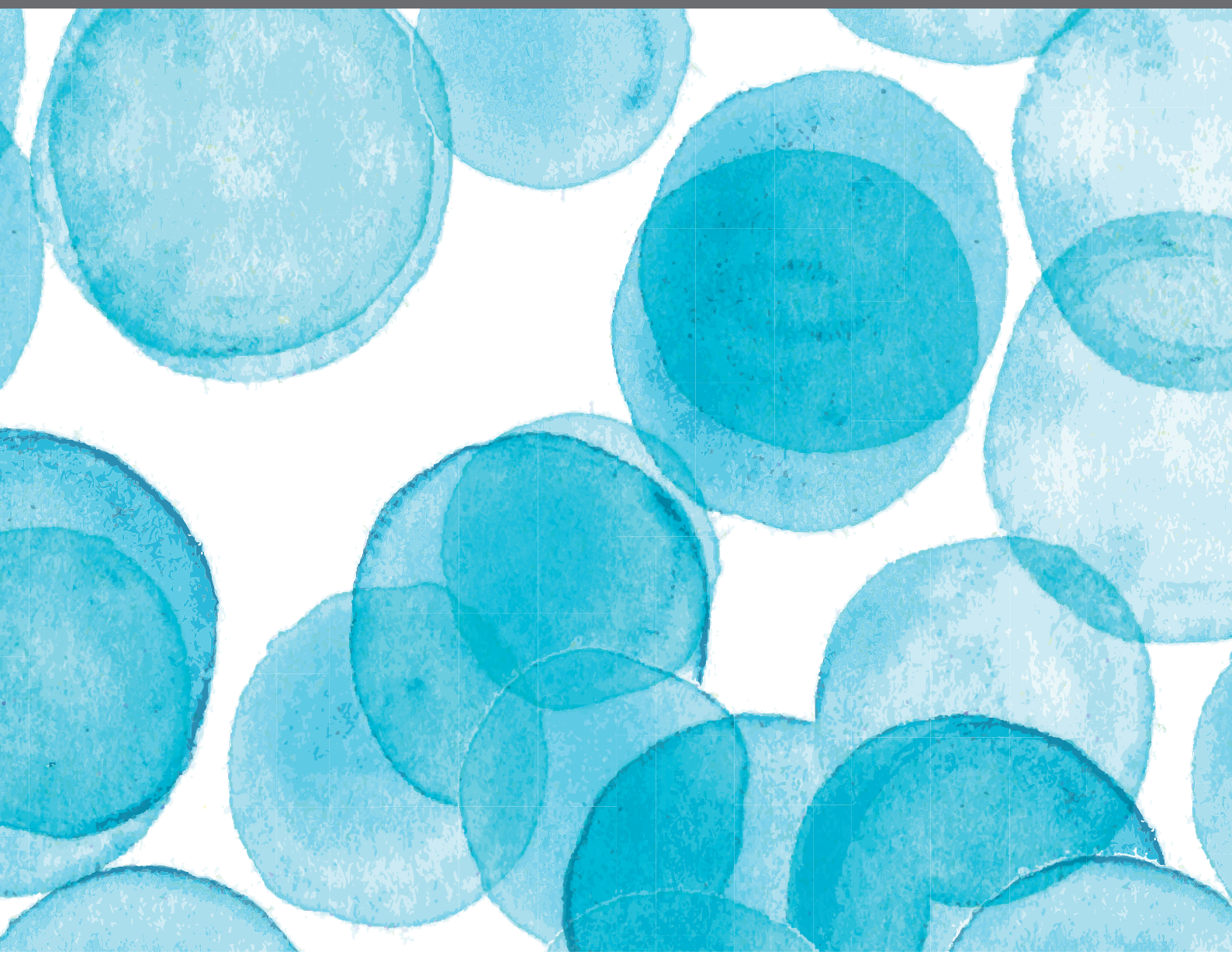
THE BORDER BETWEEN KITAVIRIDS AND NEGE-LIKE VIRUSES: TRACKING THE EVOLUTIONARY PACE OF PLANT- AND ARTHROPOD-INFECTING VIRUSES

EDITED BY: Juliana Freitas-Astua, Nikos Vasilakis, Arvind Varsani, Hideki Kondo, Pedro Luis Ramos-González, Sergey Morozov and Mengji Cao

PUBLISHED IN: Frontiers in Microbiology and Frontiers in Plant Science



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-368-9
DOI 10.3389/978-2-88976-368-9

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

THE BORDER BETWEEN KITAVIRIDS AND NEGE-LIKE VIRUSES: TRACKING THE EVOLUTIONARY PACE OF PLANT- AND ARTHROPOD-INFECTING VIRUSES

Topic Editors:
Juliana Freitas-Astua, Brazilian Agricultural Research Corporation (EMBRAPA), Brazil
Nikos Vasilakis, University of Texas Medical Branch at Galveston, United States
Arvind Varsani, Arizona State University, United States
Hideki Kondo, Okayama University, Japan
Pedro Luis Ramos-González, Biological Institute of São Paulo, Brazil
Sergey Morozov, Lomonosov Moscow State University, Russia
Mengji Cao, Southwest University, China

Citation: Freitas-Astua, J., Vasilakis, N., Varsani, A., Kondo, H., Ramos-González, P. L., Morozov, S., Cao, M., eds. (2022). The Border Between Kitavirids and Nege-Like Viruses: Tracking the Evolutionary Pace of Plant- and Arthropod-Infecting Viruses. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-368-9

Table of Contents

frontiers | Frontiers in Plant Science

EDITORIAL
published: 24 May 2022
doi: 10.3389/fpls.2022.932523

Editorial: The Border Between Kitavirids and Nege-Like Viruses: Tracking the Evolutionary Pace of Plant- and Arthropod-Infecting Viruses

Pedro Luis Ramos-González[1†], Hideki Kondo[2†], Sergey Morozov[3†], Nikolaos Vasilakis[4†], Arvind Varsani[5†], Mengji Cao[6†] and Juliana Freitas-Astúa[7†]*

[1] Applied Molecular Biology Laboratory, Biological Institute of São Paulo, São Paulo, Brazil, [2] Institute of Plant Science and Resources, Okayama University, Kurashiki, Japan, [3] A. N. Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, Russia, [4] Institute for Human Infections and Immunity, University of Texas Medical Branch at Galveston, Galveston, TX, United States, [5] Center for Fundamental and Applied Microbiomics, The Biodesign Institute, Arizona State University, Tempe, AZ, United States, [6] Citrus Research Institute, Chinese Academy of Agricultural Sciences, Beibei, China, [7] Laboratório de Fitopatologia, Embrapa Mandioca e Fruticultura, Cruz das Almas, Brazil

Keywords: ***Kitaviridae***, ***Cilevirus***, ***Higrevirus***, **negeviruses**, ***Blunervirus***, ***Brevipalpus***

OPEN ACCESS

Edited and reviewed by:
Junmin Li,
Zhejiang Academy of Agricultural Sciences, China

***Correspondence:**
Pedro Luis Ramos-González
plrg1970@gmail.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Pathogen Interactions, a section of the journal Frontiers in Plant Science

Received: *29 April 2022*
Accepted: *10 May 2022*
Published: *24 May 2022*

Citation:
Ramos-González PL, Kondo H, Morozov S, Vasilakis N, Varsani A, Cao M and Freitas-Astúa J (2022) Editorial: The Border Between Kitavirids and Nege-Like Viruses: Tracking the Evolutionary Pace of Plant- and Arthropod-Infecting Viruses. Front. Plant Sci. 13:932523. doi: 10.3389/fpls.2022.932523

Editorial on the Research Topic

The Border Between Kitavirids and Nege-Like Viruses: Tracking the Evolutionary Pace of Plant- and Arthropod-Infecting Viruses

Since the disclosure of the citrus leprosis virus C (CiLV-C) genome and the creation of the family *Kitaviridae*, viruses of the eight accepted species within three genera of kitavirids have become intriguing pieces of the plant-infecting virome. With genomes split into two, three, or four single-stranded (+) RNA molecules, most of the kitaviruses produce non-systemic infections that, in the case of citrus leprosis disease, entail serious economic losses to the citrus crop. Phylogenetic analyses based on the RNA-dependent RNA polymerase proteins indicate that kitavirus closest relatives are a group of unclassified arthropod-infecting viruses tentatively known as negeviruses and nege/kita-like viruses.

Contributions gathered in this Research Topic offer data that broaden the boundaries of the diversity of kitaviruses and nege/kita-like viruses, provide, with a special emphasis on the cilevirus CiLV-C, new insight into the interaction between kitavirids and plants, and increase the understanding of the role of some viral open reading frames (ORF) typical to these viruses. They also demonstrate the vectorial activity of mites of the genus *Brevipalpus* for some of the newly described kitaviruses and postulate a hypothesis describing the movement of CiLV-C virions in mites. Particular attention to the forces modulating the population structure of CiLV-C, and the likely evolutionary pathways that could give rise to the diversity of viruses in the family *Kitaviridae* are also addressed.

Five new kitaviruses are reported. The description of three strains of passion fruit green spot virus representing a typical member of the genus *Cilevirus* revealed, however, that their RNA2 molecules have highly variable 5′-ends, with putative orphan ORFs exclusively found in some of these isolates (Ramos-González, dos Santos, et al.). Hibiscus yellow blotch virus, recently accepted as a cilevirus, has a chimeric genome with an RNA1 similar to that of hibiscus green spot virus (HYBV, genus *Higrevirus*), and its RNA2 displays a novel organization in the family, truncated at its 5′-end but incorporating an extra ORF at its 3′-end (Olmedo-Velarde et al.).

The remaining three kitavirids were found infecting ornamental plants and are indicated as tentative cileviruses, but their genomes also show divergences from the recognized members of the genus. The genomic organization of Solanum violifolium ringspot virus, Ligustrum chlorotic spot virus, and Ligustrum leprosis virus resembles that in cileviruses, excluding the absence of the 5′-end region of the RNA2, upstream of the ORF *p61*, reminiscent of the HYBV RNA2 (Ramos-González, Chabi-Jesus, et al.). Studies on virion morphology, transmission by *Brevipalpus* mites, and morphopathology of the infected plant cells complement the molecular characterization of these viruses and improve the integral understanding of the family *Kitaviridae*.

In parallel, a meta-transcriptomic approach to study the virome of 15 aphids populations infesting barley fields in combination with searches on the transcriptome shotgun assembly (TSA) libraries of several arthropods and plants reveal viruses of a dozen of RNA families, including some novel nege/kita-like viruses (Kondo et al.). Barley aphid RNA viruses 1–4 (BARV 1–4) have non-segmented genomes comprising a large ORF (RdRp) followed by two putative structural protein genes for a predicted glycoprotein and the membrane protein SP24, also present in kitavirids. BARV 1–4 cluster with other nege/kita-like viruses forming new genera tentatively designated as Centivirus and Aphiglyvirus. Phylogenetically, they contribute to filling the gaps between the lineages of negeviruses and kitaviruses.

The interplay of plant-kitavirids is addressed in two studies. The global plant response to the infestation by viruliferous *Brevipalpus* mites shows that the CiLV-C infection triggers a progressive reprogramming of the plant transcriptome; namely the activation of the SA-mediated pathway, including ROS burst and HR, and the downregulation of the JA/ET-mediated pathways (Arena et al.). The transient expression of the CiLV-C P61 protein mimics the responses typically observed during CiLV-C-plant interaction indicating that P61 is a putative viral effector causing the HR-like symptoms associated with the infection, likely the outcome of an incompatible interaction. Furthermore, using several strategies to drive the ectopic expression of CiLV-C ORFs, the RNA silencing suppressor (RSS) activity of the viral proteins is investigated (Leastro et al.). Although local RSS activities were not found, the expression of P29 and P15 enhances the pathogenicity of potato virus X (PVX) resulting in the death of the infected tobacco plants. Meanwhile, PVX-p61 infection resulted in HR, corroborating the evidence that the expression of this protein activates plant defense mechanisms.

CiLV-C population study presented in this Research Topic includes the largest collection of samples from leprosis-affected citrus trees ever considered, including herbarium samples collected since the early twentieth century (Chabi-Jesus et al.). Eighteen complete or near-complete viral genomes were recovered, and the examination of the dataset suggests that the CiLV-C population consists of the major lineages CRD and SJP, unevenly distributed in citrus orchards, plus a third one called ASU, represented by a single isolate found in an herbarium sample collected in Asuncion, Paraguay, in 1937. Viruses from these three lineages show signs of inter-clade recombination events. The viral origin by recombination of the 5′-end of the RNA2 in cileviruses and their putative acquisition from a heterologous source is also suggested in a study based on the prediction and computational analyses of 3D protein structures of P15 proteins. These small and poorly conserved proteins are encoded by ORFs unique to cileviruses (Ramos-González, Pons, et al.). Beyond cileviruses, recombination seems to be a common feature in the evolution of kitaviral genomes. The recombinant evolutionary origin of kitaviral genomes is emphasized in a thorough study that includes kitaviruses, nege/kita-like viruses, and also virus-like RNA assemblies found by searching TSA, as the quadripartite blunervirus-like RNA detected in *Paulownia tomentosa* (Morozov et al.).

Finally, the Research Topic presents a paper compiling what is known about cilevirus transmission by *Brevipalpus* mites (Tassi et al.). It describes the presence of CiLV-C particles aligned in the intercellular spaces between adjacent cells at the caeca and the podocephalic gland of mites, which supports the hypothesis of the virion paracellular movement.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

The authors were grateful to CAPES, CNPq, and FAPESP for scholarships, research fellowships, and research grants associated with this work (CAPES PNPD20132154-33141010001P4-PNPD-IBSP, 88882.157041/2017-01, CNPq 312528/2020-5, FAPESP 2017/50222-0, 2017/50334-3, and 2019/25078-9).

ACKNOWLEDGMENTS

The editors thank the authors for their contributions and hope that this Research Topic will support and encourage the expansion of knowledge about kitaviruses, negeviruses, and kita/nege-like viruses.

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 14 February 2020
doi: 10.3389/fmicb.2020.00206

Passion Fruit Green Spot Virus Genome Harbors a New Orphan ORF and Highlights the Flexibility of the 5′-End of the RNA2 Segment Across Cileviruses

Pedro Luis Ramos-González[1], Gustavo Francisco dos Santos[1], Camila Chabi-Jesus[1,2], Ricardo Harakava[1], Elliot W. Kitajima[3] and Juliana Freitas-Astúa[1,4]**

[1] Instituto Biológico, Unidade Laboratorial de Referência em Biologia Molecular Aplicada, São Paulo, Brazil, [2] PPG Microbiologia Agrícola Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, Brazil, [3] Núcleo de Apoio à Pesquisa em Microscopia Eletrônica Aplicada a Agricultura, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, Brazil, [4] Embrapa Cassava and Fruits, Cruz das Almas, Brazil

OPEN ACCESS

Edited by:
Jesús Navas-Castillo, Institute of Subtropical and Mediterranean Hortofruticultura La Mayora (IHSM), Spain

Reviewed by:
Nobuhiro Suzuki, Okayama University, Japan
Beatriz Navarro, Istituto per la Protezione Sostenibile delle Piante, Sede Secondaria di Bari, Italy
Michael Melzer, University of Hawai'i, United States

***Correspondence:**
Pedro Luis Ramos-González
plrg1970@gmail.com
Juliana Freitas-Astúa
juliana.astua@embrapa.br

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *22 October 2019*
Accepted: *29 January 2020*
Published: *14 February 2020*

Citation:
Ramos-González PL, Santos GF, Chabi-Jesus C, Harakava R, Kitajima EW and Freitas-Astúa J (2020) Passion Fruit Green Spot Virus Genome Harbors a New Orphan ORF and Highlights the Flexibility of the 5′-End of the RNA2 Segment Across Cileviruses. Front. Microbiol. 11:206. doi: 10.3389/fmicb.2020.00206

Passion fruit green spot and passion fruit sudden death are two reportedly distinct viral diseases that recurrently affect passion fruit (*Passiflora* spp.) groves in Brazil. Here we used a systematic approach that interconnects symptoms, transmission electron microscopy, RT-PCR detection assays followed by Sanger sequencing, and high-throughput sequencing of the RNA of affected passion fruit plants to gain insights about these diseases. Our data confirmed not only the involvement of cileviruses in these two pathologies, as previously suggested, but also that these viruses belong to the same tentative species: passion fruit green spot virus (PfGSV). Results revealed that PfGSV has a positive-sense RNA genome split into two molecules of approximately 9 kb (RNA1) and 5 kb (RNA2), which share about 50–70% nucleotide sequence identity with other viruses in the genus *Cilevirus*. Genome sequences of five PfGSV isolates suggest that they have more conserved RNA1 (<5% of nucleotide sequence variability) compared to RNA2 (up to 7% of variability) molecules. The highest nucleotide sequence divergence among PfGSV isolates and other cileviruses is in the genomic segment covering from the 5′-end of the RNA2 until the 5′-end of the open reading frame (ORF) *p61*, which includes the ORF *p15* and the intergenic region. This genomic stretch also harbors a novel orphan ORF encoding a 13 kDa protein presenting a cysteine-rich domain. High variability of 5′-end of the RNA2 in cileviruses is discussed in an evolutionary context assuming that they share putative common ancestors with unclassified arthropod-infecting single-strand positive RNA viruses, including mosquito-specific viruses of the group Negevirus (clades Nelorpivirus and Sandwavirus), and other viruses in the family *Kitaviridae.*

Keywords: *Kitaviridae*, Negevirus-like lineage, Nelorpivirus, HTS, arthropod-infecting single-strand positive RNA viruses, *Passiflora* spp.

INTRODUCTION

Passion fruit (*Passiflora* spp.) crops can be severely affected by viral infections that may cause up to 100% of production losses and limit their commercial expansion around the tropical and near-tropical regions of the world (Fischer and Rezende, 2008; Santos et al., 2015; Atukunda et al., 2018). South America supports the greatest collection of *Passiflora* spp. where Colombia and Brazil are

considered as *Passiflora* diversity hot spots (Cerqueira-Silva et al., 2016). Several *Passiflora* spp. are well valued as human food, in the cosmetic industry, for the treatment of some human illnesses, and as ornamentals (Yockteng et al., 2011; Zeraik et al., 2011). In the period 2015–2017, the Brazilian passion fruit harvest represented approximately 65% of the worldwide production (Altendorf, 2018). Passion fruit woodiness disease, caused by a potyvirus, is the major disease affecting the passion fruit crop in that country (Nascimento et al., 2006; Rodrigues et al., 2015; de Oliveira Freitas et al., 2016).

Since the identification in the 1990s, two seemingly different viral diseases known as passion fruit green spot (PfGS) and passion fruit sudden death (PfSD) intermittently occur in passion fruit orchards in Brazil (Kitajima et al., 1997, 2003; Santos Filho et al., 1999). In PfGS disease, spots of approximately 5 mm showing brilliant green borders with or without a central necrotic dip develop on the peel of yellow mature fruits, and necrotic lesions appear on the stems. These stem wounds, which sometimes exhibit deep slits, may also coalesce and encircle the branch leading to the death of the distal end. In leaves, randomly distributed spots intersperse with uneven patches commonly observed alongside the veins. Usually, yellowish foliar lesions arising during the initial stages of the infection gradually turn green island-like spots on the senescent leaves. As a consequence of PfSD disease, affected plants display the major symptoms of PfGS, but they rapidly progress toward death, passing through abundant necrosis, branch death, and finally, the orchard's collapse (Antonioli-Luizon et al., 2009).

Both PfGS and PfSD diseases are transmitted by false-spider mites of the genus *Brevipalpus* (Acari: *Tenuipalpidae*) (Kitajima et al., 2003). Experimental reproduction of PfGS symptoms on leaves and stems was achieved by transferring brevipalpus mites collected from affected field passion fruit plants onto healthy ones (Kitajima et al., 1997, 2003). In addition, the ubiquity of short, bacilliform virus particles (50–70 nm × 100–120 nm) in the cisternae of the endoplasmic reticulum, and viroplasms in the cytoplasm of the infected plant cells suggested the infection by putative cileviruses as their causal agents (Kitajima et al., 2003). From isolated viral dsRNA molecules from PfGS-affected passion fruit plants, a PCR-based molecular detection method allowed the specific detection of a virus tentatively named as passion fruit green spot virus (PfGSV) (Antonioli-Luizon, 2010). This approach revealed the prevalence of PfGSV in PfGS affected plants (Kitajima et al., 2003; Antonioli-Luizon et al., 2009). Amplicon-derived sequences (GenBank accession numbers HM002746 and HM002747) showed < 60% pairwise nucleotide sequence identity to the cilevirus citrus leprosis virus C (CiLV-C) (Antonioli-Luizon, 2010). However, despite the growing body of evidence, PfGSV has not been classified due to the lack of genome sequence data and, consequently, its unclear evolutionary history with other members of the genus *Cilevirus*.

Genus *Cilevirus*, family *Kitaviridae*[1] has viruses with short bacilliform and enveloped particles that encapside two molecules of single-strand positive sense [ss(+)] RNA as genome (Freitas-Astúa et al., 2018). CiLV-C and citrus leprosis virus C2 (CiLV-C2) are the only recognized members of this genus (Locali-Fabris et al., 2006; Roy et al., 2013). Both CiLV-C and CiLV-C2 show a narrow range of natural hosts, produce non-systemic diseases and are persistently transmitted by *Brevipalpus yothersi* mites in a circulative manner (Bastianel et al., 2010; Roy et al., 2015; Ramos-González et al., 2016). RNA genomic molecules of CiLV-C, type-member of the genus, are 3′-polyadenylated. RNA1 molecule, of ∼9 kb, has two open reading frames (ORFs) that encode the RNA-dependent RNA polymerase (RdRp) and the putative 29 kDa capsid protein. The ∼5 kb RNA2 has three ORFs encoding proteins with unknown functions, i.e., the two taxonomically restricted ORFs *p61* and *p24*, and the orphan ORF *p15*, in addition to the *mp* that encodes the putative movement protein (MP). ORFs *p15* and *p61* are separated by a stretch of ∼1000 nts known as the intergenic region (IR). CiLV-C proteins can form homo- and heterodimers, and associate with plant cell membranes producing severe remodeling of the endoplasmic reticulum and the Golgi complex (Leastro et al., 2018).

Phylogenetic analyses including plant-infecting viruses of the families *Kitaviridae, Bromoviridae*, *Closteroviridae*, and *Virgaviridae,* reveal a distant but consistent relationship with an increasing number of arthropod-infecting viruses [ss(+)RNA non-segmented genomes] suggesting that they share a common ancestor (Li et al., 2015; Vasilakis and Tesh, 2015; Shi et al., 2016; Kondo et al., 2019; Vinokurov and Koloniuk, 2019). The genome of some of these viruses, e.g., members of the group Negevirus (proposed genera Nelorpivirus and Sandewavirus) and from the negevirus-like lineage have, besides a large ORF encoding the viral RdRp, the ORF2, and ORF3, which encode a putative glycoprotein and a small membrane-bound protein (Vasilakis et al., 2013; Nunes et al., 2017). These two proteins show structural features, e.g., SP24 (Pfam: 16504), transmembrane domains, signal peptides, which are conserved across their orthologs (P61 and P24 in CiLV-C) in viruses of the family *Kitaviridae* (Kuchibhatla et al., 2014).

In this current work, we describe the identification and genome sequence of viruses belonging to a tentative new species of cilevirus: passion fruit green spot virus (PfGSV). Analysis by high-throughput sequencing (HTS) and Sanger sequencing of RT-PCR amplicons from PfGS and PfSD diseased plants collected in distantly and economically important growing regions in Brazil indicated the ubiquity of PfGSV in these samples. Genome analysis of the viruses revealed the presence of a putative novel ORF in the RNA2 molecule, which lacks homologous sequences in both the genome of known cileviruses and any other documented nucleotide or amino acid sequence. In this regard, it might represent a new ORF that enlarges the list of orphan and taxonomically restricted genes only detected in members of the family *Kitaviridae* and related invertebrate-infecting viruses. In parallel, we revealed the genome sequences of two potyviruses found in mixed infections with PfGSV in two out of the three samples analyzed by HTS. Finally, supported by phylogenetic and ancestral trait reconstruction analyses we hypothesize the putative origin of kitaviruses and other unassigned nege-like viruses from arthropod-infecting common ancestors.

[1] https://talk.ictvonline.org/taxonomy/

MATERIALS AND METHODS

Plant Material

Passion fruit (*Passiflora* spp.) plants showing PfSD symptoms (green spots on fruits, necrotic lesions on stems, and chlorotic strips alongside the leaf veins) were collected in counties near Sinop (Snp) and Nova Fronteira (NFo), both in the State of Mato Grosso, in Central-North Brazil, in 2016 and 2018, respectively; and in Bom Jesus da Lapa (BJL), State of Bahia, in Northeastern Brazil, in 2008 (**Figure 1**). A sample from a passion fruit plant (*Passiflora* sp.) displaying PfGS symptoms was collected in Brasilia (BSB), Distrito Federal, in Central Brazil, in 2003. Samples and viral isolates were identified accordingly with their geographic origins, i.e., Snp1, Snp2, NFo1, BJL1, and BSB1. Straight-line distances between these towns range from 430 to 1396 km (Snp-NFo: 430 km, BSB-BJL: 578 km, Snp-BSB: 880 km, BSB-Snp 934 km, BJL-Snp: 1321 km, and BJL-NFo: 1396 km). Fragments of leaves, stems, and fruits were prefixed in a modified Karnovsky solution (Kitajima and Nome, 1999) for transmission electron microscopy (TEM) analyses as previously described (Chabi-Jesus et al., 2018). Other fragments from the symptomatic samples were conserved at −80°C until RNA extraction.

Detection of Passion Fruit Green Spot Virus

Total RNA was extracted from approximately 100 mg of plant samples using TRIzol® Reagent and following the manufacturer's recommendation (Life Technologies, Foster City, CA, United States). Approximately 500 ng of the RNA templates were used to obtain the cDNA using random hexamer primers and GoScript™ Reverse Transcriptase Kit as described by the manufacturer (Promega, Madison, WI, United States). Three microliters of each cDNA solution were tested by PCR using three pairs of primers for the detection of PfGSV: RNA1 C13F: 5′-ATTCATGCGTTTCACGGTTA-3′, C13R: 5′-CGAATGCCTCTGACACAACT-3′, and PfGSV: RNA2 C6F: 5′-

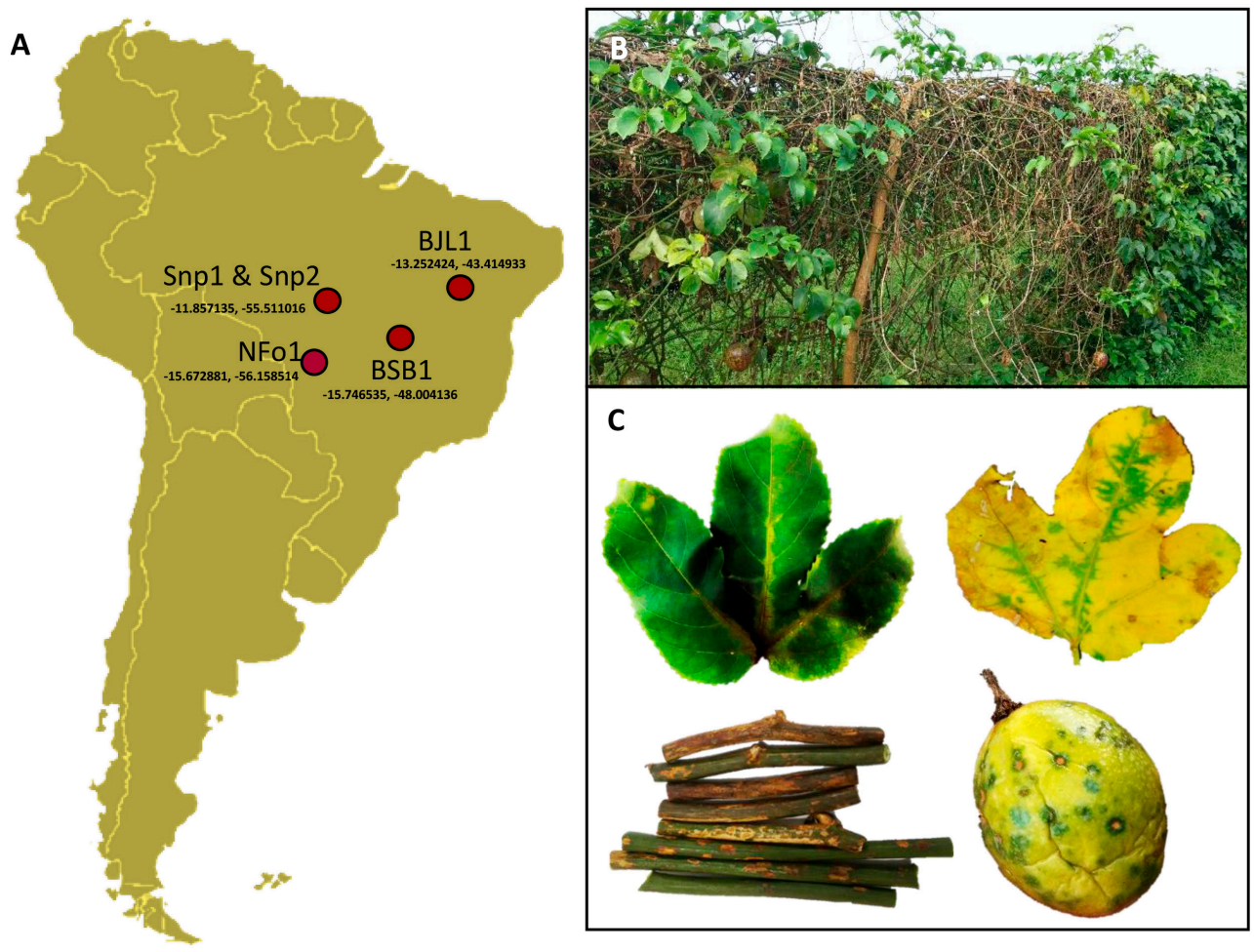


FIGURE 1 | Symptoms of passion fruit green spot virus (PfGSV)-infected passion fruit plants. **(A)** Samples of diseased passion fruit plants (*Passiflora* spp.) were collected in four regions of Brazil. **(B)** Overview of a PfGSV-affected orchard in Sinop, MT, Brazil. **(C)** Details of chlorotic and necrotic symptoms on leaves, stems, and fruits. Note the yellow strips along the veins in green young leaves (left) that turn green areas in a yellow background covering the non-infected areas in the oldest leaves (right).

CGATATTTGATCAATCCGTT-3′, C6R: 5′-CACCTTAAAATT CGAGGGTT-3′, C8F: 5′-TTCATCGCAAGTTCGTATACCT-3′, and C8R: 5′-CTGTTGTGCCAAATCATCAA-3′ (Antonioli-Luizon, 2010). Amplicons were separated on 1% agarose gels in 1X Tris-acetate-ethylenediaminetetraacetic acid (TAE) and visualized with ethidium bromide (0.1 μg/mL). Some amplicons were further purified using Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, United States) and sequenced by the Sanger method.

Sequencing and *in silico* Assembly of Viral Genomes

RNA extracts independently obtained from samples Snp1 (2016), BSB1 and BJL1 were further purified using RNeasy Mini Kit (QIAGEN, Venlo, Netherlands). Quantification and estimation of the A260/A280 ratio were carried out using a NanoDrop ND-8000 micro-spectrophotometer (Thermo Scientific, Waltham, MA, United States). Five hundred nanograms of each final extract were sent to the Animal Biotech Laboratory at Escola Superior de Agricultura Luiz de Queiroz, University of São Paulo (Piracicaba, Brazil) for HTS using HiSeq 2500 Technology (2 × 150 nt paired-end reads) (Illumina, San Diego, CA, United States). Preparation of the HTS libraries was carried out as previously defined (Ramos-González et al., 2017). Read assembly up to the generation of the viral-sequence-containing contigs were done by using both SPAdes (Bankevich et al., 2012) available in Geneious software package version 11.1.4 (Kearse et al., 2012) and Trinity (Haas et al., 2013). Contigs were annotated using BlastX, implemented in the Geneious software, against both viral genome database[2] and a customized local database including the sequences of kitavirids: cileviruses, blunerviruses (Quito-Avila et al., 2013; Hao et al., 2018), and the higrevirus hibiscus chlorotic green spot virus 2 (HGSV-2) (Melzer et al., 2013a). Those contigs producing the best E-value score (∼0) with the cilevirus RNA1 and RNA2 molecules of either CiLV-C (GenBank accession numbers DQ352194 and DQ352195) or CiLV-C2 (JX000024 and JX000025) were selected. A set of 32 primers that generate overlapping amplicons were designed with the program Primer3 (Untergasser et al., 2012) and taking as reference the largest contigs corresponding to each genomic segment of the Snp1 viral isolate (**Supplementary Table S1**). These primers and the total RNA of the sample Snp1 were subsequently used to validate the HTS and to obtain the sequences of the 5′-ends of both genomic molecules of the viral isolate Snp1 using a RACE SMARTer® RACE 5′/3′ Kit (Clontech Laboratories, Mountain View, CA, United States). For RACE analyses, amplicons were purified, ligated in pGEM-T-easy vector (Promega, Madison, WI, United States), and further sequenced with universal primer pairs M13F/M13R by the Sanger method (Instituto Biológico, SP, Brazil). At least five clones per each type of amplicons were sequenced. Snp1-specific primers were also used to generate the near-complete genome of the viruses in samples Snp2, collected in 2018, and NFo1. Sanger method-generated sequences of the amplicons were assembled using CAP3 assembly (Huang and Madan, 1999) implemented in Geneious (Kearse et al., 2012). Sequences of the assembled genomes were submitted to the GenBank database and assigned to the following GenBank accession numbers: PfGSV_Snp1 RNA1: MK804171 and RNA2: MK804172, PfGSV_BSB1 RNA1 MK804173: and RNA2 MK804174, PfGSV_BJL1 RNA1: MK804175 and RNA2: MK804176, PfGSV_Snp2 RNA1: MN746810 and RNA2: MN746811, and PfGSV_NFo1 RNA1: MN746812 and RNA2: MN746813.

[2]ftp://ftp.ncbi.nih.gov/genomes/Viruses/all.fna.tar.gz

Northern Blot Assays

Ten micrograms of denatured RNA from both PfGSV_Snp1-infected and healthy passion fruit plants were separated in 1% agarose gel containing formaldehyde. The assay also included RNA extracts from CiLV-C-infected and healthy sweet orange plants. Non-radioactive probes were generated using the PCR DIG-labeling kit (Roche Diagnostics, Mannheim, Germany) according to the manufacturer's instructions. Each probe was obtained in independent reactions adding 1 ng of DNA fragments corresponding to each PfGSV_Snp1 or CiLV-C_CRD ORF (*p29* or *p24*) and the appropriate primer pairs. DNA labeling was verified comparing the amplicon sizes of DIG PCR products with the unlabeled DNA fragments. Hybridization and detection were carried out as recommended by the Dig Luminescent Detection Kit manual (Roche Diagnostics, Mannheim, Germany). The presence of bands in the blots was revealed using the BCIP-NBT substrate.

Genome, Protein, and Phylogenetic Analyses

Viral ORFs were identified *in silico* by the ORF finder[3]. Conserved domain architecture, presence of signal peptide, and predictions of transmembrane helices, subcellular localization and isoelectric points of viral proteins were done by using SPARCLE[4] (Marchler-Bauer et al., 2017), Pfam v 32.0 EMBL-EBI[5], SignalP 5.0 (Almagro Armenteros et al., 2019), TMHMM Server 2.0[6] (Sonnhammer et al., 1998), Deeploc v1 (Almagro Armenteros et al., 2017), and ExPASy Compute pI/Mw tool[7]. The presence of coiled-coil regions was detected using the neural network-based tool DeepCoil (Ludwiczak et al., 2019). Database searching for remote protein homology detection was carried out using the MPI Bioinformatics Toolkit (HHblits, HHpred, and HMMER)[8] (Hildebrand et al., 2009; Finn et al., 2011; Remmert et al., 2012; Zimmermann et al., 2018). Nucleotide and predicted amino acid sequences were aligned using the CLUSTAL algorithm (Chenna et al., 2003) implemented in MEGA X version 10.0.5 (Kumar et al., 2018) and MAFTT (Katoh and Standley, 2013), respectively. Additionally, alignments and predicted secondary structures of deduced amino acid sequences of the putative viral proteins were obtained using PROMALS (PROfile Multiple Alignment with predicted Local Structure)

[3]https://www.ncbi.nlm.nih.gov/orffinder/
[4]https://www.ncbi.nlm.nih.gov/sparkle
[5]https://pfam.xfam.org/
[6]http://www.cbs.dtu.dk/services/TMHMM/
[7]https://web.expasy.org/compute_pi/
[8]https://toolkit.tuebingen.mpg.de

(Pei and Grishin, 2007) and PRALINE (Profile ALIgNmEnt) (Bawono and Heringa, 2014).

For phylogenetic analyses, besides the sequences from known kitaviruses and negeviruses, others from non-classified ss(+)RNA viruses were retrieved from GenBank after a BLAST search (cut-off E-value $< e^{-10}$) using PfGSV and Negev virus (NC_030294.1) sequences as the query. Some plant-infecting viruses of the family *Virgaviridae* were included in the study as an external group. Phylogenetic informative regions of the multiple sequence alignments (MSAs) were selected using BMGE software (Criscuolo and Gribaldo, 2010) implemented in https://ngphylogeny.fr (Lemoine et al., 2019). Substitution models with the lower Bayesian information criterion scores for each MSA were determined using MEGA X version 10.0.5 (Kumar et al., 2018). For the MSA of the RdRp proteins, the sequences of the domains methyltransferase and helicase, and RdRp encoded by the RNA1 and RNA2 molecules, respectively, of the blunerviruses blueberry necrotic ring blotch virus strains Giorgia and RL (BNRBV_Giorgia and BNRBV_RL, respectively) and tea plant necrotic ring blotch virus (TPNRBV) were concatenated as previously studied (Quito-Avila et al., 2013). Phylogenetic relationships were inferred by using MrBayes v3.2.6 implemented in Geneious prime software package version 2019.2.3 (Kearse et al., 2012). MCMC convergence was obtained for four independent runs of 6 million generations, which were sufficient to obtain a proper sample for the posterior probability, assessed by effective sample sizes (ESSs) above 200. The posterior probabilities of the clades were determined by a 50% majority consensus of the trees retained. Trees were edited and visualized using Interactive Tree Of Life (iTOL) v4 (Letunic and Bork, 2019).

Ancestral trait reconstruction was carried out using the Bayesian Binary MCMC (BBM) analysis implemented in RASP v4.2 (Yu et al., 2019). The 14,142 post-burn-in trees from the RdRp Bayesian Inference analysis using MrBAYES were input into RASP to estimate the probabilities of ancestral hots at each node on the condensed tree generated by RASP. Hosts were coded as Arthropods, Plants, Anemones, and Nematodes. MCMC chains were run simultaneously for 5,000,000 generations and the reconstructed state was sampled every 1,000 generations. The fixed model JC + G (Jukes-Cantor + Gamma) was used for the BBM analysis.

RESULTS

Identification of Passion Fruit Green Spot Virus in Symptomatic Passion Fruit Plants

The occurrence of PfGSV was assessed in passion fruit plants showing symptoms ascribed to the PfGS and PfSD diseases, i.e., circular necrotic spots on fruits, and chlorotic and necrotic lesions on their leaves and stems (**Figures 1B,C**). The presence of the suspected virus was first verified by RT-PCR assays using the primer pairs C6, C8, and C13 (Antonioli-Luizon, 2010). All the samples tested positive with at least one of the primer pairs used, and amplicons of the expected sizes were visualized (**Supplementary Figure S1**). However, the plant organ used for the RNA extractions apparently affected the efficiency of these tests. Only a quarter of the symptomatic stem samples were RT-PCR positive whereas 100% of the leaf samples were positive (**Supplementary Figure S1**). Amplicons obtained using the C6 and C13 primer pairs showed 98–99.2% nucleotide sequence identity with fragments of the PfGSV genome (GenBank accession numbers HM002746 and HM002747). In addition to the molecular detection, virus-like particles with a bacilliform shape of approximately 50 nm width × 100 nm length were observed in the cytoplasm of parenchymal cells of the passion fruit leaves of the sample Snp1 by TEM (**Figure 2**). Presumed

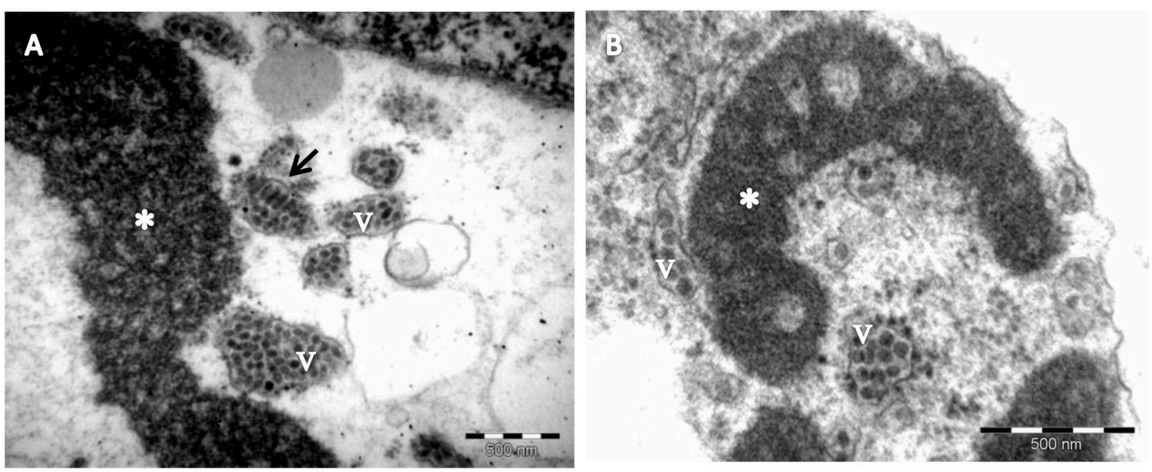


FIGURE 2 | Electron micrograph of passion fruit green spot virus (PfGSV)-infected passion fruit leaves. **(A,B)** Micrographs of thin sections of tissues from green patches on senescent passion fruit leaves, collected in commercial fields near Sinop, State of Mato Grosso, Brazil. Large, electron-dense and vacuolated inclusions (viroplasms) (*) are present in the cytoplasm of parenchymal cells. Short, bacilliform particles, presumed PfGSV virions, can be seen accumulating within cisternae of the endoplasmic reticulum, either in cross- (v) or longitudinal (arrow) sections.

viral particles were always found in the lumen of cytosolic vesicles whereas electron-dense and vacuolar viroplasms occupied large areas of the cytoplasm.

Passion Fruit Green Spot Virus Genome Resembles That of Known Cileviruses

Genome sequences of five isolates of PfGSV were obtained from distinct diseased plants. Isolates Snp1, BSB1, and BJL1 were assembled from HTS reads and Snp2 and NFo1 by Sanger sequencing of overlapping amplicons obtained by RT-PCR.

HTS libraries from the BSB1 and BJL1 isolates contained the highest PfGSV RNA matched reads with 3.9 and 4.2%, respectively, out of the total recovery reads (**Supplementary Table S2**). Snp1's library contained the lowest percentage of viral-derived reads (2.8%), probably indicating a lower viral load in that sample. Regardless of the origin of libraries, the numbers of reads corresponding to the RNA2 molecules surpassed those derived from the RNA1 molecules. The minimum value of mean coverage per viral base reached 9,369X and it corresponded to the RNA1 molecule of the Snp1 isolate. The 5′-ends of the RNA1 and RNA2 molecules of the Snp1 isolate were obtained by RACE.

Excluding the poly-A tails at their 3′-end, the genome of PfGSV_Snp1 consists of 13,523 splits into two molecules designated as RNA1 and RNA2 with 8,740 and 4,783 nts, respectively (accession numbers MK804171 and MK804172). Near-complete RNA1 molecules of the isolates BSB1 (MK804173), BJL1 (MK804175), Snp2 (MN746810), and NFo1 (MN746812) include 8,694; 8,717; 8,541 and 8,554 nts; whereas their near-complete RNA2 consist of 4,752; 4,769, 4,578 and 4,545 nts (MK804174, MK804176, MN746811, and MN746813), respectively. Shorter molecules in the isolates BSB1, BJL1, Snp2, and NFo1 may be a direct consequence of the strategies deployed during their sequencing. Despite this, when compared with PfGSV_Snp1 molecules, more than 95% of the genome of each isolate was recovered.

Pairwise comparisons of the genomic molecules of the isolate Snp1 with those from Snp2, NFo1, and BSB1 revealed a high nucleotide sequence identity (RNA1 > 99.1% and RNA2 > 97.3), whilst with the BJL1 isolate the values were slightly lower (RNA1: 95.5% and RNA2: 93.3%) (**Table 1**). Due to the high nucleotide sequence identity between the PfGSV isolates, Snp1 will be considered the type variant of PfGSV. Therefore, further analysis will be focused on Snp1, and also on isolates BSB1, and BJL1, which collectively were collected from distant geographic regions and whose genomes were more extensively obtained.

Seven putative ORFs with more than 200 nts were detected in the PfGSV genomes, two in the RNA1 and five in the RNA2 (**Figure 3A**). ORFs *p15* and *p13*, in the 5′-end of the RNA2, encompass the largest discrepancies among the coding nucleotide sequences of BJL1 with the other isolates. Comparisons of PfGSV_Snp1 with known cileviruses revealed the highest nucleotide and deduced amino acid sequences identities with CiLV-C2 isolates Hw and Co (**Table 1**). At the genomic level, the highest identity values corresponded to RNA1 molecules (∼72.0%); while the RNA2 segments are more divergent (∼64.0%), even though they harbor the *p24* ORF that is the most conserved ORF across cileviruses. In the comparisons with the CiLV-C2 isolates, P29, P15, and P61 proteins accounted for the lower levels of amino acid identities. In general, regardless of the evaluated gene, deduced protein or virus, the lowest values of identity always corresponded to the comparisons between PfGSV and CiLV-C.

From the seven putative polypeptides encoded by the genome of PfGSV, the RdRp (RNA1) is the largest one, encoding a 2506 aa protein (285.9 kDa). In PfGSV_Snp1 this protein shows the viral methyltransferase domain (Pfam: PF01660) located near to the NH_2-terminus of the protein. RdRp also includes the viral helicase_1 (Pfam: PF01443) and the RdRP_2 (Pfam: PF00978) domains, both localized in the second half of the protein, being the RdRP_2 motif the nearest to the COOH-terminus. P29_Snp1 (RNA1: 258 aa; 26.6 kDa), the putative coat protein, although with a relatively low level of amino acid identity with isolates of the citrus infecting-cileviruses (<33% with CiLV-C and <65% with CiLV-C2, **Table 1**), shows stretches of highly conserved residues among cileviruses. Eighteen out of the first 32 amino

TABLE 1 | Nucleotide (nt) and deduced amino acid (aa) identities of passion fruit green spot virus (PfGSV).

