<i>Oxera neriifolia</i>	CAI5383848	NA	NA	NA
RTL1 ORF4	173	19.6	50.91	52.00	Coat protein	Plant associated tobamo-like virus 1	Unclassified (<i>Virgaviridae</i>)	<i>Solanum lycopersicum</i>	UTQ50505	85	28.86	6E-04
RTL2 ORF4	191	22.0				Goji berry chlorosis virus	Unclassified	<i>Lycium chinense</i>	AYO99570	78	28.76	1E-04
RTL1 ORF5	119	13.0	65.29	64.17	Unknown	Oxera neriifolia associated virus*	Unclassified	<i>Oxera neriifolia</i>	CAI5383850	NA	NA	NA
RTL2 ORF5	119	12.8				Oxera neriifolia associated virus*	Unclassified	<i>Oxera neriifolia</i>	CAI5383850	NA	NA	NA
RTL1 ORF6	289	32.5	60.57	60.00	Movement protein	<i>Pinus nigra</i> virus 1	Unclassified (<i>Caulimoviridae</i>)	<i>Pinus nigra</i>	AXB54820	54	34.36	3E-10
RTL2 ORF6	274	30.4				<i>Pinus nigra</i> virus 1	Unclassified (<i>Caulimoviridae</i>)	<i>Pinus nigra</i>	AXB54820	40	33.04	2E-06
RTL1 ORF7	112	12.4	47.49	30.36	Unknown	No matches	NA	NA	NA	NA	NA	NA
RTL2 ORF7	109	12.2				No matches	NA	NA	NA	NA	NA	NA

The best matched plant virus for each ORF was determined by blastx max scores in search against the NCBI NR database, except for items marked with an asterisk (*), that were inferred by sequence alignment of aa sequences as no significant results returned. (Refer to Supplementary Figures S1, S2 and Figure 2 for ORF2, ORF3, and ORF5, respectively.) CDS, coding sequence; MW, molecular weight; NA, not available (no returned significant hits).