PfGSV_Snp1	PfGSV_BSB1	PfGSV_BJL1	PfGSV_Snp2	PfGSV_NFo1	CiLV-C_CRD	CiLV-C_SJP	CiLV-C2_Co	CiLV-C2_Hw
	nt/aa[a]	nt/aa	nt/aa	nt/aa	nt/aa	nt/aa	nt/aa	nt/aa
RNA1	99.3/–	95.5/–	99.3/–	99.1/–	60.0/–	59.8/–	71.5/–	**72.0**/–
RdRp	99.4/99.8	95.1/99.5	99.3/99.7	99.1/99.7	61.1/57.8	60.8/57.8	72.5/81.0	**73.1**/**81.0**
p29	99.1/99.2	97.5/100.0	99.4/100	99.4/100	48.8/32.7	49.9/31.7	**62.8**/**63.4**	62.0/60.2
RNA2	99.2/–	93.3/–	97.3/–	97.3/–	52.6/–	53.0/–	**63.7**/–	62.8/–
p15	99.7/99.2	87.0/83.3	96.4/95.4	93.9/84.7	48.0/14.1	46.7/14.1	**60.3**/**56.7**	57.2/55.6
p13/p11	99.7/100.0	88.8/36.9	98.8/98.2	99.7/99.1	–	–	**–**	**–**
IR[b]	99.0/–	91.3/–	98.7/–	98.6/–	37.6/–	37.4/–	**44.8**/**–**	**44.8**/**–**
p61	99.1/99.4	92.1/94.8	96.8/98.0	98.7/98.9	50.0/31.6	51.1/31.8	63.7/59.3	**63.9**/**60.4**
mp	99.4/99.7	97.6/99.3	97.5/99.7	96.7/99.3	56.2/51.2	56.3/48.8	**71.5**/**74.3**	68.0/70.5
p24	99.7/100.0	96.5/100.0	96.8/99.5	95.7/99.5	63.8/60.8	64.6/59.6	**78.2**/**88.0**	77.2/86.1

Comparisons between PfGSV_Snp1 and the isolates BSB1, BJL1, Snp2, NFo1, and the citrus-infecting cileviruses. The highest values are highlighted in bold. [a]*Percentage of nucleotide sequence identity/amino acid sequence identity.* [b]*IR: Intergenic region as identified in CiLV-C. This region encompasses the nucleotide sequences between 3′-end of p15 ORF and 5′-end of p61 ORF in the RNA2 molecule.*

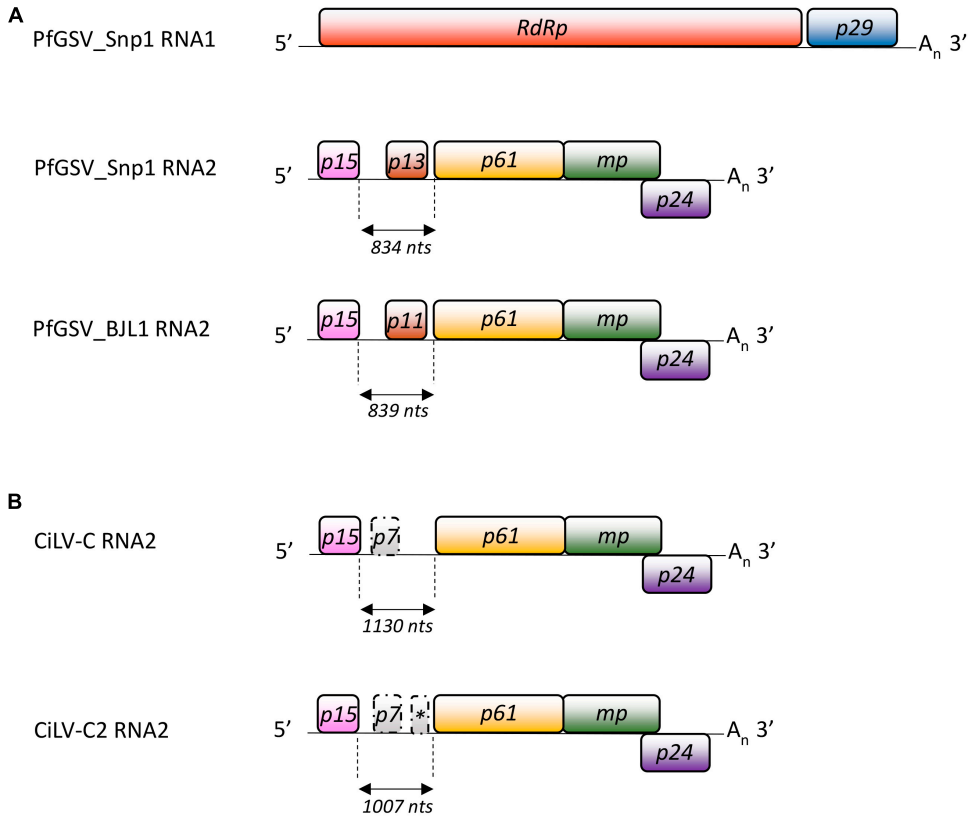


FIGURE 3 | Genomic structure of passion fruit green spot virus. Graphic representation of **(A)** the PfGSV-Snp1 genome and its comparisons with the RNA2 molecules from PfGSV_BJL1 and **(B)** the cileviruses CiLV-C and CiLV-C2. Solid bars indicate the ORFs. *RdRp*: RNA-dependent RNA polymerase; *p29*: putative coat protein; *MP*: putative movement protein; *p15*, *p13*, *p11*, *p61,* and *p24*: proteins with unknown functions. ORFs *p7* and that identified with an asterisk in the IR of CiLV-C and CiLV-C2 have been poorly described. A_n: poly(A) tail.

acids at the NH_2-terminus of these proteins are conserved, while five other positions display residues with similar biochemical characteristics (**Supplementary Figure S2**). Alignment of P15 (RNA2: 131 aa; 14.8 kDa) from the PfGSV isolates with those from cileviruses displays 20 invariable amino acid residues. Some of them are conserved Cys residues likely essential for the P15 functioning, as previously described for CiLV-C (Ramos-González et al., 2016). Nonetheless, P15_Snp1 has an amino acid identity of < 15% in comparison with its cognates from CiLV-C strains, being, in general, the most divergent protein among cileviruses (**Table 1**).

The largest protein encoded by the RNA2 is the putative glycoprotein P61 (543 aa; 61.0 kDa). *In silico* analyses of P61_Snp1 revealed the presence of a signal peptide, which is likely cleaved between the positions 25 and 26, residues Ser-Lys-Gly/Arg-Phe. Besides three putative N-glycosylation sites at positions 249, 310, 340 in the mature peptide, P61_Snp1 also shows two adjacent transmembrane domains located near the COOH-terminus of the protein (positions 442–466 and 471–498). Similar structural features are conserved through the P61 of the isolates BJL1 and BSB1. MP_Snp1 (RNA2: 295 aa; 32.6 kDa) analysis revealed the presence of the 3A motif (Pfam: 00803) involved in the cell-to-cell movement of some plant-infecting viruses of the 30K superfamily (Melcher, 2000). This protein also has the typical "D motif", which occurs in the MPs encoded by a diverse group of plant viruses (Mushegian and Elena, 2015).

P24 protein is the most conserved among those encoded by cileviruses. P24_Snp1 (207 aa; 23.8 kDa) has 88.0 and 60.8% amino acids sequence identity, respectively, with its cognate proteins from CiLV-C2_Co and CiLV-C_CRD (**Table 1**). All the studied isolates of PfGSV has the SP24 motif (Pfam: 16504), which, besides in cileviruses, is found in several arthropod-infecting viruses such as chroparaviruses, negeviruses, negevirus-like viruses, and in plant-infecting viruses of the genera *Higrevirus* and *Blunervirus* (Kuchibhatla et al., 2014; Nunes et al., 2017; Viljakainen et al., 2018). The alignment of P24 proteins from cileviruses reveals a less conserved NH_2-terminus (approximately first 30 aa) that in the case of P24_Snp1 also shows a high density of positively charged residues (five Arg and three Lys) (**Supplementary Figure S3**). P24_Snp1 has several helices that might determine four transmembrane domains: TM1 (residues 46–62), TM2 (82–100), TM3 (112–130), and TM4 (150–169). Particularly, residues comprised between positions 48 and 59 are predicted to form a coiled-coil domain (**Supplementary Figures S3C,D**).

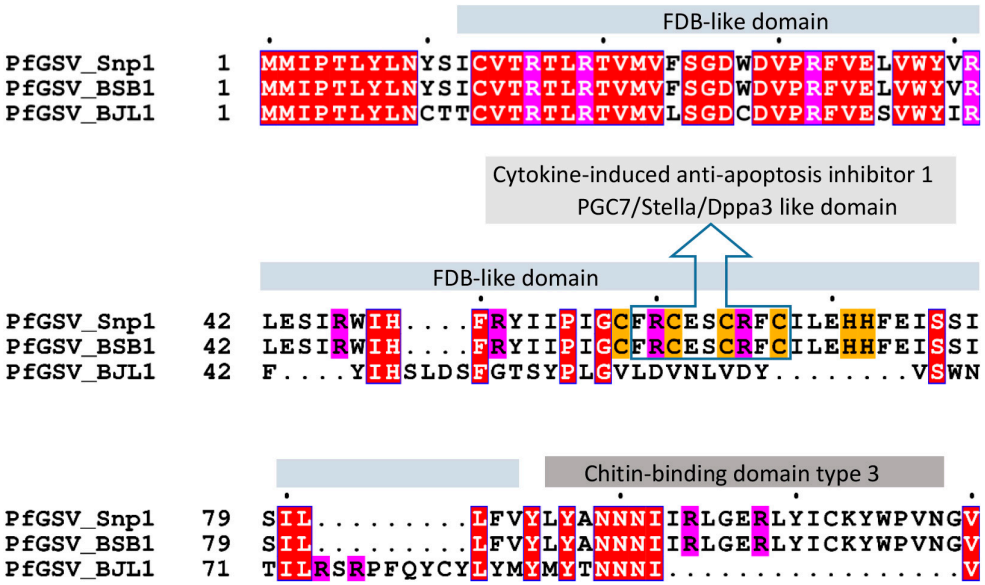


FIGURE 4 | Deduced amino acid sequence alignment of proteins P13 and P11 from passion fruit green spot virus isolates Snp1, BSB1, and BJL1. Solids boxes on top of alignments label domains with remote identity in proteins P13, as identified by *HHblits*, *HHpred*, and *Smart* tools. White letters represent conserved amino acids across the three proteins. Magenta and mustard backgrounds highlight arginine (R), cysteine (C), and histidine (H) residues.

The Intergenic Region in the RNA2 From PfGSV Is Shorter Than Those of Known Cileviruses and Harbors a Novel Orphan ORF

In their RNA2 molecules, between the ORFs *p15* and *p61,* the cileviruses CiLV-C and CiLV-C2 have a stretch of non-coding nucleotide sequences known as the IRs. In the PfGSV strains Snp1, BJL1, BSB1, these regions are 834, 839, and 834 nts in length, respectively (**Figure 3A**). Following the trend observed in the comparisons among the ORFs and deduced proteins encoded by the PfGSV genomes (**Table 1**), the nucleotide sequence identity between IRs from the isolates Snp1 and BSB1 are higher (99.0%) than that from the isolates Snp1 and BJL1 (91.3%). When compared with other cileviruses, the stretch of nucleotide sequences between *p15* and *p61* in PfGSV isolates are 17–35% shorter than those displayed in CiLV-C2 (strains Co: 1007 nts, Hw: 1078 nts, and Fla: 1048 nts), and 25–43% smaller than those in CiLV-C (strains CRD: 1130 nts and SJP: 1135 nts) (**Figure 3B**). The alignments of regions between the ORFs *p15* and *p61* from PfGSV, CiLV-C and CiLV-C2 show, at best, 45% nucleotide sequence identity (**Table 1**).

Besides the relatively low sequence identities, the stretch between the ORFs *p15* and *p61* in the PfGSV isolates Snp1, Snp2, NFo1, and BSB1 display an ORF (*p13*) with 330 nts that encodes a putative polypeptide of 110 residues, highly conserved, and ∼13 kDa (**Figure 3A**) (**Table 1**). It is noteworthy that due to the existence of the ORF *p13* the use of the term IR to designate the stretch of nucleotide sequences between *p15* and *p61* in PfGSV, at least conceptually, is no longer conceivable.

P13 protein has no apparent homology with any protein available in public databases. A thorough analysis using the *HHblits*, *HHpred,* and *Smart* tools revealed amino acid motifs that may be essential for the understanding of the putative origin and function of this protein. P13, estimated pI: 8.59, displays 11 positively charged residues (10 Arg + 1 Lys) and four regularly distributed Cys residues in a stretch of 15 amino acids that also contains two His residues: **Cys**$_{58}$-Phe-Arg-**Cys**-Glu-Ser-**Cys**-Arg-Phe-**Cys**-Ile-Leu-Glu-**His**-**His**$_{72}$ (**Figure 4**). This sequence shows a slight identity with motifs identified in the cytokine-induced anti-apoptosis inhibitor 1 (CIAPIN1, PF05093), and PGC7/Stella/Dppa3-like domain (PGC7_Stella, PF15549). The Cys-rich motif is embedded within a larger region that, according to *Smart,* shows a remote identity with the FDB domain found in FBox and BRCT domain-containing plant proteins (Doerks et al., 2002). A sequence of 26 residues in the COOH-terminal region of P13 shows a moderate identity with the Chitin-binding domain type 3 (InterPro IPR003610). In PfGSV_BJL1, the locus of its putative ORF *p13* is shorter. With only 282 nts, it encodes a smaller protein of 94 aa, ∼11.1 kDa, that shares almost the same sequence over the first 40 residues with P13_Snp1 (**Figure 4**).

The Pattern of Subgenomic RNAs From PfGSV Resembles That Resultant From CiLV-C Transcription

Total RNA extract from the PfGSV_Snp1-infected plant was analyzed by Northern blot assay. The test also included RNA extract from CiLV-C-infected sweet orange leaves whose detected bands were used as molecular weight markers. Hybridizations were carried out using probes derived from the ORFs *p29* and *p24* of each virus. The identity of bands in each profile was assigned after a comparison between our results and those

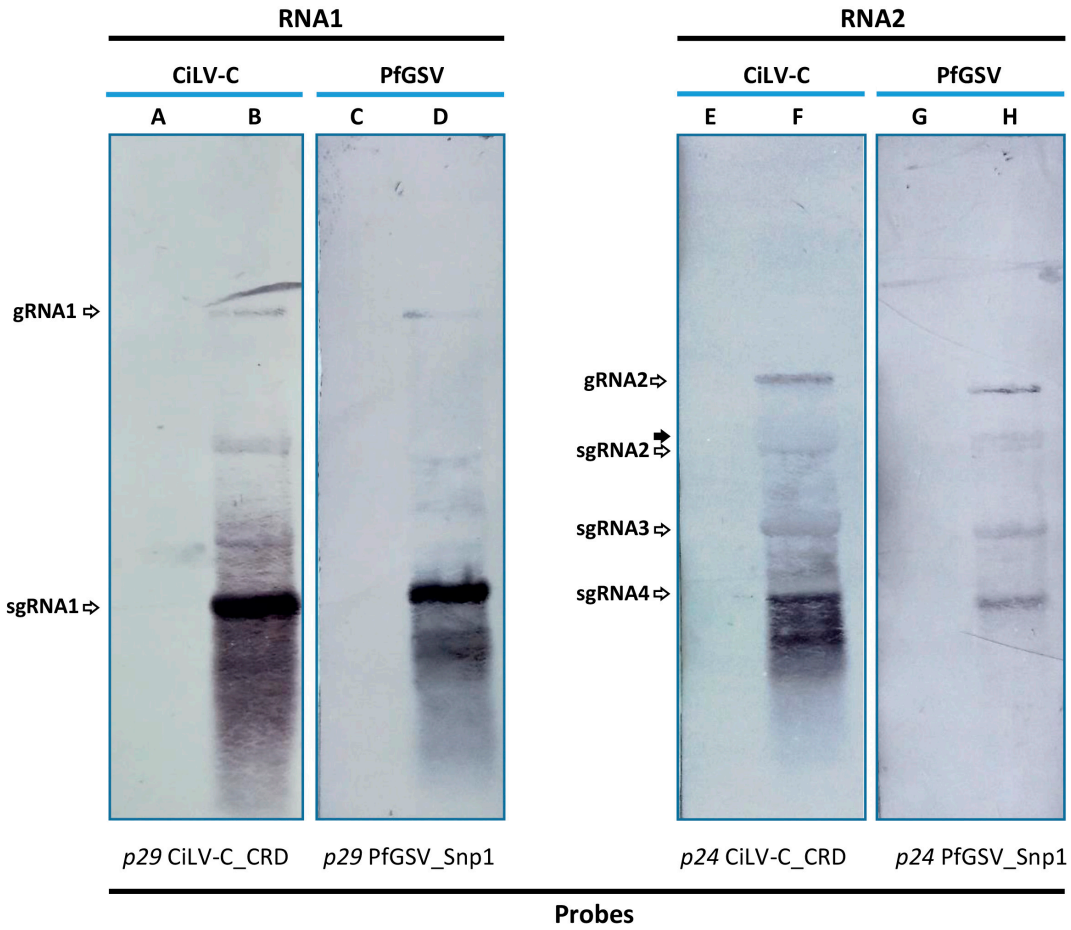


FIGURE 5 | Passion fruit green spot virus RNA's profile in infected plants. Northern blot hybridization using 10 μg of total RNA extracts from PfGSV-infected passion fruit plant **(D,H)**, which was compared with that of CiLV-C in sweet orange plant **(B,F)**. Digoxigenine (DIG-11-dUTP)- labeled probes were individually generated by PCR using specific primers for *p29* and *p24* ORFs from CiLV-C isolate CRD and PfGSV isolate Snp1. In lanes **A** and **C**, and **E** and **G** 10 μg of total RNA extracts from healthy sweet orange and passion fruit plants were assayed, respectively. The solid arrow indicates a band in the lane H that may correspond to sgRNA molecules for the putative expression of P13 protein.

shown in a Northern blot analysis of CiLV-C previously obtained (Pascon et al., 2006).

Two bands were identified in the blots corresponding to the RNA1: gRNA1 and sgRNA1 (**Figure 5**). Relative amounts of these bands indicated much more accumulation of that corresponding to the sgRNA1, which harbors the ORF *p29* according to the probe used in this test. In blots hybridized with the *p24*-derived probes, besides the band of the gRNA2, three other bands were identified as sgRNA2, sgRNA3, and sgRNA4. sgRNA4, the most abundant and with the lowest molecular weight of the three bands, comprises the ORF *p24* region. sgRNA2 and sgRNA3 likely act as the mRNA for the expression of the ORFs *p61* and *mp*, respectively. New hybridizations using an ORF *p13* -specific probe need to be carried out to confirm whether a weak band with a molecular weight slightly higher than, and as abundant as, the sgRNA2 band in the PfGSV blots is the sgRNA from which the P13 protein may be expressed (**Figure 5**, lane H). No signal was detected in any of the lanes where RNA extracts from healthy plants were run (lanes A, C, E, and G). Other bands that remained unidentified seem to be the results of either partial degradation or defective products derived from the viral replication or transcriptional processes.

Combined Phylogenetic and Ancestral Reconstruction Analyses Suggest That Direct Ancestors of Kitavirids Likely Colonized Arthropods

Phylogenetic analyses of the RdRp and MP sequences revealed a closer relationship between PfGSV and strains of CiLV-C2, CiLV-C, and, although to a lesser extent, with other viruses in the family *Kitaviridae* (**Figure 6** and **Supplementary Figure S4**). In the RdRp tree, viruses of the genera *Cilevirus* and *Higrevirus* shared the same clade with Saiwaicho virus, an unassigned monopartite virus found in *Drosophila suzukii* (Medd et al., 2018), while blunerviruses divided their clade with Beihai barnacle virus 2, an unassigned monopartite found infecting a crustacean (Shi et al., 2016). These two clades besides the clade

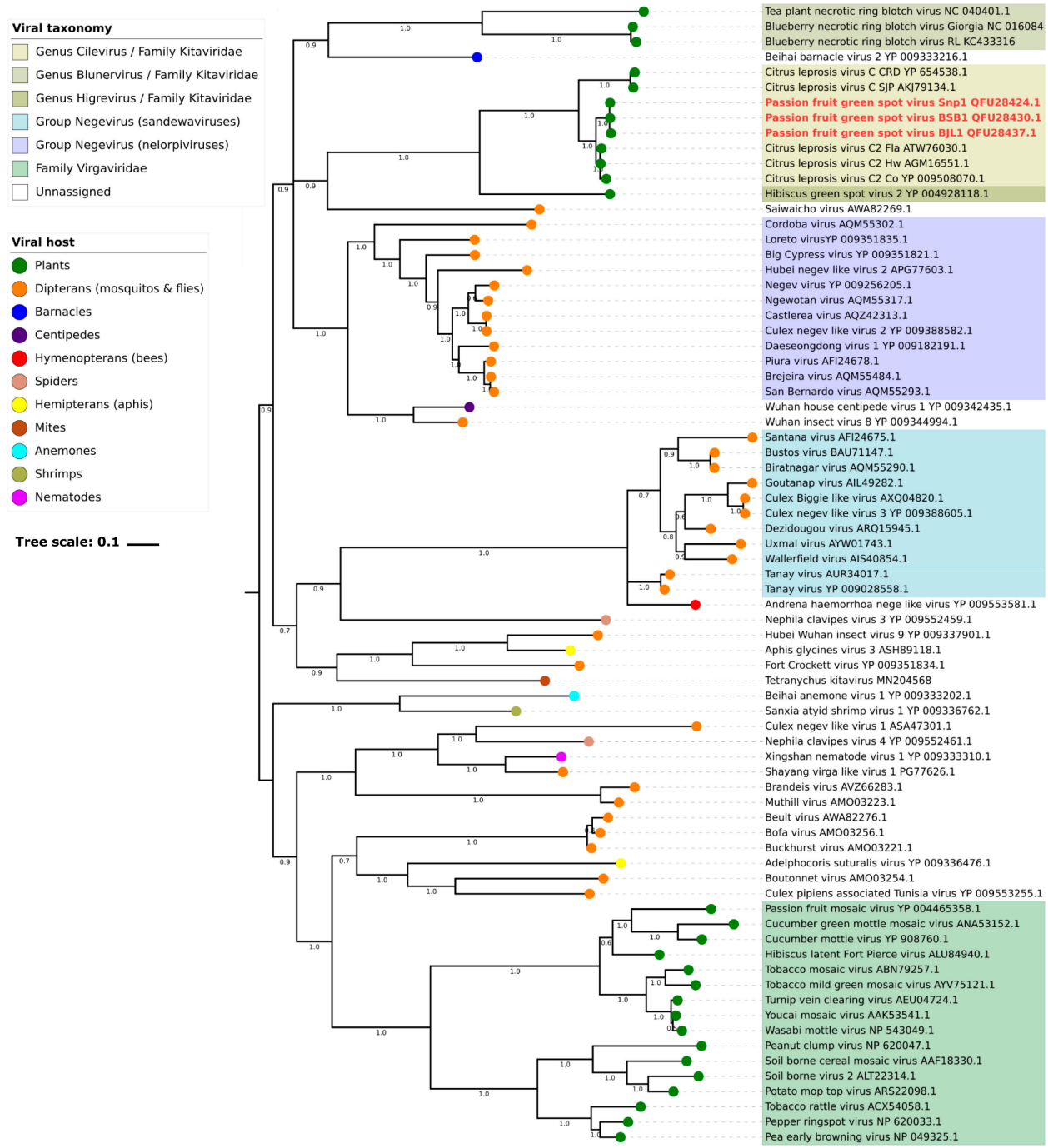


FIGURE 6 | Molecular phylogenetic analysis of plant-infecting viruses of the family *Kitaviridae*. The midpoint-rooted Bayesian maximum clade credibility tree was inferred using a Markov Chain Monte Carlo (MCMC) of 6 million generations using the amino acid sequences of RNA dependent RNA polymerase (RdRp) of members of the family *Kitaviridae* (genera *Cilevirus*, *Higrevirus,* and *Blunervirus*), arthropod-infecting viruses, and viruses of the family *Virgaviridae*. Dataset included 443 positions and its evolutionary history was inferred based on the model LG + G + I + F (Le and Gascuel, 2008). Figures next to nodes indicate values of posterior probability branch support. Scale bar indicates the average number of amino acid substitutions per site. Passion fruit green spot virus sequences are highlighted in red.

comprising nelorpiviruses and two other unassigned ss(+)RNA viruses also found in arthropods, partook a polytomy, suggest either simultaneous speciation events or, more likely, absence of enough data to figure out how those lineages are related. Regardless of this, the inferred phylogenetic tree based on RdRp suggested the probable existence of common ancestors between

plant-infecting kitaviruses and a number of arthropod-infecting viruses. In the MP tree, which as expected only included plant-infected viruses, a closer phylogenetic relationship was observed between cileviruses, then with the blunerviruses, and finally with the furoviruses oat golden stripe virus and sorghum chlorotic spot virus (**Supplementary Figure S4**).

Ancestral trait reconstruction analysis based on the consensus RdRp phylogenetic tree indicated the likely hosts of most recently ancestors of kitaviruses (**Figure 7**). Nodes 84 and 101, which represent the common ancestors between viruses of the genera *Cilevirus-Higrevirus* and *Blunervirus*, respectively, with their closest arthropods-infecting relatives, are supported by more than 90% of trees, and they most likely infected arthropods (95%), as inferred by the RASP analysis. Similarly, the analysis of node 102, representing the most recent common ancestor of nelorpiviruses (group Negevirus) and kitaviruses, also indicated with a high probability (>99%) an arthropod-infecting virus. Node 98, although represented in the topology of the consensus tree, is not taken into account due to its relatively low support value (42%).

Mixed Infections of Potyviruses and PfGSV in Passion Fruit Plants Affected by PfGS Disease

The near-complete genome sequences of two potyviruses, soybean mosaic virus (SMV) and cowpea aphid-borne virus (CABMV), were assembled from reads obtained from the analyses of the samples collected in Brasilia and Bom Jesus da Lapa, respectively. The genome of SMV_BSB1 (GB accession number MN124783) comprises 9,510 nts and shows the highest nucleotide sequence identity (97.7%) with SMV strain G4 (FJ640979) (Seo et al., 2009). Translated polyprotein of SMV_BSB1, 3,069 residues, shares 97.85% amino acid sequence identity with the SMV_Gulupa found infecting passion fruit plants in Colombia (Jaramillo-Mesa et al., 2018). The genome of CABMV_BJL1 (GB accession number MN124782) consists of 9,936 nts and encodes a polyprotein of 3,182 aa. Comparisons of CABMV_Snp1 with known potyviruses revealed both the best nucleotide (88.61%) and deduced amino acid (92.14%) sequence identities with CABMV_MG-Avr (HQ880243) (Barros et al., 2011). In the analysis of the HTS library obtained from the sample collected in Sinop, virus sequences other than that from PfGSV were not detected.

DISCUSSION

We have revealed the genome of five novel bipartite viruses found in passion fruit plants affected by either passion fruit green spot (PfGS) or passion fruit sudden death (PfSD) diseases. Our results not only complement previous works that suggested the presence of cileviruses as causal agents of these pathologies (Kitajima et al., 2003; Antonioli-Luizon, 2010), but also indicate that these viruses belong to the same tentative cilevirus species, previously named as PfGSV (Kitajima et al., 1997). In addition, we detected mixed infections between PfGSV and the potyviruses SMV and CABMV in two out of the three samples analyzed by HTS. Accordingly, in the current discussion relevant aspects involving PfGSV-caused diseases emanates from the evaluation of the symptoms showed by the sample Snp1, in which PfGSV was the only virus detected in the HTS analysis, and by other passion fruit plants experimentally infected by PFGSV using viruliferous brevipalpus mites under controlled conditions in a greenhouse (*unpublished results*).

Both the shape and size of detected virus particles, as well as the cytopathic effects observed in the cells of the diseased passion fruit tissues, matched with those previously noted in PfSD-affected plants (Kitajima et al., 2003) and, in general, with those described for cileviruses, and the anatomy of the cells affected by their infections (Freitas-Astúa et al., 2018). PfGSV genomes encompass the six typical cilevirus ORFs subdivided into two RNA molecules with ∼9 and 5 kb (**Figure 3**) (Freitas-Astúa et al., 2018). Besides the two genomic RNA molecules, four sgRNA molecules were detected in passion fruit infected plants (**Figure 5**), whereas, a fifth sgRNA, from which P13 protein might be translated, still needs more in wet analyses to confirm its identity. Despite this, the array and number of PfGSV-specific sgRNAs, particularly those derived from its RNA2, resemble more the sgRNA expression pattern revealed by CiLV-C (Pascon et al., 2006), than that shown by CiLV-C2, whose RNA2 sgRNA profile comprises only two molecules (Roy et al., 2013). Moreover, comparisons of nucleotide and deduced amino acid sequences from PfGSV with the cognates from CiLV-C and CiLV-C2 (**Table 1**) validated that PfGSV isolates are distinct from known viruses and, taken together, support their definitive assignment to a new species of the genus *Cilevirus*, family *Kitaviridae*.

RNA2 of PfGSV isolates Snp1, Snp2, NFo1, and BSB1 harbors a seventh ORF (*p13*) (**Figure 3**) potentially encoding a protein lacking recognizable homologs. P13 displays 10 residues of Arg, intermingled with other amino acids, rendering a positively charged protein that might account for its likely interaction with either proteins or nucleic acids (Chandana and Venkatesh, 2016). P13 also has Cys residues that make it resemble small cysteine-rich proteins (CRPs) encoded by hordei-, furo-, tobra-, beny-, peclu-, and carlaviruses (Koonin et al., 1991; Yelina et al., 2002). However, unlike those CRP encoded in the 3′-termini of viral genomes, ORF*p13* is placed in the 5′ half of the RNA2 molecule probably suggesting its different transcriptional regulation and accumulation pattern. In-depth bioinformatic analyses of P13 protein indicated it harbors motifs showing the best, although weak, identities with those in proteins involved in transcriptional regulation, e.g., FDB and PGC7/Stella/Dppa3 -like domains (**Figure 4**). Remarkably, besides P13, P15 also shows Cys residues, which likely contribute to a putative Zn-finger structure (Ramos-González et al., 2016). Studies with CRP from, e.g., tobacco rattle virus and beet necrotic yellow vein virus revealed that Cys residues are essential for protein stability and/or their activity as RNA silencing suppressors (Chiba et al., 2013; Sun et al., 2013; Fernández-Calvino et al., 2016). While P13 expression from the RNA2 of PfGSV_Snp1 still needs to be proven, a thought-provoking question is what would be the likely selective advantage provided by this gene. Clearly contrasting with CiLV-C and CiLV-C2, PfGSV has been identified naturally infecting more than a dozen species of

FIGURE 7 | Ancestral state reconstructions performed by the Bayesian Binary MCMC analysis as implemented in RASP v4.2 using the RdRp MrBayes rooted tree. Pie charts at each node represent ancestral host estimations. Each node is internally identified with a number and the figures immediately at their right indicate their frequency in the consensus tree. The probabilities of particular hosts in major nodes are as follows: Node 84 – Arthropods (Arth) 96.15%, Arth + Plants (P) 3.40%, P 0.26%; Node 101– Arth 95.89%, Arth + P 3.55%, P 0.31%; Node 102 – Arth 99.08%, Arth + P 0.80%.

plants belonging to nine families (*unpublished data*), and their lesions in infected plants are less restricted than those produced by the citrus-infecting cileviruses (Freitas-Astúa et al., 2018; Ramos-González et al., 2018) (**Figure 1C**). It should be noted that the putative allele of the *p13* ORF in the isolate BJL1 (BJL1 allele) produces a truncated version of the protein with ∼11 kDa. Further variability studies including more than 30 isolates are underway in our laboratory aiming to assess the prevalence of the *p13* alleles in the PfGSV population. Besides its presence in four out of the five isolates included in this study, preliminary results indicate the predominance of the Snp1 type (*unpublished data*).

A broader examination of the first third of the PfGSV RNA2 molecule highlights the variability of this region in PfGSV and in general across cileviruses. Considering PfGSV as a reference, this stretch of nucleotide sequence (*i*) contains two divergent alleles for the *p13* ORF (represented by the alleles Snp1 or BJL1), (*ii*) is 17–43% shorter than those in other cileviruses, (*iii*) includes the *p13* ORF that is absent in CiLV-C and CiLV-C2, and (*iv*) accommodates the *p15* ORF, the less conserved among the cileviruses (**Table 1**). Variability between the RNA2 IR in CiLV-C and CiLV-C2 has been previously emphasized (Melzer et al., 2013b; Roy et al., 2013; Ramos-González et al., 2016). CiLV-C and CiLV-C2 harbor a putative small ORF encoding a 7 kDa protein whose position inside the IR reveals a contrasting structural array between the IR of these two viruses. Early studies on the IR led us to hypothesize the occurrence of continuous illegitimate (non-homologous) recombination processes inter- or intra-species of cileviruses (Ramos-González et al., 2016).

Phylogenetic inferences using RdRp confirmed previous reports indicating close genetic and evolutionary links between plant-infecting viruses of the family *Kitaviridae* and several arthropod-infecting viruses (**Figure 6**) (Vasilakis et al., 2013; Kallies et al., 2014; Nunes et al., 2017; Kondo et al., 2019; Vinokurov and Koloniuk, 2019). Particularly, the analysis of the RdRp tree indicated that kitavirids might have emerged via a divergent lineage from which other plant-infected viruses, e.g., members of the family *Virgaviridae,* have likely evolved, further underpinning that the diversity of arthropod viruses is the source from where several plants and animal viruses evolved (Shi et al., 2016; Dolja and Koonin, 2018). Through combined analyses based on the RdRp tree and the genome structure of known negeviruses, negevirus-like viruses, and members of the family *Kitaviridae*, it might be speculated that their putative direct ancestors colonized arthropods and displayed unsegmented ss(+)RNA genome with a poly-A tail and three major ORFs arranged in the following order: 5′-ORF1 (RdRp)-ORF2 (putative glycoprotein)-ORF3 (SP24 motif-containing protein)-3′. While circulating negeviruses and negevirus-like viruses keep this general genome structure (Vasilakis et al., 2013), kitavirids show segmented genomes where the orthologs of ORF1 and ORFs 2–3 lie in distinct RNA molecules. In particular, cileviruses might epitomize one of the outgrowths of this evolutionary process in which the high-variable 5′-end of the RNA2, placed upstream the *p61* ORF (putative glycoprotein), represents a block of sequence that lacks in the genome of the known negeviruses and negevirus-like viruses.

The shaping of current cilevirus genomes might have undergone the action of evolutionary forces biased not only by their plant hosts but also by their vectors. Horizontal gene transfer among viruses, or even between viruses and plants, has been suggested as a suitable mechanism for the acquisitions of new genes by plant virus genomes, as properly explained in the case of the binary movement block genes of the kitavirus hibiscus green spot virus 2 (Solovyev and Morozov, 2017). Moreover, blunerviruses and cileviruses encode a 30K MP displaying their best identities with those from virgavirids (**Supplementary Figure S4**) from which it could likely have been integrated by recombination (Quito-Avila et al., 2013). These two examples suggest that gene acquisition processes could have happened before the diversification of cileviruses and blunerviruses or, otherwise, they came about more than twice during the evolution of current kitavirids. This latter scenario seems to fit better with the speciation processes from two different nodes suggested by the RdRp-based analyses (**Figures 6**, **7**).

Moreover, brevipalpus mites, the known vectors of cileviruses, also act as hosts and vectors of viruses grouped in the genus *Dichorhavirus,* one of the two genera of rhabdoviruses showing segmented genomes (Dietzgen et al., 2018). Genome segmentation of an ancestral mononegavirus into dichorhaviruses has been suggested as likely linked to their unique vector (Kondo et al., 2017). If any, monopartite nege-like viruses infecting brevipalpus mites have not been detected yet. However, besides a number of negeviruses (Agboli et al., 2019), Saiwaicho virus (Medd et al., 2018), aphis glycine virus 3 (Feng et al., 2017), and *Tetranychus urticae* kitavirus (Niu et al., 2019), have been recovered from herbivore arthropods. In such a way, the role of arthropods acting as a bridge between ancestral nege-like viruses and plants seems to be probable.

In this work, we have revealed the genomic sequence of viruses belonging to a putative new species of cilevirus. Deeper studies on molecular variability and prevalence of PfGSV isolates, degree of susceptibility of passion fruit cultivars to virus infection and vector infestation, specificity of the virus-mite interactions, relationship between dynamic of mite populations, e.g., density and seasonal fluctuations, and plant symptom severity will provide us with robust tools to detect and control the infection by PfGSV and to reveal the factors that determine the outbreak of two different diseases with the same etiological agent. Our current study has expanded our knowledge about cileviruses since PfGSV is only the third virus of the genus with a complete genome revealed and, so far, the only one able to infect passion fruit plants. This latter fact is of particular significance because differently from CiLV-C and CiLV-C2, whose primary host is sweet orange (*Citrus sinensis*) trees introduced in America only after the colonization ca. 500 years ago, the mean host of PfGSV is native to South America (Ocampo Pérez and Coppens d'Eeckenbrugge, 2017). If, as it is speculated, South America is the center of origin of brevipalpus-transmitted viruses (Freitas-Astúa et al., 2018), a group that besides the cileviruses includes the ss(−)RNA bipartite viruses of the genus

Dichorhavirus (Dietzgen et al., 2018), virus-host co-evolutionary history of PfGSV is much larger than that of the CiLV-C or CiLV-C2. More importantly, they might shed light on the complex evolutionary processes involving kitavirids and related arthropod-infecting viruses. Comprehensively, high variability involving the presence of orphan or taxonomically restricted genes in the RNA2 of cileviruses launch these viruses as an excellent system to study the mechanisms determining the birth of new genes, as well as, the role of these genes enabling the viral fitness to novel conditions (Khalturin et al., 2009; Schlötterer, 2015; Johnson, 2018; Van Oss and Carvunis, 2019).

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the PfGSV_Snp1 RNA1: MK804171 and RNA2: MK804172, PfGSV_BSB1 RNA1: MK804173 and RNA2: MK804174, PfGSV_BJL1 RNA1: MK804175 and RNA2: MK804176, PfGSV_Snp2 RNA1: MN746810 and RNA2: MN746811, PfGSV_NFo1 RNA1: MN746812 and RNA2: MN746813, SMV_BSB1: MN124783, and CABMV_BJL1: MN124782.

AUTHOR CONTRIBUTIONS

PR-G conceptualized and wrote the original draft. PR-G and JF-A worked on the formal analysis, supervised the study, wrote reviewed and edited the manuscript. JF-A and EK were responsible for funding acquisition. PR-G, CC-J, EK, and GS worked on the investigation and methodology. JF-A, RH, and EK were responsible for the resources.

FUNDING

The authors are grateful to CAPES and FAPESP for scholarships and research grants associated with this work (CAPES PNPD20132154 – 33141010001P4 – PNPD – IBSP, 88882.157041/2017-01; FAPESP 2017/50222-0, 2014/08458-9, 2016/01960-6, and 2017/50334-3. This study was financed in part the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) Finance Code 001.

ACKNOWLEDGMENTS

The authors thank Dr. Marcelo Eiras (Plant Virology Laboratory, Instituto Biológico de São Paulo) for providing hybridization reagents.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.00206/full#supplementary-material

REFERENCES

Agboli, E., Leggewie, M., Altinli, M., and Schnettler, E. (2019). Mosquito-specific viruses—transmission and interaction. *Viruses* 11:873. doi: 10.3390/v11090873

Almagro Armenteros, J. J., Sønderby, C. K., Sønderby, S. K., Nielsen, H., and Winther, O. (2017). DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics* 33, 3387–3395. doi: 10.1093/bioinformatics/btx431

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. doi: 10.1038/s41587-019-0036-z

Altendorf, S. (2018). *Minor Tropical Fruits: Mainstreaming a Niche Market*. Rome: FAO.

Antonioli-Luizon, R. (2010). *Sequenciamento Parcial do Vírus da Pinta Verde do Maracujazeiro (Passion Fruit Green Spot Virus-PFGSV), Desenvolvimento de Métodos Para Sua Detecção e Estudos Sobre sua Variabilidade Genética*. Master's dissertation, University of Sao Paulo, Piracicaba. 1–70.

Antonioli-Luizon, R., Jesus-Brabosa, C., Ferraz-Laranjeira, F., Kitajima, E. W., and Freita-Astua, J. (2009). *Diagnóstico da Pinta Verde: Definhamento Precoce do Maracujazeiro*. Available at: https://www.embrapa.br/busca-de-publicacoes/-/publicacao/711786/diagnostico-da-pinta-verde-definhamento-precoce-do-maracujazeiro (accessed December 27, 2018).

Atukunda, R., Semakula, P., Karungi, J., Kyamanwa, S., Erbaugh, M., and Ochwo-Ssemakula, M. (2018). Farmers' knowledge of passion fruit virus diseases and their management in central Uganda. *Afr. J. Hort. Sci.* 13, 53–64.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Barros, D. R., Alfenas-Zerbini, P., Beserra, J. E. A., Antunes, T. F. S., and Zerbini, F. M. (2011). Comparative analysis of the genomes of two isolates of cowpea aphid-borne mosaic virus (CABMV) obtained from different hosts. *Arch. Virol.* 156, 1085–1091. doi: 10.1007/s00705-011-0962-7

Bastianel, M., Novelli, V. M., Kitajima, E. W., Kubo, K. S., Bassanezi, R. B., Machado, M. A., et al. (2010). Citrus leprosis: centennial of an unusual mite–virus pathosystem. *Plant Dis.* 94, 284–292. doi: 10.1094/PDIS-94-3-0284

Bawono, P., and Heringa, J. (2014). *PRALINE: A Versatile Multiple Sequence Alignment Toolkit*. Totowa, NJ: Humana Press, 245–262.

Cerqueira-Silva, C. B. M., Faleiro, F. G., de Jesus, O. N., dos Santos, E. S. L., and de Souza, A. P. (2016). "The genetic diversity, conservation, and use of passion fruit (*Passiflora* spp.)," in *Genetic Diversity and Erosion in Plants. Sustainable Development and Biodiversity*, eds M. Ahuja, and S. Jain, (Cham: Springer), 215–231. doi: 10.1007/978-3-319-25954-3_5

Chabi-Jesus, C., Ramos-González, P. L., Tassi, A. D., Guerra-Peraza, O., Kitajima, E. W., Harakava, R., et al. (2018). Identification and characterization of citrus chlorotic spot virus, a new dichorhavirus associated with citrus leprosis-like symptoms. *Plant Dis.* 102, 1588–1598. doi: 10.1094/PDIS-09-17-1425-RE

Chandana, T., and Venkatesh, Y. P. (2016). Occurrence, functions and biological significance of arginine-rich proteins. *Curr. Protein Pept. Sci.* 17, 507–516. doi: 10.2174/1389203717666151201192348

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., et al. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* 31, 3497–3500. doi: 10.1093/nar/gkg500

Chiba, S., Hleibieh, K., Delbianco, A., Klein, E., Ratti, C., Ziegler-Graff, V., et al. (2013). The *Benyvirus* rna silencing suppressor is essential for long-distance movement, requires both zinc-finger and NoLS basic residues but not a nucleolar localization for its silencing-suppression activity. *Mol. Plant Microbe Interact.* 26, 168–181. doi: 10.1094/MPMI-06-12-0142-R

Criscuolo, A., and Gribaldo, S. (2010). BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from

multiple sequence alignments. *BMC Evol. Biol.* 10:210. doi: 10.1186/1471-2148-10-210

de Oliveira Freitas, J. C., Pio Viana, A., Santos, E. A., Paiva, C. L., de Lima e Silva, F. H., and Souza, M. M. (2016). Sour passion fruit breeding: strategy applied to individual selection in segregating population of *Passiflora* resistant to Cowpea aphid-born mosaic virus (CABMV). *Sci. Hortic.* 211, 241–247. doi: 10.1016/J.SCIENTA.2016.09.002

Dietzgen, R. G., Freitas-Astúa, J., Chabi-Jesus, C., Ramos-González, P. L., Goodin, M. M., Kondo, H., et al. (2018). Dichorhaviruses in their host plants and mite vectors. *Adv. Virus Res.* 102, 119–148. doi: 10.1016/BS.AIVIR.2018.06.001

Doerks, T., Copley, R. R., Schultz, J., Ponting, C. P., and Bork, P. (2002). Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res.* 12, 47–56. doi: 10.1101/gr.203201

Dolja, V. V., and Koonin, E. V. (2018). Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus Res.* 244, 36–52. doi: 10.1016/j.virusres.2017.10.020

Feng, Y., Krueger, E. N., Liu, S., Dorman, K., Bonning, B. C., and Miller, W. A. (2017). Discovery of known and novel viral genomes in soybean aphid by deep sequencing. *Phytobiomes J.* 1, 36–45. doi: 10.1094/PBIOMES-11-16-0013-R

Fernández-Calvino, L., Martínez-Priego, L., Lakatos, L., Llave, C., González, I., Canto, T., et al. (2016). Tobacco rattle virus 16K silencing suppressor binds ARGONAUTE 4 and inhibits formation of RNA silencing complexes. *J. Gen. Virol.* 97, 246–257. doi: 10.1099/jgv.0.000323

Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, W29–W37. doi: 10.1093/nar/gkr367

Fischer, I. H., and Rezende, J. A. M. (2008). *Diseases of Passion Flower (Passiflora spp.)*. Singapore: GSB.

Freitas-Astúa, J., Ramos-González, P. L., Arena, G. D., Tassi, A. D., and Kitajima, E. W. (2018). Brevipalpus-transmitted viruses: parallelism beyond a common vector or convergent evolution of distantly related pathogens? *Curr. Opin. Virol.* 33, 66–73. doi: 10.1016/J.COVIRO.2018.07.010

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi: 10.1038/nprot.2013.084

Hao, X., Zhang, W., Zhao, F., Liu, Y., Qian, W., Wang, Y., et al. (2018). Discovery of plant viruses from tea plant (*Camellia sinensis* (L.) O. Kuntze) by metagenomic sequencing. *Front. Microbiol.* 9:2175. doi: 10.3389/fmicb.2018.02175

Hildebrand, A., Remmert, M., Biegert, A., and Söding, J. (2009). Fast and accurate automatic structure prediction with HHpred. *Proteins Struct. Funct. Bioinform.* 77, 128–132. doi: 10.1002/prot.22499

Huang, X., and Madan, A. (1999). CAP3: a DNA sequence assembly program. *Genome Res.* 9, 868–877. doi: 10.1101/gr.9.9.868

Jaramillo-Mesa, H., Marín Montoya, M., and Gutiérrez, P. A. (2018). Molecular characterization of Soybean mosaic virus (SMV) infecting Purple passion fruit (*Passiflora edulis* f. edulis) in Antioquia, Colombia. *Arch. Phytopathol. Plant Prot.* 51, 617–636. doi: 10.1080/03235408.2018.1505411

Johnson, B. R. (2018). Taxonomically restricted genes are fundamental to biology and evolution. *Front. Genet.* 9:407. doi: 10.3389/fgene.2018.00407

Kallies, R., Kopp, A., Zirkel, F., Estrada, A., Gillespie, T. R., Drosten, C., et al. (2014). Genetic characterization of goutanap virus, a novel virus related to negeviruses, cileviruses and higreviruses. *Viruses* 6, 4346–4357. doi: 10.3390/v6114346

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Khalturin, K., Hemmrich, G., Fraune, S., Augustin, R., and Bosch, T. C. G. (2009). More than just orphans: are taxonomically-restricted genes important in evolution? *Trends Genet.* 25, 404–413. doi: 10.1016/j.tig.2009.07.006

Kitajima, E. W., and Nome, C. F. (1999). "Microscopia eletrônica en virologia vegetal," in *Métodos Para Detectar Patógenos Sistémicos*, eds D. M. Do Campo, and S. L. Lenardon (Córdoba: Instituto Nacional de Tecnologia Agropecuaria), 59–87.

Kitajima, E. W., Rezende, J. A. M., Rodrigues, J. C. V., Chiavegato, L. G., Piza Júnior, C. T., and Morozini, W. (1997). Green spot of passion fruit, a possible viral disease associated with infestation by the mite Brevipalpus phoenicis. *Fitopatol. Bras.* 22, 555–559.