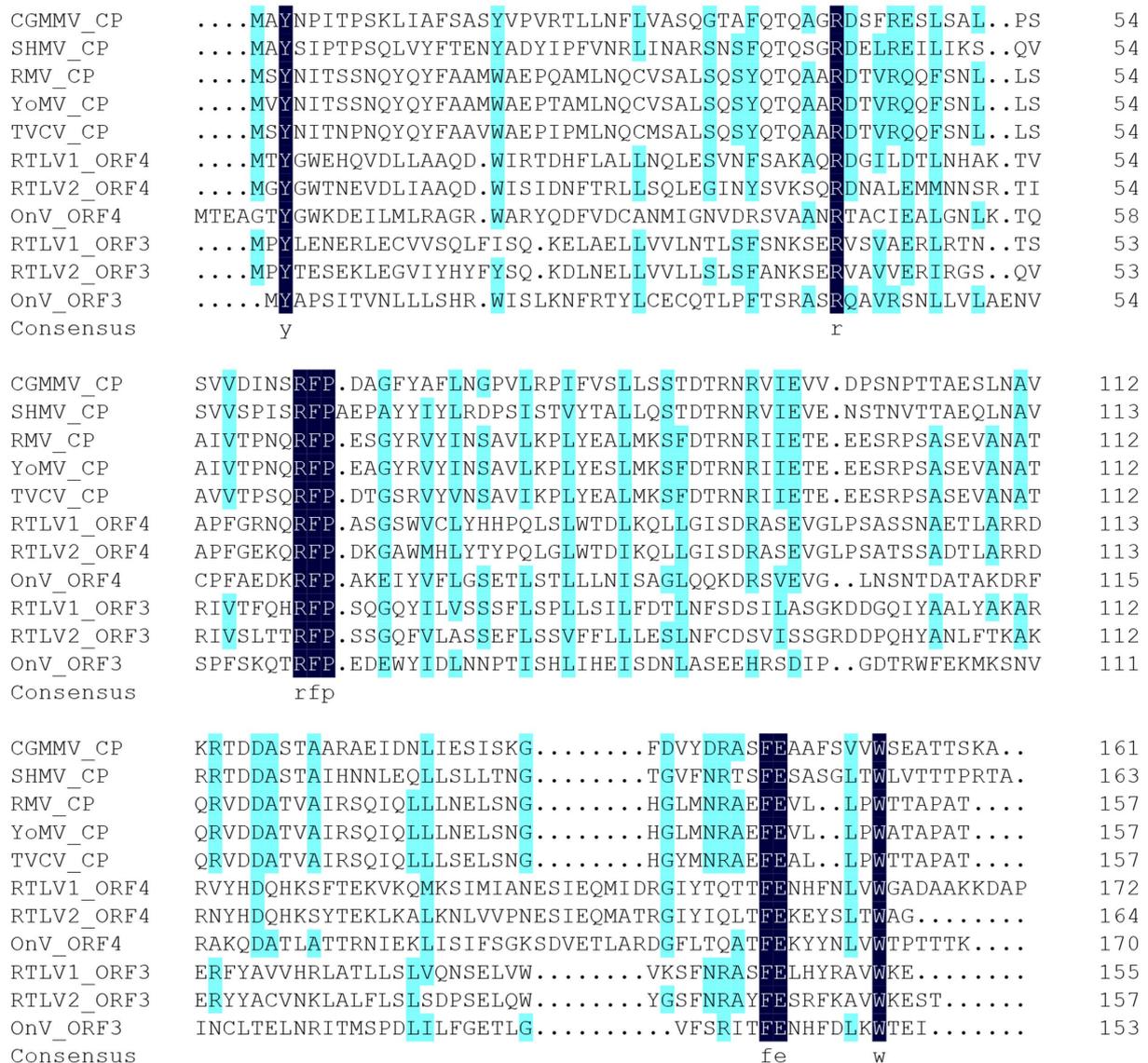


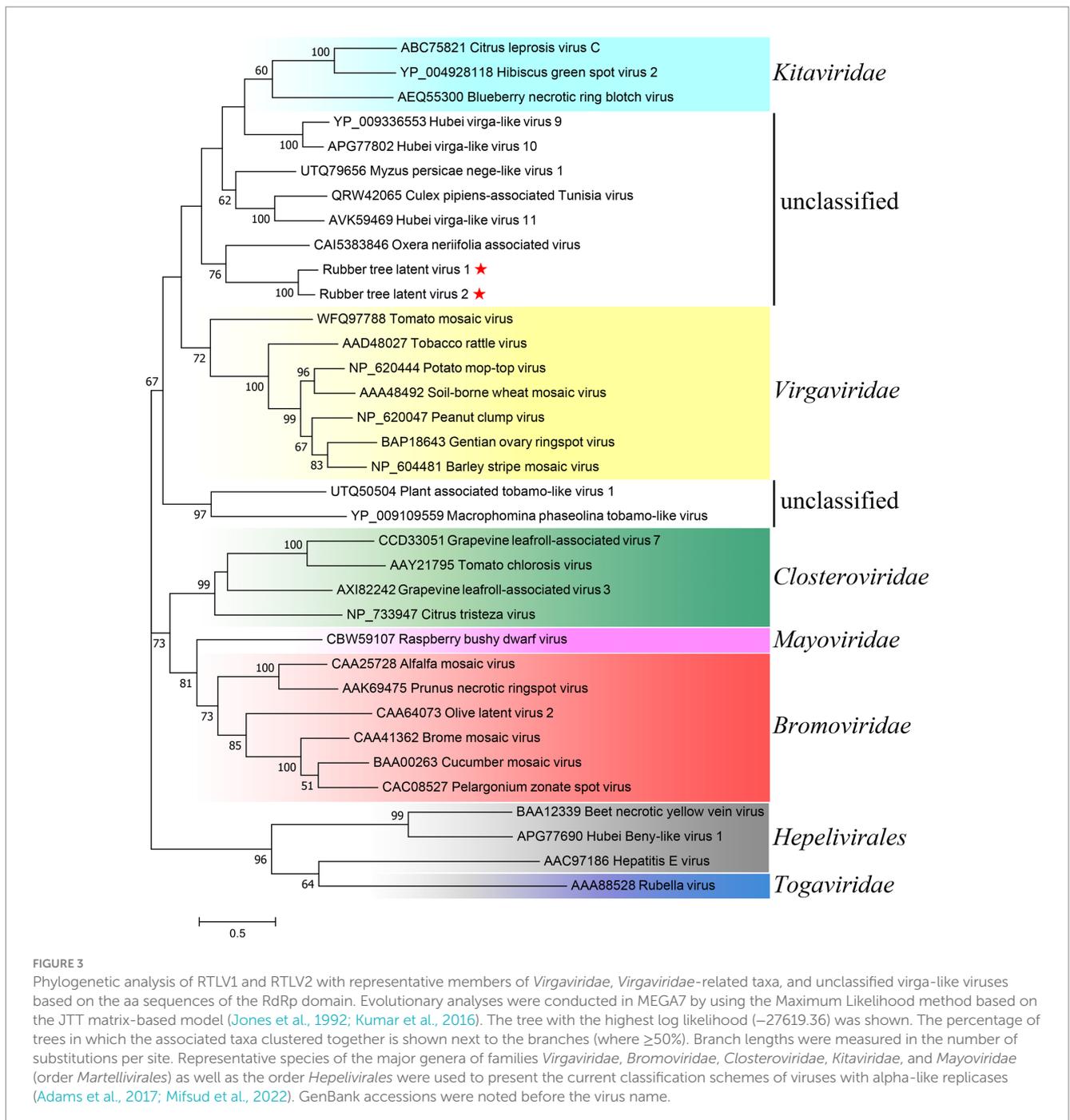
FIGURE 2

Amino acid sequence alignment of the putative CPs of RTLV1, RTLV2, OnV and reviewed CPs of Tobamoviruses. Identical residues were highlighted by dark blue, including two motifs RFP and FE which extensively exist in the TMV-like CP domain family (pfam00721). aa residues with homology level $\geq 50\%$ were highlighted by light blue (calculated by DNAMAN). Virus species and the Swissprot accession of reviewed CPs of assigned members of genus *Tobamovirus*, family *Virgaviridae* are as follows: CGMMV, Cucumber green mottle mosaic virus, P69475; SHMV, Sunn-hemp mosaic virus, P03581; RMV, Ribgrass mosaic virus, Q9WDG7; YoMV, Youcai mosaic virus, Q66222; TVCV, Turnip vein-clearing virus, Q88922. OnV, *Oxera nerifolia* associated virus. aa sequences of OnV ORF3 and ORF4 were translated from nt 8,695–9,156 and nt 9,222–9,734 of the full-length genome (GenBank accession OX380366), respectively.

(Supplementary Table S3). However, the majority of homologues found in Blastx of ORF4 containing the TMV CP-like domain were putative CPs of unclassified invertebrate-infecting viruses, while the CPs of ratified members of *Virgaviridae* were absent among the returned hits (Supplementary Table S4). To investigate the taxonomic position of RTLV1 and RTLV2, aa sequences of the RdRp domain of representative species from *Virgaviridae*, *Virgaviridae*-related taxa (Adams et al., 2017) and unclassified viruses with homologous replicase found in Blastx were used in phylogenetic analysis (Figure 3). The generated tree conformed to the general phylogenetic topology of *Virgaviridae* with its related taxa in orders *Martellivirales* and *Hepelivirale* (Adams

et al., 2017), while the RdRp of unclassified viruses formed two clusters in the *Kitaviridae-Virgaviridae* branch. Consistent with the Blastx results (Supplementary Table S3), RTLV1 and RTLV2 clustered with OnV and formed a monophyletic clade separated from the invertebrate viruses. This clade could form a paraphyletic group together with the *Virgaviridae* branch, however the node support was only moderate (42%) for this subjective partition.

As ORF3 and ORF4 both contained TMV CP-like motifs (Figure 2), aa sequences of CPs of representative species from *Virgaviridae* and CP-like homologues found in Blastx were used to investigate the phylogenetic relationship of RTLV1 and RTLV2