Kitajima, E. W. W., Rezende, J. A. M. A. M., and Rodrigues, J. C. V. C. V. (2003). Passion fruit green spot virus vectored by *Brevipalpus phoenicis* (Acari: Tenuipalpidae) on passion fruit in Brazil. *Exp. Appl. Acarol.* 30, 225–231. doi: 10.1023/B:APPA.0000006551.74604.84

Kondo, H., Chiba, S., Maruyama, K., Andika, I. B., and Suzuki, N. (2019). A novel insect-infecting virga/nege-like virus group and its pervasive endogenization into insect genomes. *Virus Res.* 262, 37–47. doi: 10.1016/j.virusres.2017.11.020

Kondo, H., Hirota, K., Maruyama, K., Andika, I. B., and Suzuki, N. (2017). A possible occurrence of genome reassortment among bipartite rhabdoviruses. *Virology* 508, 18–25. doi: 10.1016/j.virol.2017.04.027

Koonin, E. V., Boyko, V. P., and Dolja, V. V. (1991). Small cysteine-rich proteins of different groups of plant RNA viruses are related to different families of nucleic acid-binding proteins. *Virology* 181, 395–398. doi: 10.1016/0042-6822(91)90512-a

Kuchibhatla, D. B., Sherman, W. A., Chung, B. Y. W., Cook, S., Schneider, G., Eisenhaber, B., et al. (2014). Powerful sequence similarity search methods and in-depth manual analyses can identify remote homologs in many apparently "orphan" viral proteins. *J. Virol.* 88, 10–20. doi: 10.1128/JVI.02595-13

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Le, S. Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25, 1307–1320. doi: 10.1093/molbev/msn067

Leastro, M. O., Kitajima, E. W., Silva, M. S., Resende, R. O., and Freitas-Astúa, J. (2018). Dissecting the subcellular localization, intracellular trafficking, interactions, membrane association, and topology of citrus leprosis virus C proteins. *Front. Plant Sci.* 9:1299. doi: 10.3389/fpls.2018.01299

Lemoine, F., Correia, D., Lefort, V., Doppelt-Azeroual, O., Mareuil, F., Cohen-Boulakia, S., et al. (2019). NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Res.* 47, W260–W265. doi: 10.1093/nar/gkz303

Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239

Li, C.-X., Shi, M., Tian, J.-H., Lin, X.-D., Kang, Y.-J., Chen, L.-J., et al. (2015). Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *eLife* 4:e05378. doi: 10.7554/eLife.05378

Locali-Fabris, E. C., Freitas-Astúa, J., Souza, A. A., Takita, M. A., Astúa-Monge, G., Antonioli-Luizon, R., et al. (2006). Complete nucleotide sequence, genomic organization and phylogenetic analysis of citrus leprosis virus cytoplasmic type. *J. Gen. Virol.* 87, 2721–2729. doi: 10.1099/vir.0.82038-0

Ludwiczak, J., Winski, A., Szczepaniak, K., Alva, V., and Dunin-Horkawicz, S. (2019). DeepCoil—a fast and accurate prediction of coiled-coil domains in protein sequences. *Bioinformatics* 35, 2790–2795. doi: 10.1093/bioinformatics/bty1062

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., et al. (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* 45, D200–D203. doi: 10.1093/nar/gkw1129

Medd, N. C., Fellous, S., Waldron, F. M., Xuéreb, A., Nakai, M., Cross, J. V., et al. (2018). The virome of Drosophila suzukii, an invasive pest of soft fruit. *Virus Evol.* 4:9. doi: 10.1093/ve/vey009

Melcher, U. (2000). The '30K' superfamily of viral movement proteins. *J. Gen. Virol.* 81, 257–266. doi: 10.1099/0022-1317-81-1-257

Melzer, M. J., Sether, D. M., Borth, W. B., and Hu, J. S. (2013a). Characterization of a virus infecting Citrus volkameriana with citrus leprosis-like symptoms. *Phytopathology* 102, 122–127. doi: 10.1094/PHYTO-01-11-0013

Melzer, M. J., Simbajon, N., Carillo, J., Borth, W. B., Freitas-Astúa, J., Kitajima, E. W., et al. (2013b). A cilevirus infects ornamental hibiscus in Hawaii. *Arch. Virol.* 158, 2421–2424. doi: 10.1007/s00705-013-1745-0

Mushegian, A. R., and Elena, S. F. (2015). Evolution of plant virus movement proteins from the 30K superfamily and of their homologs integrated in plant genomes. *Virology* 476C, 304–315. doi: 10.1016/j.virol.2014.12.012

Nascimento, A. V. S., Santana, E. N., Braz, A. S. K., Alfenas, P. F., Pio-Ribeiro, G., Andrade, G. P., et al. (2006). Cowpea aphid-borne mosaic virus (CABMV) is widespread in passionfruit in Brazil and causes passionfruit woodiness disease. *Arch. Virol.* 151, 1797–1809. doi: 10.1007/s00705-006-0755-6

Niu, J., Zhang, W., Sun, Q.-Z., and Wang, J.-J. (2019). Three novel RNA viruses in the spider mite Tetranychus urticae and their possible interactions with the host RNA interference response. *J. Invertebr. Pathol.* 166:107228. doi: 10.1016/j.jip.2019.107228

Nunes, M., Contreras-Gutierrez, M. A., Guzman, H., Martins, L. C., Barbirato, M. F., Savit, C., et al. (2017). Genetic characterization, molecular epidemiology, and phylogenetic relationships of insect-specific viruses in the taxon Negevirus. *Virology* 504, 152–167. doi: 10.1016/j.virol.2017.01.022

Ocampo Pérez, J., and Coppens d'Eeckenbrugge, G. (2017). Morphological characterization in the genus *Passiflora L.*: an approach to understanding its complex variability. *Plant Syst. Evol.* 303, 531–558. doi: 10.1007/s00606-017-1390-2

Pascon, R. C., Kitajima, J. P., Breton, M. C., Assumpção, L., Greggio, C., Zanca, A. S., et al. (2006). The complete nucleotide sequence and genomic organization of citrus leprosis associated virus, cytoplasmatic type (CiLV-C). *Virus Genes* 32, 289–298. doi: 10.1007/s11262-005-6913-1

Pei, J., and Grishin, N. V. (2007). PROMALS: towards accurate multiple sequence alignments of distantly related proteins. *Bioinformatics* 23, 802–808. doi: 10.1093/bioinformatics/btm017

Quito-Avila, D. F., Brannen, P. M., Cline, W. O., Harmon, P. F., and Martin, R. R. (2013). Genetic characterization of Blueberry necrotic ring blotch virus, a novel RNA virus with unique genetic features. *J. Gen. Virol.* 94, 1426–1434. doi: 10.1099/vir.0.050393-0

Ramos-González, P. L., Chabi-Jesus, C., Arena, G. D., Tassi, A. D., Kitajima, E. W., and Freitas-Astúa, J. (2018). Citrus leprosis: a unique multietiologic disease. *Citrus Am.* 1, 4–19.

Ramos-González, P. L., Chabi-Jesus, C., Guerra-Peraza, O., Breton, M. C., Arena, G. D., Nunes, M. A., et al. (2016). Phylogenetic and molecular variability studies reveal a new genetic clade of citrus leprosis virus C. *Viruses* 8:153. doi: 10.3390/v8060153

Ramos-González, P. L., Chabi-Jesus, C., Guerra-Peraza, O., Tassi, A. D., Kitajima, E. W., Harakava, R., et al. (2017). Citrus leprosis virus N: a new dichorhavirus causing Citrus Leprosis disease. *Phytopathology* 107, 963–976. doi: 10.1094/PHYTO-02-17-0042-R

Remmert, M., Biegert, A., Hauser, A., and Söding, J. (2012). HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods* 9, 173–175. doi: 10.1038/nmeth.1818

Rodrigues, L. K., Silva, L. A., Garcêz, R. M., Chaves, A. L. R., Duarte, L. M. L., Giampani, J. S., et al. (2015). Phylogeny and recombination analysis of Brazilian yellow passion fruit isolates of Cowpea aphid-borne mosaic virus: origin and relationship with hosts. *Austral. Plant Pathol.* 44, 31–41. doi: 10.1007/s13313-014-0308-5

Roy, A., Choudhary, N., Guillermo, L. M., Shao, J., Govindarajulu, A., Achor, D., et al. (2013). A novel virus of the genus Cilevirus causing symptoms similar to citrus leprosis. *Phytopathology* 103, 488–500. doi: 10.1094/PHYTO-07-12-0177-R

Roy, A., Hartung, J. S., Schneider, W. L., Shao, J., Leon, M. G., Melzer, M. J., et al. (2015). Role bending: complex relationships between viruses, hosts, and vectors related to citrus leprosis, an emerging disease. *Phytopathology* 105, 872–884. doi: 10.1094/PHYTO-12-14-0375-FI

Santos, E. A., Viana, A. P., de Oliveira Freitas, J. C., de Lima e Silva, F. H., Rodrigues, R., and Eiras, M. (2015). Resistance to Cowpea aphid-borne mosaic virus in species and hybrids of *Passiflora*: advances for the control of the passion fruit woodiness disease in Brazil. *Eur. J. Plant Pathol.* 143, 85–98. doi: 10.1007/s10658-015-0667-y

Santos Filho, H. P., Lima, A., de, A., Barbosa, C., de, J., Borges, A. L., et al. (1999). *Definhamento Precoce do Maracujazeiro*. Cruz das Almas,BA: Brazilian Agricultural Research Corporation.

Schlötterer, C. (2015). Genes from scratch–the evolutionary fate of de novo genes. *Trends Genet.* 31, 215–219. doi: 10.1016/j.tig.2015.02.007

Seo, J.-K., Ohshima, K., Lee, H.-G., Son, M., Choi, H.-S., Lee, S.-H., et al. (2009). Molecular variability and genetic structure of the population of Soybean mosaic virus based on the analysis of complete genome sequences. *Virology* 393, 91–103. doi: 10.1016/j.virol.2009.07.007

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

Solovyev, A. G., and Morozov, S. Y. (2017). Non-replicative integral membrane proteins encoded by plant alpha-like viruses: emergence of diverse orphan ORFs and movement protein genes. *Front. Plant Sci.* 8:1820. doi: 10.3389/fpls.2017.01820

Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 6, 175–182.

Sun, L., Andika, I. B., Kondo, H., and Chen, J. (2013). Identification of the amino acid residues and domains in the cysteine-rich protein of Chinese wheat mosaic virus that are important for RNA silencing suppression and subcellular localization. *Mol. Plant Pathol.* 14, 265–278. doi: 10.1111/mpp.12002

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., et al. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Res.* 40:e115. doi: 10.1093/nar/gks596

Van Oss, S. B., and Carvunis, A.-R. (2019). De novo gene birth. *PLoS Genet.* 15:e1008160. doi: 10.1371/journal.pgen.1008160

Vasilakis, N., Forrester, N. L., Palacios, G., Nasar, F., Savji, N., Rossi, S. L., et al. (2013). Negevirus: a proposed new taxon of insect-specific viruses with wide geographic distribution. *J. Virol.* 87, 2475–2488. doi: 10.1128/JVI.00776-12

Vasilakis, N., and Tesh, R. B. (2015). Insect-specific viruses and their potential impact on arbovirus transmission. *Curr. Opin. Virol.* 15, 69–74. doi: 10.1016/J.COVIRO.2015.08.007

Viljakainen, L., Holmberg, I., Abril, S., and Jurvansuu, J. (2018). Viruses of invasive Argentine ants from the European main supercolony: characterization, interactions and evolution. *J. Gen. Virol.* 99, 1129–1140. doi: 10.1099/jgv.0.001104

Vinokurov, K. S., and Koloniuk, I. (2019). Discovery and characterization of a novel alphavirus-like RNA virus from the red firebug *Pyrrhocoris apterus* L. (Heteroptera). *J. Invertebr. Pathol.* 166:107213. doi: 10.1016/J.JIP.2019.107213

Yelina, N. E., Savenkov, E. I., Solovyev, A. G., Morozov, S. Y., and Valkonen, J. P. T. (2002). Long-distance movement, virulence, and RNA silencing suppression controlled by a single protein in hordei- and potyviruses: complementary functions between virus families. *J. Virol.* 76, 12981–12991. doi: 10.1128/jvi.76.24.12981-12991.2002

Yockteng, R., d'Eeckenbrugge, G. C., and Souza-Chies, T. T. (2011). *Wild Crop Relatives: Genomic and Breeding Resources*. Berlin: Springer, 129–171.

Yu, Y., Blair, C., and He, X. (2019). RASP 4: ancestral state reconstruction tool for multiple genes and characters. *Mol. Biol. Evol.* doi: 10.1093/molbev/msz257 [Epub ahead of print].

Zeraik, M. L., Serteyn, D., Deby-Dupont, G., Wauters, J.-N., Tits, M., Yariwake, J. H., et al. (2011). Evaluation of the antioxidant activity of passion fruit (*Passiflora edulis* and *Passiflora alata*) extracts on stimulated neutrophils and myeloperoxidase activity assays. *Food Chem.* 128, 259–265. doi: 10.1016/j.foodchem.2011.03.001

Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., et al. (2018). A Completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* 430, 2237–2243. doi: 10.1016/j.jmb.2017.12.007

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 07 April 2020
doi: 10.3389/fmicb.2020.00509

Virome Analysis of Aphid Populations That Infest the Barley Field: The Discovery of Two Novel Groups of Nege/Kita-Like Viruses and Other Novel RNA Viruses

Hideki Kondo[1], Miki Fujita[1], Hiroshi Hisano[1], Kiwamu Hyodo[1], Ida Bagus Andika[2] and Nobuhiro Suzuki[1]*

[1] Institute of Plant Science and Resources (IPSR), Okayama University, Kurashiki, Japan, [2] College of Plant Health and Medicine, Qingdao Agricultural University, Qingdao, China

OPEN ACCESS

Edited by:
Jesús Navas-Castillo, Institute of Subtropical and Mediterranean Horticulture La Mayora, Spain

Reviewed by:
Francis Onono Wamonje, University of Saskatchewan, Canada
Martin Drucker, INRAE, Centre de recherche Grand Est-Colmar, France

***Correspondence:**
Hideki Kondo
hkondo@okayama-u.ac.jp

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *27 January 2020*
Accepted: *09 March 2020*
Published: *07 April 2020*

Citation:
Kondo H, Fujita M, Hisano H, Hyodo K, Andika IB and Suzuki N (2020) Virome Analysis of Aphid Populations That Infest the Barley Field: The Discovery of Two Novel Groups of Nege/Kita-Like Viruses and Other Novel RNA Viruses. Front. Microbiol. 11:509. doi: 10.3389/fmicb.2020.00509

Aphids (order Hemiptera) are important insect pests of crops and are also vectors of many plant viruses. However, little is known about aphid-infecting viruses, particularly their diversity and relationship to plant viruses. To investigate the aphid viromes, we performed deep sequencing analyses of the aphid transcriptomes from infested barley plants in a field in Japan. We discovered virus-like sequences related to nege/kita-, flavi-, tombus-, phenui-, mononega-, narna-, chryso-, partiti-, and luteoviruses. Using RT-PCR and sequence analyses, we determined almost complete sequences of seven nege/kitavirus-like virus genomes; one of which was a variant of the Wuhan house centipede virus (WHCV-1). The other six seem to belong to four novel viruses distantly related to Wuhan insect virus 9 (WhIV-9) or Hubei nege-like virus 4 (HVLV-4). We designated the four viruses as barley aphid RNA virus 1 to 4 (BARV-1 to -4). Moreover, some nege/kitavirus-like sequences were found by searches on the transcriptome shotgun assembly (TSA) libraries of arthropods and plants. Phylogenetic analyses showed that BARV-1 forms a clade with WHCV-1 and HVLV-4, whereas BARV-2 to -4 clustered with WhIV-9 and an aphid virus, Aphis glycines virus 3. Both virus groups (tentatively designated as Centivirus and Aphiglyvirus, respectively), together with arthropod virus-like TSAs, fill the phylogenetic gaps between the negeviruses and kitaviruses lineages. We also characterized the flavi/jingmen-like and tombus-like virus sequences as well as other RNA viruses, including six putative novel viruses, designated as barley aphid RNA viruses 5 to 10. Interestingly, we also discovered that some aphid-associated viruses, including nege/kita-like viruses, were present in different aphid species, raising a speculation that these viruses might be distributed across different aphid species with plants being the reservoirs. This study provides novel information on the diversity and spread of nege/kitavirus-related viruses and other RNA viruses that are associated with aphids.

Keywords: negevirus, kitavirus, aphid, virome, RNA seq, barley, diversity, horizontal transmission

INTRODUCTION

Aphids (order Hemiptera: superfamily Aphidoidae) are phloem feeders and important insect pests that affect a wide range of crops. Aphid infestation can reduce both the yield and quality of crops by direct feeding and transmitting symptomatic plant viruses (Powell et al., 2006). Many aphid species act as vectors for a number of agriculturally important plant viruses (Ng and Perry, 2004; Brault et al., 2010). Most aphid-vectored plant viruses do not replicate within the aphids and are transmitted by either a non-circulative (non-persistent) or a circulative (persistent) manner (Ng and Perry, 2004; Whitfield et al., 2015; Dader et al., 2017); however, some plant rhabdoviruses (negative-sense RNA genome, family *Rhabdoviridae*) are transmitted by aphids through the circulative (persistent) manner and replicate within their vectors (Whitfield et al., 2018).

In cereal crops (within the family Poaceae), such as wheat (*Triticum aestivum* L.) and barley (*Hordeum vulgare* L.), aphids can cause great yield losses by directly feeding on the phloem sap and more significantly by transmitting several yellow dwarf disease-causing viruses (Jarosova et al., 2013; Krueger et al., 2013), such as barley yellow dwarf, cereal yellow dwarf and wheat yellow dwarf viruses (positive-sense RNA viruses, genera *Luteovirus* and *Polerovirus*, family *Luteoviridae*) (King et al., 2011). These viruses are transmitted in a circulative (persistent) and non-propagative manner by common aphid species that infest cereals, such as the bird cherry-oat aphid (*Rhopalosiphum padi*), corn leaf aphid (*R. maidis*), greenbugs (*Schizaphis graminum*), English grain aphid (*Sitobion avenae*), and rose-grain aphid (*Metopolophium dirhodum*) (Miller et al., 2002; Parry et al., 2012).

In addition to harboring numerous plant viruses, aphids are known to host several insect-specific single stranded DNA viruses [Myzus persicae densovirus (MpDNV) and Dysaphis plantaginea densovirus (DplDNV)] (van Munster et al., 2003; Ryabov et al., 2009) and positive-sense RNA viruses, the latter of which includes three picornaviruses: the Rhopalosiphum padi virus (RhPV) and aphid lethal paralysis virus (ALPV), both of the family *Dicistroviridae*, and the Brevicoryne brassicae virus (BrBV) of the family *Iflaviridae* (Moon et al., 1998; van Munster et al., 2002; Ryabov, 2007); and two unclassified RNA viruses: rosy apple aphid virus (a calici-like virus) and *Acyrthosiphon pisum* virus (APV) (distantly related to the members of the family *Solinviviridae*) (vandenHeuvel et al., 1997; Ryabov et al., 2009). Many of these viruses were discovered because of their phenotype alternations of the host aphids (Teixeira et al., 2016). More recently, analyses using next-generation sequencing (NGS) have identified several novel aphid-infecting viruses, including six positive-sense RNA viruses: three flavi-like viruses [Macrosiphum euphorbiae virus 1 (MeV-1); two jingmenviruses, Wuhan aphid virus 1 and 2 (WhAV-1 and -2)] (Shi et al., 2016b; Teixeira et al., 2016), a tetra-like virus (Aphis glycine virus 2), a nege-like virus [Aphis glycine virus 3 (ApGlV-3)] (Feng et al., 2017) and a sobemo-like virus (Macrosiphum euphorbiae virus 3) (Teixeira et al., 2018); a negative-sense RNA phlebo-like virus (Aphis citricidus bunyavirus) (Zhang et al., 2019); and a DNA densovirus (Macrosiphum euphorbiae virus 2) (Teixeira et al., 2018). Nevertheless, little is known about the viromes of aphid populations, particularly those that infest cereal plants in the field.

During the past decades, an increasing number of the novel "insect–infecting viruses" have been identified from mosquitoes (order Diptera). These viruses have a host range that is restricted to insects and are related to human arthropod-borne viral pathogens (so called arboviruses), such as dengue, Zika and West Nile viruses (family *Flaviviridae*), chikungunya virus (family *Togaviridae*) and Rift Valley fever virus (family *Phenuiviridae*) (Bolling et al., 2015). One of these insect–infecting virus groups has been proposed as the taxon "Negevirus," consisting of alpha-like viruses from mosquitos and sandflies (order Diptera) (Vasilakis et al., 2013; Vasilakis and Tesh, 2015). Negevirus particles are likely "spherical or elliptical" in shape with diameter of 45–55 nm (Vasilakis et al., 2013; Nabeshima et al., 2014; Kawakami et al., 2016; O'Brien et al., 2017; Zhao et al., 2019). Their genomes consist of a non-segmented, positive-sense RNA, approximately 9 to 10 kb in length and comprises three open reading frames (ORF). The second and third ORFs encode two structural proteins, a predicted glycoprotein (ORF2), and a predicted membrane protein SP24 (ORF3) (Kuchibhatla et al., 2014; Fujita et al., 2017; Zhao et al., 2019; Colmant et al., 2020). Based on phylogenetic analysis, this taxon could be separated into two groups at the genus level, namely "Nelorpivirus" and "Sandewavirus" (Kallies et al., 2014). Negeviruses are distantly related to members of three plant virus genera, *Cilevirus*, *Higrevirus*, and *Blunervirus* (Vasilakis et al., 2013; Kallies et al., 2014; Nunes et al., 2017; Ramos-González et al., 2020). These genera have recently been assigned as the members of the family *Kitaviridae* (Walker et al., 2019). Kitaviruses have bi-, tri- or tetra-partite positive-sense RNA genomes (Locali-Fabris et al., 2006; Melzer et al., 2013; Hao et al., 2018), and some of them, such as citrus leprosis virus cytoplasmic type (CiLV-C, a cilevirus, which is prevalent in several countries of the American continent), have non-enveloped bacilliform particles and are transmitted by false spider mites, *Brevipalpus* spp. (class Arachnida) (Tassi et al., 2017; Freitas-Astua et al., 2018). Recently, a meta-transcriptomic approach on invertebrates expanded the diversity of nege- and nege-like viruses infecting or associated with mosquitos and flies (Webster et al., 2015, 2016; Shi et al., 2017; Medd et al., 2018; Sadeghi et al., 2018; Pettersson et al., 2019) and other invertebrates (Shi et al., 2016a; Debat, 2017; Feng et al., 2017; Nunes et al., 2017; Schoonvaere et al., 2018; Kondo et al., 2019). Nevertheless, there is still very limited information about the relationships between negeviruses and kitaviruses and also the evolutionary history of these two virus groups.

In this study, a meta-transcriptomic approach was used to investigate the viromes of aphid populations collected from a barley field for three consecutive years, 2016–2018. We identified at least 60 virus-like sequence contigs related to RNA viruses such as nege/kita-, flavi/jingmen-, tombus-, phenui- and luteoviruses, including eight putative novel RNA viruses. In particular, our study reveals the new nege/kita-like virus lineages with the members that are associated with aphid species and, thus,

enhances our knowledge on the diversity and evolution of insect viruses that are related to plant viral pathogens.

MATERIALS AND METHODS

Collection of Aphids and Species Identification

Collection of aphids were performed in the experimental field at the Institute of Plant Science and Resources (IPSR) in Okayama University, Kurashiki, Japan (34°59' N and 133°77' E) on April and May in 2016–2018. Aphids (total 15 aphid samples) were obtained from the colonies that infest leaves and ears of barley cultivar "Golden Promise", "Morex", "Kikai Hadaka", "Minori Mugi", "Ryofu", "Kobin Katagi", and three landraces (OUK327, OUU094, and OUU659), and then aphids samples were stored at −80°C until further analysis.

To determine the aphid species, a 658-bp fragment of mitochondrial DNA from the 5′ region of the mitochondrial cytochrome c oxidase subunit 1 (COI) gene, was analyzed (Foottit et al., 2008). Total genomic DNA was extracted from a portion of each aphid sample by using DNeasy Blood and Tissue Kit (Qiagen, Venlo, Netherlands). The COI sequence of mitochondrial DNA was amplified by PCR using QuickTaq HS Dye Mix (Toyobo, Osaka, Japan) with the primer pair "LepF and LepR" (**Supplementary Table S1**; Foottit et al., 2008). PCR fragments were purified using the Wizard SV Gel and PCR Clean-Up System (Promega, United States) and then sequenced using the conventional Sanger sequencing method with an ABI3100 DNA sequencer (Applied Biosystems, Foster City, CA, United States). The amplified COI fragments were subjected to a BlastN search and Neighbor Joining (NJ) tree construction (see below).

RNA Extraction and RT-PCR

Total RNA from each aphid sample (pools of roughly 20–40 individual aphids from each colony) was extracted using TaKaRa RNAiso Plus Reagent (TaKaRa Biotech. Co., Shiga, Japan), following the manufacturer's instructions. The total RNA fraction was analyzed by electrophoresis in a 1% (W/V) agarose gel in 1 × TAE buffer. For reverse transcription (RT)-PCR to detect virus-like sequences, cDNAs were synthesized using MMLV reverse transcriptase (Thermo Fisher Scientific, Waltham, MA, United States) with random hexamers, following the manufacturer's instructions. The cDNAs were then used as templates for PCR amplification with QuickTaq HS Dye Mix. The primer sets used to amplify virus-like sequences are provided in **Supplementary Table S1**. The 16S ribosomal RNA of aphids was used as a reference target gene for RT-PCR using a primer set "Apisum16S_F and Apisum16S_R" for *Acyrthosiphon pisum* (**Supplementary Table S1**; Koramutla et al., 2016). To confirm the presence of primary and secondary parasitoid wasps, we selected two primer sets "L.testa_RpL3_F and L.testa_RpL3_R" and "DcarpF and DcarpR" for amplification of the ribosomal protein L3 (RpL3) gene of a putative primary parasitoid species (*Lysiphlebus testaceipes*, Hymenoptera: Braconidae) (this study) and the 16S rRNA gene of a putative secondary parasitoid species (*Dendrocerus carpenteri*, Hymenoptera: Megaspilidae) (**Supplementary Table S1**; Chen et al., 2006). PCR conditions used were as follows: an initial denaturation step at 94°C for 2 min; followed by 30 or 35 cycles of: denaturing at 94°C for 10 seconds, annealing at 53°C or 55°C for 30 seconds, then extension at 72°C for 1 min, then finishing with a final extension step at 72°C for 10 min. PCR products were then purified and subjected to Sanger sequencing.

To determine the major uncertain or unknown regions of virus-like sequences, RT-PCR was performed using the specific primer sets, and PCR fragments were then sequenced using the Sanger method. Sequences of the primers used in the RT-PCR are provided listed in **Supplementary Table S1** or are available upon request.

Next-Generation Sequencing and Read Assembly

Total RNA fractions were prepared separately from 15 aphid samples (each population from a single colony) and according to the year of aphid collection, pooled into three groups: pool-1 (total 81.9 μg, RNA integrity number: RIN = 8.1) consisted of samples from 2016 (5 aphid populations, BaA1-16-1 to -5), pool-2 (total 12.4 μg, RIN = 9.1) consisted of samples from 2017 (2 aphid populations, BaA2-17-6 and -7) and pool-3 (total 36.9 μg, RIN = 9.1) consisted of samples from 2018 (8 aphid populations, BaA3-18-8 to -16) (**Table 1**). Each RNA sample pool was depleted of rRNA using a Ribo-Zero kit (Illumina, San Diego, CA, United States) and then used for synthesis of a cDNA library using the TruSeq RNA Sample Preparation kit v2 (Illumina). The cDNA library obtained was subjected to deep sequencing using the Illumina HiSeq 2000 or 4000 platform (Illumina, pair-end 100 bp reads). Deep-sequencing was performed by Macrogen Inc (Tokyo, Japan). Raw reads (pool-1: 58,056,740; poo1-2: 58,030,114 and pool-3: 53,125,246) were trimmed by removing the adapter sequences and *de novo* assembled using the CLC Genomics Workbench version 11 (CLC Bio-Qiagen, Aarhus, Denmark). The assembled-contigs derived from each sample pool (over 1.0 kb contigs) were subjected to local Blast searches against the viral reference sequence (RefSeq) dataset of NCBI[1] (the *E*-value cut-off is >0.05 for local BlastN). To roughly estimate the other sequence read origins for each NGS data set, we also conducted local Blast analyses against the following genome sequences: (1) aphid (*R. padi*) transcript data obtained from AphidBase[2]; (2) insect parasitoid wasp (*M. demolitor,* a larval lepidopteran endoparasitoid) transcript data from NCBI; (3) an aphid endosymbiotic bacterium (*Buchnera aphidicola* str. Ak, a blue alfalfa aphid *Acyrthosiphon kondoi* endosymbiont, accession no. CP002645) genome sequences; and (4) *R. padi* mitochondrion (accession no. KT447631) genome sequences. Mapping of sequence reads to each virus-like contigs or gap-filling virus-like sequences determined using RT-PCR was done using the Read Mapping algorithm of the CLC Genomics Workbench.

[1]https://www.ncbi.nlm.nih.gov

[2]https://bipaa.genouest.org/is/aphidbase/

TABLE 1 | List of the aphid samples collected from the barley field.

Year sample no.	Aphid species[1] (BlastN best hit)	Barley cultivar or variety[2]	NGS group
2016			
BaA1-16-1	*Rhopalosiphum padi*	cv. Morex	pool-1
BaA1-16-2	*Rhopalosiphum maidis*	OUU659	pool-1
BaA1-16-3	*Rhopalosiphum padi*	cv. Kikai Hadaka	pool-1
BaA1-16-4	*Sitobion avenae*	cv. Kikai Hadaka	pool-1
BaA1-16-5	*Rhopalosiphum padi*	cv. Golden Promise	pool-1
2017			
BaA2-17-6	*Rhopalosiphum maidis*	cv. Golden Promise	pool-2
BaA2-17-7	*Sitobion avenae*	cv. Kikai Hadaka	pool-2
2018			
BaA3-18-8	*Rhopalosiphum maidis*	OUK327	pool-3
BaA3-18-9	*Rhopalosiphum maidis*	OUU094	pool-3
BaA3-18-10	*Rhopalosiphum maidis*	cv. Minori Mugi	pool-3
BaA3-18-11	*Rhopalosiphum maidis*	cv. Minori Mugi	pool-3
BaA3-18-12	*Rhopalosiphum maidis*	cv. Kobin Katagi	pool-3
BaA3-18-13	*Rhopalosiphum maidis*	cv. Kobin Katagi	pool-3
BaA3-18-14	*Rhopalosiphum maidis*	cv. Ryofu	pool-3
BaA3-18-15	*Rhopalosiphum maidis*	cv. Golden Promise	pool-3

[1]The species of the aphid colonies were identified by sequencing the nucleotides of fragments of a mitochondrial gene cytochrome c oxidase 1 (COI) using RT-PCR. Rhopalosiphum padi (the bird cherry-oat aphid); R. maidis (the corn leaf aphid); Sitobion avenae (the English grain aphid). [2]The five aphid samples (BaA3-18-10 to BaA3-18-14) were obtained from the barley plants grown in a different area of the IPSR field.

Sequence Analysis and Database Search

Open reading frames were identified using Enzyme X v3.3.3[3]. Pairwise sequence comparisons were performed using the Sequence Demarcation Tool version 1.2, with the MUSCLE alignment (Muhire et al., 2014). The conserved protein domains were searched using the NCBI conserved domain database[4]. Putative transmembrane domains were predicted using the TMHMM server version 2.04[5] (Krogh et al., 2001). Blast or reverse Blast searches were run on the non-redundant (nr) DNA and protein databases from NCBI (nucleotide collection, nr/nt; transcriptome shotgun assembly, TSA).

Phylogenetic Analyses

Maximum-likelihood (ML) tree construction was carried out as described previously with minor modifications (Kondo et al., 2015, 2016). Multiple amino acid alignments were obtained by using MAFFT online version 7[6], set to the default parameters (Katoh and Standley, 2013). Ambiguous portions of the alignment were removed using Gblocks online version 0.91b[7] with the stringency levels lowered for all parameters (Talavera and Castresana, 2007). ML trees were then generated using the PhyML 3.0 online program[8] with automatic model selection by Smart Model Selection (Guindon et al., 2010; Lefort et al., 2017). The NJ trees (Saitou and Nei, 1987) were constructed based on the multiple alignments using MAFFT. Then the phylogenetic trees were refined using FigTree version 1.3.1 software[9].

RESULTS AND DISCUSSION

Identification of Aphid Species Collected From the Barley Field

In April and May in 2016–2018, we collected a total of 15 aphid population samples from the colonies on the barley plants that were grown in the experimental field of IPSR, Okayama (**Table 1**). Analysis of the mitochondrial COI sequence using genomic PCR and sequencing indicated that the aphid species collected were *R. padi* (the bird cherry-oat aphid), *R. maidis* (the corn leaf aphid), and *S. avenae* (the English grain aphid) (**Table 1** and **Supplementary Figure S1** and data not shown). *R. maidis* may be the predominant species of aphid populations in this barley field because *R. maidis* was obtained during each year of sampling and was also the main species collected of the aphid populations (10 out of 15 samples) (**Table 1**). Nevertheless, further intensive surveys are required to determine the actual aphid population and their dynamics in the field.

Aphid Transcriptomic Analysis Using NGS

To investigate the virome of these aphid samples, total RNA samples were pooled according to the year of sampling (pool-1 to pool-3) and subjected to transcriptomic analysis using NGS. The *de novo* assembled contigs that were larger than 1.0 kb

[3]https://nucleobytes.com/enzymex/index.html
[4]https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
[5]http://www.cbs.dtu.dk/services/TMHMM-2.0/
[6]http://mafft.cbrc.jp/alignment/server/
[7]http://molevol.cmima.csic.es/castresana/Gblocks_server.html
[8]http://www.atgc-montpellier.fr/phyml-sms/
[9]http://tree.bio.ed.ac.uk/software/

(pool-1: 10,133; pool-2: 15,562; and pool-3: 10,324 contigs) were subjected to local Blast (tBlastX) analyses (**Figure 1A**). Most of the contigs were associated with aphids (63–91% contigs), with much fewer portions being associated with wasps (3–22%) and endosymbionts (2 or 3%). Aphid samples of the pool-1 might be highly parasitized by parasitoids (22% contigs), such as *L. testaceipes* compared with those of other pools (7% or 3%) (**Figure 1A** and data not shown). Among the total contigs (>1 kb), at least 60 contigs (pool-1: ∼43 contigs; pool-2: 5 contigs; and pool-3: 12 contigs) were virus-related sequences, with the length ranging from 1,102 to 23,141 nt (**Figures 1A,B** and see below). The proportions of sequence reads related to viral sequences largely differed among three libraries; 13%, 0%, and 2% for pool-1, pool-2, and pool-3, respectively (**Figure 1B**).

Identification of Virus-Related Sequences From the Aphid Transcriptomes

Local Blast analyses revealed that the virus-like contigs present in the aphid transcriptomes have sequence similarity to relatives or members of the virus families, the *Kitaviridae* (or proposed Nelorpivirus and Sandewavirus), *Flaviviridae*, *Tombusviridae*, *Phenuiviridae*, *Narnaviridae*, *Chrysoviridae*, *Partitiviridae*, *Rhabdoviridae*, and *Luteoviridae* (**Figure 1B** and **Table 2**). Notably, virus-like reads were predominantly mapped to virus-like contigs associated with nege/kitaviruses (25%), flavi/jingmenviruses (49%), and tombus-like virus (25%) in pool-1; and nege/kitaviruses (39%), flavi/jingmenviruses (35%), and luteoviruses (26%) in pool-3 (**Figure 1B**). The three luteovirus-like contig sequences in pools 1 and 3 are likely derived from plant viruses including barley yellow dwarf virus (a well-known luteovirus, family *Luteoviridae*). The detailed analysis of luteovirus, or luteo-like virus sequences identified in the current study will be reported elsewhere as a part of the ongoing barley virome project conducted in this barley field.

We performed RT-PCR to verify the presence of virus-related RNAs in the aphid samples using the specific primer sets (**Supplementary Table S1**). The virus targets were successfully amplified from the aphid RNA samples (**Supplementary Figures S2**).

Characterization of Nege/Kita-Like Virus Sequences

We identified four contigs that resembled almost complete (coding complete) nege/kitavirus-like genome sequences (BaA1_c12; BaA3_c89; BaA3_c1889; and BaA3_c133,

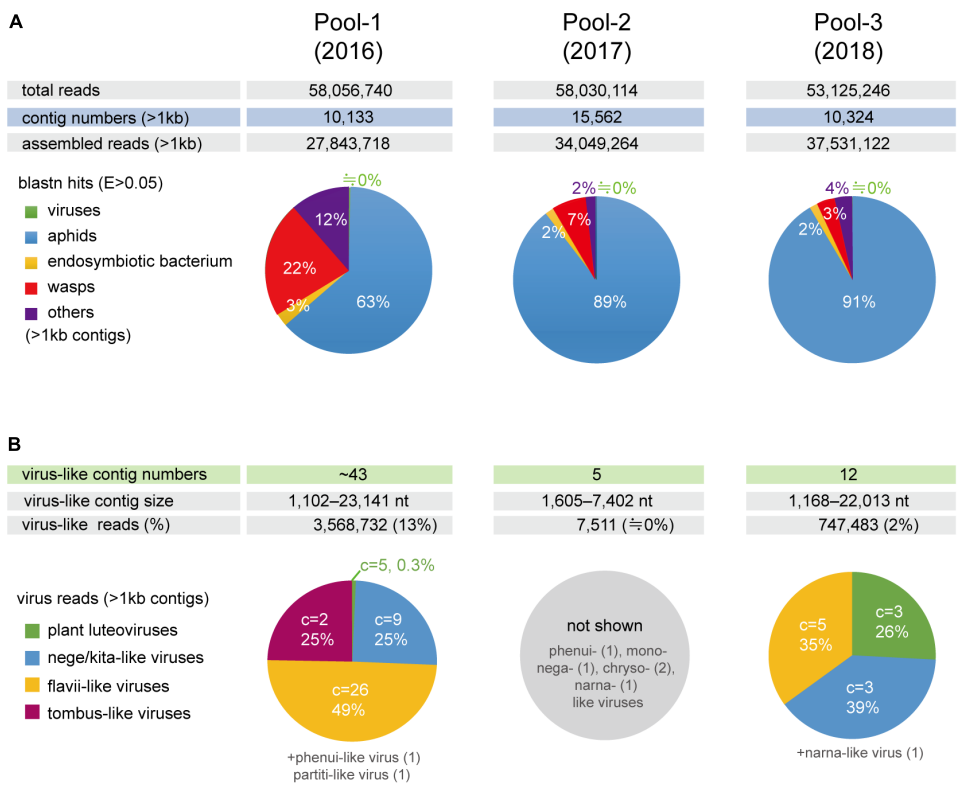


FIGURE 1 | Number of *de novo*-assembled contig sequences in the three libraries (pools 1–3) derived from the barley aphid transcriptomic analysis using NGS. **(A)** Distribution of the contigs based on organisms (aphids, parasitoid wasps, endosymbiont bacterium, viruses, and others) analyzed using BlastN matches with an *e*-Value < 0.05 for the selected reference sequence data (transcriptomic or genomic sequences) (see Section "Materials and Methods"). **(B)** The number of raw reads mapping to virus-like contig sequences related to different viral groups (families or proposed taxa). Charts show the percentage of raw read, estimated using a read mapping approach, according to the virus groups in each library, except for pool-2, where only a small number of virus-like contigs were identified. C = the number of virus-like contigs.

TABLE 2 | Major virus-like sequences from the barley aphid transcriptomes.

Contig name[1]	Size (nt)	Read no.	Virus or tentative[2] virus name	Aphid[3]	Accession no.
Nege/kita-like virus sequences					
BaA1_c12	10,318	328,424	Wuhan house centipede virus 1	Rp, Sa	LC516834
BaA3_c89[4a]	8,890	253,596	barley aphid RNA virus 1	Rm	LC516835
BaA1_c637[4a]	6,421	1,390	barley aphid RNA virus 1	Rp, Sa	
BaA3_c1889	9,398	22,346	barley aphid RNA virus 2	Rm	LC516836
BaA3_c133[4b]	9,321	18,472	barley aphid RNA virus 3	Rm	LC516837
BaA1_c42/165[4b,5]	9,322	303,799	barley aphid RNA virus 3	Rp, Rm, Sa	
BaA1_c14/63[5]	8,395	283,449	barley aphid RNA virus 4	Rp,	LC516838
Flavi-like virus sequences					
BaA3_c535	3,140	43,197	Wuhan aphid virus 1 S1	Rp, Rm, Sa	LC516839
BaA3_c172	1,168	57,745	Wuhan aphid virus 1 S2		LC516840
BaA3_c296	2,839	77,321	Wuhan aphid virus 1 S3		LC516841
BaA3_c581	2,758	77,906	Wuhan aphid virus 1 S4		LC516842
BaA1_c359	23,141	21,307	Sitobion miscanthi flavi-like virus 1	Rp, Sa	LC516843
BaA1_c652	15,288	5,199	barley aphid RNA virus 9	Rp, Sa	LC516844
BaA1_c1578	3,691	1,025	barley aphid RNA virus 9		
BaA1_c891	3,749	742	barley aphid RNA virus 9		
BaA3_c3690[4c]	21,997	5,333	barley aphid RNA virus 10	Rm	LC516845
BaA1_c326[4c]	21,210	18,335	barley aphid RNA virus 10	Rm	
BaA1_c33	10,551	1,440	barley aphid RNA virus 10	Rp	LC516846
BaA1_c194	5,869	877	barley aphid RNA virus 10		
Tombus-like virus sequences					
BaA1_c277[5]	3,702	198,247	Wuhan insect virus 21, S1	Rp, Sa	LC516847
BaA1_c31[5]	2,201	690,152	Wuhan insect virus 21, S2		LC516848
Phenui-like virus sequences					
BaA1_c346	7,403	2,040	barley aphid RNA virus 5	Rp, Rm	LC516849
BaA2_c4524	7,402	1,329	barley aphid RNA virus 5	Rm	LC516850
Other insect virus-like sequences					
BaA2_c1487	2,636	4,820	barley aphid RNA virus 6	Rm	LC516851
BaA3_c16152	2,057	239	barley aphid RNA virus 7	Rm	LC516852
BaA2_c19028	3,552	657	barley aphid RNA virus 8, S1	Rm	LC516853
BaA2_c18060	3,220	575	barley aphid RNA virus 8, S4		LC516854
BaA1_c21296	1,259	105	a partitivirus-like contig	Rm	LC516855
BaA2_c20409	1,605	127	a mononegavirus-like contig	Rm	LC516856

[1]Pool-1: BaA1; Pool-2: BaA2; Pool-3: BaA3. Some other variants of virus-like contigs are shown in ***Supplementary Figure S7****, or not shown. Plant luteovirus-related sequences are not listed here. [2]Nucleotide and/or amino acid sequence identity of the contig sequences are shown within* ***Table 3*** *and* ***Supplementary Table S4****. Ten discovered novel viruses were designated as barley aphid RNA viruses 1 to 10 (see main text). S, RNA or dsRNA segment. [3]Aphid species in discovery colonies assessed by RT-PCR detection (see* ***Supplementary Figure S1A****). Rp: Rhopalosiphum padi; Rm, R. maidis; Sa, Sitobion avenae. [4]These sequences are slightly different from each other at the nucleic acid level ([4a]: 3%, [4b]: 1%, and [4c]: 5% difference, respectively). [5]The gap or unassembled regions were sequenced using RT-PCR (see* ***Figure 2D*** *and* ***Supplementary Figure S7C****).*

8,890–10,318 nt in length) with 18,472–328,424 mapped reads (**Figure 2** and **Table 1**). Some other smaller fragments of nege/kitavirus-like sequences were also present in pools 1 and 3, but not in pool-2 (**Figure 2** and **Table 1**). Each of the three sets of contigs, "BaA1_c42 and BaA1_c165," "BaA1_c14 and BaA1_c63," and "BaA1_c637," and "BaA1_c28720" are most likely derived from three different variants of a single viral strain, respectively (**Figures 2B,D**). RT-PCR and sequencing were carried out to determine their sequence gaps and, thus, semi-entire virus-like sequences were obtained (namely, BaA1_c42/165 and BaA1_c14/63, 9,322 and 8,395 nt with 303,799 and 283,449 mapped reads, respectively) (**Figure 2D**). In the case of contigs BaA1_c637 (1,390 reads) and BaA1_c28720 (210 reads), an attempt to connect their sequences using RT-PCR failed, possibly because of a very low level of virus titer in the samples (data not shown). These two newly assembled contigs derived from four above-mentioned virus-like sequences have large replicase-like ORFs (ORF1) followed by two putative structural protein genes for predicted glycoprotein and membrane protein SP24 (ORFs 2 and 3) (Kuchibhatla et al., 2014; Solovyev and Morozov, 2017), except for the BaA1_c12 contig, which has an additional small ORF in its 3′-proximal region, or the BaA1_c14/63 contig, which lacks the 3′-terminal region sequence (Shi et al., 2016a; **Figure 2** and **Supplementary Figures S3**, **S4**).

The nucleotide sequence and encoded proteins of BaA1_c12 are almost identical (>99% nucleotide and 99–100% amino

FIGURE 2 | Genome organization of nege/kita-like virus sequences identified from the transcriptomes of barley aphids, *Rhopalosiphum padi; R. maidis*, and *Sitobion avenae*. **(A)** Wuhan house centipede virus 1 genome. **(B)** Hubei nege-like virus 4 genome (a reference sequence lacking the 3′ terminal part) and a related novel aphid virus (named barley aphid RNA virus 1, BARV-1) **(C)** Wuhan insect virus 9 genome (a reference sequence) and the second novel aphid virus (named barley aphid RNA virus 2, BARV-2). **(D)** The third and fourth novel aphid viruses (named barley aphid RNA virus 3 and 4, BARV-3 and -4) genome. The open boxes in the genomic RNA represent open reading frames (ORFs). The conserved domains in viral replicase are shown (methyltransferase, Mtr; FtsJ-like RNA ribosomal methyltransferase, AdoMet/FtsJ; RNA helicase, HEL; RNA-dependent RNA polymerase, RdRp). Other putative conserved domains (putative glycoprotein and SP24 protein) are indicated as gray highlights within the ORFs. Virus-like contigs identified from this study are shown as blue bars or blue lines. The sequence gaps between the two contigs for BARV-3 (BaA1_c42 and BaA1_c165) and BARV-4 (BaA1_c14 and BaA1_c63) were determined using RT-PCR and sequencing. The nucleotide sequence identities between a reference virus and a virus-like contig are shown in parentheses with black color. The nucleotide sequence identities between two virus-like contigs are shown in blue color with no parentheses. Read depth coverage throughout the assembled virus-like contigs together with two gap-filling virus-like sequences determined using RT-PCR (BaA3_c42/165 and BaA1_c14/63) is presented below the schematic structures of the viral genomes. The Y-axis shows the mapping-read coverage with the maximum read depth (read number) on each virus-like sequence. Green and red colors indicate positive and negative strand of sequence reads, respectively.

acid identities) to those of the Wuhan house centipede virus 1 (WHCV-1) (**Supplementary Table S3**), suggesting that the BaA1_c12 sequence is derived from a strain of WHCV-1, infecting two aphid species (*R. padi* and *S. avenae*) that infested barley plants. The two contigs BaA3_c89 and BaA1_c637 (lacking 3′-terminal region sequence) showed 97% nucleotide identity with each other and 71% nucleotide identity with Hubei virga-like virus 4 (HVLV-4) (**Figure 2B** and **Table 3**). A BlastP search revealed that the two contigs share moderate levels of amino acid sequence identity (45–58%) with HVLV-4 (**Table 3**). Although taxonomic criteria for species demarcation have not been established for nege-like viruses, both of the two HVLV-4-like sequences most likely represent a novel nege/kita-like virus that infects *R. padi* and *S. avenae* (BaA1_c637) or *R. maidis* (BaA3_c89) [our tentative cut-off value for the species classification criteria for nege-like viruses is <90% amino acid identities of RNA-dependent RNA polymerase (RdRp) sequences], rather than being a strain of HVLV-4; therefore we tentatively designated the virus as barley aphid RNA virus 1 (BARV-1).