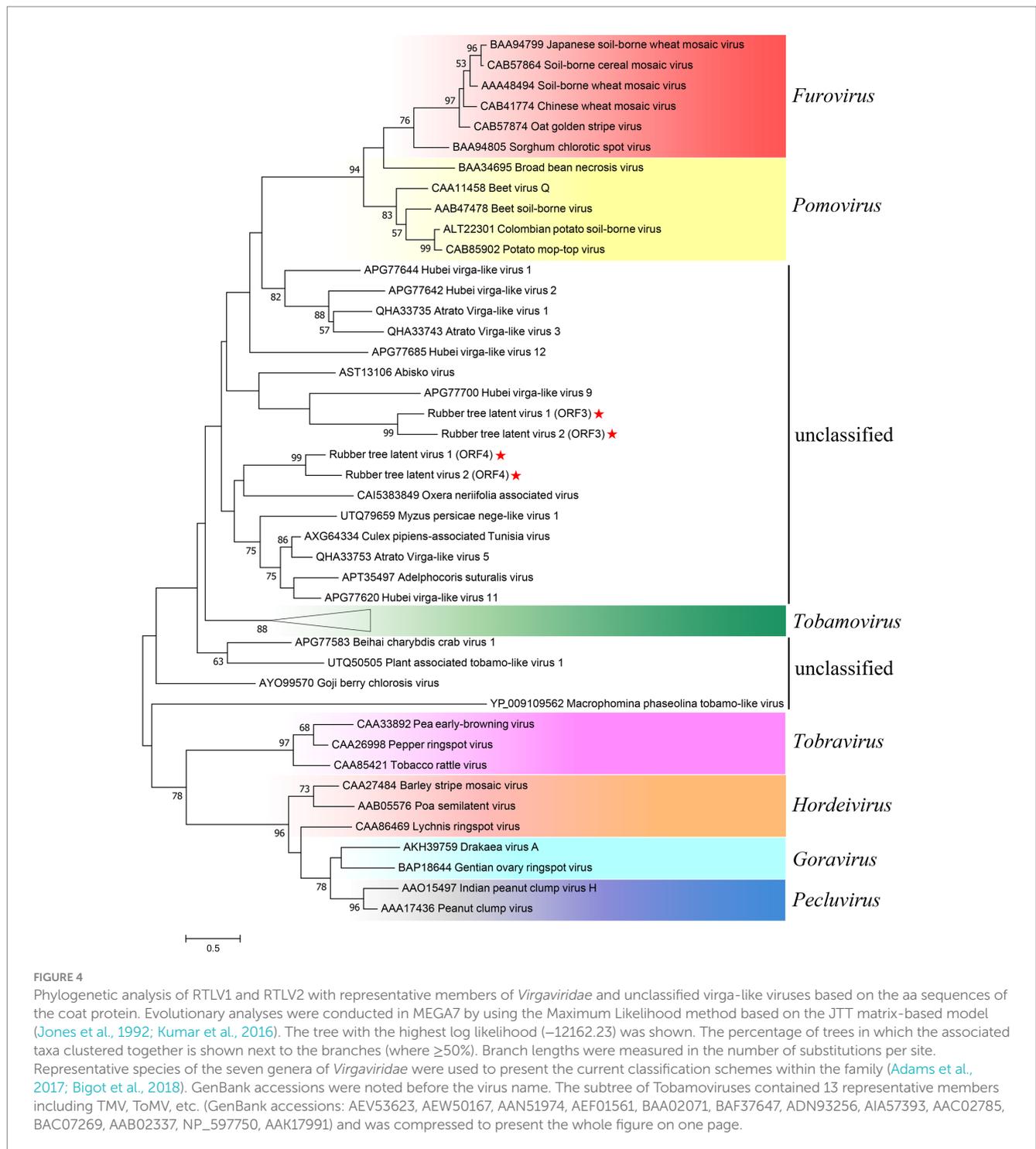


with *Virgaviridae* (Figure 4). CPs from the seven genera of *Virgaviridae* clustered in a genus-specific way, forming distinct monophyletic or paraphyletic groups reconstructing the CP phylogenetic topology within *Virgaviridae* (Adams et al., 2017). Most CPs of unclassified viruses were placed between the *Tobamovirus* and *Furovirus-Pomovirus* clades and could be divided into two groups which contained ORF3 and ORF4, respectively. Nevertheless, bootstrap support for most nodes in this heterogeneous group was low, probably due to the lack of evolutionary links between them. Phylogenetic relationship based on the RdRp domain suggested a common evolutionary origin with families of the order *Martellivirales*, within

which RTLV1 and RTLV2 may potentially constitute a new virus family.

Virus prevalence in rubber tree plantations

A limited survey of prevalence of RTLV1 and RTLV2 was conducted in rubber tree plantations in Qionghai County and Dingan County of Hainan, China. We have previously detected RTLV1 in 35 out of 44 asymptomatic trees collected from Dingan County (Chen et al., 2022). Using the same sample pool, RTLV2 was detected in 32 trees by RT-PCR with two primer sets



(Supplementary Table S1A). A second batch of 30 trees showing no viral symptom were collected from Qionghai County and analyzed following the same protocol. One tree and four trees tested positive for RTLV1 and RTLV2, respectively. Most remarkably, a high rate of co-infection of the two viruses was observed (Table 2). In the Dingnan samples pool, 32 trees (91.4%) out of 35 RTLV1-infected trees were co-infected by RTLV2, and all 32 RTLV2-infected trees were co-infected by RTLV1.

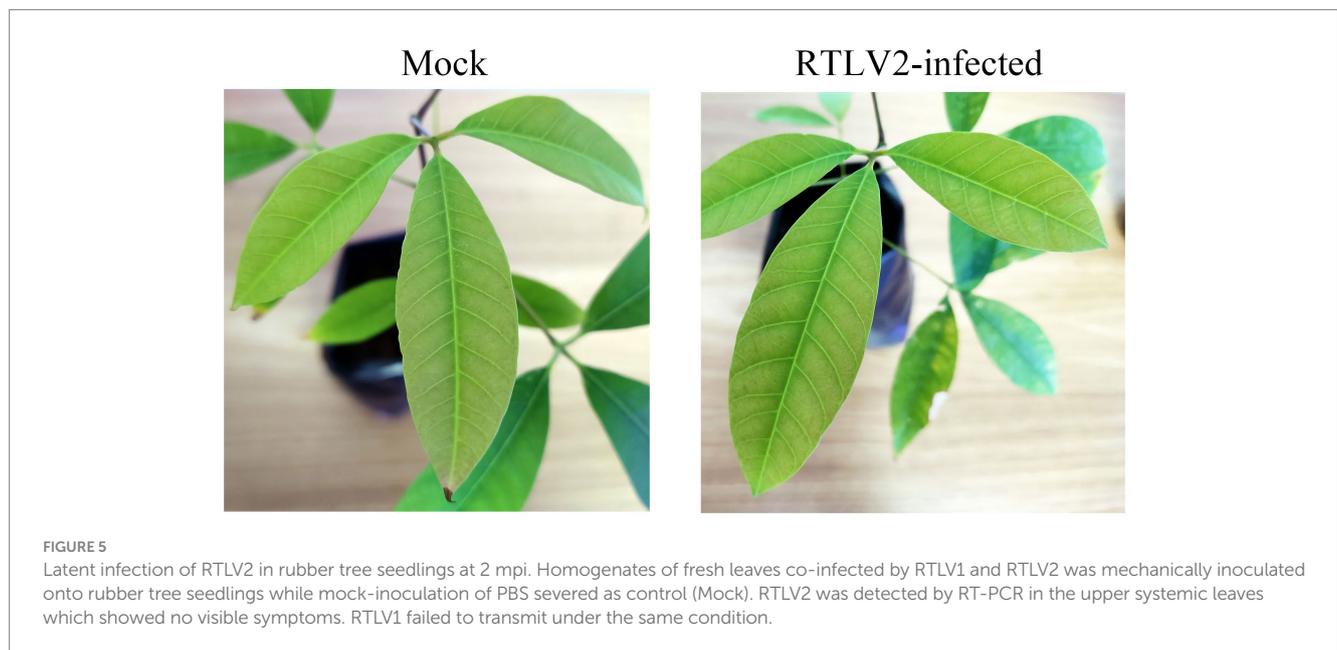
Mechanical transmissibility assay

Mechanical inoculation of RTLV1 and RTLV2 was carried out using homogenates of fresh leaves co-infected by the two viruses. Thirty seedlings of rubber trees tested virus-free and *N. benthamiana* were manually inoculated and continuously observed until detection at 2 mpi, respectively. No visible difference or viral symptom was observed amongst plants inoculated with viruliferous inoculum and

TABLE 2 Limited incidence survey of RTLVL1 and RTLVL2 in commercial plantations in Hainan, China.

Sampling location	Sample size	Uninfected	Infected by RTLVL1	Infected by RTLVL2	Infected by RTLVL1 only	Infected by RTLVL2 only	Co-infected by RTLVL1 and RTLVL2
Dingan	44	9	35	32	3	0	32
		(20.5%)	(79.5%)	(72.7%)	(6.8%)	(0.0%)	(72.7%)
Qionghai	30	26	1	4	0	3	1
		(86.7%)	(3.3%)	(13.3%)	(0.0%)	(10.0%)	(3.3%)

Two batches of 44 and 30 healthy trees showing no viral symptom and continually producing latex were collected from Dingan and Qionghai County of Hainan, respectively, for RT-PCR detection of the presence of RTLVL1 and RTLVL2. Frequent co-infection of the two viruses was observed in the Dingan sample pool. The percentage within brackets was calculated relative to the sample size of each location.



mock-inoculated control plants. Top systemic leaves were sampled for RT-PCR detection and all 30 *N. benthamiana* plants remained uninfected at 2 mpi. Eighteen out of 30 (60%) rubber tree seedlings tested positive of RTLVL2, while no plants tested positive of RTLVL1. The presence of RTLVL2 was further confirmed by Sanger sequencing of the RT-PCR amplicons, which only had a few single nucleotide polymorphisms (SNP) compared with the consensus genome. Close-up observation of newly emerged leaves of RTLVL2-infected plants and mock-inoculated plants confirmed the latent infection of RTLVL2 in rubber tree seedling (Figure 5).

Discussion

The TPD syndrome of rubber trees has been reported for over a century (Chandler, 1922). Numerous studies have proposed TPD as a physiological syndrome (Gebelin et al., 2013; Zhang et al., 2017). However, the search for the causal agent of TPD has persisted. While the identification of TPD-related pathogens has not yielded conclusive results, investigations have unveiled several novel life forms within the rubber tree, including viroids and previously unidentified viruses (Ramachandran et al., 2000; Pellegrin et al., 2004; Li et al., 2020; Chen

et al., 2022). Among these, RTLVL1 and RTLVL2 are two closely related viruses that have been rarely reported in plants and potentially belong to a novel viral family.

Our current understanding of virus biodiversity and classification is limited and fragmented because virology studies have primarily concentrated on disease-causing agents (Shi et al., 2016). Extensive transcriptome mining across multiple phyla has uncovered hundreds of novel viruses that infect animals and plants, some of which exhibit unique features, potentially leading to the establishment of new virus families (Shi et al., 2016; Mifsud et al., 2022). The genomes of RTLVL1 (11,923 nt) and RTLVL2 (11,866 nt) exhibit a similar organization, distinct from the known plant virus families. ORF1 encompasses a viral MTR, HEL, and RdRp domain, conservatively arranged in *Virgaviridae* and related viruses with alpha-like replicases (Koonin and Dolja, 1993). The presence of a 2OG-Fe(II) oxygenase domain and an OTU-like domain between MTR and HEL (Figure 1) unmistakably places RTLVL1 and RTLVL2 in the Hepe-Virga virus group (Shi et al., 2016; Kondo et al., 2019). However, ORF3 and ORF4 were predicted to encode dual TMV-like CPs, contrary to the sole CP in ratified members of *Virgaviridae* (Figures 2, 4). Phylogenetic analysis showed the distinct origins for ORF3 and ORF4 that might be derived from independent acquisitions. The divergence event of RTLVL1 and RTLVL2 might occur

after acquisition of the second CP. Parallel acquisition of multiple copies of structural proteins by the Hepe-Virga viruses is not uncommon, some of which contain triple CPs (Shi et al., 2016). The CP duplex excludes RTL1 and RTL2 from assignment to current genera of *Virgaviridae*. Furthermore, members of *Virgaviridae* have no poly(A) tails and usually possess a tRNA-like at the 3'-terminus, while RTL1 and RTL2 have a poly(A) tail.

RTL1 and RTL2 diverged from current plant virus families and were clustered together with OnV in phylogenetic analyses based on RdRp amino acid sequences (Figure 3). OnV was proposed to constitute a new virus family as its RdRp formed a distinct and well-supported outgroup to related plant virus families (Mifsud et al., 2022). The OnV genome is quite similar to those of RTL1 and RTL2 in terms of structure and phylogeny of coding proteins as all of the five ORFs encoded by OnV were homologous to corresponding ORFs of RTL1 and RTL2. The assembled genome of OnV has not been verified by RACE. Nevertheless, OnV is the closest relative of RTL1 and RTL2 among currently recorded viruses, and together they may represent the existence of a novel, uncharacterized virus group.