Contig BaA3_c1889, derived from *R. maidis,* appeared to be a novel nege/kita-like virus related to Wuhan insect virus 9 (WhIV-9) because they share moderate levels of identities at the nucleotide (72%) and amino acid (66%–80%) sequence (**Tables 2** and **3**). We named this putative nege/kita-like virus as barley aphid RNA virus 2 (BARV-2). Two contigs, BaA3_c133 and BaA1_c14/63 (lacking 3′-terminal region sequence), together with BaA1_c42/165 (almost identical to the BaA3_c133 contig) also showed similar levels of nucleotide identity (72%) with WhIV-9 and its relatives, Hubei Wuhan insect virus 9 (HWIV-9) (**Table 3**). Their proteins showed lower levels of amino acid sequence identity (41%–58%) than those of the BaA3_c1889 contig (**Table 3**). In addition, they shared 77% nucleotide and 69% amino acid sequence identities between each other. Thus, contigs BaA3_c133 and BaA1_c42/165, and BaA1_c14/63 appear to represent two novel nege/kita-like viruses, which we tentatively designated as barley aphid RNA virus 3 and 4 (BARV-3 and -4), and it associated with *R. padi* (BaA1_c14/63, BARV-4); *R. maidis* (BaA3_c133, BARV-3); or *R. padi*, *R. maidis*, and *S. avenae* (BaA1_c42/165, BARV-3) (**Table 2** and **Supplementary Figure S2A**).

Search for Nege/Kita-Like Virus Sequences Using the TSA Database

Recently, the presence of several nege/kita-like viruses and related virus-like sequences, has been reported from invertebrate meta-transcriptome analyses or surveys of insect transcriptomic

TABLE 3 | Nege/kita-like virus sequences identified from the barley aphid transcriptomes using Blast search.

Query contig	Top hit virus	QC[1]	*E*-value	identify	Accession
BlastN search (discontiguous megablast)					
Query sequence: entire nucleotide sequences					
BaA1_c12	Wuhan house centipede virus 1	99%	0.0	99%	KX883740
BaA3_c89	Hubei virga-like virus 4	52%	0.0	71%	KX883814
BaA3_c1889	Wuhan insect virus 9	99%	0.0	72%	KX883816
BaA1_c42/165	Wuhan insect virus 9	53%	0.0	72%	KX883816
BaA1_c14/63	Hubei Wuhan insect virus 9	54%	0.0	73%	KX883782
BlastP search					
Query sequence: RdRp					
BaA1_c12	Wuhan house centipede virus 1	99%	0.0	99%	APG77795
BaA3_c89	Hubei virga-like virus 4	52%	0.0	58%	APG77770
BaA3_c1889	Wuhan insect virus 9	98%	0.0	78%	APG77775
BaA3_c133	Wuhan insect virus 9	53%	0.0	57%	APG77775
BaA1_c14/63	Hubei Wuhan insect virus 9	54%	0.0	58%	APG77665
Query sequence: p2 (putative glycoprotein)					
BaA1_c12	Wuhan house centipede virus 1	99%	0.0	99%	APG77796
BaA3_c89	Hubei virga-like virus 4	77%	2e-112	50%	APG77771
BaA3_c1889	Wuhan insect virus 9	96%	0.0	66%	APG77776
BaA3_c133	Wuhan insect virus 9[2]	87%	3e-86	41%	APG77776
BaA1_c14/634	Wuhan insect virus 9[2]	91%	1e-92	48%	APG77776
Query sequence: p3 (SP24 family)					
BaA1_c12	Wuhan house centipede virus 1	99%	4e-139	100%	APG77797
BaA3_c89	Wuhan house centipede virus 1	69%	3e-46	45%	APG77797
BaA3_c1889	Hubei Wuhan insect virus 9	90%	5e-117	80%	APG77670
BaA13_c133	Hubei Wuhan insect virus 9[2]	98%	3e-68	52%	APG77670
BaA1_c14/63	not available				

*[1]QC, query coverage (%). [2]The partial genome sequence of a Hubei Wuhan insect virus 9 variant (APG77666/APG77667) shows higher aa identity to each query sequence (59%–85%) (see **Supplementary Figures S4B–E**).*

datasets (Shi et al., 2016a; Kondo et al., 2019). To further explore the presence of unknown nege/kitavirus-like sequences, we conducted tBlastN searches against publicly available TSA datasets using BARV-1 and BARV-2 sequences as queries. The search identified several TSA sequences derived from hemipteran, hymenopteran, and dipteran species with a few plant species that represent nearly complete (>9.0 kb lengths) or partial viral genome sequences (**Supplementary Table S2** and data not shown). Two plant TAS sequences from a tree species, *Paulownia tomentosa* (family Paulowniaceae) (Fan et al., 2016), most likely represented the segments of a novel RNA virus belonging to the blunervirus-lineage (**Supplementary Table S2** and also see **Figure 3**).

Phylogenetic Relationships Among Nege/Kita- and Nege/Kita-Like Viruses

To understand the relationships between newly discovered nege/kita-like viruses, viral-like TSAs and reported nege/kitaviruses, we conducted ML phylogenetic analysis based on amino acid alignment of their replicase and replicase-like sequences. The ML tree showed that the BARV-1 variants clustered with HVLV-4 and its relatives, WHCV-1 and Wuhan insect virus 8 (WhIV-8) (referred to as "group 1"); whereas the BARV-2, -3, and -4 form a well-supported clade with the relatives of WhIV-9, and HWhIV-9, and ApGIV-3 from *Aphis glycines* (the soybean aphid) (referred to as "group 2") (**Figure 3**). These two virus groups together with other nege/kita-like viruses and arthropod virus-like TSAs, fill the phylogenetic gaps between negevirus-related (nelorpiviruses and sandewaviruses) and kitavirus-related (cile/higreviruses and blunerviruses) lineages (Vasilakis et al., 2013; Kallies et al., 2014; Nunes et al., 2017; **Figure 3** and **Supplementary Table S2**). Our phylogenetic analysis, thus, suggests that novel groups 1 and 2 of nege/kita-like viruses including aphid-infecting viruses belong to two new distinct genera, tentatively designated as "Centivirus" (after Wuhan house centipede virus 1) and "Aphiglyvirus" (after Aphis glycines virus 3), respectively. Members of these two proposed genera, together with two proposed taxa of arthropod-restrictive viruses, "Nelorpivirus" and "Sandewavirus," which all have non-segmented genomes, could be classified into a novel viral family or be assigned to the family *Kitaviridae*. However, it is to too early to establish a family for these viruses or to assign them into the *Kitaviridae*, because there is no reliable statistical support for these lineages within the currently available phylogenetic trees (Kallies et al., 2014; Nunes et al., 2017; Kondo et al., 2019; Ramos-González et al., 2020; **Figure 3**). The result of molecular phylogenetics indicate that the two major lineages of kitaviruses (cile/higreviruses and bulunerviruses) are not monophyletic, which all have two or multi-segmented RNA genomes and separately encode their replicase (RdRp) and SP24 family protein on the different segments (Locali-Fabris et al., 2006; Melzer et al., 2013; Hao et al., 2018), suggesting that genome segmentation possibly occurred independently during the evolution of kitaviruses. As previously proposed for plant rhabdoviruses (Kondo et al., 2017), the genome segmentation event may be associated with the vector mites (*Brevipalpus* spp.), at least in the case of the cile/higreviruses lineage.

Host Ranges and Spread of Nege/Kita-Like Viruses

WHCV-1 (a centivirus) was identified from the house centipede (*Scutigeridae* sp., phylum Arthropoda) and also from three other arthropods, spiders, and insects (WHCCII, insect mix 4) (Shi et al., 2016a). Because the last sample WHCCII included two aphid species (*Hyalopterus pruni* and *Aulacorthum magnolia*), a parasitoid wasp (*Aphelinus* sp., Hymenoptera), and some other insects such as dipterans, an orthopteran, and a coleopteran (Shi et al., 2016a), WHCV-1 may have a wide host range within arthropods (centipedes, spiders, and insects, including aphids as described in this study). Other centiviruses, WhIV-8 and HVLV-4, were also identified from the same insect pool (WHCCII) along with WhIV-9 (an aphiglyvirus), but their insect host ranges are still uncertain (Shi et al., 2016a). WHCV-1 and the two novel viruses, BARV-1 and BARV-3, were detected in four aphid colonies from the 2016 sample (WHCV-1 and BARV-1) or six aphid colonies from both 2016 and 2018 samples (BARV-3); whereas other viruses or variants were detected in a single aphid colony from the 2016 (BARV-4) and 2018 samples (BARV-2) (**Supplementary Figure S2A**). Notably, WHCV-1 and BARV-3 were detected in three different aphid species, *R. padi, R. maidis* and *S. avenae* collected in 2016 (**Supplementary Figure S2A**). Intriguingly, our preliminarily results showed the presence of these aphid virus-related sequences in the barley plant samples from the same field in the same year (HK and MF unpublishded results). Therefore, barley and/or other plants may potentially reserve aphid viruses and facilitate the horizontal transmission of the viruses among aphid populations, similarly to what was reported previously for an aphid picornavirus (RhPV) (Gildow and Darcy, 1988; Ban et al., 2007). WHCV-1-, BARV-3- and BARV-4-related sequences were identified in the TSA datasets derived from the Polish wheat (*Triticum polonicum*) (**Supplementary Figure S5**). WhIV-8- and BARV-2-like sequences were found in an allopolyploid bean species (*Glycine dolichocarpa*, Fabaceae) or the pomegranate (*Punica granatum,* Lythraceae) (**Supplementary Figure S5** and **Supplementary Table S2**), and some insect-specific viruses (densoviruses and dicistroviruses) have also been found from plant materials, such as cucumber, bean, *Brassica oleracea*, red cestrum (Solanaceae), maize and sea barley (van Munster et al., 2005; François et al., 2014; Maina et al., 2017; Wamonje et al., 2017; Roumi et al., 2020). Thus, these findings support the notion that plants serve as passive vectors for the transmission of aphid viruses.

WhIV-9, an aphiglyvirus from the WHCCII insect pool, has been identified from a dragonfly (*Odonata* sp., order Odonata) (Shi et al., 2016a), whereas WhIV-9-like TSA sequences was also found in *D. carpenter*, an aphid hyperparasitoid wasp (**Figure 3** and **Supplementary Table S4**). The presence of WhIV-9-like sequences in the hyperparasitoid wasp may suggest that the virus could be horizontally transmitted between this wasp and an aphid host through a primary parasitoid. In cereal aphids,

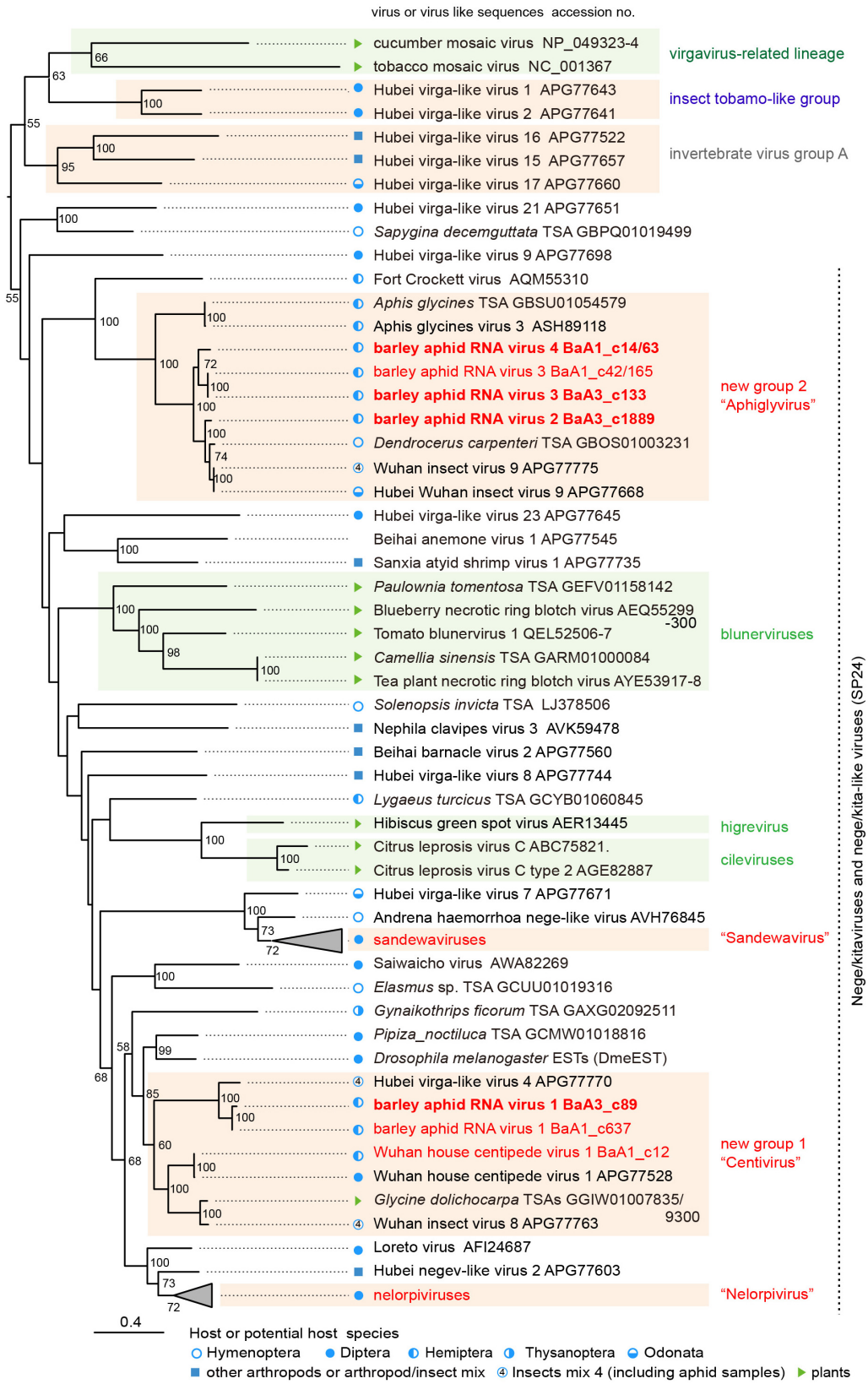


FIGURE 3 | Continued

FIGURE 3 | Phylogenetic relationships of the nega/kitaviruses and their related viruses and other virus-like sequences. The maximum likelihood (ML) phylogenetic tree was constructed using PhyML 3.0 based on a multiple amino acid sequence alignment of the replicase protein or its candidate sequences. A model LG + I + G + F was selected as a best-fit model for the alignment. The tree was rooted using the midpoint rooting method. Virus names referring to plant-infecting viruses (genera *Cilevirus, Higrevirus*, and *Blunervirus*, family *Kitaviridae*) and insect-infecting or associated viruses are followed by the GenBank accession numbers of their sequences. The selected members of the proposed groups "Sandewavirus" and "Nelorpivirus" are displayed as collapsed triangles (see **Supplementary Table S3** for the virus names in the collapsed triangles). Nege/kitavirus-like transcriptome shotgun assembly (TSA) sequences, discovered by this study (see **Supplementary Table S2**) or our previous study (Kondo et al., 2019), were also included in this analysis. A library of insect mixtures (WHCCII, insect virus 4) for the meta-transcriptomic analysis included two aphid species, a parasitoid wasp, and several other insects (Shi et al., 2016a). The hosts, or possible hosts, for nege/kitaviruses or virus-like sequences present in TSAs are marked with symbols (circles, squares, and triangles, respectively). Vertical dashed lines indicate that most of those viruses or virus-like sequences potentially encode conserved small proteins (SP24 family) (Kuchibhatla et al., 2014). The selected members of distantly related virus lineage, belonging to the plant virgavirus-related lineage (tobacco mosaic and cucumber mosaic viruses), insect tobamo-like virus group (Hubei virga-like viruses 1 and 2), and invertebrate virus group A (Hubei virga-like viruses 15, 16, and 17) (Kondo et al., 2019), were used as the outgroups. The scale bar represents amino acid distances. The numbers at the nodes are bootstrap values of >50%.

such as *S. avenae,* the attack of hyperparasitoids against primary parasitoids has been observed in the field (Holler et al., 1993). In our meta-transcriptome analyses, some aphid populations might be infected with the primary parasitoids, although it is still unknown whether the hyperparasitoids are also present in the IPSR field (**Supplementary Figure S2A**). A primary parasitoid (*L. testaceipes*), but not a secondary parasitoid (*D. carpenteri*), was detected in the aphid samples using RT-PCR (**Supplementary Figure S6**). It has been reported that an RNA virus, Lysiphlebus fabarum virus, related to BrBV (an aphid iflavirus), was found in an aphid parasitoid wasp (*Lysiphlebus fabarum*) (Lüthi et al., 2019). Its closest related virus, Venturia canescens picorna-like virus, replicates in both a parasitoid wasp and a lepidopteran host (Reineke and Asgari, 2005). An aphid picornavirus ALPV has been detected in the honey bees and its predator hornet *Vespa velutina* (Hymenoptera) (Granberg et al., 2013; Yang et al., 2019). Therefore, further studies are necessary to understand the dynamics and spread of nege/kita-like viruses among barley aphid populations, including their parasitoid and hyperparasitoid hymenopterans.

Characterization of Flavi-Like Virus Sequences

In addition to the nege/kita-like viruses (centi- and aphiglyviruses), we identified several contigs related to other RNA viruses, such as flavi/jingmen- and tombus-like virus sequences (**Figure 1** and **Table 2**). WhAV-1, a jingmenvirus, which belongs to a recently identified group of the four-segmented RNA viruses related to unsegmented flaviviruses (Shi et al., 2016b; Temmam et al., 2019), was one of the major constituents of flavi/jingmenvirus-like reads (**Figure 1** and **Table 2**). Sequence homology analyses suggest that at least three variants of WhAV-1 (1a, 1b, and 1c variants) were associated with barley aphids (**Supplementary Figure S7A** and **Supplementary Table S4**). The "1b" variant of WhAV-1 (the BaA3_c535 contig and some others), which is related to the reference viral sequence sharing ∼87% nucleotide and ∼95% amino acid sequence identities, was found in five aphids colonies containing all three species from the 2016 sample and *R. maidis* from the 2018 sample (**Supplementary Figure S2B**). In contrast, the "1a" (the BaA1_c139 contig) and "1c" (the BaA1_c13 contig) variants, which are related to the reference sequence sharing 91% and 78% nucleotide and 92% and 87% amino acid sequence identities, were only detected in an *R. padi* colony from the 2016 sample (BaA1_c139) and in two aphid colonies (*R. padi* and *R. maidis*) from the 2016 sample (BaA1_c13), respectively (**Supplementary Figure S2B** and data not shown). WhAV-1b sequences (BaA3_c535, c172, c296, and c581 for RNA segments 1–4) have been deposited into the DDBJ data bank as a representative genome sequence (**Table 2**). In the phylogenetic tree, based on the non-structural protein NSP1 (NS5-like, RdRp) encoded in segment 1, jingmenviruses are divided into two different clades; the WhAV-1 variants and a second aphid-associated jingmenvirus, WhAV-2, form a subclade within the clade consisting of insect and some other arthropod jingmenviruses; whereas the second clade contains jingmenviruses infecting ticks (class Arachnida), such as jingmen tick virus, which may be a novel tick-borne arbovirus that infects humans and other mammals (Jia et al., 2019; **Supplementary Figure S7B**). WhAV-1 sequences were identified in the aphid-free barley samples as mentioned above (HK and MF unpublishded results), and interestingly, a variant of WhAV-2 was found in a pea plant transcriptome (*Pisum sativum*, accessions MK948535–8, deposited by Y.Z.A. Gaafar and H. Ziebell) (**Supplementary Figure S2B** and **Supplementary Table S2**), raising a speculation that jingmenviruses that infect aphids may have been horizontally transmitted among aphid populations, with plants as the reservoirs (Gildow and Darcy, 1988; Ban et al., 2007).

Another type of flavivirus-like sequences seems to be the minor populations among aphid flavi-related viruses (**Table 2**). The BaA1_c359 contig is derived from a variant Sitobion miscanthi flavi-like virus 1 (SmFLV-1, MH778148 deposited by T. Li), showing 92% nucleotide and 98% amino acid sequence identities (**Supplementary Table S4**). The BaA1_c652 contig (probably together with BaA1_c1578 and c891 contigs) sharing 73% nucleotide and 62% amino acid sequence identities with that of MeV-1 from European *M. euphorbiae* populations (Teixeira et al., 2016) and the BaA3_c3690 together with its two variants BaA1_c326 and c33 (a partial sequence) contigs showing 67–68% nucleotide and 44–59% amino acid sequence identities with that of SmFLV-1, are likely to represent two novel flavi-like viruses. Thus, we tentatively named barley aphid RNA virus 9 (BARV-9) and barley aphid RNA virus 10 (BARV-10) (**Figure 4A** and **Table 2**). The complete coding sequence

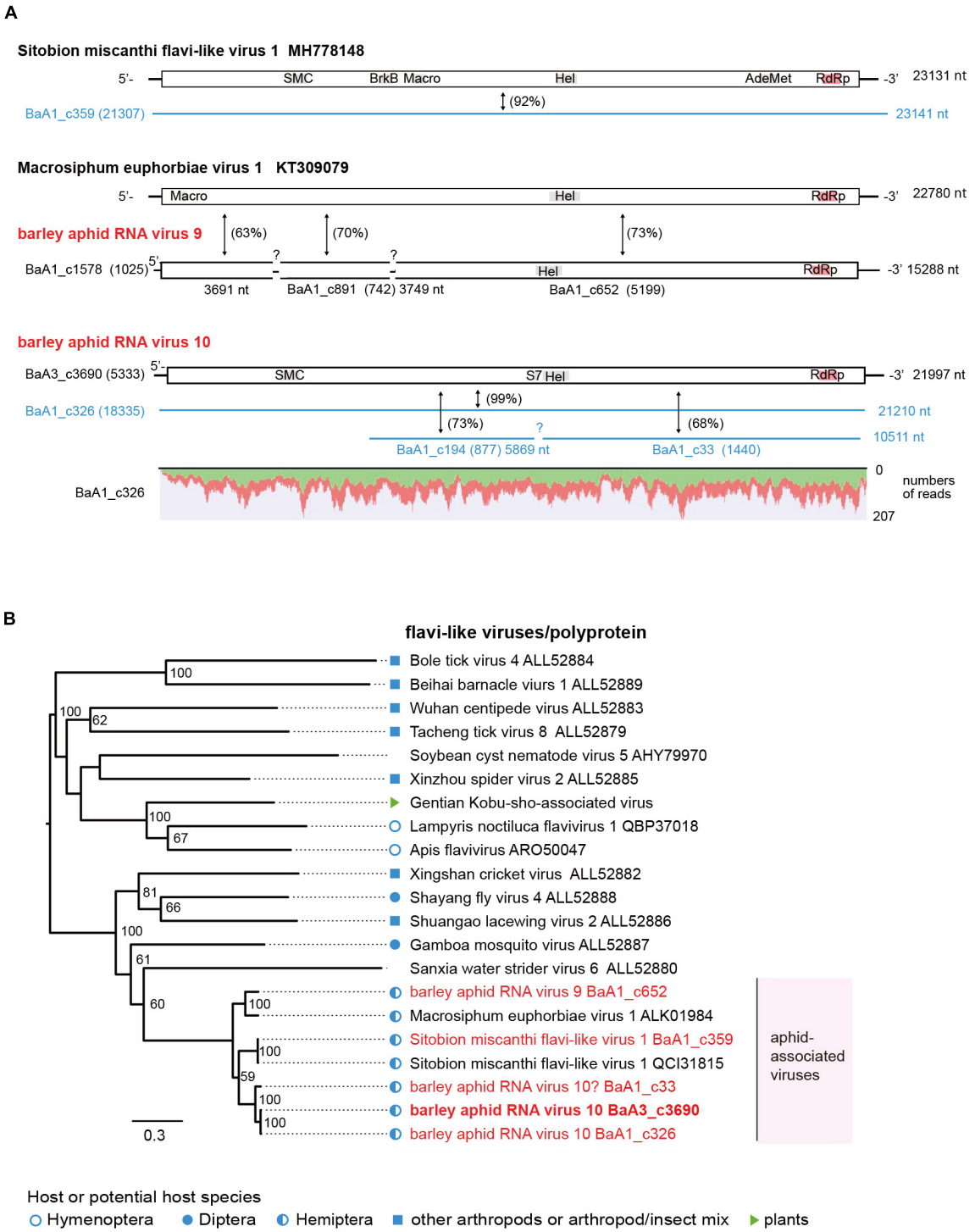


FIGURE 4 | Genome organizations and phylogenetic relationships of flavi-like viruses. **(A)** The genome structures of flavi-like viruses, Sitobion miscanthi flavi-like virus 1, Macrosiphum euphorbiae virus 1, and two novel flavi-like viruses (named barley aphid RNA virus 9 and 10, BARV-9 and -10) from *Rhopalosiphum padi, R. maidis*, and *Sitobion avenae*. Conserved domains for viral polyproteins or other proteins are shown as gray or red highlights within the ORFs (Macro domain, a high-affinity ADP-ribose binding module; SMC super family, chromosome segregation ATPase; BrkB, This family may act as a virulence factor; S7, Flavivirus NS3 serine protease; HEL, RNA helicase; AdoMet, S-adenosylmethionine-dependent methyltransferases; RdRp, RNA-dependent RNA polymerase). Virus-like contigs derived from this study are represented as black bars or blue lines and other information is the same as shown in **Figure 2**. Read depth coverage throughout the assembled virus contig is shown below the schematic structure of the viral genome. **(B)** Phylogenetic relationships of flavi-like viruses. The ML phylogenetic tree was constructed using a multiple amino acid sequence alignment of the polyprotein or its candidate sequences. The scale bar represents amino acid distances. The numbers at the nodes are bootstrap values of >50%.

of BARV-3 (BaA3_c3690) is 21,997 nt in length, similar to the genomes of MeV1 (22,780 nt) and SmFLV-1 (23,131 nt) (**Figure 4A**). The single large ORF of BARV-9 encodes a putative polyprotein of 7,081 amino acids, containing putative helicase, NS3 serine protease, and RdRp motifs, but not the motifs predicted for structural proteins and methyltransferase (**Figure 4A**). In the phylogenetic tree based on polyproteins (**Figure 4B**), aphid-associated flavi-like viruses form a clade and clustered with some invertebrate flavi-like viruses, whereas its sister clade contains several invertebrate flavi-like viruses and a plant virus, Gentian Kobu-sho-associated virus (Kobayashi et al., 2013; Bekal et al., 2014; Shi et al., 2016b; Remnant et al., 2017), that still have an uncertain taxonomic status. SmFLV-1 and BARV-9 variants were found in two aphid colonies (*R. padi* and *S. avenae*) from the 2016 sample and SmFLV-1 probably also in a *R. maidis* colony from the 2018 sample in which low levels of viral RNA accumulation was detected using RT-PCR, whereas no virus reads were found in the NGS data, thus expanding its aphid host range. BARV-9-like TSA sequences were also found in *S. avenae* (**Supplementary Figure S5**). BARV-10 variants were detected in two *R. maidis* colonies from the 2016 and 2018 samples (BaA1_c326) or an *R. padi* colony from the 2016 sample (BaA1_c33) (**Supplementary Figure S2B**). Similar to the nege/kita-like viruses, multiple flavi-like viruses belonging to at least three viruses (BARV-9, BARV-10 and SmFLV-1) were present in barley aphid populations that infest the IPSR field; however, their prevalence in the aphid populations seems to be much lower than that of nege/kita-like viruses and a jingmenvirus (WhAV-1) (**Table 2** and **Supplementary Figure S2**).

Characterization of Tombus-Like Virus Sequences

Recently, the diversity of the tombusvirus-like lineage has been greatly expanded by the discovery of a large number of tombus-like viruses from invertebrates (Shi et al., 2016a). One of these viruses, Wuhan insect virus 21 (WhIV-21), was detected in the WHCCII insect pool; the same pool in which the above mentioned viruses were detected. In this study, we found a variant of WhIV-21 (BaR1_c277 and c31 contigs, sharing 83% and 73% nucleotide and 90% and 50% amino acid sequence identities with WhIV-21) in *R. padi* and *S. avenae* aphids from the 2016 sample, but not from *R. maidis* (**Table 2**, **Supplementary Figure S2B**, and **Supplementary Table S4**). The complete coding sequences of WhIV-21 were verified using RT-PCR and sequencing (**Supplementary Figure S7C**). The read numbers of WhIV-21 (198,247 reads for BaR1_c277, and 690,152 reads for BaR1_c31), whose sequences were also detected in the barley samples (HK and MF unpublishded results), were higher than those of jinmenviruses and similar to some nege/kitaviruses (**Table 2**). The ML phylogenetic tree, based on RdRp sequences encoded by the RNA1 segments, showed that WhIV-21 is clustered with a bee pathogen, Chronic bee paralysis virus (CBPV, proposed genus "Chroparavirus"), and several invertebrate tombus-like viruses (Shi et al., 2016a; Bigot et al., 2017; **Supplementary Figure S7D**). Similar to CBPV, which may form a ellipsoidal virion, the WhIV-21 RNA2 segment seems to encode two proteins; a putative virion glycoprotein and SP24 homolog that probably have features of the major structural components, similar to what was previously proposed for CBPV and some negeviruses (Kuchibhatla et al., 2014; Solovyev and Morozov, 2017).

Characterization of Other Novel Aphid-Associated Virus Sequences

Two negative-sense RNA virus-like sequences (BaA1_c346 and BaA2_c4524, both minor read numbers within the data set) represent phenuiviruses (family *Phenuiviridae*) (**Figure 5A** and **Table 2**). These two sequences are closely related to each other (82% nucleotide and 94% amino acid sequence identities) and their encoded proteins (2,344 or 2,345 amino acids) are distantly related to L proteins (RdRp) of phenui-like viruses, such as Hubei bunya-like virus 2 from a dragonfly (35% amino acid sequence identity) (Shi et al., 2016a; **Supplementary Table S4**). We tentatively named this potentially novel virus as barley aphid RNA virus 5 (BARV-5). The BaA1_c346 related sequences were detected in all aphid colonies (*R. padi, R. maidis* and probably *S. avenae*) from the 2016 sample, whereas the BaA2_c4524 sequence was only detected in *R. maidis* aphids from the 2017 sample (**Supplementary Figure S2B**). In the phylogenetic tree, based on L proteins, BARV-5 was clustered with some invertebrate phenui-like viruses that will probably be classified as the members of a novel genus (or genera). A neighboring clade consisted of recently discovered phenui-like viruses that have bi- or tripartite RNA genomes and are associated with plants, fungi, and invertebrates, belonging to the proposed viral genera (Navarro et al., 2018; Tokarz et al., 2018; Lin et al., 2019; **Figure 6B**). Similar to these phenui-like viruses, BARV-5 may also have additional RNA segment(s) encoding other viral proteins, such as the nucleocapsid protein. In addition to the phenui-like virus, one negative-sense mononega-like virus sequence (BaA2_c20409, a 1.6 kb fragment) was detected in *R. maidis* aphids from the 2017 sample (**Table 2** and **Supplementary Figure S2B**). This contig sequence showed a moderate level of sequence similarity (40% amino acid identity) with the L protein of *Linepithema humile* rhabdo-like virus from the queen samples of the Argentine ant (*Linepithema humile*) (Viljakainen et al., 2018; **Supplementary Table S4**) and forms a clade of unassigned non-segmented RNA viruses (order *Mononegavirales*), together with Hubei rhabdo-like virus 3 and Tacheng tick virus 6 (Shi et al., 2016a) (data not shown).

Two positive sense RNA virus-like sequences (BaA2_c1487 and BaA3_c16152) were identified in *R. maidis* aphids from the 2017 sample (BaA2_c1487, a complete coding 2.6 kb sequence) and the 2018 sample (BaA3_c16152, a fragment of 2.0 kb) (**Table 2**, **Figure 6A** and **Supplementary Figures S2B**). Both sequences potentially encoded an RdRp related to that of a narna-like virus (Aedes angustivittatus narnavirus) sharing 34% (BaA2_c1487) or 30% (BaA3_c16152) amino acid sequence identities (Fauver et al., 2019; **Supplementary Table S4**). We named these two putative narna-like viruses as barley aphid RNA virus 6 and 7 (BARV-6 and -7). Although the BaA3_c16152 contig lacks terminal region sequences, BARV-7 may have a genome

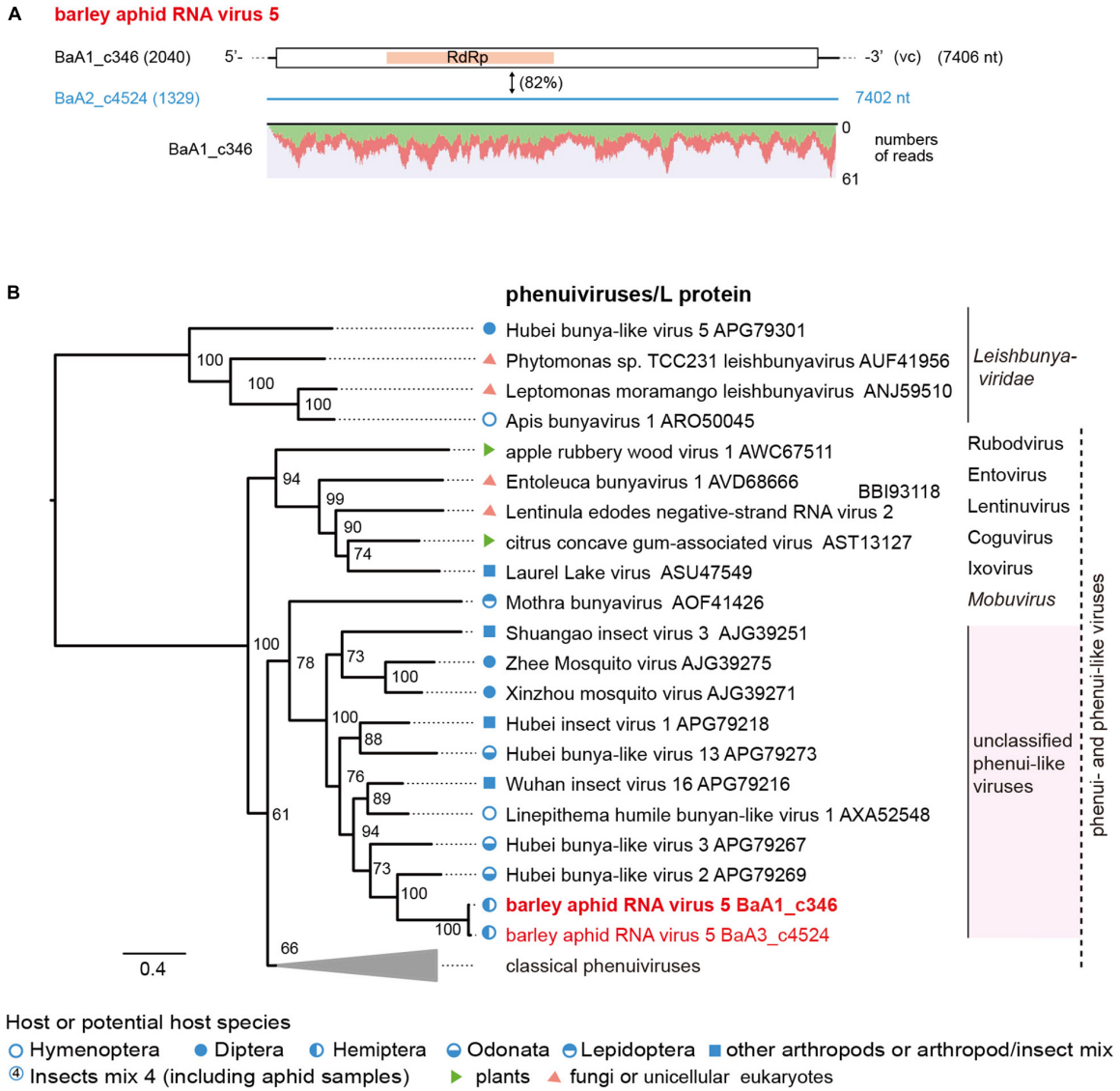


FIGURE 5 | The genome organization and phylogenetic relationships of phenui- and phenui-like viruses. **(A)** The genome structure of the putative L segments of a phenui-like virus (named barley aphid RNA virus 5, BARV-5) identified from *Rhopalosiphum padi*, *R. maidis*, and *Sitobion avenae* or *R. maidis* transcriptomes. The conserved domain for viral L proteins is shown as a red highlight within the ORF (RNA-dependent RNA polymerase, RdRp). Virus-like contigs in this study are shown as a black bar or a blue line, together with their read numbers in parentheses, and the nucleotide sequence identity between two virus-like contigs is shown. Read depth coverage throughout the assembled virus contig is shown below the schematic structure of the viral genome. **(B)** Phylogenetic relationships of BARV-5 and other selected phenuiviruses (family *Phenuiviridae*) and phenui-like viruses. The ML phylogenetic tree was constructed using a multiple amino acid sequence alignment of the L proteins or its candidate sequences. The selected members of the genera within the family *Phenuiviridae* are displayed as a collapsed triangle (see **Supplementary Table S3** for the virus names in the collapsed triangle). Some members of family *Leishbuniyaviridae* were used as the outgroups. The scale bar represents amino acid distances. The numbers at the nodes are bootstrap values of >50%.

that potentially encodes a reverse-ORF, as previous studies have proposed for some narna-like viruses, such as Hubei narna-like virus 21 and Zhejiang mosquito virus 3 (Cook et al., 2013; DeRisi et al., 2019; Dinan et al., 2020; **Figure 6A**). In the phylogenetic tree, BARV-6 and -7 grouped together with yeast narnaviruses (family *Narnaviridae*) and many invertebrate-associated narna-like viruses (**Figure 6B**).

Two dsRNA virus-like contigs (BaA2_c19028 and BaA2_c18060) were derived from an *R. maidis* colony from the 2016 sample (**Table 2**, **Figure 6C**, and **Supplementary Figures S2B** and data not shown). According to BlastP search results, both sequences were most closely related to the segments S1 and S4 of Shuangao chryso-like virus 1, with 37% and 24% amino acid sequence identities, respectively (Shi et al., 2016a) (**Supplementary Table S4**). However, we did not find the possible segments corresponding to S2 and S3. We tentatively named this chryso-like virus as barley aphid RNA virus 8 (BARV-8). Phylogenetically, chrysoviruses are grouped into two distinct

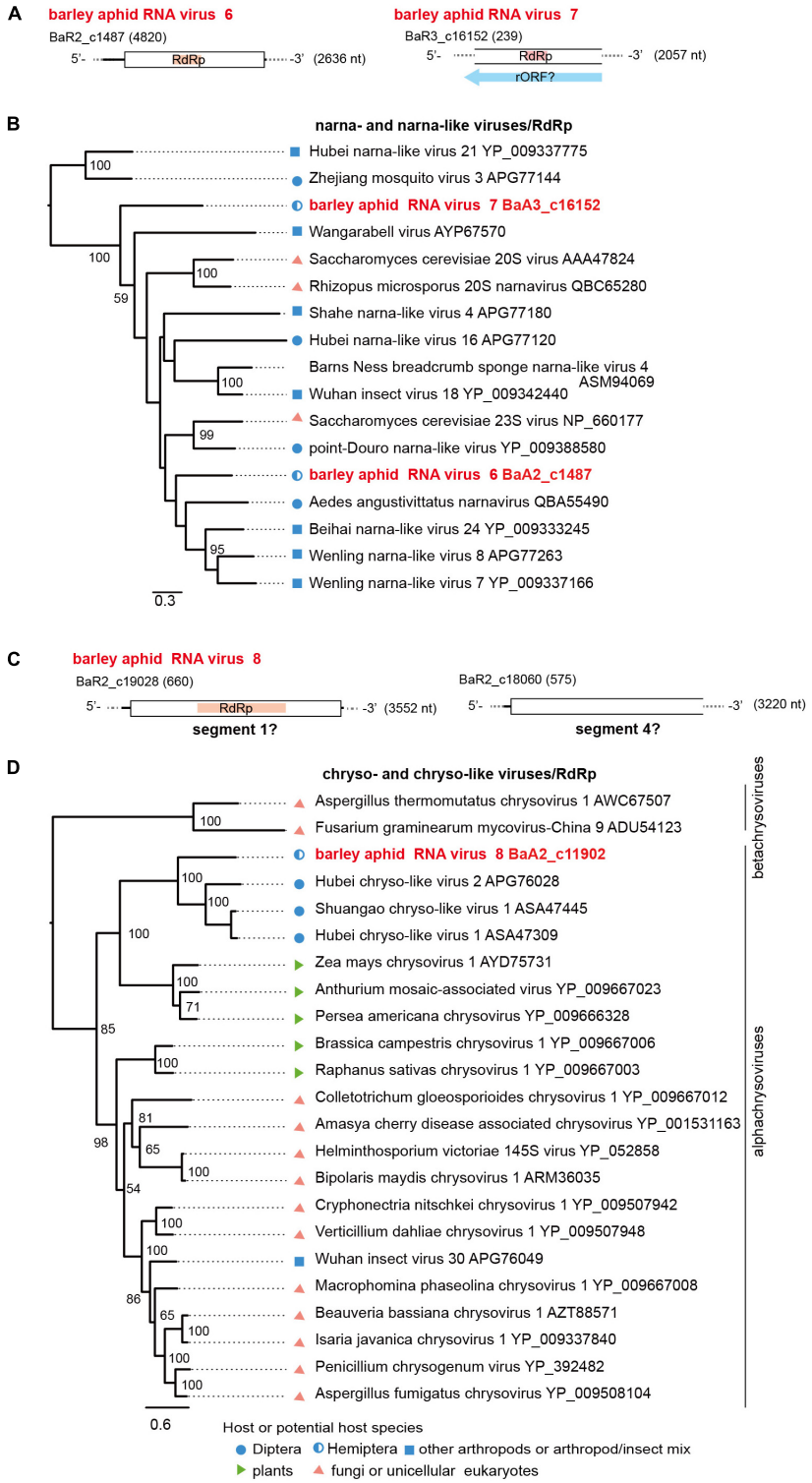


FIGURE 6 | Genome organizations and phylogenetic relationships of narna- and narna-like viruses **(A,B)** and chryso-and chryso-like viruses **(C,D)**. **(A,C)** The genome structures of two novel narna-like viruses (named barley aphid RNA virus 6 and 7, BARV-6 and -7) in panel **(A)** and a chryso-like virus (named barley aphid RNA virus 8, BARV-8) in panel **(C)**. rORF: reverse-frame ORF. **(B,D)** Phylogenetic relationships of two novel narna-like viruses BARV-6 and -7, yeast narnaviruses and invertebrate narna-like viruses **(B)**, and BARV-8, chrysoviruses (the members of genus *Alphachrysovirus*) and chryso-like viruses **(D)**. The ML trees were constructed using multiple amino acid sequence alignments of the RdRp sequences. Two narna-like viruses (Zhejiang mosquito virus 3 and Hubei narna-like virus 21) **(B)** or some members of genus *Betachrysovirus* **(D)** were used as the outgroups, respectively. The scale bar represents amino acid distances. The numbers at the nodes are bootstrap values of >50%.

clusters that are currently classified into *Alphachrysovirus* and *Betachrysovirus* (Kotta-Loizou et al., 2020; **Figure 6D**). BARV-8 (S1 segment, RdRp) is placed within the clade of alphachrysoviruses and forms a subclade with chryso-like viruses derived from dipterans (Shi et al., 2016a). In addition, in two aphid colonies (*R. maidis* and probably *R. padi*) from the 2016 sample, we also identified another dsRNA virus-like sequence (BaA2_21290, a 1.3 kb fragment) (**Table 2** and **Supplementary Figure S2B**) whose RdRp sequence was closely related to that of invertebrate alphapartiti-like viruses (Hubei partiti-like virus 25–28, 55%–66% amino acid identities) from spiders or dragonflies (Shi et al., 2016a) and alphapartitiviruses that infect plant and filamentous fungi (**Supplementary Table S4** and data not shown).

CONCLUSION

In the current study, we discovered at least 60 virus-like sequences related to nege/kita-, flavi/ jingmen-, tombus-, phenui-, mononega-, narna-, chryso-, partiti-, and luteoviruses from 15 aphid population samples (*R. padi*, *R. maidis*, and *S. avenae*) that were collected from a barley field in the spring of 2016, 2017, and 2018. From these sequences we identified eight potentially novel RNA viruses belonging to nege/kita-, flavi-, phenui-, mononega-, narna- and chrysovirus lineages, as well as some previously described RNA viruses. Moreover, based on the phylogenetic analyses, we proposed novel genera, Centivirus and Aphiglyvirus, for aphid associated nege/kita-like viruses and their relatives. Our data provide novel information on the diversity of aphid-associated viruses in aphid populations infesting the barley field. Our study, along with others, has discovered that some aphid-associated viruses are present in different aphid species and in plant-derived samples. This raises the speculation that aphid-associated viruses may be distributed across varieties of aphid species, with plants being the reservoirs. Since aphid virus-like sequences are also present in the hyperparasitoid wasp, it would also be interesting to explore the possible horizontal transmission of viruses between aphids and primary parasitoids and/or hyperparasitoids. To deepen our understanding on the population dynamics and spread of aphid viruses in the field, investigation of horizontal virus transmissions between insects and plants as well as between insects and their parasitoids is an important topic for future research. Some aphid viruses cause diseases in their hosts, for example, rapid population decline and phenotype alternations of insect performances (Williamson et al., 1988; vandenHeuvel et al., 1997; Moon et al., 1998; van Munster et al., 2002), implying their potentiality to be used as biological control agents against aphid pests (van Munster et al., 2005; Feng et al., 2017). On the other hand, a densovirus (DplDNV) that induces the production of winged morphs in the rosy apple aphids (Ryabov et al., 2009) is regarded as a conditionally mutualistic symbiont because this relationship facilitates the movement of aphids between the host plants (Roossinck, 2011, 2015). Moreover, a recent study reported that an insect virus (APV) facilitates host aphid adaptation to the host plant by suppressing the plant's defense response, demonstrating the unique interactions among virus, aphid and plant (Lu et al., 2020). Our data lay the foundation for further exploration of the ecological roles of aphid-associated viruses in barley ecosystem, either in a beneficial or harmful way as mentioned above. The physiological and ecological roles of aphid-specific viruses in aphid populations in the field warrant further investigation.

DATA AVAILABILITY STATEMENT

The virus and virus-like sequences derived from this study can be found in GenBank under the accession numbers LC516834 – LC516856.

AUTHOR CONTRIBUTIONS

HK designed the experiments and analyzed the data and wrote the manuscript. HH and HK collected the samples. HK, MF, HH, KH, IA, and NS performed the experimental work. IA and NS were involved in discussion and manuscript revision. All authors have given approval to the final version of the manuscript.

FUNDING

This study was supported by the Grants-in-Aids for Scientific Research (C) (KAKENHI 15K07312) to HK and NS; and Scientific Research on Innovative Areas (16H06436, 16H06429, and 16H21723) to NS and HK from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan; and a Joint Research Grant to HH and HK by the Ohara Foundation for Agriculture Research, Kurashiki, Japan.

ACKNOWLEDGMENTS

We would like to thank Sakae Hisano, Kazuyuki Maruyama, Hideki Nishimura, Takakazu Matsuura, and Makoto Ishii for their helpful technical assistance. We also thank the handling editor and two reviewers for their assistances and valuable suggestions. *Rhopalosiphum padi* DNA sequencing data were obtained from AphidBase, which acknowledges funding from the European Research Council (Grant No. APHIDHOST-310190 to Dr. Jorunn Bos at the James Hutton Institute). The barley samples were provided by the National BioResource Project for barley (Principal Investigator Prof. Kazuhiro Sato), run by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. We are also grateful to all researchers who have shared their sequence data deposited in publicly available databases before publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.00509/full#supplementary-material

REFERENCES

Ban, L., Didon, A., Jonsson, L. M. V., Glinwood, R., and Delp, G. (2007). An improved detection method for the *Rhopalosiphum padi* virus (RhPV) allows monitoring of its presence in aphids and movement within plants. *J. Virol. Methods* 142, 136–142. doi: 10.1016/j.jviromet.2007.01.014

Bekal, S., Domier, L. L., Gonfa, B., McCoppin, N. K., Lambert, K. N., and Bhalerao, K. (2014). A novel flavivirus in the soybean cyst nematode. *J. Gen. Virol.* 95, 1272–1280. doi: 10.1099/vir.0.060889-0

Bigot, D., Dalmon, A., Roy, B., Hou, C. S., Germain, M., Romary, M., et al. (2017). The discovery of *Halictivirus* resolves the *Sinaivirus* phylogeny. *J. Gen. Virol.* 98, 2864–2875. doi: 10.1099/jgv.0.000957

Bolling, B. G., Weaver, S. C., Tesh, R. B., and Vasilakis, N. (2015). Insect-specific virus discovery: significance for the arbovirus community. *Viruses* 7, 4911–4928. doi: 10.3390/v7092851

Brault, V., Uzest, M., Monsion, B., Jacquot, E., and Blanc, S. (2010). Aphids as transport devices for plant viruses. *C. R. Biol.* 333, 524–538. doi: 10.1016/j.crvi.2010.04.001

Chen, Y., Pike, K. S., Greenstone, M. H., and Shufran, K. A. (2006). Molecular markers for identification of the hyperparasitoids *Dendrocerus carpenteri* and *Alloxysta xanthopsis* in *Lysiphlebus testaceipes* parasitizing cereal aphids. *Biocontrol* 51, 183–194. doi: 10.1007/s10526-005-1518-0

Colmant, A. M., O'Brien, C. A., Newton, N. D., Watterson, D., Hardy, J., Coulibaly, F., et al. (2020). Novel monoclonal antibodies against Australian strains of negeviruses and insights into virus structure, replication and host-restriction. *J. Gen. Virol.* doi: 10.1099/jgv.0.001388 [Epub ahead of print],

Cook, S., Chung, B. Y. W., Bass, D., Moureau, G., Tang, S. Y., McAlister, E., et al. (2013). Novel virus discovery and genome reconstruction from field RNA samples reveals highly divergent viruses in dipteran hosts. *PloS One* 8:e80720. doi: 10.1371/journal.pone.0080720

Dader, B., Then, C., Berthelot, E., Ducousso, M., Ng, J. C. K., and Drucker, M. (2017). Insect transmission of plant viruses: multilayered interactions optimize viral propagation. *Insect Sci.* 24, 929–946. doi: 10.1111/1744-7917.12470

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

DeRisi, J. L., Huber, G., Kistler, A., Retallack, H., Wilkinson, M., and Yllanes, D. (2019). An exploration of ambigrammatic sequences in narnaviruses. *Sci. Rep.* 9:17982. doi: 10.1038/s41598-019-54181-3

Dinan, A. M., Lukhovitskaya, N. I., Olendraite, I., and Firth, A. E. (2020). A case for a negative-strand coding sequence in a group of positive-sense RNA viruses. *Virus Evol.* 6:veaa007. doi: 10.1093/ve/veaa007

Fan, G., Wang, L., Deng, M., Zhao, Z., Dong, Y., Zhang, X., et al. (2016). Changes in transcript related to osmosis and intracellular ion homeostasis in *Paulownia tomentosa* under salt stress. *Front. Plant Sci.* 7:384. doi: 10.3389/fpls.2016.00384

Fauver, J. R., Akter, S., Morales, A. I. O., Black, W. C., Rodriguez, A. D., Stenglein, M. D., et al. (2019). A reverse-transcription/RNase H based protocol for depletion of mosquito ribosomal RNA facilitates viral intrahost evolution analysis, transcriptomics and pathogen discovery. *Virology* 528, 181–197. doi: 10.1016/j.virol.2018.12.020

Feng, Y., Krueger, E. N., Liu, S., Dorman, K., Bonning, B. C., and Miller, W. A. (2017). Discovery of known and novel viral genomes in soybean aphid by deep sequencing. *Phytobiomes* 1, 36–45. doi: 10.1094/PBIOMES-11-16-0013-R

Foottit, R. G., Maw, H. E. L., Von Dohlen, C. D., and Hebert, P. D. N. (2008). Species identification of aphids (Insecta: Hemiptera: Aphididae) through DNA barcodes. *Mol. Ecol. Resour.* 8, 1189–1201. doi: 10.1111/j.1755-0998.2008.02297.x

François, S., Bernardo, P., Filloux, D., Roumagnac, P., Yaverkovski, N., Froissart, R., et al. (2014). A novel itera-like densovirus isolated by viral metagenomics from the sea barley *Hordeum marinum*. *Genome Announc.* 2, e01196-14. doi: 10.1128/genomeA.01196-14

Freitas-Astua, J., Ramos-Gonzalez, P. L., Arena, G. D., Tassi, A. D., and Kitajima, E. W. (2018). *Brevipalpus*-transmitted viruses: parallelism beyond a common vector or convergent evolution of distantly related pathogens? *Curr. Opin. Virol.* 33, 66–73. doi: 10.1016/j.coviro.2018.07.010

Fujita, R., Kuwata, R., Kobayashi, D., Bertuso, A. G., Isawa, H., and Sawabe, K. (2017). Bustos virus, a new member of the negevirus group isolated from a *Mansonia* mosquito in the Philippines. *Arch. Virol.* 162, 79–88. doi: 10.1007/s00705-016-3068-4

Gildow, F. E., and Darcy, C. J. (1988). Barley and oats as reservoirs for an aphid virus and the influence on barley yellow dwarf virus transmission. *Phytopathology* 78, 811–816.

Granberg, F., Vicente-Rubiano, M., Rubio-Guerri, C., Karlsson, O. E., Kukielka, D., Belak, S., et al. (2013). Metagenomic detection of viral pathogens in Spanish honeybees: co-infection by aphid lethal paralysis, Israel acute paralysis and Lake *Sinai* viruses. *PloS One* 8:e57459. doi: 10.1371/journal.pone.0057459

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Sys. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Hao, X. Y., Zhang, W. F., Zhao, F. M., Liu, Y., Qian, W. J., Wang, Y. C., et al. (2018). Discovery of plant viruses from tea plant (*Camellia sinensis* (L.) *O. Kuntze*) by metagenomic sequencing. *Front. Microbiol.* 9:2175. doi: 10.3389/fmicb.2018.02175

Holler, C., Borgemeister, C., Haardt, H., and Powell, W. (1993). The relationship between primary parasitoids and hyperparasitoids of cereal aphids – an analysis of field data. *J. Anim. Ecol.* 62, 12–21. doi: 10.2307/5478

Jarosova, J., Chrpova, J., Sip, V., and Kundu, J. K. (2013). A comparative study of the barley yellow dwarf virus species PAV and PAS: distribution, accumulation and host resistance. *Plant Pathol.* 62, 436–443. doi: 10.1111/j.1365-3059.2012.02644.x

Jia, N., Liu, H. B., Ni, X. B., Bell-Sakyi, L., Zheng, Y. C., Song, J. L., et al. (2019). Emergence of human infection with Jingmen tick virus in China: a retrospective study. *EBioMedicine* 43, 317–324. doi: 10.1016/j.ebiom.2019.04.004

Kallies, R., Kopp, A., Zirkel, F., Estrada, A., Gillespie, T. R., Drosten, C., et al. (2014). Genetic characterization of goutanap virus, a novel virus related to negeviruses, cileviruses and higreviruses. *Viruses* 6, 4346–4357. doi: 10.3390/v6114346

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kawakami, K., Kurnia, Y. W., Fujita, R., Ito, T., Isawa, H., Asano, S., et al. (2016). Characterization of a novel negevirus isolated from *Aedes* larvae collected in a subarctic region of Japan. *Arch. Virol.* 161, 801–809. doi: 10.1007/s00705-015-2711-9

King, A. M. Q., Adams, M. J., Carstens, E. B., and Lefkowitz, E. J. (2011). *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*. San Diego: Elsevier Academic Press.

Kobayashi, K., Atsumi, G., Iwadate, Y., Tomita, R., Chiba, K., Akasaka, S., et al. (2013). Gentian Kobu-sho-associated virus: a tentative, novel double-stranded RNA virus that is relevant to gentian Kobu-sho syndrome. *J. Gen. Plant Pathol.* 79, 56–63. doi: 10.1007/s10327-012-0423-5

Kondo, H., Chiba, S., Maruyama, K., Andika, I. B., and Suzuki, N. (2019). A novel insect-infecting virga/nege-like virus group and its pervasive endogenization into insect genomes. *Virus Res.* 262, 37–47. doi: 10.1016/j.virusres.2017.11.020

Kondo, H., Chiba, S., and Suzuki, N. (2015). "Detection and analysis of non-retroviral RNA virus-like elements in plant, fungal, and insect genomes," in *Plant Virology Protocols -New Approaches to Detect Viruses and Host Responses*, eds I. Uyeda and C. Masuta (New York, NY: Springer), 73–88. doi: 10.1007/978-1-4939-1743-3_7

Kondo, H., Hirota, K., Maruyama, K., Andika, I. B., and Suzuki, N. (2017). A possible occurrence of genome reassortment among bipartite rhabdoviruses. *Virology* 508, 18–25. doi: 10.1016/j.virol.2017.04.027

Kondo, H., Hisano, S., Chiba, S., Maruyama, K., Andika, I. B., Toyoda, K., et al. (2016). Sequence and phylogenetic analyses of novel totivirus-like double-stranded RNAs from field-collected powdery mildew fungi. *Virus Res.* 213, 353–364. doi: 10.1016/j.virusres.2015.11.015

Koramutla, M. K., Aminedi, R., and Bhattacharya, R. (2016). Comprehensive evaluation of candidate reference genes for qRT-PCR studies of gene expression in mustard aphid, *Lipaphis erysimi* (Kalt). *Sci. Rep.* 6:25883. doi: 10.1038/srep25883

Kotta-Loizou, I., Castón, J. R., Coutts, R. H. A., Hillman, B. I., Jiang, D., Kim, D. H., et al. (2020). ICTV virus taxonomy profile: *Chrysoviridae*. *J. Gen. Virol.* 101, 143–144. doi: 10.1099/jgv.0.001383

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315

Krueger, E. N., Beckett, R. J., Gray, S. M., and Miller, W. A. (2013). The complete nucleotide sequence of the genome of barley yellow dwarf virus-RMV reveals it to be a new *Polerovirus* distantly related to other yellow dwarf viruses. *Front. Microbiol.* 4:205. doi: 10.3389/fmicb.2013.00205

Kuchibhatla, D. B., Sherman, W. A., Chung, B. Y. W., Cook, S., Schneider, G., Eisenhaber, B., et al. (2014). Powerful sequence similarity search methods and in-depth manual analyses can identify remote homologs in many apparently "orphan" viral proteins. *J. Virol.* 88, 10–20. doi: 10.1128/JVI.02595-13

Lefort, V., Longueville, J. E., and Gascuel, O. (2017). SMS: smart model selection in PhyML. *Mol. Biol. Evol.* 34, 2422–2424. doi: 10.1093/molbev/msx149

Lin, Y. H., Fujita, M., Chiba, S., Hyodo, K., Andika, I. B., Suzuki, N., et al. (2019). Two novel fungal negative-strand RNA viruses related to mymonaviruses and phenuiviruses in the shiitake mushroom (*Lentinula edodes*). *Virology* 533, 125–136. doi: 10.1016/j.virol.2019.05.008

Locali-Fabris, E. C., Freitas-Astua, J., Souza, A. A., Takita, M. A., Astua-Monge, G., Antonioli-Luizon, R., et al. (2006). Complete nucleotide sequence, genomic organization and phylogenetic analysis of *Citrus leprosis* virus cytoplasmic type. *J. Gen. Virol.* 87, 2721–2729. doi: 10.1099/vir.0.82038-0

Lu, H., Zhu, J., Yu, J., Chen, X., Kang, L., and Cui, F. (2020). A symbiotic virus facilitates aphid adaptation to host plants by suppressing jasmonic acid responses. *Mol. Plant Microbe Interact.* 31, 55–65. doi: 10.1094/MPMI-01-19-0016-R

Lüthi, M. N., Vorburger, C., and Dennis, A. B. (2019). A novel RNA virus in the parasitoid wasp *Lysiphlebus fabarum*: genomic structure, prevalence, and transmission. *Viruses* 12:59. doi: 10.3390/v12010059

Maina, S., Edwards, O. R., de Almeida, L., Ximenes, A., and Jones, R. A. C. (2017). Metagenomic analysis of cucumber RNA from east timor reveals an Aphid lethal paralysis virus genome. *Genome Announc.* 5:e01445-16. doi: 10.1128/genomeA.01445-16

Medd, N. C., Fellous, S., Waldron, F. M., Xuereb, A., Nakai, M., Cross, J. V., et al. (2018). The virome of *Drosophila suzukii*, an invasive pest of soft fruit. *Virus Evol.* 4:vey009. doi: 10.1093/ve/vey009

Melzer, M. J., Simbajon, N., Carillo, J., Borth, W. B., Freitas-Astua, J., Kitajima, E. W., et al. (2013). A cilevirus infects ornamental hibiscus in Hawaii. *Arch. Virol.* 158, 2421–2424. doi: 10.1007/s00705-013-1745-0

Miller, W. A., Liu, S. J., and Beckett, R. (2002). Barley yellow dwarf virus: *Luteoviridae* or *Tombusviridae*? *Mol. Plant Pathol.* 3, 177–183. doi: 10.1046/j.1364-3703.2002.00112.x

Moon, J. S., Domier, L. L., McCoppin, N. K., D'Arcy, C. J., and Jin, H. (1998). Nucleotide sequence analysis shows that *Rhopalosiphum padi* virus is a member of a novel group of insect-infecting RNA viruses. *Virology* 243, 54–65. doi: 10.1006/viro.1998.9043

Muhire, B. M., Varsani, A., and Martin, D. P. (2014). SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PloS One* 9:e108277. doi: 10.1371/journal.pone.0108277

Nabeshima, T., Inoue, S., Okamoto, K., Posadas-Herrera, G., Yu, F. X., Uchida, L., et al. (2014). Tanay virus, a new species of virus isolated from mosquitoes in the Philippines. *J. Gen. Virol.* 95, 1390–1395. doi: 10.1099/vir.0.061887-0

Navarro, B., Minutolo, M., De Stradis, A., Palmisano, F., Alioto, D., and Di Serio, F. (2018). The first phlebo-like virus infecting plants: a case study on the adaptation of negative-stranded RNA viruses to new hosts. *Mol. Plant Pathol.* 19, 1075–1089. doi: 10.1111/mpp.12587

Ng, J. C. K., and Perry, K. L. (2004). Transmission of plant viruses by aphid vectors. *Mol. Plant Pathol.* 5, 505–511. doi: 10.1111/j.1364-3703.2004.00240.x

Nunes, M. R. T., Contreras-Gutierrez, M. A., Guzman, H., Martins, L. C., Barbirato, M. F., Savit, C., et al. (2017). Genetic characterization, molecular epidemiology, and phylogenetic relationships of insect-specific viruses in the taxon Negevirus. *Virology* 504, 152–167. doi: 10.1016/j.virol.2017.01.022

O'Brien, C. A., McLean, B. J., Colmant, A. M. G., Harrison, J. J., Hall-Mendelin, S., van den Hurk, A. F., et al. (2017). Discovery and characterisation of Castlerea virus, a new species of Negevirus isolated in Australia. *Evol. Bioinform.* 13:1176934317691269. doi: 10.1177/1176934317691269

Parry, H. R., Macfadyen, S., and Kriticos, D. J. (2012). The geographical distribution of yellow dwarf viruses and their aphid vectors in Australian grasslands and wheat. *Australas. Plant Pathol.* 41, 375–387. doi: 10.1007/s13313-012-0133-7

Pettersson, J. H. O., Shi, M., Eden, J. S., Holmes, E. C., and Hesson, J. C. (2019). Meta-transcriptomic comparison of the RNA viromes of the mosquito vectors *Culex pipiens* and *Culex torrentium* in Northern Europe. *Viruses* 11:E1033. doi: 10.3390/v11111033

Powell, G., Tosh, C. R., and Hardie, J. (2006). Host plant selection byaphids: behavioral, evolutionary, and applied perspectives. *Ann. Rev. Entomol.* 51, 309–330. doi: 10.1146/annurev.ento.51.110104.151107

Ramos-González, P. L., Santos, G. F. D., Chabi-Jesus, C., Harakava, R., Kitajima, E. W., and Freitas-Astúa, J. (2020). Passion fruit green spot virus genome harbors a new orphan ORF and highlights the flexibility of the 5'-end of the RNA2 segment across cileviruses. *Front. Microbiol.* 11:206. doi: 10.3389/fmicb.2020.00206

Reineke, A., and Asgari, S. (2005). Presence of a novel small RNA-containing virus in a laboratory culture of the endoparasitic wasp *Venturia canescens* (Hymenoptera : Ichneumonidae). *J. Insect Physiol.* 51, 127–135. doi: 10.1016/j.jinsphys.2004.05.005

Remnant, E. J., Shi, M., Buchmann, G., Blacquiere, T., Holmes, E. C., Beekman, M., et al. (2017). A diverse range of novel RNA viruses in geographically distinct honey bee populations. *J. Virol.* 91:e00158-17. doi: 10.1128/JVI.00158-17

Roossinck, M. J. (2011). The good viruses: viral mutualistic symbioses. *Nat. Rev. Microbiol.* 9, 99–108. doi: 10.1038/nrmicro2491

Roossinck, M. J. (2015). Plants, viruses and the environment: ecology and mutualism. *Virology* 479, 271–277. doi: 10.1016/j.virol.2015.03.041

Roumi, V., Caglayan, K., Gazel, M., and Li, S. (2020). Complete genome sequence of *Aphid lethal paralysis virus* from metagenomic analysis of *Cestrum elegans* small RNAs. *Gene Rep.* 18:100566. doi: 10.1016/j.genrep.2019.100566

Ryabov, E. V. (2007). A novel virus isolated from the aphid *Brevicoryne brassicae* with similarity to Hymenoptera picorna-like viruses. *J. Gen. Virol.* 88, 2590–2595. doi: 10.1099/vir.0.83050-0

Ryabov, E. V., Keane, G., Naish, N., Evered, C., and Winstanley, D. (2009). Densovirus induces winged morphs in asexual clones of the rosy apple aphid, *Dysaphis plantaginea*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8465–8470. doi: 10.1073/pnas.0901389106

Sadeghi, M., Altan, E., Deng, X. T., Barker, C. M., Fang, Y., Coffey, L. L., et al. (2018). Virome of > 12 thousand *Culex* mosquitoes from throughout California. *Virology* 523, 74–88. doi: 10.1016/j.virol.2018.07.029

Saitou, N., and Nei, M. (1987). The neighbor-joining method – a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425. doi: 10.1093/oxfordjournals.molbev.a040454

Schoonvaere, K., Smagghe, G., Francis, F., and de Graaf, D. C. (2018). Study of the metatranscriptome of eight social and solitary wild bee species reveals novel viruses and bee parasites. *Front. Microbiol.* 9:177. doi: 10.3389/fmicb.2018.00177

Shi, M., Lin, X. D., Tian, J. H., Chen, L. J., Chen, X., Li, C. X., et al. (2016a). Redefining the invertebrate RNA virosphere. *Nature* 7634, 539–543. doi: 10.1038/nature20167

Shi, M., Lin, X. D., Vasilakis, N., Tian, J. H., Li, C. X., Chen, L. J., et al. (2016b). Divergent viruses discovered in arthropods and vertebrates revise the evolutionary history of the *Flaviviridae* and related viruses. *J. Virol.* 90, 659–669. doi: 10.1128/JVI.02036-15

Shi, M., Neville, P., Nicholson, J., Eden, J. S., Imrie, A., and Holmes, E. C. (2017). High-resolution metatranscriptomics reveals the ecological dynamics of mosquito-associated RNA viruses in Western Australia. *J. Virol.* 91:e00680-17. doi: 10.1128/JVI.00680-17

Solovyev, A. G., and Morozov, S. Y. (2017). Non-replicative integral membrane proteins encoded by plant alpha-like viruses: emergence of diverse orphan ORFs and movement protein genes. *Front. Plant Sci.* 8:1820. doi: 10.3389/fpls.2017.01820

Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56, 564–577. doi: 10.1080/10635150701472164

Tassi, A. D., Garita-Salazar, L. C., Amorim, L., Novelli, V. M., Freitas-Astua, J., Childers, C. C., et al. (2017). Virus-vector relationship in the *Citrus leprosis* pathosystem. *Exp. Appl. Acarol.* 71, 227–241. doi: 10.1007/s10493-017-0123-0

Teixeira, M., Sela, N., Ng, J., Casteel, C. L., Peng, H. C., Bekal, S., et al. (2016). A novel virus from *Macrosiphum euphorbiae* with similarities to members of the family *Flaviviridae*. *J. Gen. Virol.* 97, 1261–1271. doi: 10.1099/jgv.0.000414

Teixeira, M. A., Sela, N., Atamian, H. S., Bao, E., Chaudhary, R., MacWilliams, J., et al. (2018). Sequence analysis of the potato aphid *Macrosiphum euphorbiae*

transcriptome identified two new viruses. *PloS One* 13:e0193239. doi: 10.1371/journal.pone.0193239

Temmam, S., Bigot, T., Chrétien, D., Gondard, M., Pérot, P., Pommelet, V., et al. (2019). Insights into the host range, genetic diversity, and geographical distribution of jingmenviruses. *mSphere* 4:e00645-19. doi: 10.1128/mSphere.00645-19

Tokarz, R., Sameroff, S., Tagliafierro, T., Jain, K., Williams, S. H., Cucura, D. M., et al. (2018). Identification of novel viruses in *Amblyomma americanum*, *Dermacentor variabilis*, and *Ixodes scapularis* ticks. *mSphere* 3:e00614-17. doi: 10.1128/mSphere.00614-17

van Munster, M., Dullemans, A. M., Verbeek, M., van den Heuvel, J., Clerivet, A., and van der Wilk, F. (2002). Sequence analysis and genomic organization of Aphid lethal paralysis virus: a new member of the family *Dicistroviridae*. *J. Gen. Virol.* 83, 3131–3138. doi: 10.1099/0022-1317-83-12-3131

van Munster, M., Dullemans, A. M., Verbeek, M., van den Heuvel, J., Reinbold, C., Brault, V., et al. (2003). A new virus infecting *Myzus persicae* has a genome organization similar to the species of the genus *Densovirus*. *J. Gen. Virol.* 84, 165–172. doi: 10.1099/vir.0.18650-0

van Munster, M., Janssen, A., Clérivet, A., and Van den Heuvel, J. (2005). Can plants use an entomopathogenic virus as a defense against herbivores? *Oecologia* 143, 396–401. doi: 10.1007/s00442-004-1818-6

vandenHeuvel, J., Hummelen, H., Verbeek, M., Dullemans, A. M., and vanderWilk, F. (1997). Characteristics of *Acyrthosiphon pisum* virus, a newly identified virus infecting the pea aphid. *J. Invertebr. Pathol.* 70, 169–176. doi: 10.1006/jipa.1997.4691

Vasilakis, N., Forrester, N. L., Palacios, G., Nasar, F., Savji, N., Rossi, S. L., et al. (2013). Negevirus: a proposed new taxon of insect-specific viruses with wide geographic distribution. *J. Virol.* 87, 2475–2488. doi: 10.1128/JVI.00776-12

Vasilakis, N., and Tesh, R. B. (2015). Insect-specific viruses and their potential impact on arbovirus transmission. *Curr. Opin. Virol.* 15, 69–74. doi: 10.1016/j.coviro.2015.08.007

Viljakainen, L., Holmberg, I., Abril, S., and Jurvansuu, J. (2018). Viruses of invasive Argentine ants from the European main supercolony: characterization, interactions and evolution. *J. Gen. Virol.* 99, 1129–1140. doi: 10.1099/jgv.0.001104

Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Dempsey, D. M., Dutilh, B. E., et al. (2019). Changes to virus taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2019). *Arch. Virol.* 164, 2417–2429. doi: 10.1007/s00705-019-04306-w

Wamonje, F. O., Michuki, G. N., Braidwood, L. A., Njuguna, J. N., Mutuku, J. M., Djikeng, A., et al. (2017). Viral metagenomics of aphids present in bean and maize plots on mixed-use farms in Kenya reveals the presence of three dicistroviruses including a novel big Sioux River virus-like dicistrovirus. *Virol. J.* 14, 188. doi: 10.1186/s12985-017-0854-x

Webster, C. L., Longdon, B., Lewis, S. H., and Obbard, D. J. (2016). Twenty-five new viruses associated with the *Drosophilidae* (Diptera). *Evol. Bioinform.* 12, 13–25. doi: 10.4137/EBO.S39454

Webster, C. L., Waldron, F. M., Robertson, S., Crowson, D., Ferrari, G., Quintana, J. F., et al. (2015). The discovery, distribution, and evolution of viruses associated with *Drosophila melanogaster*. *PloS Biol.* 13:e1002210. doi: 10.1371/journal.pbio.1002210

Whitfield, A. E., Falk, B. W., and Rotenberg, D. (2015). Insect vector-mediated transmission of plant viruses. *Virology* 479, 278–289. doi: 10.1016/j.virol.2015.03.026

Whitfield, A. E., Huot, O. B., Martin, K. M., Kondo, H., and Dietzgen, R. G. (2018). Plant rhabdoviruses-their origins and vector interactions. *Curr. Opin. Virol.* 33, 198–207. doi: 10.1016/j.coviro.2018.11.002

Williamson, C., Rybicki, E. P., Kasdorf, G. G. F., and Vonwechmar, M. B. (1988). Characterization of a new picorna-like virus isolated from aphids. *J. Gen. Virol.* 69, 787–795. doi: 10.1099/0022-1317-69-4-787

Yang, D. H., Zhao, H. X., Shi, J. M., Xu, X., Wu, Y. Y., Guo, R., et al. (2019). Discovery of aphid lethal paralysis virus in *Vespa velutina* and *Apis cerana* in China. *Insects* 10, 157. doi: 10.3390/insects10060157

Zhang, W., Wu, T. F., Guo, M. M., Chang, T. Y., Yang, L., Tan, Y., et al. (2019). Characterization of a new bunyavirus and its derived small RNAs in the brown citrus aphid, *Aphis citricidus*. *Virus Genes* 55, 557–561. doi: 10.1007/s11262-019-01667-x

Zhao, L., Mwaliko, C., Atoni, E., Wang, Y. J., Zhang, Y. Z., Zhan, J. B., et al. (2019). Characterization of a novel Tanay virus isolated from *Anopheles sinensis* mosquitoes in Yunnan, China. *Front. Microbiol.* 10:1963. doi: 10.3389/fmicb.2019.01963

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 09 June 2020
doi: 10.3389/fmicb.2020.01231

Citrus Leprosis Virus C Encodes Three Proteins With Gene Silencing Suppression Activity

Mikhail Oliveira Leastro[1,2†], Deibis Yorlenis Ortega Castro[2†], Juliana Freitas-Astúa[1,3], Elliot Watanabe Kitajima[4], Vicente Pallás[2] and Jesús Ángel Sánchez-Navarro[2*]*

[1] Unidade Laboratorial de Referência em Biologia Molecular Aplicada, Instituto Biológico, São Paulo, Brazil, [2] Instituto de Biología Molecular y Celular de Plantas, Universidad Politécnica de Valencia-Consejo Superior de Investigaciones Cientifcas (CSIC), Valencia, Spain, [3] Embrapa Mandioca e Fruticultura, Cruz das Almas, Brazil, [4] Departamento de Fitopatologia e Nematologia, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, Brazil

OPEN ACCESS

Edited by:
Jesús Navas-Castillo, Institute of Subtropical and Mediterranean Horticulture La Mayora, Spain

Reviewed by:
Aiming Wang, Agriculture and Agri-Food Canada (AAFC), Canada
Beatriz Navarro, Istituto per la Protezione Sostenibile delle Piante, Sede Secondaria di Bari, Italy

***Correspondence:**
Mikhail Oliveira Leastro
m.leastro@gmail.com
Jesús Ángel Sánchez-Navarro
jesanche@ibmcp.upv.es

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: 14 March 2020
Accepted: 14 May 2020
Published: 09 June 2020

Citation:
Leastro MO, Castro DYO, Freitas-Astúa J, Kitajima EW, Pallás V and Sánchez-Navarro JÁ (2020) Citrus Leprosis Virus C Encodes Three Proteins With Gene Silencing Suppression Activity. Front. Microbiol. 11:1231. doi: 10.3389/fmicb.2020.01231

Citrus leprosis virus C (CiLV-C) belongs to the genus *Cilevirus*, family *Kitaviridae,* and is considered the most devastating virus infecting citrus in Brazil, being the main viral pathogen responsible for citrus leprosis (CL), a severe disease that affects citrus orchards in Latin America. Here, proteins encoded by CiLV-C genomic RNA 1 and 2 were screened for potential RNA silencing suppressor (RSS) activity by five methods. Using the GFP-based reporter agroinfiltration assay, we have not found potential local suppressor activity for the five CiLV-C encoded proteins. However, when RSS activity was evaluated using the alfalfa mosaic virus (AMV) system, we found that the p29, p15, and p61 CiLV-C proteins triggered necrosis response and increased the AMV RNA 3 accumulation, suggesting a suppressive functionality. From the analysis of small interfering RNAs (siRNAs) accumulation, we observed that the ectopic expression of the p29, p15, and p61 reduced significantly the accumulation of GFP derived siRNAs. The use of the RSS defective turnip crinkle virus (TCV) system revealed that only the *trans*-expression of the p15 protein restored the cell-to-cell viral movement. Finally, the potato virus X (PVX) system revealed that the expression of p29, p15, and p61 increased the PVX RNA accumulation; in addition, the p29 and p15 enhanced the pathogenicity of PVX resulting in the death of tobacco plants. Furthermore, PVX-p61 infection resulted in a hypersensitive response (HR), suggesting that p61 could also activate a plant defense response mechanism. This is the first report describing the RSS activity for CiLV-C proteins and, moreover, for a member of the family *Kitaviridae*.

Keywords: RNA silencing suppressor, citrus leprosis virus C, RSS activity, hypersensitive response, family *Kitaviridae*

INTRODUCTION

Organisms have a primary cellular defense mechanism known as RNA silencing. RNA silencing has a fundamental "sequence specific gene regulatory feature" (Kakumani et al., 2013) and plays an important role in defense against invading microorganisms (pathogens), especially viruses (Pumplin and Voinnet, 2013). This defense mechanism is activated by double-stranded RNA from high genome amplification of invasive microorganisms, transposons or ectopic expressed genes

(Aravin et al., 2001; Pumplin and Voinnet, 2013; Guo et al., 2016). Briefly, double-stranded RNAs are processed by dicer-like RNases (DCLs) in small RNAs (siRNAs) of 20–24 nt in size (Hamilton and Baulcombe, 1999; Matranga and Zamore, 2007; Borges and Martienssen, 2015) then the small RNAs are loaded onto Argonaute (AGO) proteins to guide the silencing of DNA or RNA elements by a specific recognition of sequence complementarity (Hammond et al., 2000; Ding and Voinnet, 2007; Pumplin and Voinnet, 2013; Nakanishi, 2016; Pisacane and Halic, 2017). On the other side, in counter-defense against the RNA silencing antiviral defense, viruses evolved to encode suppressors of RNA silencing proteins (RSS) (Voinnet et al., 1999; Roth et al., 2004; Li and Ding, 2006; Burgyan and Havelda, 2011; Csorba et al., 2015; Moon and Park, 2016) which may differ by their ability to suppress intracellular and/or intercellular silencing machinery (Li and Ding, 2001; Lu et al., 2004, 2005; Merai et al., 2006; Garcia and Pallas, 2015; Samuel et al., 2016).

As reviewed by Martinez-Perez et al. (2019), several viral RSSs have been identified using different procedures (Moissiard and Voinnet, 2004; Roth et al., 2004; Qu and Morris, 2005; Li and Ding, 2006; Gupta et al., 2018; Yang et al., 2018; Martinez-Perez et al., 2019). The most common assay is the "patch" technique (Voinnet et al., 1998) a system that uses *Agrobacterium tumefaciens* cultures harboring the putative RSS and a reporter gene infiltrated on *Nicotiana benthamiana* leaves, in which the natural silencing process of the overexpressed reporter gene is delayed by the presence of an RSS. The RSS screening based on viral vector, such as potato X virus (PVX), in which the expression of an RSS is associated with a more aggressive viral infection (Voinnet et al., 1999), is also widely used. Functional complementation of defective viral mutants (Chiba et al., 2006; Powers et al., 2008) or viral vectors in which the RSS is correlated with symptoms appearance (Guilley et al., 2009) have been used to identify RSS in the last decade.

Recently, a new approach to screen for RSS activity based on a viral system derived from alfalfa mosaic virus (AMV) has been presented, which revealed a correlation between the presence of necrotic lesions on inoculated leaves and RSS activity (Martinez-Perez et al., 2019).

Citrus leprosis virus C is the type member of the genus *Cilevirus,* family *Kitaviridae* (Locali-Fabris et al., 2012; Freitas-Astua et al., 2018; Quito-Avila et al., 2020). It is the main viral pathogen responsible for the citrus leprosis, a re-emergent disease that considerably affects citrus production in Latin America. "Its genome is composed of two linear positive sense ssRNA segments with the presence of 5'cap structure and a 3' poly(A) tail, organized in six open reading frames (ORFs). The first segment (RNA1) has two ORFs that code for a replication-associated protein containing conserved domains of methyl transferase, helicase, and RNA dependent RNA polymerase, and the capsid protein (p29)" (Locali-Fabris et al., 2006; Pascon et al., 2006; Leastro et al., 2018). The second segment (RNA2) encodes four proteins: (i) p15, a small protein for which a specific function needs to be determined, but potentially involved with virus replication as suggested by its involvement in the formation of vesicles thought ER remodeling (Leastro et al., 2018); (ii) p61, which exhibits features of glycoprotein (Kuchibhatla et al., 2014; Leastro et al., 2018) with biological activity on remodeling the ER system and redistributing the Golgi apparatus (Leastro et al., 2018); (iii) p32 which has the function of a movement protein (MP) (Leastro et al., unpublished); and (iv) p24, an integral membrane protein with the ability to form vesicle-like spherical structures in association with the ER, which strongly suggests the involvement of this protein in viral replication and assembly (Leastro et al., 2018). Those features, in addition to membrane topology presented for the p24 protein (Leastro et al., 2018) combined with its homology to virion membrane proteins of plant and arthropod viruses (Kuchibhatla et al., 2014; Leastro et al., 2018) suggest a potential structural role as a matrix protein. The natural infection by citrus leprosis-associated viruses results only in localized lesions in field conditions; for a yet unknown process, these viruses seem to be unable to infect phloem tissues and, therefore, are unable to become systemic in their hosts.

Molecular aspects related to the mechanism of cileviruses infection have been elucidated in recent years; however, some important features remain to be explored. Currently, there is no information about the RSS activity of the proteins encoded by the cileviruses. In this work, we provide a step further in the molecular understanding of how these viruses cause infection. Here, we screened the CiLV-C proteins to identify potential RSS activity by using five systems well-described in the literature. We show that the CiLV-C p29, p15, and p61 proteins have suppressor activity, thus elucidating the functionality of p15, and characterizing additional functions for the capsid protein (p29) and the putative glycoprotein (p61). Furthermore, the p61 expression on the PVX infection context generates a hypersensitive response (HR), suggesting that this protein could also activate a plant defense response mechanism.

MATERIALS AND METHODS

DNA Manipulation

An infectious cDNA 3 construct of AMV that expresses green fluorescent protein-GFP (pGFP/MP/CP) (Sanchez-Navarro et al., 2001) was used to express the potential RSS proteins by exchanging the GFP gene. The introduction of the Human influenza hemagglutinin epitope (HA) at the N- or C-terminus of the GFP was performed as previously described (Martinez-Perez et al., 2019). The resultant pGFP:HA/MP/CP construct expressed the GFP with the HA epitope at the C-terminus and allowed the exchange of the GFP gene by using the *Nco*I and *Nhe*I sites. The CiLV-C genes *p29*, *p15*, *p61*, *MP*, and *p24* (GenBank accessions YP_654539.1, YP_654540.1, YP_655441.1, YP_654542.1, and YP_6545543.1, respectively) were amplified by PCR with specific primers containing the *BspHI*/*Nhe*I (*p29*), *Pci*I/*Nhe*I (*p15* and *p24*), or *Nco*I/*Nhe*I (*p61* and *MP*) restriction sites. The corresponding fragments were inserted into an AMV RNA 3 clone. The insertion of the p61 with the HA at its N-terminus (pHA:p61/MP/CP) was performed as previously described (Martinez-Perez et al., 2019). The specific primers for p61 frame shift amplification generated an amplicon carrying the 5'*Nco*I and 3'*Nhe*I restrict sites plus a stop (TAA) sequence after the fifth codon. This PCR product, previously digested, was

cloned into the pGFP:HA/MP/CP construct, as aforementioned, to generate the construction pp61(fs)stop:HA/MP/CP. The AMV construction harboring the RSS gene of tobacco etch virus (TEV) *HCPro* (GenBank accession DQ986288) was obtained from Martinez-Perez et al. (2019).

For transient expression of the proteins, the amplified genes above described carrying stop codons were introduced in the expression cassette of the plasmid pSK35S-MP_{TSWV}:HA-PoPit (Leastro et al., 2015) by exchanging of the tomato spotted wilt virus (TSWV) MP. The cassettes resulting were under the control of 35S constitutive promoter from cauliflower mosaic virus (CaMV) and the terminator from the potato proteinase inhibitor (PoPit) (Leastro et al., 2015). Then, the corresponding expression cassettes (35S-ORF_{stop}:HA-PoPit) were subcloned into the $pMOG_{800}$ binary vector by using the restriction sites *Hind*III (for *p15* and *p61*) and *Xho*I (for *p29*, *MP,* and *p24*). The leader peptidase (Lep) construct used as negative control was obtained from Peiro et al. (2014).

For the PVX assay, the heterologous viral genes were introduced in a PVX expression vector (Lu et al., 2003). For that, the CiLV-C corresponding genes were amplified by PCR with primers carrying *Sal*I restriction site. The amplicons were digested and cloned into the plasmid pGR107, previously digested with *Sal*I and dephosphorylated. *HCPro* and coat protein (CP) of carnation mottle virus (CarMV) (GenBank acc.: AJ304989) were obtained from Martinez-Perez et al. (2019).

N. benthamiana Wild Type and 16c RNA Silencing Suppression Assay

To study the effect of different viral factors in intracellular and intercellular RNA silencing suppression, we performed the agroinfiltration technique using wild type or GFP transgenic *N. benthamiana* plants (line 16c) (Hamilton et al., 2002; Burgyan and Havelda, 2011). $pMOG_{800}$ binary constructions harboring the *HCPro* and *p29*, *p15*, *p61*, *MP*, and *p24* CiLV-C genes were introduced into C58C1 cells and kept overnight at 28°C in Luria-Bertani (LB) broth with rifampicin and kanamycin. pMOG(GFP) construct carrying the eGFP gene was used to trigger the silencing of the GFP transgene of 16c tobacco plants or to generate small RNAs (siRNAs) in both wt and 16c line of *N. benthamiana*. In co-infiltration experiments, we performed a mixture in equal volume (OD_{600} = 0.5) of *A. tumefaciens* culture containing pMOG-GFP binary plasmid and individual cultures harboring each of the above mentioned viral factors. A mixture consisting of equal volume of *A. tumefaciens* cultures carrying pMOG(GFP) or empty $pMOG_{800}$ were used as controls. Three independent experiments were performed, each one included the infiltration of 10 plants per construct. The plants were grown under two-step cycle of 10 h of darkness at 18°C and 14 h of light at 20°C or under conditions of 23°C day 18°C night and 16 h light/8 h dark regime maintained in FITOTRON® plant chamber. Agroinfiltrated leaves were photographed at 6 days post-inoculation (dpi) under long-wavelength UV light (UVGL-58 Handheld UV lamp; UV Products) by using a tripod and a Nikon D3000 digital camera at F11 aperture value and 1/10 s shutter speed (Yaegashi et al., 2012; Martinez-Perez et al., 2019).

Alfalfa Mosaic Virus Necrotic Response Assay

Plasmids of the pGFP:HA/MP/CP chimeric AMV RNA 3 constructs harboring the *HCPro*, *p29*, *p15*, *p61*, *MP*, and *p24* viral factors were linearized with *Pst*I and transcribed with T7 RNA polymerase (Takara Bio USA, Inc.) following the manufacturer's instructions. Transgenic *N. tabacum* plants that express the polymerase proteins P1 and P2 of AMV (P12 plants) (Van Dun et al., 1988) were grown and inoculated with RNA transcripts, as described previously (Taschner et al., 1991). P12 protoplasts were extracted and 2.5×10^5 protoplasts were inoculated by the polyethylene glycol method (Loesch-Fries et al., 1985) with 15 μL of the transcription mixture.

On the surface of the P12 mechanically inoculated leaves, the presence of necrotic lesions was monitored for 2 weeks post-inoculation with onset of lesions observed at 4 dpi. Three independent experiments were performed, each one included the infiltration of three or four leaves from three plants per construct.

Turnip Crinkle Virus Complementation and Co-infiltration Assays

The turnip crinkle virus (TCV) assay was performed in two different approaches: (i) as an infectious RNA transcript referred to as TCV-sGFP complementation assay or (ii) via agroinfiltration referred to as PZP-TCV-sGFP co-infiltration assay (Powers et al., 2008). In the first method, the movement-deficiency phenotype of a TCV CP deletion mutant that expresses GFP (TCV-sGFP) is complemented in *trans*. In the second approach, the presence of an RSS is identified by an increase of the GFP signal in the whole leaf previously agroinfiltrated with PZP-TCV-sGFP constructs (Martinez-Perez et al., 2019). For the TCV-sGFP complementation assay, three *N. benthamiana* leaves per plant and three plants per construct were agroinfiltrated with *A. tumefaciens* cultures (strain C58) carrying the pMOG-empty or expressing the *Lep*, as negative controls, or the $pMOG_{800}$ constructs expressing *HCPro*, *p29*, *p15*, *p61*, *MP*, and *p24* at OD_{600} = 1 (Powers et al., 2008). Next, the pTCV-sGFP plasmid was linearized with *Xba*I and, 1 day post-infiltration, TCV-sGFP infectious RNA transcripts were mechanically inoculated onto the abaxial surfaces of the infiltrated leaves as referred by Martinez-Perez et al. (2019). Local movement was evaluated at 3 dpi, with the aid of a Leica MZ16F fluorescence stereomicroscope. For the co-infiltration assay, the *A. tumefaciens* cultures aforementioned were mixed with an *Agrobacterium* culture carrying the PZP-TCV-sGFP construct at an OD_{600} = 0.0025 (Powers et al., 2008) and then infiltrated in *N. benthamiana* leaves. The GFP signal of the infiltrated leaves was monitored at 5 dpi with a Leica MZ16F fluorescence stereomicroscope. Each assay was repeated three times by inoculation of three *N. benthamiana* leaves per plant and three plants per construct.

Potato Virus X Pathogenicity Assay

PVX-derivatives either containing the *HCPro*, CarMV *CP*, *p29*, *p15*, *p61*, *MP*, or *p24* were constructed by cloning each ORF into the *Sma*I (for *HCPro* and CarMV *CP*) or *Sal*I (for

CiLV-C ORFs) sites of pGR107 downstream of the duplicated PVX CP promoter (Jones et al., 1999). *N. benthamiana* plants were agroinfiltrated (OD_{600} = 0.5) (Martinez-Perez et al., 2019) with *A. tumefaciens* strain C58C1 harboring the helper plasmid pSoup and electroporated with each of the recombinant construct. The empty pGR107 was used as negative control. Total RNA extraction was performed at 4 days post-infiltration from upper non-infiltrated leaves using VWR Life Science AMRESCO RiboZol™ RNA Extraction Reagent following the manufacturer's instructions. Three independent experiments were performed, each one included the infiltration of three plants per construct. Three weeks later, entire plants were photographed. The symptom development was monitored for up to 30 days post-infiltration.

Northern Blot Assays

Total RNA was extracted from P12 protoplasts at 16 h post-inoculation and from *N. benthamiana* leaves expressing transiently tested proteins in combination with the TCV (at 5 dpi), PVX (at 4 dpi) and 16c agroinfiltration systems (at 4 dpi), using VWR Life Science AMRESCO RiboZol™ RNA Extraction Reagent. After electrophoresis through formaldehyde-denatured gel, the RNAs were transferred to positively charged nylon membranes (Roche Mannheim, Germany) (Leastro et al., 2017) and fixed with a UV cross-linker (700 × 100 μJ/cm^2). Hybridization and detection was conducted as previously described (Pallas et al., 1998) using a dig-riboprobe (Roche, Mannheim, Germany) complementary to the 3'UTRs of the AMV RNA 3 and TCV. For detection of the genomic (g) and sub genomic (sg) PVX RNAs, northern blot assays were performed by overnight hybridization with a 500 nt length digoxigenin (DIG)-labeled-riboprobe complementary to the 3' end region of the PVX *CP* gene.

For the analysis of the small RNAs, total RNA was extracted from 0.1 g of *N. benthamiana* infiltrated leaves at 4 days post-infiltration, using TRI Reagent (Sigma-Aldrich, Steinheim, Germany). Nine micrograms of the total RNAs were electrophoresed through a 17% denaturing polyacrylamide gel and transferred to positively charged nylon membranes (Roche, Mannheim, Germany). RNAs were fixed, hybridized and detected as described above with the difference that the hybridization was performed at 38°C using a mix of three 50 nt dig-riboprobe complementary to nt: 707–756, 761–810, 881–930 of the GFP gene (GenBank accession U76561).

For analysis of mGFP expression, RNAs were extracted at 4 days post-infiltration and hybridized using dig-riboprobe complementary to the complete sequence of the GFP gene.

Statistical Analysis

Each assay reported here was performed in triplicate unless specifically noted otherwise. Standard deviation (±SD) represents data from three biological samples with at least three replicates for each sample. Student's *t*-test was performed to determine the significant differences between control specified and viral factor (sample) at each experiment. Significant difference is demonstrated by values lower than $p < 0.05$. *, **, or n.s represent $p < 0.05$, 0.01, or no significant difference, respectively. The graphs represent the relative accumulation of the RNAs corresponding to the average of three northern blot analyses from three independent experiments. The mean values, obtained from the band quantification, were normalized to the control mean values. The bands were quantified using ImageJ version 2.0cr software with ISAC plugin.

Western Blot Analysis

P12 leaves inoculated with RNA3 chimeric constructions at 3 days post-inoculation were processed with 250 μL of Laemmli loading buffer 1X (Laemmli, 1970). After boiling for 5 min, 25 μL of the mixture were subjected to 12% SDS-PAGE. Proteins were detected on Western blots using a mouse monoclonal anti-HA antibody (Sigma-Aldrich, Steinheim, Germany) and a secondary anti-mouse peroxidase labeled antibody (Sigma-Aldrich, Steinheim, Germany) together with a chemiluminescence substrate (Amersham™ ECL™ Prime Western Blotting Detection Reagent). The chemiluminescence detection was performed with a Fujifilm LAS-3000 detector and the membranes were exposed for 5 min.

RESULTS

The CiLV-C Encoded Proteins Do Not Suppress Local GFP Silencing Using the 16c RSS System

N. benthamiana 16c plants maintain a visual phenotype of overexpression of the green fluorescent protein (GFP) when the co-transiently expressed tested proteins have RSS activity (Voinnet et al., 1998). To test if any of the CiLV-C proteins have the capacity to suppress local RNA silencing, *N. benthamiana* 16c leaves were simultaneously co-infiltrated with individual *A. tumefaciens* cultures containing the binary vector pMOG-GFP, as gene silencing inducer, and $pMOG_{800}$ constructs containing the viral factors tested for potential RSS. pMOG-HCPro and pMOG-empty constructs were used as positive and negative controls, respectively. When co-infiltrated pMOG-GFP plus pMOG-HCPro, the epidermal leaf cells of 16c plants maintained GFP visual expression under UV illumination after 6 dpi (**Figure 1A**). In contrast, the leaves co-infiltrated with pMOG-CiLV-C ORFs plus pMOG-GFP, showed decreased GFP expression at 6 dpi, similar to that visualized for the negative control (pMOG-empty) (**Figure 1A**).