High-throughput sequencing pipelines typically produce fragmented viral sequences, sometimes even recombined, partially due to low transcript abundance and intrapopulation viral diversity (Studholme, 2012; Rose et al., 2016). CpATV and HVLV11, two unclassified virga-like viruses, clustered in the same clade with RTL1 and RTL2, were initially reported in a highly truncated form. These reports were later revised with substantial genome extensions of approximately 4 kb (Shi et al., 2016; Debat, 2017; Bigot et al., 2018; Debat and Ribeiro, 2018). In this study, the Trinity assembly of clean reads generated multiple fragments for RTL1, RTL2, and RTV1 (Supplementary Table S2). These fragments served as templates for designing primers used in the amplification of complete genomes. However, the assembled sequences of RTL1 and RTL2 contained a missing gap of 3 kb near the 3' end. This gap led to the previously reported truncated RTL1 genome (Chen et al., 2022). In the initial RACE experiments aimed at verifying the 3' end of RTL1, the 72°C extension was set at 1 min per cycle because the primers were predicted to be within 1 kb upstream of the genomic end. However, this insufficient extension generated amplicons derived from broken ends rather than the intact genomic end. The incomplete genome of RTL1 was corrected in the present work as the missing gap in assembled sequences was found, and RACE experiments were adjusted accordingly. It is worth noting that similar truncated ends were found in the initially reported genomes of CpATV and HVLV11, which are highly AT-rich regions (Debat and Ribeiro, 2018). In the case of RTL1 and RTL2, the upstream of ORF4 and the 3' UTR were AT-rich regions that were mistakenly joined in some of the assembled sequences, resulting in a large missing gap of 3 kb containing ORF4 to ORF7.

Apart from OnV, the rest of the close relatives of RTL1 and RTL2 revealed by phylogenetic analyses were primarily unclassified invertebrate viruses (Figures 3, 4). To exclude potential contamination, a transmissibility assay was conducted, in which RTL2 was successfully inoculated into rubber tree seedlings (Figure 5), indicating that RTL2 is indeed an infectious plant virus. A limited field survey of 44 and 30 trees collected from two counties confirmed the natural occurrence of RTL1 and RTL2, with a high rate of co-infection of the two viruses observed. This co-infection could be a sign of synergistic virus-virus interactions. However, RTL1 failed to

transmit to rubber tree seedlings using an inoculum of homogenates of fresh leaves co-infected by the two viruses, possibly due to an incompatible inoculation buffer or native features of the viruses. Mixed infections of plant viruses are common in nature and lead to various intrahost virus-virus interactions. Mixtures of synergistic and antagonistic interactions, often resulting in unpredictable biological and epidemiological consequences, are likely to occur in plants (Zhang et al., 2001; Syller, 2012). The virus-virus interaction or association between RTL1 and RTL2 requires further detailed studies.

At present, breeding approaches of rubber trees for latex yield improvement is restrained by the lack of efficient gene transforming methods for rubber trees (Priyadarshan, 2017). Viruses have been developed as vector tools for virus-induced gene silencing (VIGS) or gain-of-functions for many economic crops (Pasin et al., 2019; Wang et al., 2020). Tobacco rattle virus (TRV)-derived vector was reported to cause VIGS in rubber tree seedlings (Li et al., 2021), but the work has not been repeated. Application of viral vectors in perennial woody crops is still facing great challenge due to the common instability of foreign inserts in long-term trials (Dawson and Folimonova, 2013; Ding et al., 2018). RTL1 and RTL2 have the potential to be developed as viral vector tools and possess advantages over cross-host viral vectors as they are environmentally native and have no pathogenic effect on rubber trees.

In summary, we have characterized two novel virga-like viruses with distinct genome architectures incompatible with hallmarks of current virus families, and confirmed their natural occurrence in cultivated rubber trees. Our findings contribute to the knowledge of viral diversity in rubber trees and lay the groundwork for developing viral vector tools for rubber tree breeding and functional genomics.

Data availability statement

RNA seq raw data are available through the central BioProject database at NCBI under project accession PRJNA1040157 and BioSamples accessions SAMN38246073, SAMN38246086, SAMN38246087, and SAMN38246088. The complete genome sequences of RTL1 and RTL2 are available at GenBank under accessions OR500510 and OR500511.

Author contributions

RZ: Data curation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. XS: Data curation, Formal analysis, Investigation, Writing – review & editing. FY: Investigation, Writing – review & editing. ZL: Supervision, Writing – review & editing. XH: Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Chinese National Natural Science Foundation (grant no. 31860194).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2023.1286369/full#supplementary-material>