To confirm the results obtained after visual observation, we analyzed the accumulation of GFP mRNA in all combinations tested at 4 dpi. Northern blot analyses revealed a clear positive correlation between the visual GFP expression with abundant mRNA accumulation in leaves expressing the HCPro RSS. For the other tested proteins, the mGFP accumulation was significantly lower (**Figure 1B**). We also co-expressed, in all possible combinations, groups of two, three or four CiLV-C proteins plus the GFP inducer in 16c plants. However, no increment of the GFP fluorescence was observed in any of the analyzed combination (data not shown).

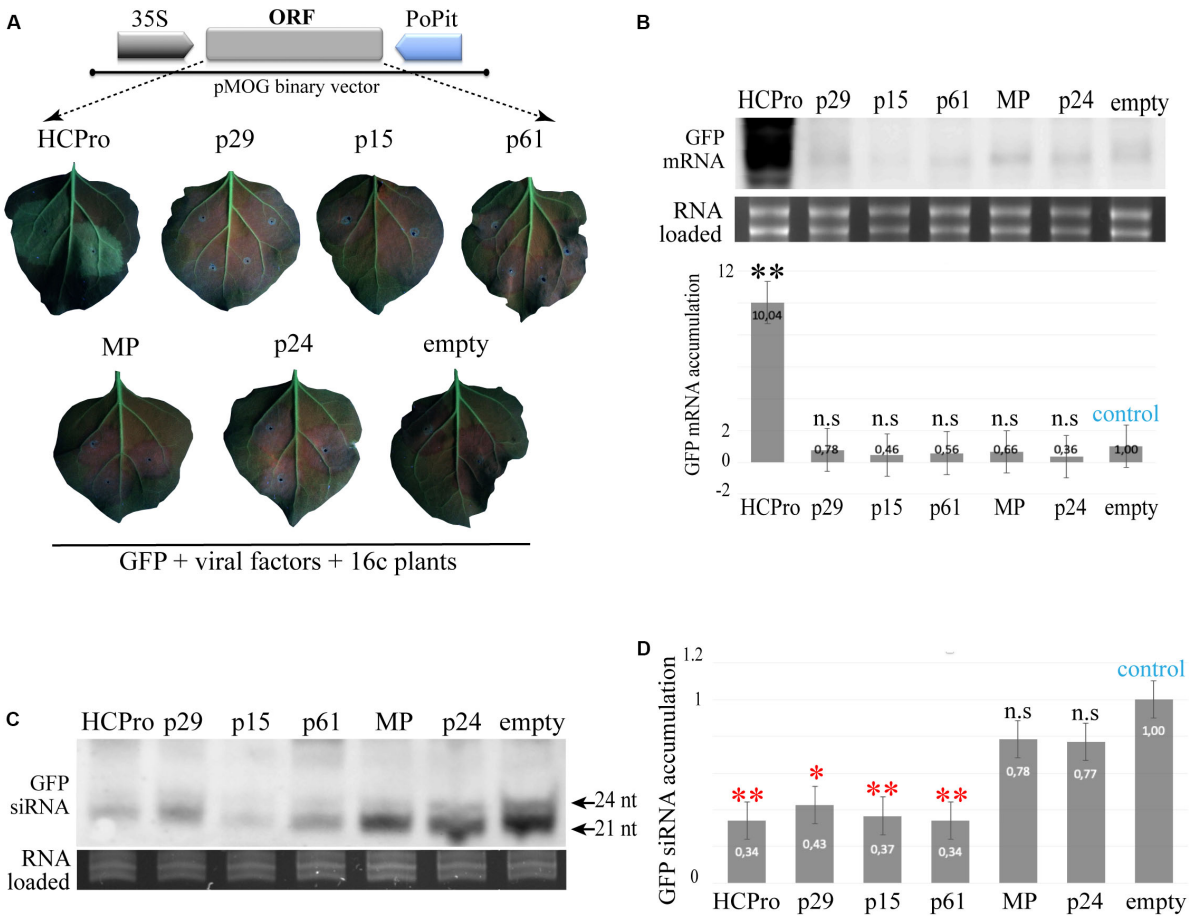


FIGURE 1 | Transient expression of CiLV-C encoded proteins in *Nicotiana benthamiana* 16c and northern blot analysis of the accumulation of GFP derived small RNAs (siRNAs). **(A)** Fluorescent images of *N. benthamiana* 16c leaves co-infiltrated with agrobacterium suspension carrying $pMOG_{800}$ binary vector containing individual CiLV-C ORFs and eGFP used to trigger the silencing of the GFP transgene. The schematic representation corresponds to the expression cassette used for transient expression of the CiLV-C proteins indicated, HCPro and GFP. The ORFs were cloned between CaMV 35S promoter and potato proteinase inhibitor terminator (PoPit). pMOG-HCPro and pMOG empty co-infiltrated with pMOG-GFP were used as positive and negative controls, respectively. The GFP fluorescence was not observed in leaves infiltrated with each of the CiLV-C proteins at 6 days post-infiltration. Three independent experiments were performed, each one included the infiltration of 10 plants per construct. **(B)** Northern blot analysis of GFP mRNA extracted at 4 dpi from the leaves infiltrated with the combination constructs above referred were performed using a dig-riboprobe complementary to the GFP gene. rRNA stained with ethidium bromide indicates equal loading of samples. The graph represents the relative accumulation of mGFP RNAs corresponding to the average of three northern blot from three independent experiments. The value of the averages obtained from the band quantification was normalized in relation to the respective negative control (pMOG-empty). The bands were quantified using the ImageJ version 2.0cr software with ISAC plugin and error bars represent standard deviation. Statistical analysis were done using Student's *t*-test. Black asterisks (*) indicate significantly increased viral gene accumulation compared to the control. **, $p < 0.01$; n.s, no significant difference. *P*-values were obtained from pairwise comparations between control vs. viral factor. **(C)** Northern blot of GFP-specific siRNA extracted from *N. benthamiana* wt and 16c patches at 4 dpi co-infiltrated with constructs expressing the GFP plus constructs expressing the HCPro (positive control) and p29, p15, p61, MP, and p24 CiLV-C proteins, using a mix of three 50 nt riboprobes complementary to the GFP gene (Martinez-Perez et al., 2019). Negative control (empty) corresponds to patches co-infiltrated with agrobacterium cultures carrying the GFP construct and the empty binary plasmid. Small RNAs of 21 and 24 nt are indicated. rRNA stained with ethidium bromide indicates equal loading of samples. **(D)** The graph represents the relative accumulation of GFP siRNA, referred to the construct carrying the GFP (empty) from three independent experiments. The value of the averages obtained from the band quantification was normalized in relation to the respective negative control (empty). Error bars represent standard deviation. Statistical analyses were done using Student's *t*-test. Red asterisks (*) indicate significantly decreased viral gene accumulation compared to the control. **, $p < 0.01$; *, $p < 0.05$; n.s, no significant difference.

Transient Expression of the p29, p15, and p61 Proteins in *N. benthamiana* (wt and 16c Line) Alters siRNA GFP Accumulation

Viral RNA silencing suppressors act in different key components of the RNA silencing, including the block on the production of small RNAs (siRNAs) (Anandalakshmi et al., 1998; Silhavy et al., 2002; Mann et al., 2016). We evaluated whether the CiLV-C proteins were able to interfere in the yield of the small RNA molecules originated from the transiently expressed GFP. To do this, transgenic 16c and wild type *N. benthamiana* leaves were infiltrated with a binary construct carrying the GFP together with constructs carrying a well-known RSS, the HCPro (positive

control) or the different CiLV-C ORFs. Northern blot analysis of the accumulation of the GFP-derived siRNAs at 4 dpi revealed that the p29, p15, and p61 reduced significantly the accumulation levels of the small RNAs in both 16 and wt *N. benthamiana* plants (**Figure 1D**). The northern blot image is representative of all replicates obtained by the infiltration of 16c and wt *N. benthamiana* leaves. Similar reduction was observed with the HCPro positive control. In contrast, a clear signal was detected in the leaves infiltrated with the MP, p24, and pMOG-empty binary vectors (**Figure 1C**). These results indicate that the p29, p15, and p61 CiLV-C proteins may act as silencing suppressors.

The Heterologous Expression of the p29, p15, and p61 CiLV-C Proteins Using AMV System Generate Necrotic Lesions on the Inoculated Leaves

Some methods are not always sensitive for identifying viral proteins with suppressive activity. To overcome that, we tested a new and sensitive method based on alfalfa mosaic virus RNA 3 expression vector and transgenic *N. tabacum* p12 plants. For this method, the RSS activity correlates with the appearance of necrotic lesions and increased accumulation of the AMV RNAs (Martinez-Perez et al., 2019). All CiLV-C genes were cloned into the AMV RNA3 carrying the HA epitope fused at their C-termini (**Figure 2A**). Chimeric AMV RNA 3 transcripts were inoculated on P12 leaves and the phenotypic lesions were monitored for 2 weeks. Chimeric AMV RNA 3 carrying the HCPro and GFP genes were used as positive and negative controls, respectively. P12 plants inoculated with AMV transcripts expressing the p29, p15, and p61 proteins showed necrotic lesions such as those observed for the positive control (**Figure 2A**). When P12 leaves were inoculated with transcripts expressing the MP and p24 proteins, no symptoms were observed. Absence of symptoms was also observed in the plants inoculated with the negative control (data not shown). No extra symptoms were visualized with longer infection time (30 dpi) in systemic leaves. In order to confirm the expression and stability of the corresponding proteins in absence of any necrotic lesions (Martinez-Perez et al., 2019), extraction of total proteins was performed at 3 dpi from P12 leaves inoculated with the corresponding chimeric AMV constructs. All proteins, except the p61, were detected using a monoclonal anti-HA antibody (**Figure 2B**). In a previous study, we also failed to detect the p61 protein by Western blot assay (Leastro et al., 2018). To confirm that the necrotic phenotype observed with the AMV construct carrying the p61 protein correlated with its expression, P12 leaves with an RNA 3 variant containing a frameshift version of the p61 gene were inoculated. P12 leaves inoculated with transcripts expressing the p61 with the HA epitope fused either at its N- or C-terminus showed necrotic lesions. However, no symptoms were observed when P12 leaves were inoculated with the p61 frameshift version (**Supplementary Figure S1**), indicating the correlation between the necrotic phenotype and the p61 expression.

Next, we evaluated the capacity of the p29, p15, and p61 proteins to increment the AMV accumulation, since the presence of an RSS in AMV RNA 3 incremented its accumulation in P12 protoplasts (Martinez-Perez et al., 2019). For this purpose, the constructs described above were transfected into P12 protoplasts. The quantification of the relative AMV RNA 3 accumulation form northern blot analysis at 16 h post-transfection revealed that the expression of p29, p15, and p61 induces an increment of the AMV RNAs accumulation compared with the GFP control (**Figure 3**). Taken together, these findings further indicate potential RSS activity for the three CiLV-C encoded proteins.

The CiLV-C p15 Protein *Trans*-Complements an RSS Defective Turnip Crinkle Virus and Increases the Genomic Viral Accumulation

To gain additional insights into the RSS functionality for the CiLV-C proteins that show suppression activity, we tested an additional approach for RSS screening based on the functional complementation of a movement-defective TCV mutant (Powers et al., 2008). Leaves infiltrated with the correspondent binary constructions expressing the viral factors at 1 dpi were mechanically inoculated with the movement defective TCV-sGFP transcripts. Foci formation not limited to three or five cells was visualized by expression of the RSS HCPro and CiLV-C p15 and MP proteins (**Figure 4A**, complementation assay). The remaining CiLV-C proteins did not *trans*-complement the TCV-sGFP movement as observed in the negative controls (Lep and empty) (**Figure 4A**). To discard that the transport complementation observed for the TCV-sGFP construct could be due to an intrinsic movement function of the assayed protein, the infiltrated leaves were also inoculated with transcripts of a previously characterized TCVΔ92-sGFP construct that has a deletion into the MP ORF, which abolishes the TCV movement but still supports transcription of the subgenomic RNA (Powers et al., 2008). In all cases, except for MP, GFP expression was visualized in individual cells (**Supplementary Figure S2**), indicating that the ability of p15 and HCPro to complement TCV-sGFP movement was due to an RSS activity rather than an intrinsic movement function, as observed for the CiLV-C MP.

Next, we further tested the RSS activity of the CiLV-C proteins using the PZP-TCV-sGFP co-infiltration assay, whereby the presence of RSS activity is associated with an increase of the GFP signal (Powers et al., 2008; Martinez-Perez et al., 2019). To do this, leaves of *N. benthamiana* were co-infiltrated with individual cultures of *A. tumefaciens* transformed with PZP-TCV-sGFP and pMOG binary constructs carrying the HCPro and all CiLV-C genes. Leaves expressing the p15 protein resulted in a visual increase of GFP signal, in accordance to the HCPro positive control (**Figure 4A**, co-infiltration assay). In contrast, the other CiLV-C proteins tested did not increment the GFP fluorescent signal (**Figure 4A**), as observed for the negative controls (Lep and empty). The observation that the CiLV-C MP did not increase the visualized GFP signal further reaffirms that the capability of this protein to restore the transport of the movement defective TCV-sGFP construct correlates with its intrinsic movement function rather than to an RSS activity. Northern blot analyses of the viral RNAs derived from the TCV-sGFP construct showed a significant increase in

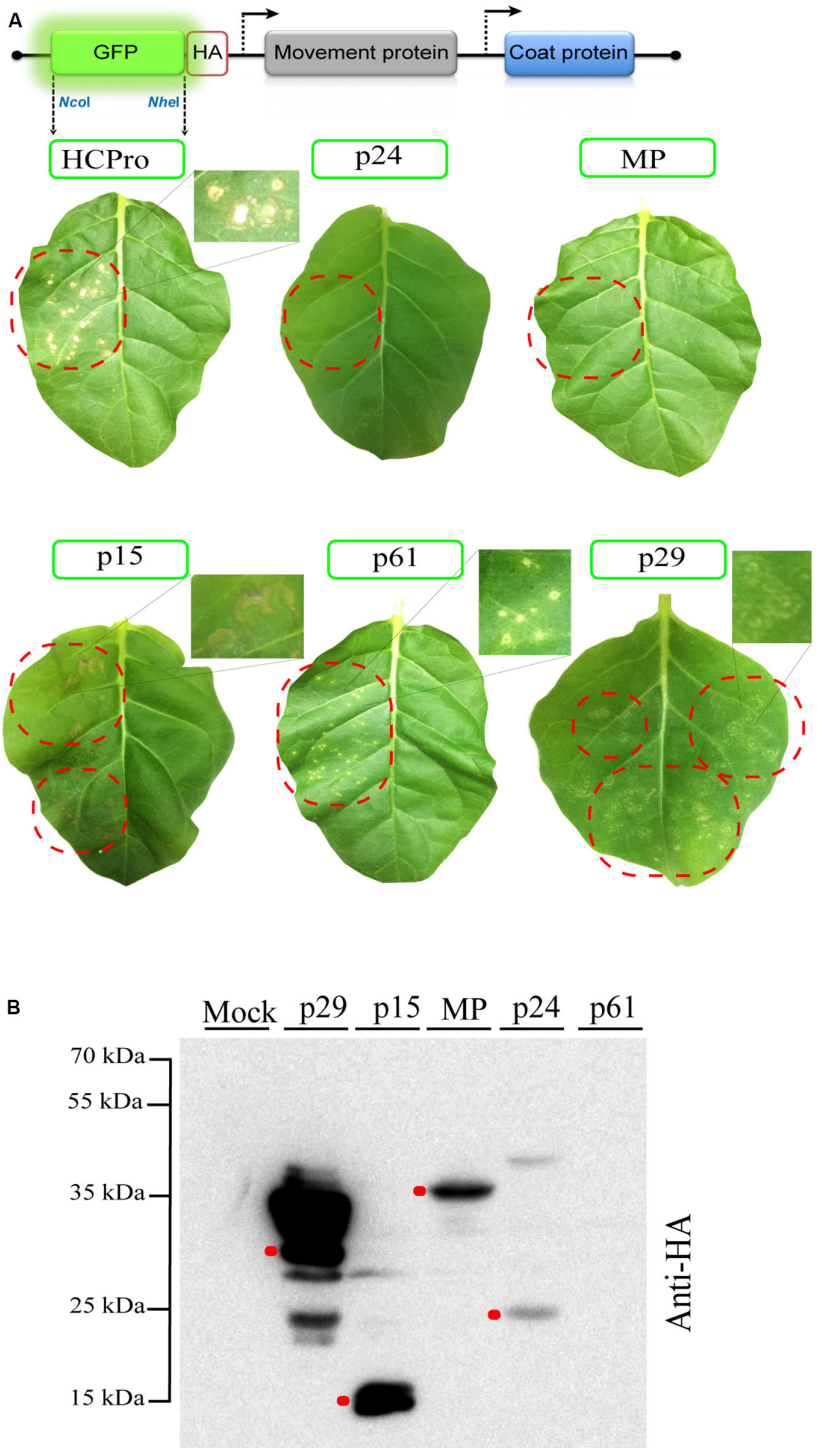


FIGURE 2 | The p29, p15, and p61 proteins trigger necrotic response in P12 *Nicotiana tabacum* leaves. **(A)** Inoculation of P12 leaves with the AMV RNA 3 derivative in which its GFP gene was exchanged by the tobacco etch virus HCPro and CiLV-C ORFs (p29, p15, p61, MP, and P24) fused at their C-terminal with the HA epitope. The schematic representation shows the GFP:HA/MP/CP AMV RNA 3, in which the open reading frames corresponding to the green fluorescent protein (GFP), the movement protein (MP), and the coat protein (CP) are represented by large boxes. Short red box corresponds to the HA epitope, meanwhile arrows represent subgenomic promoters. The *Nco*I and *Nhe*I restriction sites used for insertions of the assayed proteins are indicated. Necrotic response is observed in leaves inoculated with AMV RNA 3 construction expressing the HCPro (positive control), p29, p15, and p61 proteins at 4 days post-inoculation. The dotted red circles indicate the region of viral transcript inoculation. High magnification shows the necrotic lesions. Three independent experiments were performed, each one included the inoculation of three plants per construct; **(B)** Western blot analysis of the accumulation of proteins carrying the HA epitope in P12 leaves at 3 dpi. Red dots indicate the corresponding band for each protein.

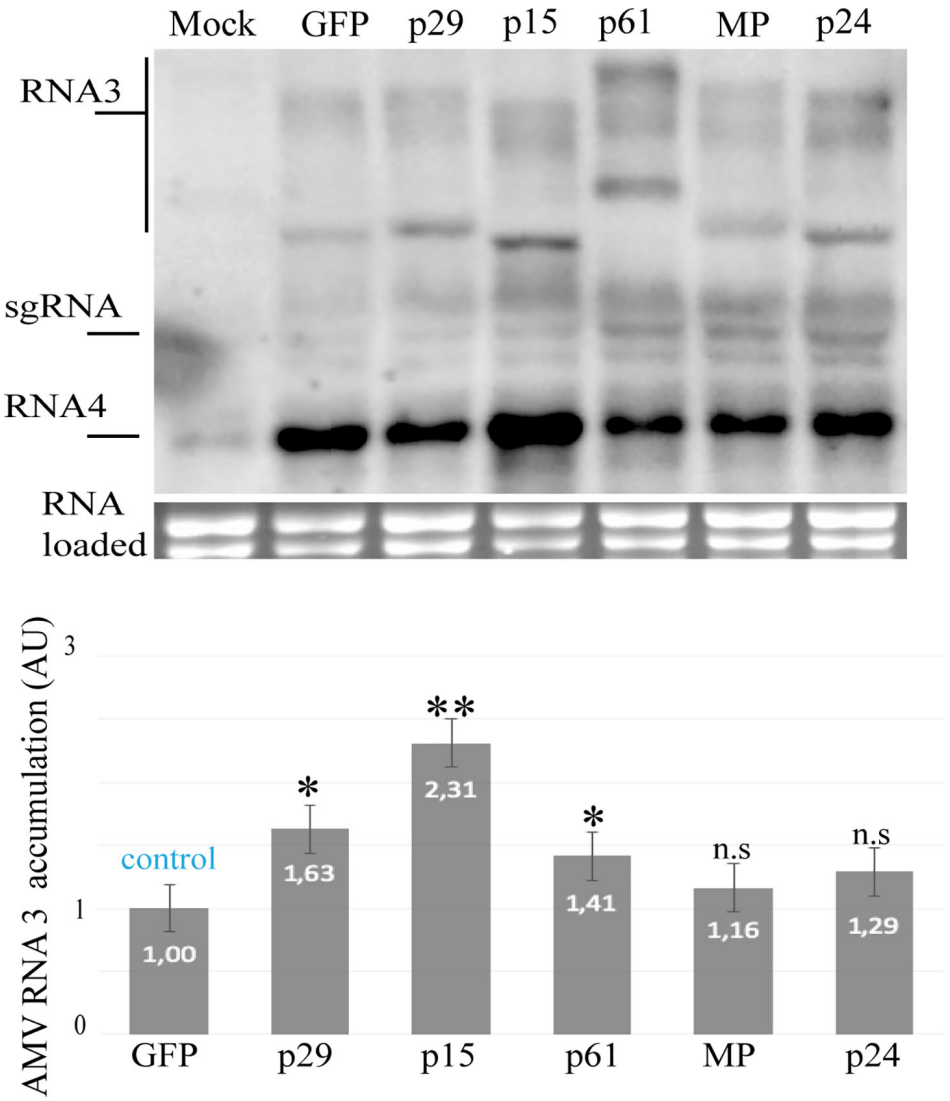


FIGURE 3 | Northern blot analysis of the accumulation of AMV RNA 3 RNA in P12 protoplasts at 16 h post-inoculation, using a dig-riboprobe complementary to the 3'UTR of the AMV. Transcripts correspond to the AMV RNA 3 derivative carrying the p29, p15, p61, MP, and p24 proteins of CiLV-C and GFP protein (negative control). Mock corresponds to the non-infected P12 plant. The localization of RNA 3 and subgenomic RNAs (sgRNA) are indicated. The graph represents the relative accumulation of AMV RNA 3, referred to the construct carrying the GFP from the average of three independent experiments. The value of the averages obtained from the band quantification was normalized in relation to the respective negative control (GFP). The bands were quantified using the ImageJ version 2.0cr software with ISAC plugin and error bars represent standard deviation. Statistical analysis were done using Student's *t*-test. Black asterisks (*) indicate significantly increased viral accumulation compared to the control. **, $p < 0.01$; *, $p < 0.05$; n.s, no significant difference. *P*-values were obtained from pairwise comparations between control vs. viral factor.

viral RNAs accumulation in leaves expressing HCPro and p15 proteins, when compared to the negative controls (empty and Lep) (**Figure 4B**).

The p29 and p15 Proteins Enhance Pathogenicity of a PVX Infectious Construct, While PVX-p61 Infection Generates a Hypersensitive Response

Using a PVX infectious system, it has been shown that an increase in the severity of the PVX infection suggests the presence of an additional gene with RSS activity (Voinnet et al., 1999; Martinez-Perez et al., 2019). To further test the RSS activity of the CiLV-C proteins we used a recombinant pGR107 PVX construct (Lu et al., 2003) to express the five CiLV-C proteins. *N. benthamiana* leaves were agroinfiltrated with the different PVX derivatives and symptoms development was monitored for 4 weeks. At 4 dpi, all PVX constructs elicited systemic mosaic, mottling, and interveinal chlorosis. At 10 dpi, plants infected with the PVX-p29 and PVX-p15 showed strong symptoms of necrosis in younger leaves, stunted growth, and leaf curling, resulting in complete death at 20 dpi. The same dead phenotype was observed

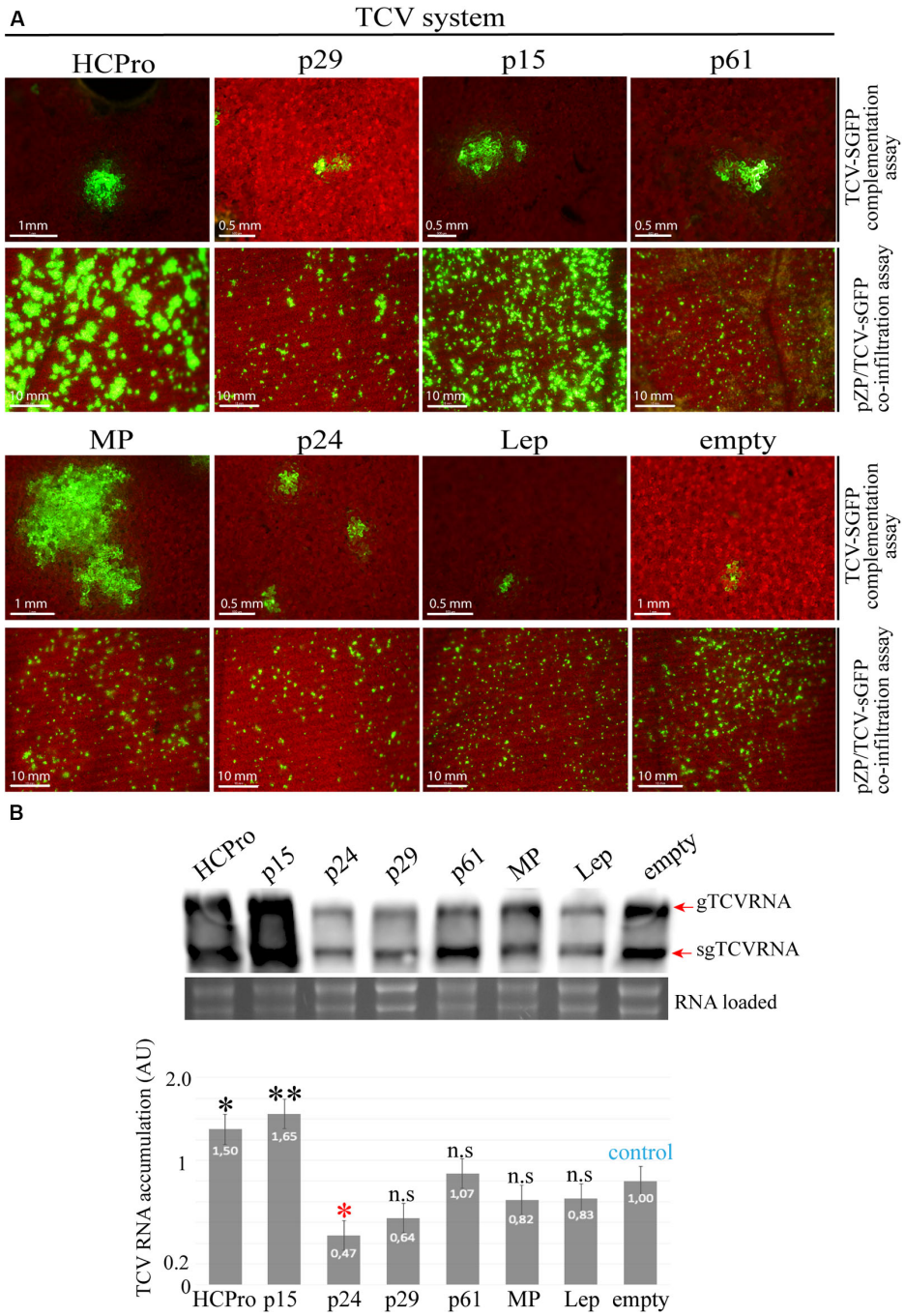


FIGURE 4 | The p15 protein complements the movement of an RSS defective turnip crinkle virus (TCV) mutant. **(A)** TCV assay based on the complementation or increment of fluorescence of the movement-deficiency phenotype of a TCV CP deletion mutant replaced by GFP (TCV-sGFP). For TCV-sGFP complementation assay, three *N. benthamiana* leaves per plant were infiltrated with $pMOG_{800}$ constructs carrying HCPro (positive control), p29, p15, p61, MP, p24, or pMOG-empty and pMOG-Lep as negative controls. Infectious RNA transcripts of the TCV-sGFP were mechanically inoculated 1 day post agroinfiltration. Cell-to-cell movement was evaluated at 3 dpi. For pZP/TCV-sGFP co-infiltration assay, pZP/TCV-sGFP was co-infiltrated with $pMOG_{800}$ empty and Lep (negative controls) or $pMOG_{800}$ carrying HCPro (positive control), p29, p15, p61, MP, and p24 CiLV-C genes. The increment of GFP fluorescence was monitored at 5 dpi. White bars correspond to 500 μm–10 mm. **(B)** Northern blot analysis showing accumulation of TCV-sGFP genomic (gTCV) and subgenomic (sgTCV) RNAs, using a dig-riboprobe complementary to the TCV 3'UTR at 5 dpi. rRNA stained with ethidium bromide indicates equal loading of samples. The graph represents the relative accumulation of gTCV and sgTCV RNAs from three independent experiments. The value of the averages obtained from the band quantification was normalized in relation to the respective negative control (empty). The bands were quantified using ImageJ version 2.0cr software with ISAC plugin and error bars represent standard deviation. Statistical analyses were done using Student's *t*-test. Black asterisks (*) indicate significantly increased viral accumulation compared to the control, while red asterisks (*) indicate significatively decreased viral accumulation. **, $p < 0.01$; *, $p < 0.05$; n.s, no significant difference. *P*-values were obtained from pairwise comparations between control vs. viral factor.

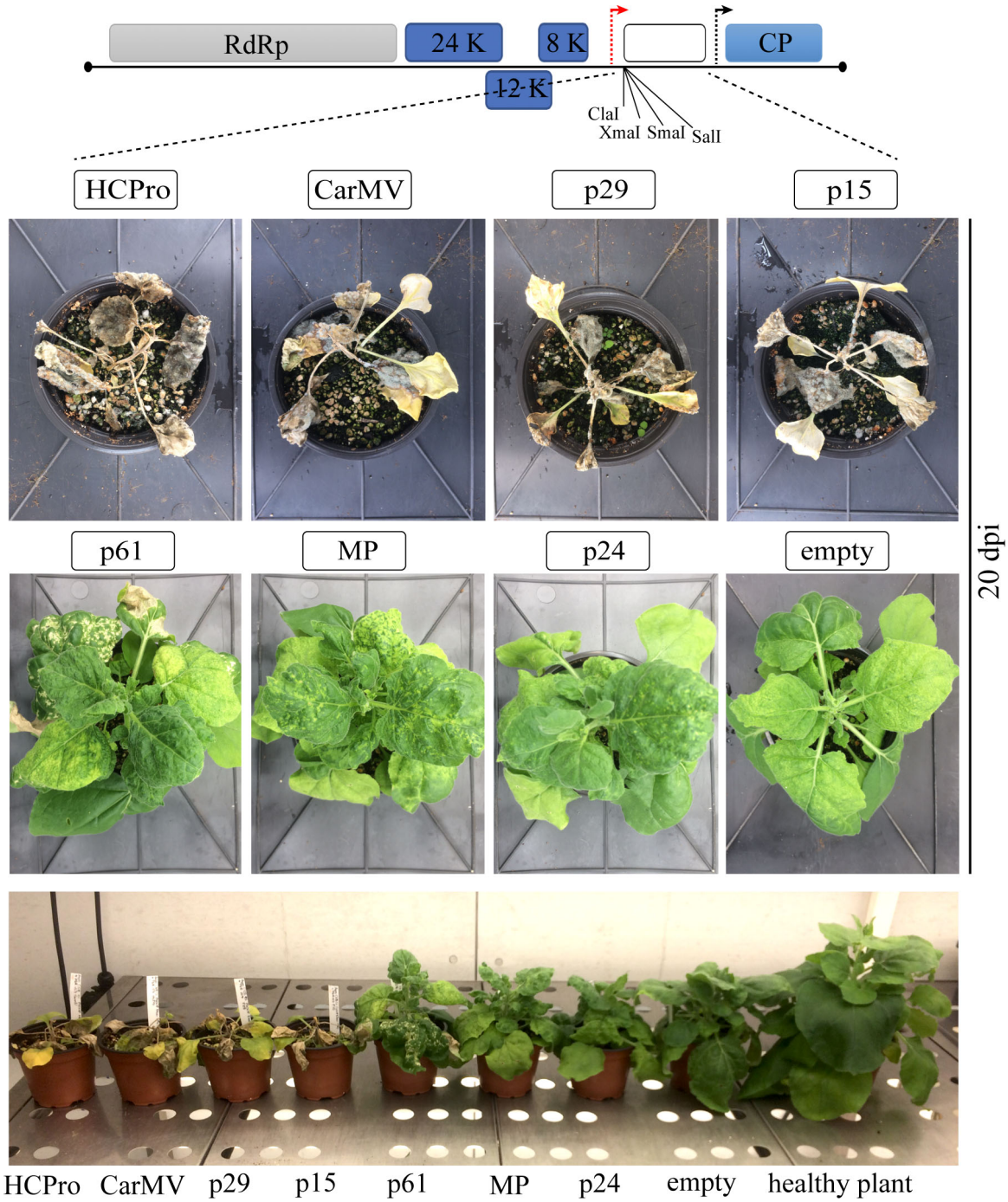


FIGURE 5 | RSS activity of CiLV-C proteins using the potato virus X (PVX) assay. Symptom phenotype of wild type PVX (negative control) and PVX expressing the HCPro, CP CarMV (positive controls), p29, p15, p61, MP, and p24 CiLV-C proteins in *N. benthamiana* plants at 20 dpi. The schematic representation corresponds the pGR107 PVX infectious construct containing *Cla*I, *Xma*I, *Sma*I, and *Sal*I sites downstream of the duplicated PVX CP promoter (red arrow), used for the insertion of the indicated viral ORFs.

with the positive controls carrying the HCPro (PVX-HCPro) or the CarMV CP (PVX-CP CarMV) (Martinez-Perez et al., 2019; **Figure 5**). At 3 dpi, the infiltrated and first upper leaves infected with PVX-p61 construct exhibited clear hypersensitive response unable to contain the pathogen and causing systemic mosaic symptoms (**Figure 5**). For the plants infected with the PVX constructs expressing the MP and p24 CiLV-C proteins, symptoms were similar to those observed for the PVX-empty negative control, but with a little more mosaic (**Figure 5**). To correlate disease severity with PVX viral accumulation, we

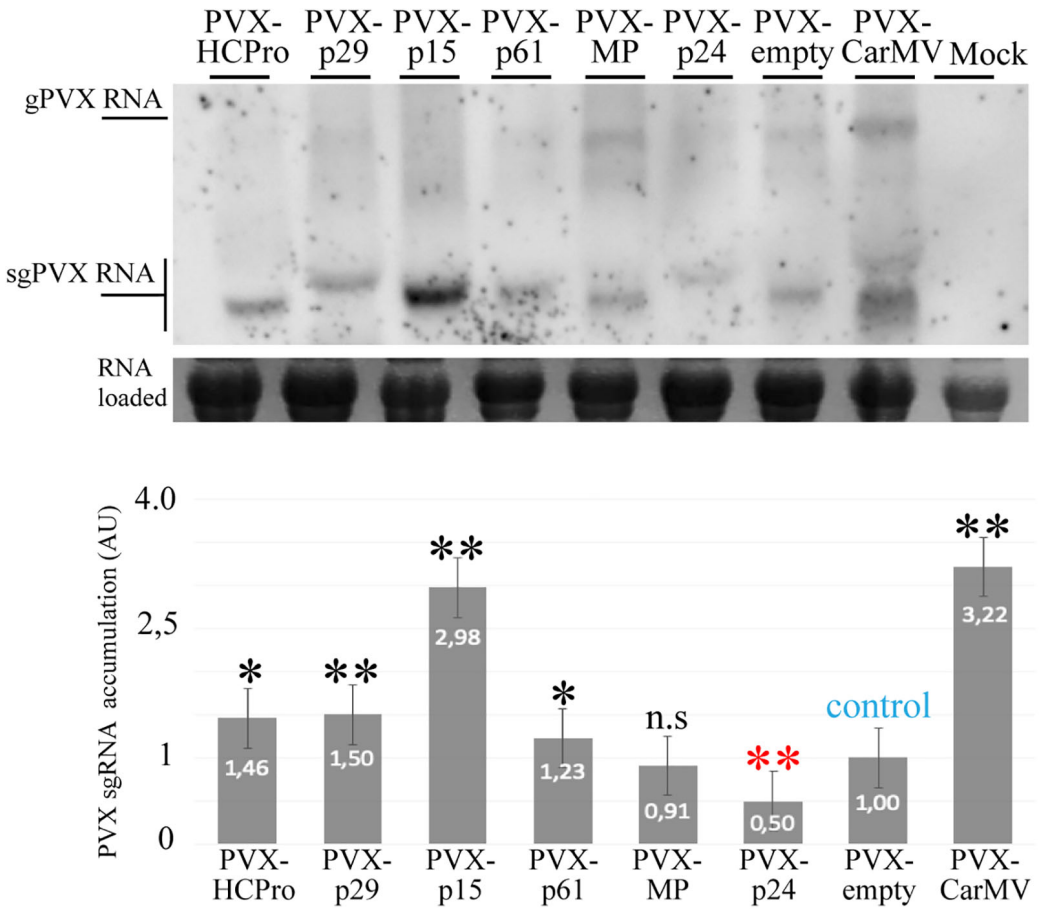


FIGURE 6 | The p29, p15, and p61 expression increase the PVX accumulation. Northern blot analysis showing accumulation of the subgenomic (sgPVX) PVX RNAs using a dig-riboprobe complementary to the 3' end region of the PVX CP gene at 4 dpi. rRNA stained with ethidium bromide indicates equal loading of samples. Mock corresponds to uninfected plant. The graph represents the relative sgPVX RNA accumulation. The values correspond to the averages of three different experiments in which the data are normalized to the negative control (PVX-empty). The bands were quantified using ImageJ version 2.0cr software with ISAC plugin and error bars represent standard deviation. Statistical analysis were done using Student's *t*-test. Black asterisks (*) indicate significantly increased viral accumulation compared to the control, while red asterisks (*) indicate significatively decreased viral accumulation. **, $p < 0.01$; *, $p < 0.05$; n.s, no significant difference. *P*-values were obtained from pairwise comparations between control vs. viral factor.

performed northern blot analysis at 4 dpi. The quantification of the relative PVX subgenomic RNA accumulation revealed that constructs carrying the p29, p15, and p61 genes and the two positive controls (HCPro and CP CarMV) induced an increment of the PVX RNA accumulation compared with the wild type PVX (**Figure 6**). These results indicate that p29, p15, and p61 significantly enhanced the pathogenicity of PVX by increasing virus accumulation.

On the other hand, the infection derived from PVX-p61 construct presented two different phenotypes in the first upper non-infiltrated leaf, showing a clear hypersensitive response (see HR, **Figure 7**) together with regions showing mosaic, mottling, and inward leaf curling symptoms (a mosaic of symptoms – MS) (see MS, **Figure 7A**). The HR phenotype derived from the PVX-p61 construct was clearly distinguished from the necrotic response derived from the PVX-HCPro and PVX-empty variants (**Figure 7**). In order to correlate the HR phenotype with the presence of the p61 gene, RT-PCR analysis was performed using specific p61 ORF primers from inoculated (IL) and upper (UL) leaves showing either HR or MS phenotypes. The expected p61 amplicon was detected only in the infected tissue showing the HR phenotype (**Figure 7B**, see lines IL and HR), indicating a direct correlation between this phenotype and the presence of p61 gene. The absence of the p61 amplicon in the infected tissue showing MS phenotype also indicates the low genetic stability of the PVX-p61 construct. For the others PVX derivatives assayed, RT-PCR analysis using specific ORF primers and performed at 14 dpi rendered the expected amplicons in upper non-inoculated leaves, indicating their genetic stability (**Supplementary Figure S3**).

DISCUSSION

In this study, we reported that citrus leprosis virus C, the type member of the genus *Cilevirus*, harbors three proteins with RNA silencing suppressor activity, using different assays for

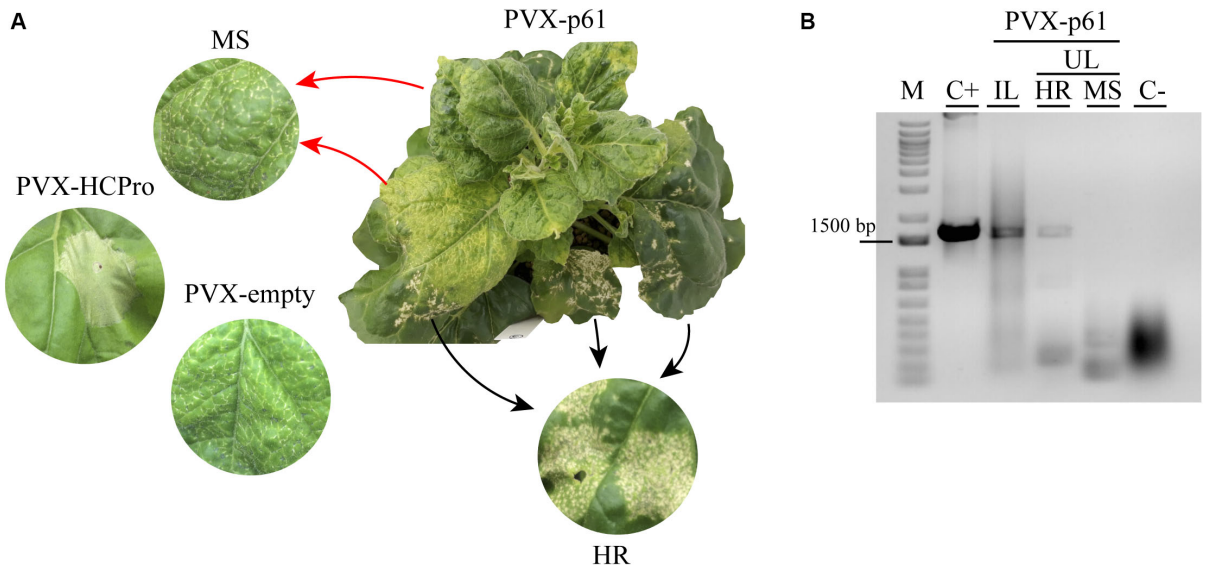


FIGURE 7 | The p61 expression on PVX infection generates a hypersensitive response (HR) phenotype. **(A)** *N. benthamiana* plant infiltrated with pGR107 PVX construct expressing p61 CiLV-C protein. Symptoms of hypersensitive response (HR) and a mosaic of symptoms (MS) are presented in the same upper leave from PVX-p61 infection at 3 dpi. The MS phenotype is also visualized in wild type PVX-empty infection. Necrotic lesions are observed in *N. benthamiana* plants infiltrated with pGR107 PVX construct expressing the RSS HCPro protein. **(B)** RT- PCR analysis to detect the p61 gene (1.650 bp) in inoculated (IL) and upper (UL) leaves of *N. benthamiana* plants inoculated with pGR107-p61 construct, exhibiting HR and MS symptoms. C+, RT-PCR from *N. benthamiama* leaves transiently expressing the p61 (positive control); C-, RT-PCR from uninfected plant sample. The marker band size of 1.500 bp is indicated. M, 1 kb DNA ladder (ThermoFisher Scientific Inc., United States).

the RSS screening (*N. benthamiana* 16c plants, AMV system, siRNA accumulation, RSS defective TCV and PVX pathogenicity system). We observed that three CiLV-C proteins showed RSS activity in almost all evaluated systems but were unable to suppress the local silencing using the 16c assay. These findings show that the method based on *N. benthamiana* 16c is not a foolproof system to screen RSS viral proteins, indicating the continuous need to use other methods to find RSS functionality of unknown viral genes. *Agrobacterium* co-infiltration assay is a relatively quick and easy method to identify RSS proteins, but some RSS are not detected due to their mode of action, sensitivity or lack of intracellular suppression activity (Lu et al., 2004). Regarding the inability of the CiLV-C proteins to suppress the local silencing in 16c plants, two hypotheses can be raised: (1) the p29, p15, and p61 could have a weak suppressive activity at the cellular level or (2) these proteins would not act at the local RNA silencing process, performing their blocking activity in another step of the silencing defense machinery. Unfortunately, it was not possible to test the ability of the CiLV-C proteins to prevent the systemic spread of the GFP transgene silencing signal in 16c plants and, therefore, further experiments are needed to address this question. It is interesting to note that although the p29, p15, and p61 CiLV-C proteins were unable to induce the local RSS in 16c plants, their expression in wild type or transgenic *N. benthamiana* 16c, reduced considerably the accumulation of small RNAs, indicating that the "patch" technique is more efficient by analyzing the small RNAs accumulation derived from the reporter gene rather than to perform the screening only by detecting the fluorescent signal. Collectively, these data suggest that these viral proteins interfere in a step after dsRNA production, whereby the decrease of siRNA accumulation can be explained by a possible binding of the proteins to 21 and 24 nt siRNAs. Revealing whether the p29, p15, and p61 have affinity for 21 or 24 nt siRNAs would help clarify this issue.

In contrast to the negative results obtained with the "patch" technique based on the 16c plants, the *cis*-expression of the p29, p15, and p61 induced a phenotype of necrotic lesions in AMV infection context, increasing the viral RNA accumulation. Also, these three cilevirus proteins caused a substantial reduction of siRNA accumulation when transiently expressed in leaf patches of 16c and wild type *N. benthamiana* plants and enhanced the PVX pathogenicity, incrementing its viral RNA accumulation. Taken together, these data indicate that the p29, p15, and p61 are silencing suppressors of CiLV-C. Examples of multiple-components RNA silencing suppression viral mechanisms have been reported for criniviruses, closteroviruses, geminiviruses, emaraviruses, and potyviruses (Lu et al., 2004; Vanitharani et al., 2004; Canizares et al., 2008; Gupta et al., 2018; Rodamilans et al., 2018). This plurality of proteins with silencing suppressor activity from a single viral entity may indicate a constant viral adaptation to counteract the plant defense mechanism.

In our previous study, we showed that p15 is localized into the nucleus of *N. benthamiana* epidermal cells with unknown function (Leastro et al., 2018). Here, we clearly identified the p15 RSS activity using four different methods. The nuclear localization of p15 strongly suggests that this protein accesses

the nucleus possibly to block RNA silencing. This same feature has been demonstrated for other RSS viral proteins, i.e., the 2b proteins of cucumber mosaic virus (CMV), which nuclear localization is a prerequisite for an efficient suppression of PTGS (Lucy et al., 2000). In arabidopsis, the siRNA processing, which encompasses RdRP activity, Dicer processing and Ago mediated target cleavage are all intimately linked in the nucleus (Castel and Martienssen, 2013).

The suppression of RNA silencing activity of unknown viral genes has been identified from screening based on functional complementation of defective viral mutants (Chiba et al., 2006; Powers et al., 2008; Martinez-Perez et al., 2019). The p15 was able to rescue the cell-to-cell movement of the RSS defective TCV and to increase the GFP signal, further supporting that p15 is a silencing suppressor of CiLV-C. In contrast, the p29 and p61 CiLV-C proteins were incapable to rescue the viral movement and to increment the GFP expression in this system. The PVX p25, a viral protein that has been previously reported incapable of suppressing local silencing, acting exclusively in steps associated with intercellular suppression silencing (Voinnet et al., 2000) also was unable to rescue the movement of the RSS TCV defective construct (Powers et al., 2008). On the other hand, p29 and p61 RSS activities were identified from small RNA accumulation analyses and with the AMV and PVX system. Taken together, these data suggest that these proteins may affect the RNAi pathway (Mallory et al., 2002) differently than p15. The TCV system was unable to identify RSS activity for those suppressors that act blocking the spread of silencing signal (Powers et al., 2008), suggesting that p15 probably acts at the level of local silencing, however, it seems to be a weak local suppressor, given its inability to suppress local silencing in 16c plants.

We recently proved that cilevirus movement protein (p32) is efficiently able to generate viral cell-to-cell and long-distance spread in heterologous systems (Leastro et al., unpublished), indicating that limitation of the cileviruses to systemically infect their hosts (Freitas-Astua et al., 2018) is not due to a functional restriction in their MPs. In this sense, although speculative, we infer that the putative weak suppression activity observed for the CiLV-C RSS proteins could explain, at least in part, the systemic movement impairment, since the RNA silencing could hinder the vascular transport of viruses, which inhibits viral entry into the phloem (Yelina et al., 2002; Vuorinen et al., 2011). Our above suggestion is reinforced from RSS studies with the proteins of orchid fleck virus (OFV-citrus), a dichorhavirus also involved in the citrus leprosis complex (Roy et al., 2013). As demonstrated for CiLV-C, the identified RSS protein encoded by OFV also shows inability to suppress local RNA silencing in 16c plants (Leastro et al., unpublished) and dichorhaviruses also have limitations to systemically infect their hosts. The observation that the incapacity of cileviruses to infect their hosts is not limited to one host, but to more than 50 different natural and experimental host species (Garita et al., 2014) opens the possibility that a general mechanism (e.g., plant RNA silencing) could be responsible for impairing the long movement of these citrus leprosis-associated viruses. However, further experiments are needed to address this question.

The p29 and p15 significantly enhanced the PVX accumulation and symptoms severity, including death of tobacco plants. Similar results were previously observed in plants agroinfiltrated with PVX recombinants expressing several viral suppressor proteins (Brigneti et al., 1998; Pfeffer et al., 2002; Thomas et al., 2003; Delgadillo et al., 2004; Canizares et al., 2008; Gupta et al., 2018; Martinez-Perez et al., 2019). We observed that the p61 also enhanced the PVX accumulation, although symptom severity was not maintained over time due to the instability of the PVX-p61 construct. However, during the onset of the PVX-p61 infection, a clear hypersensitive necrosis was observed similar to that visualized during the transient expression of the p61 protein (Leastro et al., 2018). A hypersensitive-like response has been suggested as an outcome of CiLV-C infection in arabidopsis plants (Arena et al., 2016) and our data further reinforces the idea that, together with the RSS activity, p61 could also activate a plant defense response mechanism, acting as a pathogenicity determinant. On this context, recently Arena et al. (unpublished) also showed that the expression of p61 protein triggers HR and furthermore mimics plant responses to viral infection.

Collectively, our findings establish that p29, p15, and p61 proteins possess RSS activity. Thus, we elucidated at least part of the functionality of the p15 protein, and provided some additional functions for the capsid protein (p29) and the putative glycoprotein (p61) of the citrus leprosis virus C.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

ML, DC, and JS-N conceived and designed the experiments, analyzed, and interpreted the data. ML and DC performed the experiments. ML, JF-A, EK, VP, and JS-N contributed reagents, materials, and tools. ML wrote the original draft preparation. ML, JF-A, VP, and JS-N revised and edited the manuscript.

FUNDING

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), proc. 2014/0845-9, 2017/50222-0, 2015/10249-1, and 2017/19898-8. This work was also supported by Instituto para la Formación y Aprovechamiento de Recursos Humanos, Becas IFARHU-SENACYT, contrato 270-2018-361, grant BIO2017-88321-R from the Spanish Agencia Estatal de Investigación (AEI), Fondo Europeo de Desarrollo Regional (FEDER), and the Prometeo Program GV2015/010 from the Generalitat Valenciana.

ACKNOWLEDGMENTS

We thank Dr. José Antonio Navarro (Instituto de Biología Molecular y Celular de Plantas "Primo Yúfera" – IBMCP, Valencia, Spain) for kindly providing the TCV and PVX viral vectors. We are grateful to Ms. Lorena Corachán for her excellent technical support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.01231/full#supplementary-material

FIGURE S1 | Confirmation of p61 expression on AMV infection using a frameshift version of p61 gene. *Nicotiana tabacum* P12 leaves inoculated with three variants of the AMV RNA 3 construct, which express the p61 ORF with the HA epitope fused at its N- (HA:p61) or C- (p61:HA) termini or with a frameshift mutation in the 5'terminal region (p61fs). The schematic representation shows the GFP:HA/MP/CP AMV RNA 3, in which the open reading frames correspond to the green fluorescent protein (GFP), the movement protein (MP) and the coat protein (CP) are represented by large boxes. In the assayed AMV constructs the GFP gene is changed with the different p61 indicated versions. Necrotic response observed at 4 dpi. Arrows represent subgenomic promoters and doted circles delimited the inoculated leaf area. Three independent experiments were performed, each one included the infiltration of three leaves per plants and three plants per construct.

FIGURE S2 | The CiLV-C p15 does not complement movement of the TCVΔ92-sGFP mutant. *N. benthamiana* leave were infiltrated with $pMOG_{800}$ expressing the HCPro and CiLV-C MP and p15 proteins. Infectious RNA transcript of the TCVΔ92-sGFP construct were mechanically inoculated one day post agroinfiltration. Cell-to-cell movement evaluated at 3 dpi. White bars correspond to 0.5–2 mm.

FIGURE S3 | RT- PCR analysis to confirm the genetic stability of the PVX constructs expressing the CiLV-C p29 (792 bp), p15 (393 bp), p61 (1614 bp), MP (894 bp) and p24 (645 bp) proteins. RT-PCR performed from upper leaves of *N. benthamiana* plants infected at 14 dpi using specific primers for each gene. C+, RT-PCR from *N. benthamiama* leaves transiently expressing the HCPro (positive control); C−, RT-PCR from uninfected plant sample. The marker band size of 300, 650, 1,000, and 1,500 bp are indicated. M, 1 kb DNA ladder.

REFERENCES

Anandalakshmi, R., Pruss, G. J., Ge, X., Marathe, R., Mallory, A. C., Smith, T. H., et al. (1998). A viral suppressor of gene silencing in plants. *Proc. Natl. Acad. Sci. U.S.A.* 95, 13079–13084.

Aravin, A. A., Naumova, N. M., Tulin, A. V., Vagin, V. V., Rozovsky, Y. M., and Gvozdev, V. A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr. Biol.* 11, 1017–1027. doi: 10.1016/s0960-9822(01)00299-8

Arena, G. D., Ramos-Gonzalez, P. L., Nunes, M. A., Ribeiro-Alves, M., Camargo, L. E., Kitajima, E. W., et al. (2016). *Citrus leprosis virus C* infection results in hypersensitive-like response, suppression of the JA/ET plant defense pathway and promotion of the colonization of its mite vector. *Front. Plant Sci.* 7:1757. doi: 10.3389/fpls.2016.01757

Borges, F., and Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nat. Rev. Mol. Cell Biol.* 16, 727–741. doi: 10.1038/nrm4085

Brigneti, G., Voinnet, O., Li, W. X., Ji, L. H., Ding, S. W., and Baulcombe, D. C. (1998). Viral pathogenicity determinants are suppressors of transgene silencing in *Nicotiana benthamiana*. *EMBO J.* 17, 6739–6746. doi: 10.1093/emboj/17.22.6739

Burgyan, J., and Havelda, Z. (2011). Viral suppressors of RNA silencing. *Trends Plant Sci.* 16, 265–272. doi: 10.1016/j.tplants.2011.02.010

Canizares, M. C., Navas-Castillo, J., and Moriones, E. (2008). Multiple suppressors of RNA silencing encoded by both genomic RNAs of the crinivirus, *Tomato chlorosis virus*. *Virology* 379, 168–174. doi: 10.1016/j.virol.2008.06.020

Castel, S. E., and Martienssen, R. A. (2013). RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat. Rev. Genet.* 14, 100–112. doi: 10.1038/nrg3355

Chiba, M., Reed, J. C., Prokhnevsky, A. I., Chapman, E. J., Mawassi, M., Koonin, E. V., et al. (2006). Diverse suppressors of RNA silencing enhance agroinfection by a viral replicon. *Virology* 346, 7–14. doi: 10.1016/j.virol.2005.09.068

Csorba, T., Kontra, L., and Burgyan, J. (2015). viral silencing suppressors: tools forged to fine-tune host-pathogen coexistence. *Virology* 479–480, 85–103. doi: 10.1016/j.virol.2015.02.028

Delgadillo, M. O., Saenz, P., Salvador, B., Garcia, J. A., and Simon-Mateo, C. (2004). Human influenza virus NS1 protein enhances viral pathogenicity and acts as an RNA silencing suppressor in plants. *J. Gen. Virol.* 85, 993–999. doi: 10.1099/vir.0.19735-0

Ding, S. W., and Voinnet, O. (2007). Antiviral immunity directed by small RNAs. *Cell* 130, 413–426. doi: 10.1016/j.cell.2007.07.039

Freitas-Astua, J., Ramos-Gonzalez, P. L., Arena, G. D., Tassi, A. D., and Kitajima, E. W. (2018). *Brevipalpus*-transmitted viruses: parallelism beyond a common vector or convergent evolution of distantly related pathogens? *Curr. Opin. Virol.* 33, 66–73. doi: 10.1016/j.coviro.2018.07.010

Garcia, J. A., and Pallas, V. (2015). Viral factors involved in plant pathogenesis. *Curr. Opin. Virol.* 11, 21–30. doi: 10.1016/j.coviro.2015.01.001

Garita, L. C., Tassi, A. D., Calegario, R. F., and Freitas-Astúa, J. (2014). Experimental host range of *Citrus leprosis virus C* (CiLV-C). *Trop. Plant Pathol.* 39, 43–55. doi: 10.1590/s1982-56762014005000004

Guilley, H., Bortolamiol, D., Jonard, G., Bouzoubaa, S., and Ziegler-Graff, V. (2009). Rapid screening of RNA silencing suppressors by using a recombinant virus derived from beet necrotic yellow vein virus. *J. Gen. Virol.* 90, 2536–2541. doi: 10.1099/vir.0.011213-0

Guo, Q., Liu, Q., Smith, N. A., Liang, G., and Wang, M. B. (2016). RNA silencing in plants: mechanisms, technologies and applications in horticultural crops. *Curr. Genomics* 17, 476–489. doi: 10.2174/1389202917666160520103117

Gupta, A. K., Hein, G. L., Graybosch, R. A., and Tatineni, S. (2018). Octapartite negative-sense RNA genome of high plains wheat mosaic virus encodes two suppressors of RNA silencing. *Virology* 518, 152–162. doi: 10.1016/j.virol.2018.02.013

Hamilton, A., Voinnet, O., Chappell, L., and Baulcombe, D. (2002). Two classes of short interfering RNA in RNA silencing. *EMBO J.* 21, 4671–4679. doi: 10.1093/emboj/cdf464

Hamilton, A. J., and Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286, 950–952. doi: 10.1126/science.286.5441.950

Hammond, S. M., Bernstein, E., Beach, D., and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293–296. doi: 10.1038/35005107

Jones, L., Hamilton, A. J., Voinnet, O., Thomas, C. L., Maule, A. J., and Baulcombe, D. C. (1999). RNA-DNA interactions and DNA methylation in post-transcriptional gene silencing. *Plant Cell* 11, 2291–2301. doi: 10.1105/tpc.11.12.2291

Kakumani, P. K., Ponia, S. S., Sood, V., Chinnappan, M., Banerjea, A. C., Medigeshi, G. R., et al. (2013). Role of RNA interference (RNAi) in dengue virus replication and identification of NS4B as an RNAi suppressor. *J. Virol.* 87, 8870–8883. doi: 10.1128/JVI.02774-12

Kuchibhatla, D. B., Sherman, W. A., Chung, B. Y., Cook, S., Schneider, G., Eisenhaber, B., et al. (2014). Powerful sequence similarity search methods and in-depth manual analyses can identify remote homologs in many apparently "orphan" viral proteins. *J. Virol.* 88, 10–20. doi: 10.1128/JVI.02595-13

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0

Leastro, M. O., Kitajima, E. W., Silva, M. S., Resende, R. O., and Freitas-Astua, J. (2018). Dissecting the subcellular localization, intracellular trafficking,

interactions, membrane association, and topology of *Citrus leprosis virus C* proteins. *Front. Plant Sci.* 9:1299. doi: 10.3389/fpls.2018.01299

Leastro, M. O., Pallas, V., Resende, R. O., and Sanchez-Navarro, J. A. (2015). The movement proteins (NSm) of distinct tospoviruses peripherally associate with cellular membranes and interact with homologous and heterologous NSm and nucleocapsid proteins. *Virology* 478, 39–49. doi: 10.1016/j.virol.2015.01.031

Leastro, M. O., Pallas, V., Resende, R. O., and Sanchez-Navarro, J. A. (2017). The functional analysis of distinct tospovirus movement proteins (NSM) reveals different capabilities in tubule formation, cell-to-cell and systemic virus movement among the tospovirus species. *Virus Res.* 227, 57–68. doi: 10.1016/j.virusres.2016.09.023

Li, F., and Ding, S. W. (2006). Virus counterdefense: diverse strategies for evading the RNA-silencing immunity. *Annu. Rev. Microbiol.* 60, 503–531. doi: 10.1146/annurev.micro.60.080805.142205

Li, W. X., and Ding, S. W. (2001). Viral suppressors of RNA silencing. *Curr. Opin. Biotechnol.* 12, 150–154. doi: 10.1016/s0958-1669(00)00190-7

Locali-Fabris, E. C., Freitas-Astúa, J., and And Machado, M. A. (2012). "Genus cilevirus," in *Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses*, eds A. M. Q. King, M. J. Adams, E. B. Carstens, and E. J. Lefkowitz (San Diego, CA: Elsevier), 1139–1142.

Locali-Fabris, E. C., Freitas-Astúa, J., Souza, A. A., Takita, M. A., Astua-Monge, G., Antonioli-Luizon, R., et al. (2006). Complete nucleotide sequence, genomic organization and phylogenetic analysis of Citrus leprosis virus cytoplasmic type. *J. Gen. Virol.* 87, 2721–2729. doi: 10.1099/vir.0.82038-0

Loesch-Fries, L. S., Jarvis, N. P., Krahn, K. J., Nelson, S. E., and Hall, T. C. (1985). Expression of alfalfa mosaic virus RNA 4 cDNA transcripts in vitro and in vivo. *Virology* 146, 177–187. doi: 10.1016/0042-6822(85)90002-9

Lu, R., Folimonov, A., Shintaku, M., Li, W. X., Falk, B. W., Dawson, W. O., et al. (2004). Three distinct suppressors of RNA silencing encoded by a 20-kb viral RNA genome. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15742–15747. doi: 10.1073/pnas.0404940101

Lu, R., Maduro, M., Li, F., Li, H. W., Broitman-Maduro, G., Li, W. X., et al. (2005). Animal virus replication and RNAi-mediated antiviral silencing in *Caenorhabditis elegans*. *Nature* 436, 1040–1043. doi: 10.1038/nature03870

Lu, R., Malcuit, I., Moffett, P., Ruiz, M. T., Peart, J., Wu, A. J., et al. (2003). High throughput virus-induced gene silencing implicates heat shock protein 90 in plant disease resistance. *EMBO J.* 22, 5690–5699. doi: 10.1093/emboj/cdg546

Lucy, A. P., Guo, H. S., Li, W. X., and Ding, S. W. (2000). Suppression of post-transcriptional gene silencing by a plant viral protein localized in the nucleus. *EMBO J.* 19, 1672–1680. doi: 10.1093/emboj/19.7.1672

Mallory, A. C., Reinhart, B. J., Bartel, D., Vance, V. B., and Bowman, L. H. (2002). A viral suppressor of RNA silencing differentially regulates the accumulation of short interfering RNAs and micro-RNAs in tobacco. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15228–15233. doi: 10.1073/pnas.232434999

Mann, K. S., Johnson, K. N., Carroll, B. J., and Dietzgen, R. G. (2016). Cytorhabdovirus P protein suppresses RISC-mediated cleavage and RNA silencing amplification in planta. *Virology* 490, 27–40. doi: 10.1016/j.virol.2016.01.003

Martinez-Perez, M., Navarro, J. A., Pallas, V., and Sanchez-Navarro, J. A. (2019). A sensitive and rapid RNA silencing suppressor activity assay based on alfalfa mosaic virus expression vector. *Virus Res.* 272:197733. doi: 10.1016/j.virusres.2019.197733

Matranga, C., and Zamore, P. D. (2007). Small silencing RNAs. *Curr. Biol.* 17, R789–R793.

Merai, Z., Kerenyi, Z., Kertesz, S., Magna, M., Lakatos, L., and Silhavy, D. (2006). Double-stranded RNA binding may be a general plant RNA viral strategy to suppress RNA silencing. *J. Virol.* 80, 5747–5756. doi: 10.1128/jvi.01963-05

Moissiard, G., and Voinnet, O. (2004). Viral suppression of RNA silencing in plants. *Mol. Plant Pathol.* 5, 71–82. doi: 10.1111/j.1364-3703.2004.00207.x

Moon, J. Y., and Park, J. M. (2016). Cross-talk in viral defense signaling in plants. *Front. Microbiol.* 7:2068. doi: 10.3389/fmicb.2016.02068

Nakanishi, K. (2016). Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins? *Wiley Interdiscip. Rev. RNA* 7, 637–660. doi: 10.1002/wrna.1356

Pallas, V., Mas, P., and Sanchez-Navarro, J. A. (1998). Detection of plant RNA viruses by nonisotopic dot-blot hybridization. *Methods Mol. Biol.* 81, 461–468. doi: 10.1385/0-89603-385-6:461

Pascon, R. C., Kitajima, J. P., Breton, M. C., Assumpcao, L., Greggio, C., Zanca, A. S., et al. (2006). The complete nucleotide sequence and genomic organization of citrus leprosis associated virus, cytoplasmatic type (CiLV-C). *Virus Genes* 32, 289–298. doi: 10.1007/s11262-005-6913-1

Peiro, A., Martinez-Gil, L., Tamborero, S., Pallas, V., Sanchez-Navarro, J. A., and Mingarro, I. (2014). The *Tobacco mosaic virus* movement protein associates with but does not integrate into biological membranes. *J. Virol.* 88, 3016–3026. doi: 10.1128/jvi.03648-13

Pfeffer, S., Dunoyer, P., Heim, F., Richards, K. E., Jonard, G., and Ziegler-Graff, V. (2002). P0 of beet Western yellows virus is a suppressor of posttranscriptional gene silencing. *J. Virol.* 76, 6815–6824. doi: 10.1128/jvi.76.13.6815-6824.2002

Pisacane, P., and Halic, M. (2017). Tailing and degradation of Argonaute-bound small RNAs protect the genome from uncontrolled RNAi. *Nat. Commun.* 8:15332. doi: 10.1038/ncomms15332

Powers, J. G., Sit, T. L., Qu, F., Morris, T. J., Kim, K. H., and Lommel, S. A. (2008). A versatile assay for the identification of RNA silencing suppressors based on complementation of viral movement. *Mol. Plant Microbe Interact.* 21, 879–890. doi: 10.1094/MPMI-21-7-0879

Pumplin, N., and Voinnet, O. (2013). RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. *Nat. Rev. Microbiol.* 11, 745–760. doi: 10.1038/nrmicro3120

Qu, F., and Morris, T. J. (2005). Suppressors of RNA silencing encoded by plant viruses and their role in viral infections. *FEBS Lett.* 579, 5958–5964. doi: 10.1016/j.febslet.2005.08.041

Quito-Avila, D. F., Freitas-Astúa, J., and Melzer, M. J. (2020). " Bluner-, Cile-, and higreviruses (kitaviridae)," in *Reference Module in Life Sciences*, ed. B. D. Roitberg (Amsterdam: Elsevier), 1–5.

Rodamilans, B., Valli, A., Mingot, A., San Leon, D., Lopez-Moya, J. J., and Garcia, J. A. (2018). An atypical RNA silencing suppression strategy provides a snapshot of the evolution of sweet potato-infecting potyviruses. *Sci. Rep.* 8:15937. doi: 10.1038/s41598-018-34358-y

Roth, B. M., Pruss, G. J., and Vance, V. B. (2004). Plant viral suppressors of RNA silencing. *Virus Res.* 102, 97–108. doi: 10.1016/j.virusres.2004.01.020

Roy, A., Stone, A., Otero-Colina, G., Wei, G., Choudhary, N., Achor, D., et al. (2013). Genome assembly of citrus leprosis virus nuclear type reveals a close association with orchid fleck virus. *Genome Announc.* 1:e00519-13. doi: 10.1128/genomeA.00519-13

Samuel, G. H., Wiley, M. R., Badawi, A., Adelman, Z. N., and Myles, K. M. (2016). Yellow fever virus capsid protein is a potent suppressor of RNA silencing that binds double-stranded RNA. *Proc. Natl. Acad. Sci. U.S.A.* 113, 13863–13868. doi: 10.1073/pnas.1600544113

Sanchez-Navarro, J., Miglino, R., Ragozzino, A., and Bol, J. F. (2001). Engineering of alfalfa mosaic virus RNA 3 into an expression vector. *Arch. Virol.* 146, 923–939. doi: 10.1007/s007050170125

Silhavy, D., Molnar, A., Lucioli, A., Szittya, G., Hornyik, C., Tavazza, M., et al. (2002). A viral protein suppresses RNA silencing and binds silencing-generated, 21- to 25-nucleotide double-stranded RNAs. *EMBO J.* 21, 3070–3080. doi: 10.1093/emboj/cdf312

Taschner, P. E., Van Der Kuyl, A. C., Neeleman, L., and Bol, J. F. (1991). Replication of an incomplete alfalfa mosaic virus genome in plants transformed with viral replicase genes. *Virology* 181, 445–450. doi: 10.1016/0042-6822(91)90876-d

Thomas, C. L., Leh, V., Lederer, C., and Maule, A. J. (2003). Turnip crinkle virus coat protein mediates suppression of RNA silencing in *Nicotiana benthamiana*. *Virology* 306, 33–41. doi: 10.1016/s0042-6822(02)00018-1

Van Dun, C. M., Van Vloten-Doting, L., and Bol, J. F. (1988). Expression of alfalfa mosaic virus cDNA1 and 2 in transgenic tobacco plants. *Virology* 163, 572–578. doi: 10.1016/0042-6822(88)90298-x

Vanitharani, R., Chellappan, P., Pita, J. S., and Fauquet, C. M. (2004). Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. *J. Virol.* 78, 9487–9498. doi: 10.1128/jvi.78.17.9487-9498.2004

Voinnet, O., Lederer, C., and Baulcombe, D. C. (2000). A viral movement protein prevents spread of the gene silencing signal in *Nicotiana benthamiana*. *Cell* 103, 157–167. doi: 10.1016/s0092-8674(00)00095-7

Voinnet, O., Pinto, Y. M., and Baulcombe, D. C. (1999). Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14147–14152. doi: 10.1073/pnas.96.24.14147

Voinnet, O., Vain, P., Angell, S., and Baulcombe, D. C. (1998). Systemic spread of sequence-specific transgene RNA degradation in plants is initiated by localized introduction of ectopic promoterless DNA. *Cell* 95, 177–187.

Vuorinen, A. L., Kelloniemi, J., and Valkonen, J. P. (2011). Why do viruses need phloem for systemic invasion of plants? *Plant Sci.* 181, 355–363. doi: 10.1016/j.plantsci.2011.06.008

Yaegashi, H., Isogai, M., and Yoshikawa, N. (2012). Characterization of plant virus-encoded gene silencing suppressors. *Methods Mol. Biol.* 894, 113–122. doi: 10.1007/978-1-61779-882-5_8

Yang, X., Ren, Y., Sun, S., Wang, D., Zhang, F., Li, D., et al. (2018). Identification of the potential virulence factors and RNA silencing suppressors of mulberry mosaic dwarf-associated geminivirus. *Viruses* 10:E472. doi: 10.3390/v10090472

Yelina, N. E., Savenkov, E. I., Solovyev, A. G., Morozov, S. Y., and Valkonen, J. P. (2002). Long-distance movement, virulence, and RNA silencing suppression controlled by a single protein in hordei- and potyviruses: complementary functions between virus families. *J. Virol.* 76, 12981–12991.

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 07 August 2020
doi: 10.3389/fpls.2020.01188

Plant Immune System Activation Upon Citrus Leprosis Virus C Infection Is Mimicked by the Ectopic Expression of the P61 Viral Protein

Gabriella D. Arena[1,2,3†], Pedro Luis Ramos-González[3†], Bryce W. Falk[4], Clare L. Casteel[5], Juliana Freitas-Astúa[3,6*] and Marcos A. Machado[1]*

[1] Laboratório de Biotecnologia, Centro de Citricultura Sylvio Moreira, Instituto Agronômico de Campinas, Cordeirópolis, Brazil, [2] Escola Superior de Agricultura Luiz de Queiroz (ESALQ), Universidade de São Paulo, Piracicaba, Brazil, [3] Laboratório de Biologia Molecular Aplicada, Instituto Biológico, São Paulo, Brazil, [4] Department of Plant Pathology, University of California, Davis, Davis, CA, United States, [5] School of Integrative Plant Science, Cornell University, Ithaca, NY, United States, [6] Laboratório de Virologia Vegetal, Embrapa Mandioca e Fruticultura, Cruz das Almas, Brazil

OPEN ACCESS

Edited by:
Jesus Navas-Castillo,
Institute of Subtropical and Mediterranean Horticulture
La Mayora, Spain

Reviewed by:
Francisco Tenllado,
Spanish National Research Council, Spain
Hernan Garcia-Ruiz,
University of Nebraska-Lincoln, United States

***Correspondence:**
Pedro Luis Ramos-González
plrg1970@gmail.com
Juliana Freitas-Astúa
juliana.astua@embrapa.br

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Plant Science

Received: 30 April 2020
Accepted: 22 July 2020
Published: 07 August 2020

Citation:
Arena GD, Ramos-González PL, Falk BW, Casteel CL, Freitas-Astúa J and Machado MA (2020) Plant Immune System Activation Upon Citrus Leprosis Virus C Infection Is Mimicked by the Ectopic Expression of the P61 Viral Protein. Front. Plant Sci. 11:1188. doi: 10.3389/fpls.2020.01188

Citrus leprosis virus C (CiLV-C, genus *Cilevirus*, family *Kitaviridae*) is an atypical virus that does not spread systemically in its plant hosts. Upon its inoculation by *Brevipalpus* mites, only localized lesions occur, and the infection remains limited to cells around mite feeding sites. Here, we aimed to gain insights into the putative causes of viral unfitness in plants by expanding the limited knowledge of the molecular mechanisms underlying plant/kitavirid interactions. Firstly, we quantified the CiLV-C viral RNAs during the infection in *Arabidopsis thaliana* plants using RT-qPCR and systematized it by defining three stages of distinguishing subgenomic and genomic RNA accumulation: i) 0–24 h after infestation, ii) 2–4 days after infestation (dai), and iii) 6–10 dai. Accordingly, the global plant response to CiLV-C infection was assessed by RNA-Seq at each period. Results indicated a progressive reprogramming of the plant transcriptome in parallel to the increasing viral loads. Gene ontology enrichment analysis revealed the induction of cell growth-related processes at the early stages of the infection and the triggering of the SA-mediated pathway, ROS burst and hypersensitive response (HR) at the presymptomatic stage. Conversely, infected plants downregulated JA/ET-mediated pathways and processes involved in the primary metabolism including photosynthesis. Marker genes of unfolded protein response were also induced, suggesting a contribution of the endoplasmic reticulum stress to the cell death caused by the viral infection. Finally, we transiently expressed CiLV-C proteins in *Nicotiana benthamiana* plants to undertake their roles in the elicited plant responses. Expression of the CiLV-C P61 protein consistently triggered ROS burst, upregulated SA- and HR-related genes, increased SA levels, reduced JA levels, and caused cell death. Mimicry of responses typically observed during CiLV-C–plant interaction indicates P61 as a putative viral effector causing the HR-like symptoms associated with the infection. Our data strengthen the hypothesis that symptoms of CiLV-C infection might be the outcome of a hypersensitive-like response during an incompatible interaction. Consequently, the locally restricted infection of CiLV-C,

commonly observed across infections by kitavirids, supports the thesis that these viruses, likely arising from an ancestral arthropod-infecting virus, are unable to fully circumvent plant defenses.

Keywords: *Cilevirus*, RNA-Seq, plant–virus interaction, hypersensitive response, salicylic acid, jasmonic acid, *Arabidopsis thaliana*, *Nicotiana benthamiana*

INTRODUCTION

Viruses that accomplish plant systemic infections replicate in the entry cell and use it as the source for local infections in contiguous cells, invade, and spread throughout the plant *via* the vascular system. Upon challenging a resistant host, viral multiplication and/or movement may be compromised. Differently from the majority of plant viruses, citrus leprosis virus C (CiLV-C, genus *Cilevirus*, family *Kitaviridae*) is unable to systemically infect any of its natural or experimental host species, even those belonging to distant plant families (Nunes et al., 2012; Arena et al., 2013; Garita et al., 2014). Invariably, CiLV-C remains restricted to cells around the vector-mediated inoculation sites, where symptoms of viral infection are chlorotic or necrotic spots (Bastianel et al., 2010). Despite the constraint in systemic infection for CiLV-C, it causes citrus leprosis, the most important viral disease affecting the citrus industry in Brazil, the world leader in sweet orange production. Annually, prevention and control of citrus leprosis cost approximately 50 million dollars, mainly for the chemical control of the viral vector, mites of the species *Brevipalpus yothersi* (Bassanezi et al., 2019). Endemic in the Americas, CiLV-C has spread throughout the main citrus-producing areas of the Latin American subregion (Ramos-González et al., 2018).

CiLV-C has two positive (+) sense single-stranded genomic RNA molecules with six open reading frames (ORFs). RNA1 (8,745 nts) harbors two ORFs encoding the RNA-dependent RNA polymerase (RdRp) and the putative coat protein (P29). RNA2 (4,986 nts) presents four ORFs encoding the putative movement protein (MP) and the P15, P61, and P24 proteins with unknown functions (Locali-Fabris et al., 2006; Pascon et al., 2006). While the *p15* ORF is considered an orphan gene with no homologs in any other viral species, *p61* and *p24* are taxonomically restricted ORFs also present in insect-infecting negeviruses and other nege-like viruses (Tautz and Domazet-Lošo, 2011; Kuchibhatla et al., 2014). P61 and P24 from kitaviruses and their related insect-infecting viruses show conserved structural features such as transmembrane domains and signal peptides (Kuchibhatla et al., 2014). CiLV-C RNA1 drives the transcription of one subgenomic RNA (sgRNA) of 0.7 kb for the expression of the *p29* gene, and RNA2 generates three coterminal sgRNAs of 3, 1.5, and 0.6 kb from where P61, MP, and P24 are translated, respectively (Pascon et al., 2006).

The reasons behind CiLV-C's inability to systemically infect plants are still unknown. There is speculation that the viral unfitness might be due to defective viral movement factors, an effective plant immune system that CiLV-C is unable to overcome, or even a combination of these and other ignored factors. Since a recent study revealed the functionality of cilevirus MP (Leastro et al., submitted), alternative hypotheses, such as those involving efficacious plant defenses preventing CiLV-C spread, gain strength. Plant innate defense mechanisms against pathogens are based on a two-layered immune system that uses cell surface receptors and intracellular plant resistance (R) proteins to respectively recognize pathogen-associated molecular patterns (PAMPs) or effectors (Cui et al., 2015; Couto and Zipfel, 2016; Garcia-Ruiz, 2019). Metabolic changes induced during plant defense can result in a burst of reactive oxygen species (ROS) that may culminate in a hypersensitive response (HR) (Xia et al., 2015). The transcriptional reprogramming resulting in the defense responses is mediated by the action of interconnected phytohormonal-dependent pathways and directed according to the nature of the injury. Typically, the salicylic acid (SA) pathway confers resistance to biotrophic pathogens and antagonizes the jasmonate/ethylene (JA/ET) pathways that in turn induce defenses against herbivores and necrotrophic pathogens (Pieterse et al., 2012). The SA pathway is involved in the activation of the HR, acting together with ROS molecules to trigger the resistance response (Xia et al., 2015). Furthermore, activation of the plant immune system and the induction of HR have been linked to endoplasmic reticulum (ER) stress (Poór et al., 2019). The excessive accumulation of proteins in the ER triggers the unfolded protein response (UPR), a mechanism that prevents the dangerous accumulation of unfolded proteins, but surpassed a threshold, UPR leads to a chronic stress condition that eventually triggers HR (Williams et al., 2014).

Several viral proteins have been identified as elicitors of the plant immune system. For instance, the P0 protein from poleroviruses elicits an HR that is associated with the *Nicotiana glutinosa* protein RPO1 (Resistance to Poleroviruses 1), a likely immune receptor of P0 (Wang et al., 2015). Likewise, the P38 protein from turnip crinkle virus (TCV) is recognized by the *Arabidopsis thaliana* R protein HRT (HR to TCV), which activates an HR-mediated resistance (Cooley et al., 2000; Pumplin and Voinnet, 2013). In a clear correlation between over-accumulation of viral proteins in the ER and HR, the transient expression of the P25 protein from potato virus X in *N. benthamiana* plants induces ER stress and UPR, leading to ER collapse and cell death (Aguilar et al., 2019). Besides triggering HR, viral proteins can modulate hormonal defense pathways to establish mutualism with its vector (Casteel and Falk, 2016). For example, the NIa-Pro (Nuclear inclusion a-protease) protein from turnip mosaic virus (TuMV) interferes with ET-mediated responses, resulting in defense suppression and, consequently,

the enhanced performance of its vector, the aphid *Myzus persicae* (Casteel et al., 2015). Viruses that depend on vectors to move from infected to healthy host plants use this strategy of decrease antiherbivory defense as an effective means to improve their transmissibility (Abe et al., 2012).

A previous study conducted in *A. thaliana* plants revealed that CiLV-C infection triggers ROS burst and cell death, induces the classical antiviral mechanisms of RNA silencing and SA pathway, suppresses the JA-dependent response, and favors the colonization of the mite vector. A preliminary model of the interaction using this information was depicted (Arena et al., 2016), but many underlying mechanisms of CiLV-C infection remain to be uncovered, for instance, the kinetics of viral accumulation in infected plants, the global plant response to the virus infection, and the viral effector that triggers such response. Here we aim to expand our knowledge on the molecular plant/CiLV-C interplay. Firstly, to describe the CiLV-C accumulation along the course of the infection, we quantify viral genomic and subgenomic RNAs using RT-qPCR. Then, to unravel novel mechanisms of plant response to the viral infection, we evaluate the transcriptome profile of infected *A. thaliana* plants by RNA-Seq. Finally, to test the hypothesis that specific CiLV-C proteins could trigger plant responses to the viral infection, the putative elicitor activity of each virus protein was assessed by expressing them in *Nicotiana benthamiana* plants. The current work contributes to the identification of mechanisms involved in the development of citrus leprosis disease and provides insights around the atypical restraint of the systemic movement of CiLV-C. In practical terms, we provide data supporting a comprehensive plant transcriptome analysis that can be further explored to unravel common or unique mechanisms of plant gene expression operated during the plant infection by kitavirids.

MATERIALS AND METHODS

Plant Material

Seeds from *A. thaliana* ecotype Columbia (Col-0) were obtained from the Arabidopsis Biological Resource Center (ABRC, http://www.arabidopsis.org). *A. thaliana* and *N. benthamiana* plants were grown in a controlled growth chamber Adaptis AR A1000 (Conviron, Winnipeg, Canada) at 23 ± 2°C and a 12 h photoperiod. Four-week-old plants were used in the experiments.

Mite Rearing

The population of mites was initially obtained from a single female collected from a citrus orchard and identified as *B. yothersi* using phase-contrast microscopy as reported elsewhere (Beard et al., 2015). Nonviruliferous mites were reared onto the unripe fruits of Persian lime (*Citrus latifolia* Tanaka), a genotype immune to CiLV-C. Viruliferous mites were obtained by transferring the nonviruliferous mites from the Persian lime to sweet orange fruits with citrus leprosis symptoms infected with CiLV-C strain SJP (Ramos-González et al., 2016). Fruits were prepared as described before (Rodrigues et al., 2007). Mites were reared for several generations and were evaluated for the presence of CiLV-C by RT-PCR (Locali et al., 2003) before their use in the experiments.

Kinetics of CiLV-C Accumulation Experiment

Quantification of CiLV-C RNA loads was performed in *A. thaliana* plants infested with *B. yothersi* viruliferous mites at ¼, ½, 1, 2, 4, 6, 8, and 10 dai. *A. thaliana* plants were infested with 15 mites (five per each of three rosette leaves), transferred with a brush under a stereoscopic microscope. Each time point had 10 biological replicates. Harvested leaves were flash-frozen in liquid N_2 and stored at −80°C until RNA extraction. Plant RNA was purified with the RNeasy Plant Mini Kit (Qiagen, Venlo, Netherlands); RNA concentration and purity (A_{260}/A_{280}) were determined in NanoDrop ND-8000 micro-spectrophotometer (Thermo Scientific, Waltham, USA), and cDNA was synthesized using RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Scientific, Waltham, MA, USA).

Absolute Quantification of CiLV-C *p29* and *RdRp* Genes

Absolute quantification of *p29* and *RdRp* was assessed by RT-qPCR using TaqMan assays. Reaction mixes were prepared with the TaqMan® Fast Universal PCR MasterMix 2× kit, as recommended by the manufacturer (Thermo Scientific, Waltham, MA, USA). Amplifications were carried out in a 7500 Fast Real-Time PCR System device (Thermo Scientific, Waltham, MA, USA). Samples were analyzed in triplicates and no-template controls were included to check for contaminations. *Cycle quantification* (Cq) values from infected samples were compared with the standard curves to determine absolute quantities of CiLV-C *p29* and *RdRp* molecules. Quantities of each gene at different time points were statistically compared using one-way ANOVA and Student's t-test ($\alpha \leq 0.05$).

Relative Quantification of CiLV-C Genes

Relative quantification of all CiLV-C ORFs (*p29*, *RdRp*, *p15*, *p61*, *p24*, and *MP*), the putative *p7* ORF, and the intergenic region (IR) was assessed by RT-qPCR using the GoTaq dsDNA binding dye. qPCR assays were prepared with 3 ng of cDNA, 6.5 ul of GoTaq qPCR Master Mix (Promega, Madison, WI, USA), and 120 nM of each gene-specific primer pair. Each cDNA sample was analyzed in duplicates, and melting curves were included. Primer pair efficiency (E) and Cq value were determined for each reaction using Real-time PCR Miner (Zhao and Fernald, 2005). The Cq value of each sample, expressed as the mean of the two technical replicates, was converted into relative quantities (RQs) using the function $RQ = E^{\Delta Cq}$, where ΔCq is the difference between the lowest Cq value across all samples for the evaluated gene and the Cq value of a given sample. Normalized-relative quantity (NRQ) of each sample was calculated as the ratio of the sample RQ and the reference gene (*A. thaliana SAND family protein* gene) RQ. Individual fold change values were determined by dividing the sample NRQ by the mean NRQ of samples of the calibrator, that is, plants collected at the time point with the

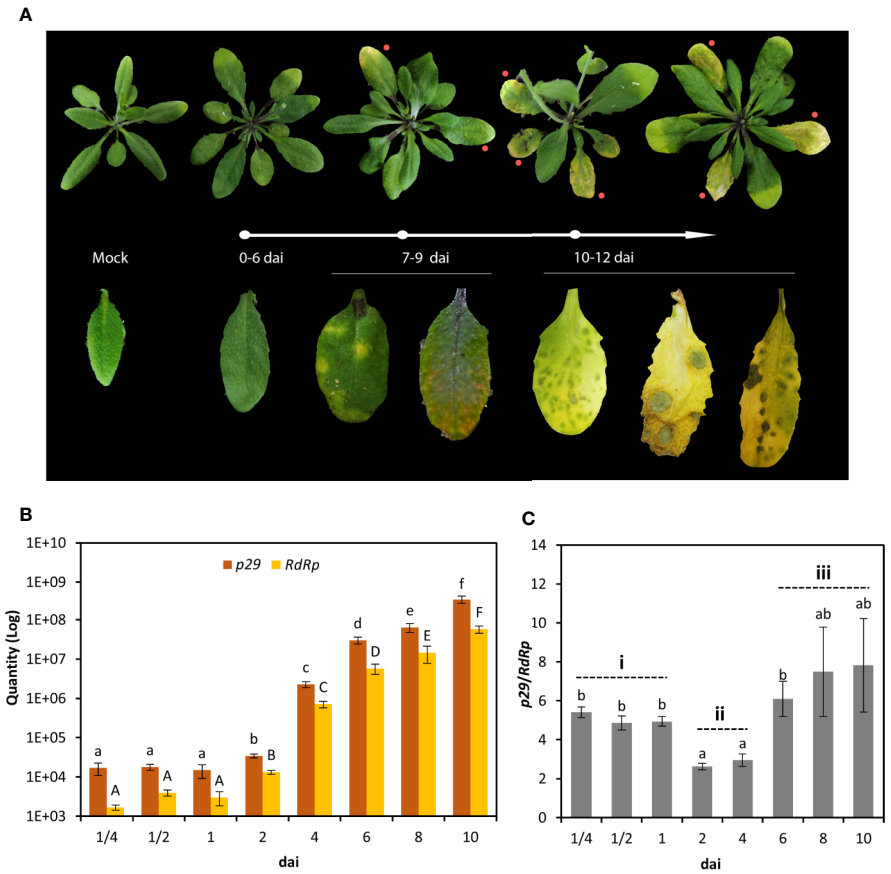


FIGURE 1 | CiLV-C infection in *Arabidopsis thaliana* plants. **(A)** Phenotypes of plants infested with CiLV-C viruliferous *B. yothersi* mites, from 0 to 12 days after infestation (dai). Red dots in the top panel indicate symptomatic leaves. From 0 to 6 dai, no symptoms are observed. Typical symptoms of CiLV-C infection in *A. thaliana* initially arise as chlorotic spots in green dark leaves (7–9 dai), evolving to green islands in yellow senescent ones (10–12 dai). Alternatively, larger areas of dead tissues frequently appear in infected plants (here, represented by the last leaf from the 10–12 dai group). No symptoms are observed in plants infested with nonviruliferous mites (mock). **(B)** Absolute quantification of CiLV-C *p29* and *RdRp* genes in eight time points after infestation with viruliferous *B. yothersi* mites. Different letters correspond to different copy numbers between the time points assessed (ANOVA and Student's t-test, $\alpha < 0.05$). **(C)** *p29/RdRp* ratio, calculated using the copy number from both molecules at each time point. Different letters correspond to different ratios between the time points assessed (ANOVA and Student's t-test, $\alpha < 0.05$). Three stages (i: 0–24 h after infestation, ii: 2–4 dai, iii: 6–10 dai) from CiLV-C infection are indicated.

lowest amount of the specific molecule. Quantities of each molecule at different time points were statistically compared using one-way ANOVA and Student's t-test ($\alpha \leq 0.05$).

RNA-Seq Time-Course Experiment

Arabidopsis thaliana plants were infested with 15 nonviruliferous or viruliferous mites (five mites per each of three rosette leaves). Infested leaves were collected at 6 h after infestation (hai) and 2 and 6 days after the infestation (dai). Sixteen plants were infested per treatment per time point, and leaves from two plants were pooled, totaling eight biological replicates. Another set of plants was kept with viruliferous mites for eight days, when symptoms were visible, to confirm virus infection. Harvested leaves were flash-frozen in liquid N_2 and stored at −80°C until RNA extraction. Plant RNA was purified with the RNeasy Plant Mini Kit (Qiagen, Venlo, Netherlands) and treated with RNAse-free DNAse (Qiagen, Venlo, Netherlands) to remoe plant DNA. RNA purity (A_{260}/A_{280} ~ 2.0) and integrity (RIN > 8) were confirmed in NanoDrop ND-8000 micro-spectrophotometer (Thermo Scientific, Waltham, USA) and Bioanalyzer 2100 (Agilent Technologies, Santa Clara, USA), respectively. CiLV-C presence in plants infested with viruliferous mites or its absence in those infested with nonviruliferous mites was confirmed by RT-PCR (Locali et al., 2003). RNA extracts from two samples (100 ng/ul each) were pooled, totaling four replicates per treatment (CiLV-C and mock) per time point for the RNA-Seq. cDNA libraries were prepared with Illumina TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, USA). Sequencing was performed in an Illumina HiSeq 2500 system (Illumina, San Diego, USA) using HiSeq SBS v4 High Output Kit (Illumina, San Diego, USA). Paired-end reads of 2 × 125 bp were generated.

Bioinformatics Analysis of RNA-Seq Data

RNA-Seq data were analyzed following the pipeline from Anders et al. (2013) with some modifications previously described (Arena et al., 2018). The biological variability of the samples

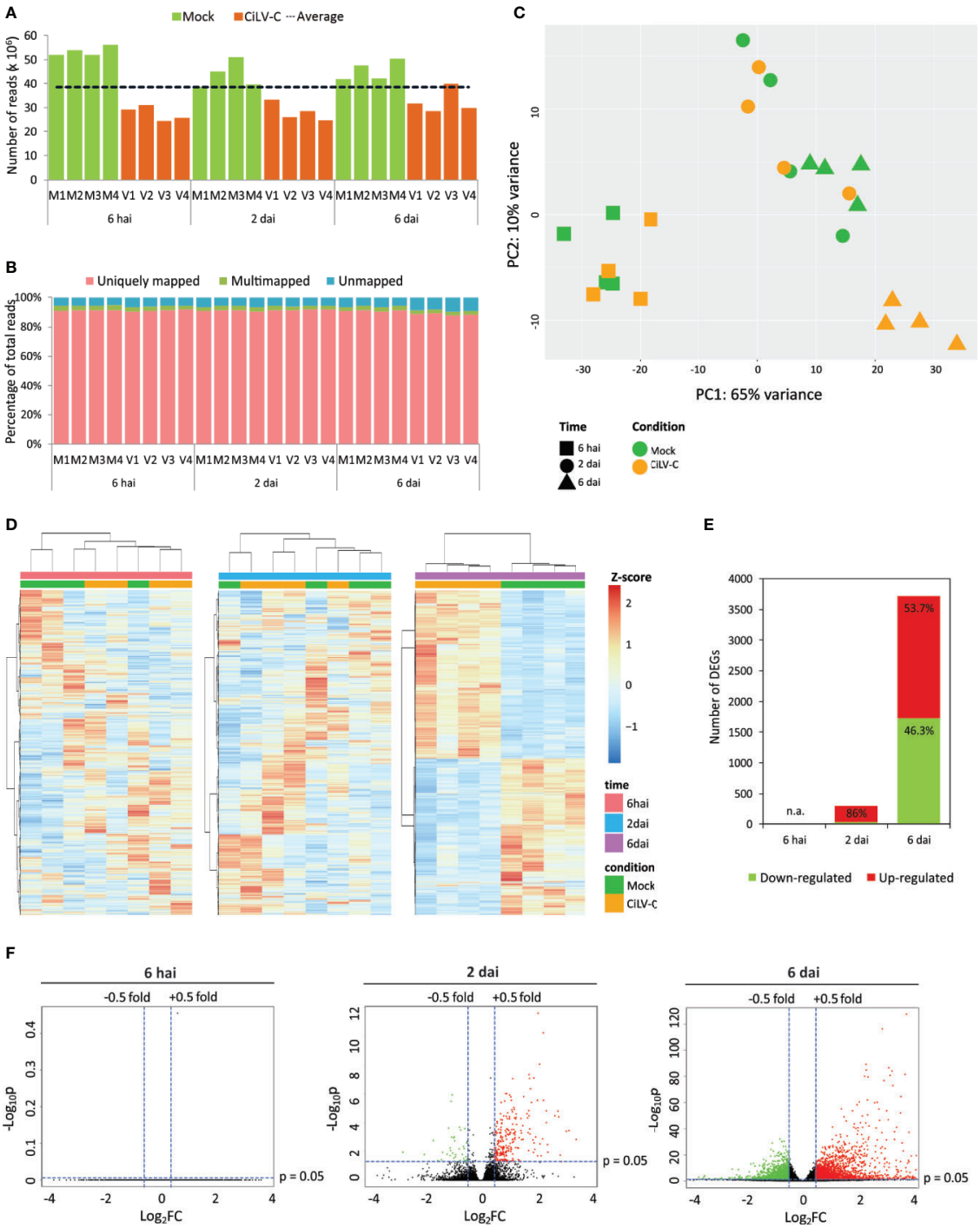


FIGURE 2 | Overview of *Arabidopsis thaliana* transcriptome upon CiLV-C infection. **(A)** Number of paired-end reads generated for each library by Illumina HiSeq sequencing. M, mock-infected (plants infested with nonviruliferous mites); V, virus-infected (plants infested with CiLV-C viruliferous mites). The dashed line represents the average of paired-end reads from all 24 libraries. **(B)** Proportion of uniquely mapped, multimapped, and unmapped reads obtained for each library. Reads were mapped in the *A. thaliana* (TAIR 10) genome using *TopHat2*. M, mock-infected; V, virus-infected plants. **(C)** Principal component analysis of normalized count data from all samples. **(D)** Hierarchical clustering analysis of normalized count data z-scores exhibited by differentially expressed genes (DEGs) of each sample within each time point. **(E)** Numbers of up- and downregulated DEGs in CiLV-C infected plants in comparison to mock controls at each time point. DEGs were identified using *DESeq2* and defined by $|\log_2 FC| \geq 0.5$ and false discovery rate (FDR)-corrected *p*-value ≤ 0.05. **(F)** Volcano-plots of $-\log_{10}p$ and $\log_2 FC$ exhibited by each gene in CiLV-C infected plants compared to mock controls at each time point. Up- and downregulated genes are presented in red and green, respectively. FC, fold-change; p, FDR-corrected *p*-value; hai, hours after infestation; dai, days after infestation.

was assessed by principal component analysis (PCA) and hierarchical clustering (using the *Euclidean* distance metric and *Ward's* clustering method). Differentially expressed genes (DEGs) between CiLV-C and mock treatments were identified at each time point using the package DESeq2 (Love et al., 2014). *False Discovery Rate* (FDR) correction for multiple comparisons was applied. DEGs with corrected p-values ≤ 0.05 and $|\log2FC| \geq 0.5$ were classified as differentially expressed. GO Enrichment Analysis was performed on DEGs to elucidate mechanisms potentially involved in the CiLV-C infection and symptoms development. A gene set was defined as all DEGs identified at each set (2 dai/upregulated, 2 dai/downregulated, 6 dai/upregulated, and 6 dai/downregulated), and the universe comprised all genes of the *A. thaliana* TAIR10 genome. Overrepresented Biological Processes (BPs) were identified by a hypergeometric test (FDR-adjusted p-values ≤ 0.001). GO networks were generated in Cytoscape using the app BinGO (Maere et al., 2005).