References

- Adams, M. J., Adkins, S., Bragard, C., Gilmer, D., Li, D., MacFarlane, S. A., et al. (2017). ICTV Virus Taxonomy Profile: Virgaviridae. *J. Gen. Virol.* 98, 1999–2000. doi: 10.1099/jgv.0.000884
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Aravind, L., and Koonin, E. V. (2001). The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. *Genome Biology* 2:Research0007. doi: 10.1186/gb-2001-2-3-research0007
- Bigot, D., Atyame, C. M., Weill, M., Justy, F., Herniou, E. A., and Gayral, P. (2018). Discovery of *Culex pipiens* associated Tunisia virus: a new ssRNA(+) virus representing a new insect associated virus family. *Virus Evol.* 4:vex040. doi: 10.1093/ve/vex040
- Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176
- Cao, X., Xu, X., Che, H., West, J. S., and Luo, D. (2019). Three Colletotrichum species, including a new species, are associated to leaf anthracnose of rubber tree in Hainan, China. *Plant Dis.* 103, 117–124. doi: 10.1094/PDIS-02-18-0374-RE
- Chandler, S. E. (1922). The Brown Bast disease of the Para rubber-tree. *Nature* 109, 357–360. doi: 10.1038/109357a0
- Chen, S., Yu, F., Li, Z., Zhang, Y., Wang, H., Zhai, J., et al. (2022). Complete genome sequence of a novel virga-like virus infecting *Hevea brasiliensis*. *Arch. Virol.* 167, 965–968. doi: 10.1007/s00705-021-05306-5
- Davidson, N. M., and Oshlack, A. (2014). Corset: enabling differential gene expression analysis for de novo assembled transcriptomes. *Genome Biol.* 15:410. doi: 10.1186/PREACCEPT-2088857056122054
- Dawson, W. O., and Folimonova, S. Y. (2013). Virus-based transient expression vectors for Woody crops: a new frontier for vector design and use. *Annu. Rev. Phytopathol.* 51, 321–337. doi: 10.1146/annurev-phyto-082712-102329
- Debat, H. J. (2017). An RNA Virome associated to the Golden orb-weaver spider *Nephila clavipes*. *Front. Microbiol.* 8:2097. doi: 10.3389/fmicb.2017.02097
- Debat, H. J., and Ribeiro, J. M. C. (2018). A divergent strain of *Culex pipiens*-associated Tunisia virus in the malaria vector *Anopheles epiroticus*. *Microbiol. Resour. Announc.* 7:7. doi: 10.1128/MRA.01026-18
- Ding, X. S., Mannas, S. W., Bishop, B. A., Rao, X., Lecoulter, M., Kwon, S., et al. (2018). An improved brant mosaic virus silencing vector: greater insert stability and more extensive VIGS. *Plant Physiol.* 176, 496–510. doi: 10.1104/pp.17.00905
- Fonseca, P. L. C., Badotti, F., de Oliveira, T. F. P., Fonseca, A., Vaz, A. B. M., Tomé, L. M. R., et al. (2018). Virome analyses of *Hevea brasiliensis* using small RNA deep sequencing and PCR techniques reveal the presence of a potential new virus. *Virol. J.* 15:184. doi: 10.1186/s12985-018-1095-3
- Gama, M., Kitajima, E., Ávila, A., and Lin, M. (1983). Um Carlavírus em seringueira (*Hevea brasiliensis*). *Fitopatol. Bras.* 3:621.
- Gebelin, V., Leclercq, J., Kuswanhadi, A., Argout, X., Chaidamsari, T., Hu, S., et al. (2013). The small RNA profile in latex from *Hevea brasiliensis* trees is affected by tapping panel dryness. *Tree Physiol.* 33, 1084–1098. doi: 10.1093/treephys/tp076
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883
- Guyot, J., and Le Guen, V. (2018). A review of a century of studies on south American leaf blight of the rubber tree. *Plant Dis.* 102, 1052–1065. doi: 10.1094/PDIS-04-17-0592-FE
- Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275
- Kondo, H., Chiba, S., Maruyama, K., Andika, I. B., and Suzuki, N. (2019). A novel insect-infecting virga-like virus group and its pervasive endogenization into insect genomes. *Virus Res.* 262, 37–47. doi: 10.1016/j.virusres.2017.11.020
- Koonin, E. V. (1991). The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J. Gen. Virol.* 72, 2197–2206. doi: 10.1099/0022-1317-72-9-2197
- Koonin, E. V., and Dolja, V. V. (1993). Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit. Rev. Biochem. Mol. Biol.* 28, 375–430. doi: 10.1019/10409239309078440
- 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. doi: 10.1093/molbev/msw054
- Li, H. L., Guo, D., Wang, Y., Zhu, J. H., Qu, L., and Peng, S. Q. (2021). Tobacco rattle virus-induced gene silencing in *Hevea brasiliensis*. *Biosci. Biotechnol. Biochem.* 85, 562–567. doi: 10.1093/bbb/zbaa085
- Li, Z., Wang, H., Zhao, R., Zhang, Z., Xia, Z., Zhai, J., et al. (2020). Complete genome sequence of a novel capillovirus infecting *Hevea brasiliensis* in China. *Arch. Virol.* 165, 249–252. doi: 10.1007/s00705-019-04459-8
- Limkaisang, S., Kom-un, S., Takamatsu, S., Furtado, E. L., Liew, K. W., Salleh, B., et al. (2005). Molecular phylogenetic and morphological analyses of *Oidium heveae*, a powdery mildew of rubber tree. *Mycoscience* 46, 220–226. doi: 10.1007/S10267-005-0238-8
- Liu, J., Shi, C., Shi, C. C., Li, W., Zhang, Q. J., Zhang, Y., et al. (2020). The chromosome-based rubber tree genome provides new insights into spurge genome evolution and rubber biosynthesis. *Mol. Plant* 13, 336–350. doi: 10.1016/j.molp.2019.10.017
- Liu, Y., Shi, Y. P., Zhuo, D., Yang, T., Dai, L., Li, L., et al. (2023). Characterization of Colletotrichum causing anthracnose on rubber trees in Yunnan: two new records and two new species from China. *Plant Dis.* 107, 3037–3050. doi: 10.1094/PDIS-11-22-2685-RE
- Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., et al. (2020). CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* 48, D265–D268. doi: 10.1093/nar/gkz991
- Melcher, U. (2000). The '30K' superfamily of viral movement proteins. *J. Gen. Virol.* 81, 257–266. doi: 10.1099/0022-1317-81-1-257
- Messick, T. E., Russell, N. S., Iwata, A. J., Sarachan, K. L., Shiekhhattar, R., Shanks, J. R., et al. (2008). Structural basis for ubiquitin recognition by the Otu1 ovarian tumor domain protein. *J. Biol. Chem.* 283, 11038–11049. doi: 10.1074/jbc.M704398200
- Mifsud, J. C. O., Gallagher, R. V., Holmes, E. C., and Geoghegan, J. L. (2022). Transcriptome mining expands knowledge of RNA viruses across the plant kingdom. *J. Virol.* 96:e0026022. doi: 10.1128/jvi.00260-22
- Pasin, F., Menzel, W., and Daros, J. A. (2019). Harnessed viruses in the age of metagenomics and synthetic biology: an update on infectious clone assembly and biotechnologies of plant viruses. *Plant Biotechnol. J.* 17, 1010–1026. doi: 10.1111/pbi.13084
- Pellegrin, F., Nandris, D., Chrestin, H., and Duran-Vila, N. (2004). Rubber tree (*Hevea brasiliensis*) bark necrosis syndrome I: still no evidence of a biotic causal agent. *Plant Dis.* 88:1046. doi: 10.1094/PDIS.2004.88.9.1046C
- Priyadarshan, P. M. (2017). Refinements to Hevea rubber breeding. *Tree Genet. Genomes* 13:20. doi: 10.1007/s12955-017-1101-8
- Ramachandran, P., Mathur, S., Francis, L., Varma, A., Mathew, J., Mathew, N. M., et al. (2000). Evidence for Association of a Viroid with tapping panel dryness syndrome of rubber (*Hevea brasiliensis*). *Plant Dis.* 84:1155. doi: 10.1094/PDIS.2000.84.10.1155C
- Rastrojo, A., Núñez, A., Moreno, D. A., and Alcamí, A. (2018). A new putative Caulimoviridae genus discovered through air metagenomics. *Microbiol. Resour. Announc.* 7:e00955-18. doi: 10.1128/MRA.00955-18
- Rivarez, M. P. S., Pecman, A., Bačnik, K., Maksimović, O., Vučurović, A., Seljak, G., et al. (2023). In-depth study of tomato and weed viromes reveals undiscovered plant virus diversity in an agroecosystem. *Microbiome* 11:60. doi: 10.1186/s40168-023-01500-6
- Rose, R., Constantinides, B., Tapinos, A., Robertson, D. L., and Prosperi, M. (2016). Challenges in the analysis of viral metagenomes. *Virus Evol.* 2:vew022. doi: 10.1093/ve/vew022

- Rozanov, M. N., Koonin, E. V., and Gorbalenya, A. E. (1992). Conservation of the putative methyltransferase domain: a hallmark of the 'Sindbis-like' supergroup of positive-strand RNA viruses. *J. Gen. Virol.* 73, 2129–2134. doi: 10.1099/0022-1317-73-8-2129
- Shi, M., Lin, X. D., Tian, J. H., Chen, L. J., Chen, X., Li, C. X., et al. (2016). Redefining the invertebrate RNA virosphere. *Nature* 540, 539–543. doi: 10.1038/nature20167
- Studholme, D. J. (2012). Deep sequencing of small RNAs in plants: applied bioinformatics. *Brief. Funct. Genomics* 11, 71–85. doi: 10.1093/bfpg/elr039
- Syller, J. (2012). Facilitative and antagonistic interactions between plant viruses in mixed infections. *Mol. Plant Pathol.* 13, 204–216. doi: 10.1111/j.1364-3703.2011.00734.x
- Tang, C., Yang, M., Fang, Y., Luo, Y., Gao, S., Xiao, X., et al. (2016). The rubber tree genome reveals new insights into rubber production and species adaptation. *Nat Plants* 2:16073. doi: 10.1038/nplants.2016.73
- Wang, M., Gao, S., Zeng, W., Yang, Y., Ma, J., and Wang, Y. (2020). Plant virology delivers diverse toolsets for biotechnology. *Viruses* 12. doi: 10.3390/v12111338
- 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. doi: 10.1146/annurev-phyto-080614-120030
- Zhang, X. S., Holt, J., and Colvin, J. (2001). Synergism between plant viruses: a mathematical analysis of the epidemiological implications. *Plant Pathol.* 50, 732–746. doi: 10.1046/j.1365-3059.2001.00613.x
- Zhang, Y., Leclercq, J., and Montoro, P. (2017). Reactive oxygen species in *Hevea brasiliensis* latex and relevance to tapping panel dryness. *Tree Physiol.* 37, 261–269. doi: 10.1093/treephys/tpw106