Identification of Enriched Transcription Factors

Enriched TFs were assessed on up and downregulated DEGs from 2 and 6 dai using two approaches. First, genes coding for TFs within DEGs were identified by searching on the PlantTFDB database (Jin et al., 2017) and overrepresented TF families on each set of genes were assessed using a hypergeometric test ($\alpha \leq 0.01$). Second, individual TFs with targets enriched within DEGs were identified using the TF enrichment tool (Jin et al., 2017), based on both the presence of *cis*-regulatory elements in the sequences of the DEGs assessed and literature mining. TFs with enriched targets were further grouped according to their families.

Validation of Gene Expression Data by RT-qPCR

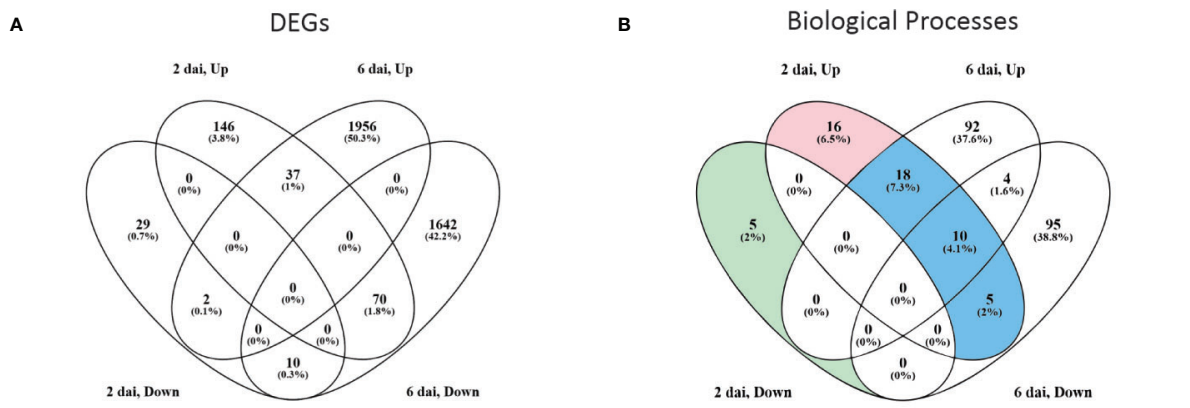
A new time-course experiment was set with *A. thaliana* Col-0 plants infested with viruliferous and nonviruliferous mites. Leaf samples were collected at 6 hai, 2 dai, and 6 dai. For each time point, plants were grouped in sets of 16 individuals assigned to each treatment (CiLV-C and mock). Plants were infested with 15 mites (five mites per each of three rosette leaves). Infested leaves were collected at each time point, and leaves from two plants were pooled, totaling eight biological replicates per treatment per time point. Leaf collection, RNA extraction, and quantification were carried out as previously indicated in this section. RNA quality was confirmed in 1.2% agarose gels. cDNA was generated using RevertAid H Minus First Strand cDNA Synthesis Kit as described by the manufacturer (Thermo Scientific, Waltham, MA, USA). qPCR assays were prepared with 3 ng of cDNA, 6.5 ul of GoTaq qPCR Master Mix (Promega, Madison, WI, USA), and 120 nM of each gene-specific primer pair (**Supplementary Table 8**). Each cDNA sample was analyzed in triplicate, and melting curves were included. Primer pairs' efficiency and Cq value were determined for each reaction using Real-time PCR Miner (Zhao and Fernald, 2005). Gene expression analyses were performed using the ΔCq model with efficiency correction and multiple reference genes (Hellemans et al., 2007) as previously described (Arena et al., 2016). The difference between infected and mock samples within each time point was assessed using Student's t-test ($\alpha \leq 0.05$).

Cloning of CiLV-C ORFs in Expression Vectors

CiLV-C ORFs were amplified from pregenerated clones using a HiFi polymerase (Thermo Scientific, Waltham, USA), and specific primers were designed to add restriction sites to the ends of each amplicon (**Supplementary Table 9**). The amplicons were digested and cloned in an intermediary vector based on the backbone of the pUC19 cloning vector (New England Biolabs, Ipswich, USA). Each transcriptional unit comprised the 35s cauliflower mosaic virus promoter driving constitutive expression of the viral ORF, the Ω fragment from TMV as a translational enhancer, the convenient CiLV-C ORF, and the *nopaline synthase* terminator. After assembly, each transcription unit was transferred from the intermediary vector to a pCambia 2300 binary vector (Marker Gene Technologies, Ipswich, USA). Final constructions were digested with the endonucleases *Xba*I and *Xho*I for validation of their identity. After the identification of the P61 elicitor activity (methods described below), new expression clones were assembled to express 3xFLAG C-terminal tagged P61 protein under the control of a dexamethasone (DEX)-inducible promoter. Somewhat similar cloning procedures were performed using the Gateway system (Thermo Scientific, Waltham, USA) to construct GFP-expressing clones as negative controls. Specific primers (**Supplementary Table 9**) were designed to include four Gs and the 25 bp of the attB regions in the 5′ end (for efficient Gateway cloning), the stop codon was removed (for fusion with the 3xFLAG), and nucleotides were added (to maintain the proper reading frame with the FLAG tag). Genes were amplified using the *Phusion High-Fidelity DNA Polymerase* (New England Biolabs Ipswich, USA). The amplicons were purified and cloned in the donor vector pDONR207 (Thermo Scientific, Waltham, USA). Entry clones were purified and confirmed by nucleotide sequencing. Transcriptional units were transferred by recombination to a Gateway-compatible version of the pTA7001 destination vector (Aoyama et al., 1997; Li et al., 2013), with a C-terminal 3xFLAG (DYKDHDGDYKDHDIDYKDDDDK) and a DEX-inducible promoter. Expression clones were purified and sequenced to confirm their identity. To generate the *p61* construct containing the frameshift mutant (p61Fs), the *p61* gene was amplified from the previously constructed vector using a specific forward primer (**Supplementary Table 9**) designed to introduce two extra nucleotides following the start codon. The amplicons were cloned in the pDONR207, and the transcriptional units were further transferred to the Gateway-compatible version of the pTA7001 destination vector. Expression clones were purified and confirmed by sequencing.

Transient Expression Assays in *Nicotiana benthamiana* Leaves

Plasmids containing the studied ORF were inserted into the *Agrobacterium tumefaciens* strain GV3101. Recombinant



C

2 dai, Up	p-value
cellular developmental process	3.50E-05
cell growth	3.50E-05
developmental growth involved in morphogenesis	9.14E-05
cell morphogenesis	9.14E-05
unidimensional cell growth	1.07E-04
cell wall organization	2.01E-04
cellular component morphogenesis	2.23E-04
growth	2.50E-04
cell differentiation	2.85E-04
developmental growth	3.65E-04

Cell growth

2 dai, Down	p-value
intracellular signal transduction	1.23E-07
phosphorelay signal transduction system	3.31E-07
ethylene-activated signaling pathway	3.90E-05
cellular response to ethylene pathway	5.77E-05
response to ethylene	5.45E-04

ET pathway

2 dai, Up x 6 dai, Up	p-value
response to chitin	2.12E-11
response to nitrogen compound	1.63E-09
response to acid chemical	8.24E-07
defense response	4.36E-06
regulation of response to stress	3.19E-05
regulation of defense response	3.50E-05
regulation of response to stimulus	3.50E-05
response to lipid	4.64E-05
cellular response to hypoxia	8.50E-05
cellular response to oxygen levels	1.07E-04

2 dai, Up x 6 dai, Down	p-value
response to hormone	4.74E-07
response to endogenous stimulus	5.16E-07
signaling	4.22E-09
cell communication	4.83E-08
response to stress	4.74E-07

Stress and Defense

FIGURE 3 | General transcriptomic changes and early responses of *Arabidopsis thaliana* plants affected by CiLV-C infection. **(A)** Venn diagram of up- and downregulated genes identified within the set of differentially expressed genes (DEGs) from each time point. DEGs were identified using *DESeq2* and defined by $|\log_2 FC| \geq 0.5$ and false discovery rate (FDR)-corrected *p*-value ≤ 0.05. The percentage value on each section of the diagram refers to the number of the corresponding DEGs relative to the total number of DEGs. **(B)** Venn diagram of overrepresented biological processes (BPs) from each set of up- and downregulated DEGs identified at each time point. Overrepresented BPs were identified based on a hypergeometric test with FDR-adjusted *p*-values ≤ 0.001. The percentage value on each section of the diagram refers to the number of the corresponding BPs relative to the total number of BPs. **(C)** Lists of overrepresented BPs exclusively modulated at 2 dai (transient early responses) or those common at 2 and 6 dai (stable early responses). The corresponding *p*-values obtained in the Gene Ontology (GO) enrichment analysis from 2 dai are included in the right column of each table. Up to ten BPs of each list are presented in each table. Complete lists of exclusive and common BPs are available in **Supplementary Table 3**. Due to the high number of BPs exclusively modulated at 6 dai, they were omitted from this figure and are presented in **Figure 4**. ET, ethylene.

A. tumefaciens was cultivated overnight in 5 ml of Luria-Bertani (LB) medium containing kanamycin (30 μg/ml) and rifampicin (50 μg/ml). Fresh media (Kn, Rif, and 150 μM acetosyringone) were inoculated with 1/10 (v/v) of the pre-inoculum until reaching a 0.8–1 OD_{600} nm. Cells were harvested by centrifugation and further incubated in the infiltration buffer (10 mM MES, pH 5.6, 10 mM $MgCl_2$, 150 μM acetosyringone) in the dark for 4 h. Infiltrations were carried out using a syringe in leaves of four- to six-leaf stage plants. In parallel, *N. benthamiana* plants were infiltrated with the empty vectors, GFP-expressing clones (negative controls), and the infiltration buffer (blank). When pTA7001 expression clones were used, infiltrated leaves were sprayed with 30 uM DEX and collected at 1, 2, and 3 days after induction. Cells were lysed in the presence of NP40-based buffer (150 mM NaCl, 1% NP40, 50 mM Tris-HCl), diluted in a protein disruption buffer (136 mM DTT, 192 mM Tris, 45 mg/ml SDS, 50 ug/ml bromophenol blue, 10 M urea), heated at 95°C, and loaded in SDS-PAGE gels. The fusion proteins were assessed by Western Blot probed with anti-DYKDDDDK-HRP conjugate (Miltenyi Biotec, Auburn, USA) at a concentration of 1:2,000 (**Figure 7F**). Leaves infiltrated with the construction for the expression of P61, empty vector, or the expression of GFP were collected for histochemical detection of H_2O_2, expression analysis of marker genes, and quantification of defense hormones. H_2O_2 was visualized by leaf staining with diaminobenzidine (DAB) as reported elsewhere (Ilarduya et al., 2003). The expression profiles of SA- and HR-related genes were assessed by RT-qPCR as described in the topic *Validation of Gene*

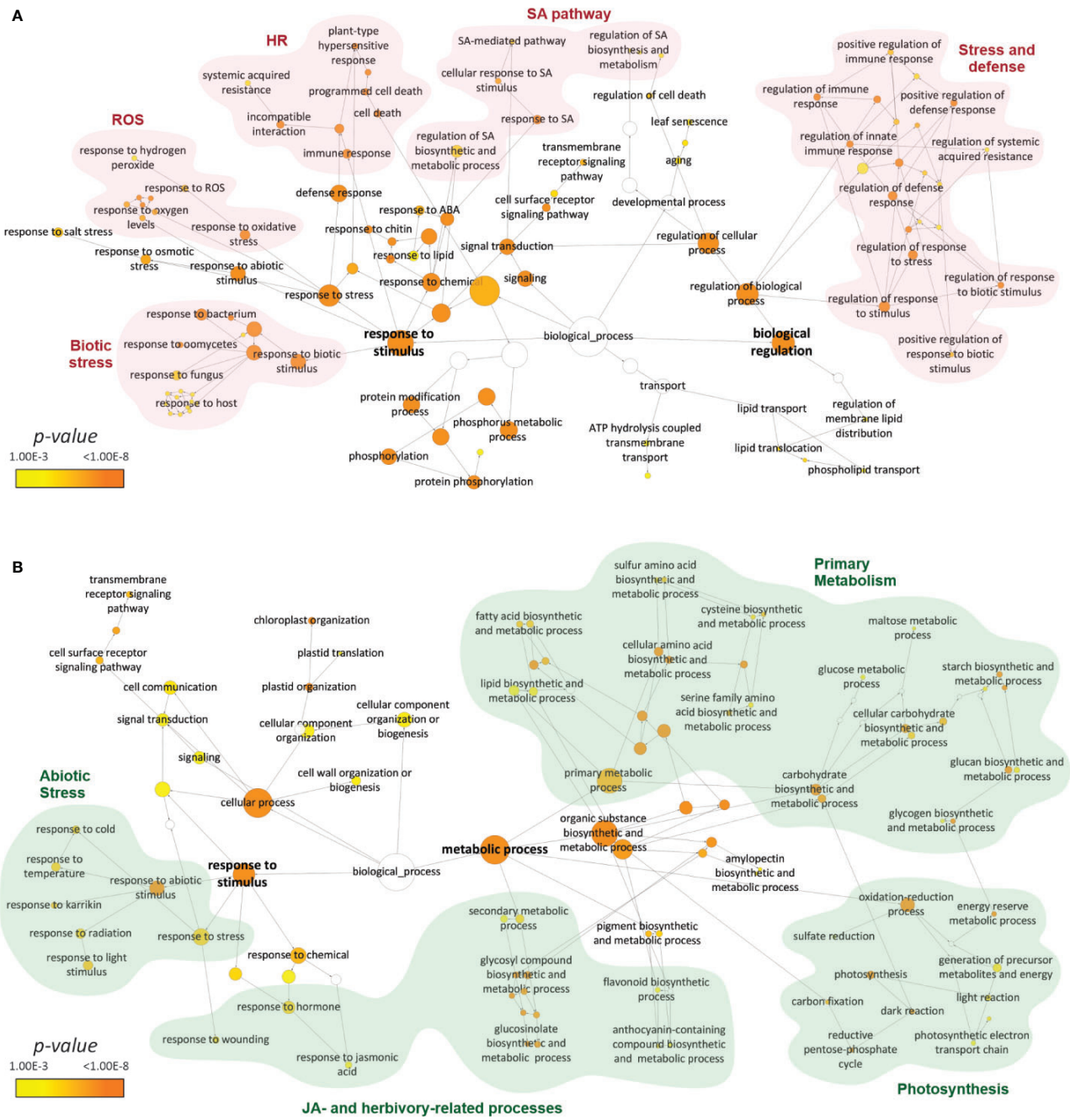


FIGURE 4 | Responses of *Arabidopsis thaliana* plants affected by CiLV-C infection at 6 dai. Induced and repressed responses are represented as networks of enriched BPs from upregulated **(A)** and downregulated **(B)** DEGs generated using the app BinGO in Cytoscape. The size of the nodes correlates with the number of DEGs, and their color reveals *p*-values of the enriched categories. Names of some BPs were simplified for clarity; full names are displayed in **Supplementary Table 3**. HR, hypersensitive response; ROS, reactive oxygen species; SA, salicylic acid; JA, jasmonic acid; ABA, abscisic acid.

Expression Data by RT-qPCR, with primer pairs described by Li et al. (2012). The SA and JA contents were quantified by LC-MS/MS as previously described (Arena et al., 2018). *N. benthamiana* plants were inoculated with CiLV-C, and symptoms were compared with those from plants agroinfiltrated with the plasmids containing the viral ORFs. Lesions on CiLV-C-infected sweet orange fruits were collected, ground in a mortar, and the sap was mechanically inoculated into carborundum-dusted leaves of four-week-old *N. benthamiana* plants.

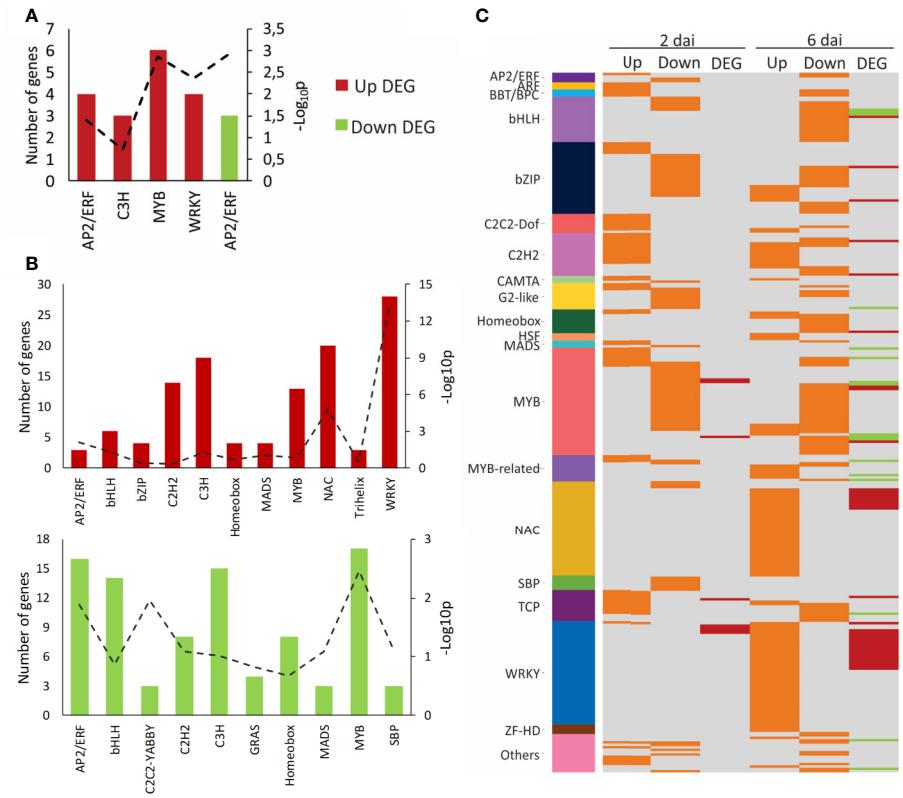


FIGURE 5 | Enriched transcription factors (TFs) and TF targets during CiLV-C infection in *Arabidopsis thaliana* plants. **(A, B)** Number of up- and downregulated genes coding for TFs from each TF family identified within the set of DEGs at 2 dai **(A)** or 6 dai **(B)**. Families encompassing two or fewer TFs were omitted. Up- and downregulated DEGs are presented in red and green, respectively. Levels of enrichment ($-\text{Log}_{10}\,p$, with p: *p*-value) of each family (hypergeometric test, $\alpha \leq 0.01$) are presented by a dashed line with its corresponding values in the secondary axis. **(C)** TFs with enriched targets within each set of up- and downregulated DEG at 2 and 6 dai identified by the TF enrichment tool. TFs are grouped according to their families. Each line identifies one TF. Orange lines correspond to TFs with enriched targets within each set. Red and green lines represent up- and downregulated DEGs, respectively, encoding TFs at each time point. Gray lines indicate the absence of enriched targets for a given TF- and/or TF not differentially expressed. Families encompassing two or fewer TFs were grouped in "Others". DEG, differentially expressed gene; dai, days after the infestation with viruliferous *Brevipalpus yothersi* mites.

RESULTS

Amounts of CiLV-C-Specific RNAs Continuously Increase Through the First Ten Days of Leaf Infection, a Period During Which Three Distinct Viral Accumulation Stages Are Distinguished

The kinetics of accumulation of CiLV-C-specific RNAs in *A. thaliana* infected leaves was evaluated by RT-qPCR. At first, the targets for TaqMan-based assays were two regions of the CiLV-C RNA1, one within the gene *p29*, which is present in the genomic and sgRNA and codes for the putative capsid protein, and another inside the ORF *RdRp* that is directly translated from the genomic RNA and codes for the viral replicase.

CiLV-C RNA loads were quantified during a time-course experiment in which leaf samples from *A. thaliana* plants were collected at ¼, ½, 1, 2, 4, 6, 8, and 10 days after infestation (dai) with viruliferous *B. yothersi* mites. Up until 6 dai, no symptoms were observed (**Figure 1A**). Symptoms of CiLV-C infection began appearing at 7 dai in 100% of the infested plants that were kept until the two latest time points. Typical symptoms of CiLV-C infection in Arabidopsis initially arise as chlorotic spots easily distinguished in green dark leaves, evolving to green islands in yellow senescent ones (**Figure 1A**) that may contain small patches of dead cells, as detected with Trypan Blue staining (Arena et al., 2016). Alternatively, areas of dead tissues with up to 5 mm in diameter appear in the infected leaves (**Figure 1A**). In both cases, these leaves undergo an accelerated senescence process leading them to death after 10–12 dai. Such a pattern of symptom development upon CiLV-C infection was previously reported in *A. thaliana* plants (Arena et al., 2013) and conforms to, but in a shorter temporal scale, those observed in CiLV-C-infected citrus species (Bastianel et al., 2010). Although not all the chlorotic spots give rise to necrotic ones, regardless of the final phenotype of the lesions, infection and symptoms are always restricted to the leaves infested by viruliferous *Brevipalpus* mites, and they are detected neither in systemic leaves nor in plants infested by nonviruliferous mites (**Figure 1A**).

Absolute quantities of *p29* and *RdRp* containing-RNA molecules were determined using the TaqMan assays and suitable standard curves (**Figure 1B**). All samples were positive in the analyses of both targets, confirming CiLV-C infection. The levels of both targets were kept invariable during the first 24 h after the infestation (hai), but afterward, they increased continuously until the last time point [Tukey's *honest significant difference* (HSD) test, $\alpha \leq 0.05$]. The highest difference between the evaluated sequential time points was obtained from 2 to 4 dai when RNA molecules containing *p29* and *RdRp* increased 66- and 53-fold, respectively (**Figure 1B**).

CiLV-C replication analysis was expanded to cover the RNA2 viral molecule by using intercalating dye-based RT-qPCR assays (**Supplementary Figure 1**). PCRs targeted sequences within ORFs *p15*, *p61*, *p24, MP, p7* (putative small ORF downstream of *p15*), and the intergenic region (IR) (Locali-Fabris et al., 2006). Additionally, specific primers for ORFs *p29* and *RdRp* were also included in this analysis to compare the data with those obtained in the TaqMan-based assays. Even though the absolute amounts of the molecules were not quantified in the intercalating dye-based RT-qPCR assays, they tended to follow the same pattern described by *p29* and *RdRp* in the TaqMan-based assays. Genes *p15*, *p61*, *p24*, and *MP* were mostly invariable within the first 24 h and continuously increased onwards (**Supplementary Figure 1**).

The ratio *p29*/*RdRp* at each time point was calculated as an indicator of the accumulation of viral subgenomic and genomic RNA (**Figure 1C**). The number of molecules containing the ORF *p29* was higher than those containing the ORF *RdRp* across the whole experiment (**Figures 1B, C**). Higher accumulation of *p29* was expected as the assay detects both genomic (also including the antigenomic) molecules and the *p29* sgRNA, while *RdRp* assay quantifies only the genomic RNA1 molecule. Within 2–4 dai, the ratio *p29*/*RdRp* reached the lowest level, whereas before and after this period it showed similar levels (**Figure 1C**). Despite the asynchronism of the viral replication process over the infected cells, a heuristic approach of the experimental data allowed us to subdivide the CiLV-C accumulation in *A. thaliana* into three main kinetic steps: i) 0–24 h after infestation (hai), ii) 2–4 dai, and iii) 6–10 dai (**Figure 1C**). Considering these steps and pathobiology features inherent to the citrus leprosis disease, time points for the evaluation of the plant response to the viral infection were further selected, *i.e.*, 6 hai, minimum inoculation access period required by viruliferous mites to obtain 100% of infected plants (Arena et al., 2016); 2 dai, first significant increase in the viral genome accumulation; and 6 dai, the presymptomatic stage when plant transcriptional responses likely take place culminating in the disease phenotype. Moreover, time points 2 and 6 dai epitomize a condition in which the ratio *p29*/*RdRp* significantly differs from the precedent kinetic steps.

CiLV-C Infection Triggers Significant Transcriptome Changes Proportionally to the Increase of Viral RNA Loads

The global transcriptomic response of *A. thaliana* plants along the course of CiLV-C infection was assessed by RNA-Seq. Plants infested with CiLV-C-viruliferous mites were compared with those infested with nonviruliferous ones (mock) at 6 hai, 2 dai, and 6 dai. Illumina sequencing generated roughly 924 million paired-end reads, with an average of 38.5 million per library and a higher number of reads from the mock treatment (**Supplementary Table 1**, **Figure 2A**). Overall, 93.5% of the reads aligned to the *A. thaliana* reference genome, with a 90.6% average of uniquely mapped reads (**Supplementary Table 1**, **Figure 2B**). The CiLV-C infected samples from 6 dai had the highest percentage of unmapped reads probably due to the higher virus titer and consequently a higher number of reads mapping to the virus genome (**Figure 2B**).

The main sources of variability within samples were assessed by principal component analysis (PCA) using the normalized count data (**Figure 2C**). The first component, which accounts for 64% of the variance, separated the samples by both variables: time after infestation and virus treatment, and they reflected the intensity of stimuli. Mock samples from different time points grouped separately, most likely due to the differential expression associated with longer mite feeding periods, as previously described (Arena et al., 2018). At 6 hai, where the lowest viral RNA loads were detected, all samples grouped regardless of the virus presence. A single group comprising both infected and mock samples was obtained for samples collected at 2 dai, where the viral RNA loads are slightly higher than at 6 hai. Differences between the expression profiles from infected and mock samples might be masked by the massive transcriptome changes in response to the mite action (Arena et al., 2018). At 6 dai, where the highest RNA virus loads were reached, infected and mock treatment formed two separated groups. The hierarchical clustering of samples within each time point confirmed the clusterization profile obtained by PCA (**Figure 2D**).

Differentially expressed genes (DEGs) in virus-infected plants compared with mock-inoculated ones were assessed within each time point using the negative binomial-based DESeq2 package and *False Discovery Rate* (FDR)-correction of *p*-values for multiple comparisons. Overall, 3,892 DEGs [$\alpha \leq 0.05$, $|\log2 \text{ fold change (FC)}| \geq 0.5$] were detected (**Supplementary Table 2**). No gene was differentially expressed at 6 hai (**Figures 2E, F**), which agrees with the similar expression profiles displayed by mock and virus-infected plants (**Figures 2C, D**). The number of DEGs progressively raised along the course of the infection (**Figures 2E, F**). At 2 dai, 294 DEGs were detected, of which, the majority (253 DEGs, $\cong$ 86%) were upregulated (**Figure 2E**). The largest number of DEGs throughout the interaction was detected at 6 dai, when CiLV-C infection deregulated 3,717 genes, evenly distributed in 1,995 ($\cong$ 53.7%) up- and 1,722 ($\cong$ 46.3%) downregulated DEGs (**Figure 2E**). This corresponds to more than 11% of all 33,602 *A. thaliana* genes being differentially expressed in response to CiLV-C at this particular time point. The analysis performed here shows that CiLV-C infection triggers a significant reprogramming on infected plants likely mirroring the course of viral replication.

CiLV-C Infection Induces Cell Growth and HR-Related Processes and Represses Both the Plant Primary Metabolism and the JA/ET-Mediated Responses

Gene ontology (GO) enrichment analyses were performed with the independent sets of up- and downregulated DEGs from each

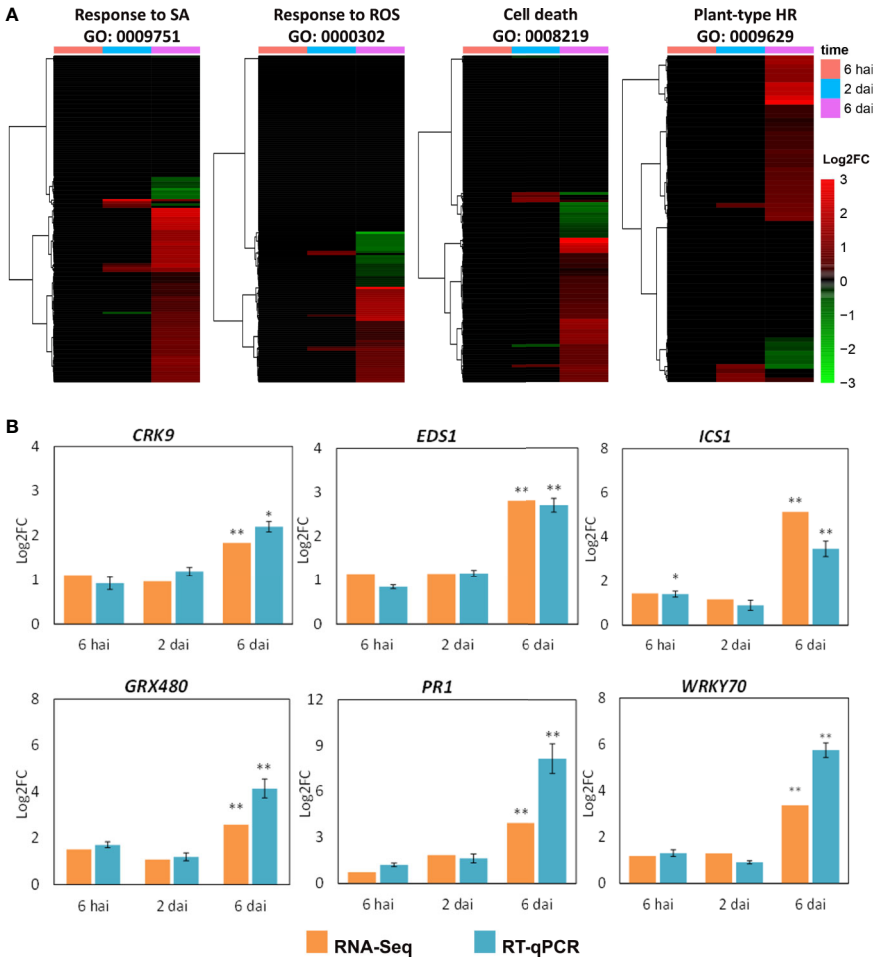


FIGURE 6 | Hypersensitive response (HR)-related genes during plant infection by CiLV-C. **(A)** Hierarchical clustering analysis of the whole set of genes of the *Arabidopsis thaliana* genome assigned to the Biological Processes "Response to SA", "Response to ROS", "Cell death", and "Plant-type HR". **(B)** Expression profile of selected *Arabidopsis thaliana* genes in virus-infected plants, quantified by RNA-Seq and RT-qPCR. Data are presented as log_2FC values in comparison with mock-infected plants (with log_2FC set to zero). Statistically significant differences of virus-infected versus mock control at *p*-values ≤ 0.01 (**) and ≤ 0.05 (*) are indicated. Hai, hours after infestation; dai, days after infestation; GO, Gene Ontology term; SA, salicylic acid; ROS, reactive oxygen species; HR, hypersensitive response; FC, fold change.

time point to identify the most relevant biological processes (BPs) disturbed during CiLV-C infection (**Supplementary Table 3**). DEGs and BPs that were either shared or exclusive to the experimental sets are presented (**Figures 3A, B**). Using the Cytoscape app BinGO, the enriched BPs were visualized as networks where the color and size of each node identify the *p*-value and number of DEG of each ontology, respectively.

The GO enrichment analysis revealed 49 and 5 overrepresented BPs (hypergeometric test, $\alpha \leq 0.001$) in the sets of DEGs that were up- and downregulated at 2 dai, respectively. Even though most of the DEGs identified at 2 dai were exclusively induced at this time point (146 DEGs, **Figure 3A**), the majority of enriched BPs obtained from the set of upregulated genes at 2 dai overlapped between the induced sets of 2 and 6 dai (28 BPs, **Figure 3B**). This suggests that several processes triggered soon at 2 dai are still modulated a few days later, although with a different number of exclusive and shared DEGs, perhaps reflecting the occurrence of early and late responses of the same process. BPs enriched in both upregulated DEGs from 2 and 6 dai (stable early responses) included general terms of plant response to stimuli such as "defense response", "regulation of response to stress", and "regulation of response to stimulus" (**Supplementary Table 3**, **Figure 3C**). Likewise, enriched BPs that were upregulated at 2 dai and downregulated at 6 dai mainly referred to broad ontologies such as "response to hormone", "response to endogenous stimulus", and "signaling" (**Supplementary Table 3**, **Figure 3C**). On the other hand, 16 and 5 BPs were exclusive to the upregulated and downregulated DEGs, respectively, at 2 dai (transient early responses) (**Figure 3B**). The BPs uniquely induced at 2 dai were predominantly related to the

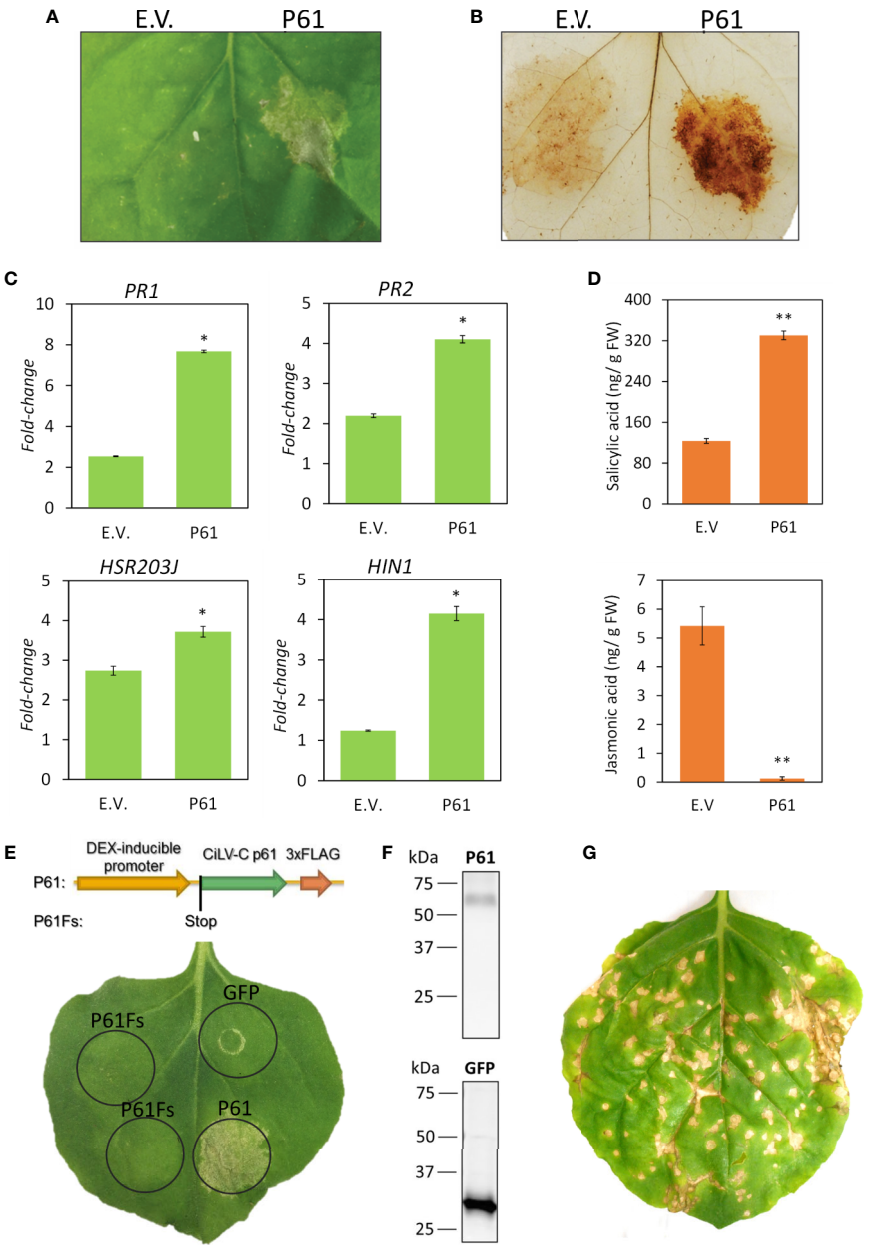


FIGURE 7 | Elicitor activity of CiLV-C P61 protein. The viral protein was transiently expressed in *Nicotiana benthamiana* leaves using *Agrobacterium*-mediated infiltration. HR-like phenotype **(A)**, ROS production **(B)**, expression of SA- and HR-related genes **(C)**, and levels of jasmonic acid and salicylic acid **(D)** were assessed at 3 days after the infiltration. E.V., empty vector. Statistically significant differences of plants infiltrated with p61 *versus* empty vector (E.V.) at *p*-values ≤ 0.01 (**) and ≤ 0.05 (*) are indicated. To verify the elicitor activity of the *p61* RNA *versus* P61 protein, a frameshift mutant (P61Fs) was produced **(E)**. P61 or GFP 3xFLAG fusion proteins were detected in *N. benthamiana* leaf extracts by Western Blot probed with an antibody for the FLAG tag conjugated with HRP (anti-DYKDDDDK-HRP); expected sizes were 30 kDa for GFP-3xFLAG and 63 kDa for p61-3xFLAG **(F)**. Symptoms of CiLV-C in *N. benthamiana*, obtained by mechanical inoculation, are presented to highlight the HR-like phenotype **(G)**.

cellular growth, *e.g.* "cell growth", "cellular developmental process", "cell wall organization", "cell morphogenesis", and "cell differentiation" (**Supplementary Table 3**, **Figure 3C**). Of the five BPs only repressed at 2 dai, three were associated with the ethylene pathway: "ethylene-activated signaling pathway", "cellular response to ethylene stimulus", and "response to ethylene" (**Supplementary Table 3**, **Figure 3C**).

Most of the detected BPs were overrepresented in those data sets with a higher number of genes, *i.e.* the ones modulated at 6 dai (**Figure 3A**). GO enrichment analysis disclosed 124 and 114

enriched BPs (hypergeometric test, $\alpha \leq 0.001$) in the groups of DEGs that were up- and downregulated, respectively, at 6 dai (**Supplementary Table 3**). Besides a few broad-term processes common to the DEGs induced at 2 dai, the vast majority of the BPs enriched at 6 dai were exclusively detected at that time point (late responses) (**Figure 3B**). Within these categories, only four were shared between the up- and downregulated-clusters corresponding to 6 dai, revealing that induced and repressed genes at that time point are mostly involved in different processes, and distinct pathways are differentially modulated in the presymptomatic stage (**Figure 3B**).

The cluster of upregulated DEGs at 6 dai was enriched in 92 exclusive categories (**Figure 3B**) and revealed a massive modulation of the plant immune system (**Figure 4A**, **Supplementary Table 3**). BP categories were mainly clustered in two groups comprising "response to stimulus" and "biological regulation" (**Figure 4A**). Processes associated with response to stress and defense preponderantly represented both groups. Particularly, the group centralized in "response to stimulus" was branched in stress-related nodes that included "response to biotic stimulus" (linked to the subcategories of response to bacteria, fungus, oomycetes, and host defenses), "response to abiotic stimulus" (represented by subcategories of response to osmotic stress and oxygen levels), and "response to oxidative stress" (specified from "response to ROS" to "response to hydrogen peroxide") (**Figure 4A**). A defense-related branch from response to stimulus group displayed general ontologies (*e.g.* "immune response" and "defense response"), and it was typified by HR-related BPs such as "plant-type HR", "defense response, incompatible interaction", "systemic acquired resistance", and "programmed cell death" (**Figure 4A**). BP group centralized in biological regulation was branched to a major subgroup comprising the ontologies related to the regulation of both responses to stress and defense (**Figure 4A**). Categories from that subgroup included general terms such as "regulation of defense response" and "regulation of response to stress" and others more specific, *e.g.* "positive regulation of response to biotic stimulus" and "regulation of systemic acquired resistance". Another branch from the biological regulation group ("regulation of cellular processes") displayed cell death- and SA-related responses including "regulation of cell death" and "regulation of SA biosynthetic and metabolic process", respectively (**Figure 4A**). A small cluster associated with the senescence process was also present in the upregulated network ("aging", "plant organ senescence", and "leaf senescence") (**Figure 4A**). Finally, the major hormonal-mediated pathway enriched in the upregulated network was the SA pathway, represented by the categories "response to SA", "cellular response to SA stimulus", and "SA mediated pathway" (**Figure 4A**).

The cluster of downregulated DEGs at 6 dai was enriched in 95 unique categories (**Figure 3B**). Most of the GOs clustered in a major group of metabolic processes harboring mainly the primary metabolism (**Figure 4B**). That subgroup included BPs associated with the metabolism of: *i*) lipids, such as "lipid biosynthetic and metabolic process" and "fatty acid biosynthetic and metabolic process"; *ii*) amino acids, whose categories included "sulfur", "cysteine", and "serine" amino acid biosynthetic and metabolic process; and *iii*) carbohydrate, with numerous broad terms (*e.g.* "cellular carbohydrate biosynthetic and metabolic process") and specific BPs associated with biosynthesis and metabolism of glucan, starch, glycogen, and maltose. Carbohydrate-related processes were connected to a cluster of photosynthesis-related categories such as "photosynthesis, light and dark reaction", "carbon fixation", and "generation of precursor metabolites and energy". Secondary metabolism formed a small branch comprising BPs directed to the biosynthesis and metabolism of glucosinolates and anthocyanins (**Figure 4B**), metabolites typically induced by JA during plant–arthropod interactions. Following the metabolism group, the downregulated GO network gathered a set of BPs, which along with many general terms shared with the upregulated network, included the response to distinct abiotic stimuli ("light", "radiation", and "temperature"), "response to wounding", and JA as the only enriched hormonal pathway within the downregulated processes. Another small group from the network was centralized in the "cellular component organization or biogenesis", with ontologies related to chloroplast and cell wall organization/biogenesis (**Figure 4B**).

Overall, the GO enrichment analysis showed that early plant responses to CiLV-C infection involve a transient induction of cell growth-related processes, transient repression of ET-responsive genes, and a stable modulation of defense and stress-related responses that kept up with the infection. At the presymptomatic stage, infected plants trigger processes related to the SA-mediated pathway, response to ROS and HR, all of which are present during incompatible interactions. Conversely, at the same stage, infected plants downregulate processes involved in the primary metabolism, JA-mediated pathway, and synthesis of glucosinolates.

Regulation of Coexpressed Genes by Specific Classes of Transcription Factors (TFs) Correlates With Modulation of Stress Defense Responses

To unravel the regulation of the transcriptional reprogramming upon viral infection, the classes of TFs associated with coexpressed DEGs were identified. First, we identified the up- and downregulated DEGs coding for TFs on each time point and their corresponding families (**Figures 5A, B**, **Supplementary Table 4**). The overrepresentation of specific families from each coexpressed set was assessed with a hypergeometric test ($\alpha \leq 0.01$). Within the set of upregulated DEGs at 2 dai, 29 TFs from 15 different families were identified. From that group, only two families were overrepresented: MYB (six genes, p-value = 1.39E-03) and WRKY (four genes, p-value = 4.46E-03) (**Figure 5A**), both typically involved in plant defense responses to stresses (Dubos et al., 2010; Phukan et al., 2016). From the downregulated genes at the same time point, only six TFs comprising three different families were detected, as expected due to the reduced number of DEGs with such expression patterns. The only overrepresented family was AP2/ERF (three genes, p-value = 1.16E-03) (**Figure 5A**), whose members are known to act as regulators of the ERF-branch of the JA/ET-mediated pathway (Pieterse et al., 2012). At 6 dai, 134 TFs from 22 families were upregulated. Only three of those families were

overrepresented, of which the largest and most significant were WRKY (28 genes, *p*-value = 3.67E-14) and NAC (20 genes, *p*-value = 1.59E-05) (**Figure 5B**). Similar to WRKY, NAC TFs are also intimately associated with immune responses and specifically to increased resistance against pathogens, including the triggering of HR against viruses (Olsen et al., 2005; Nuruzzaman et al., 2013). Within the set of downregulated DEGs at 6 dai, 121 TFs evenly distributed in 31 families were identified. Similar to the upregulated set, only three classes of TFs were enriched, and the largest and most significant ones were MYB (17 genes, *p*-value = 3.50E-03) and AP2/ERF (16 genes, *p*-value = 1.31E-02) (**Figure 5B**).

In another approach, we searched for TFs with overrepresented targets within each set of DEGs by using the TF enrichment tool (Jin et al., 2017). Potential targets were identified based on cis-regulatory elements in the promoters of the test genes and regulatory interactions described in the literature (Jin et al., 2017) (**Figure 5C**, **Supplementary Table 5**). The largest families with potential targets within DEGs induced at 2 dai were the growth-related TCP (Manassero et al., 2013) and stress-related C2H2 (Kiełbowicz-Matuk, 2012), represented by 10 and 13 TFs, respectively. On the other hand, the families MYB and bZIP—with 29 and 18 TFs, respectively—were the ones that presented the highest numbers of TFs with potential targets in the set of downregulated genes at 2 dai (**Figure 5C**). Interestingly, the analysis of 6 dai sets revealed once again WRKY (43 TFs) and NAC (37 TFs) as the largest families with targets within upregulated DEGs (**Figure 5C**), supporting the involvement of both TF classes in controlling the induction of those genes. Within the downregulated genes at 6 dai, potential targets were mainly associated with TFs from bHLH (**Figure 5C**), which includes the regulators of the MYC-branch of the JA/ABA-mediated pathway (Pieterse et al., 2012) and MYB classes, represented by 33 and 17 genes, respectively. Neither members of WRKY and NAC families had targets enriched in the downregulated DEGs from 6 dai nor TFs from the bHLH family presented potential targets within the upregulated genes at the same time point (**Figure 5C**). This data stresses the specificity of induced and repressed responses during the presymptomatic stage.

Our analyses showed that the expression of upregulated genes is mainly driven by TF of the classes WRKY and NAC, followed by those of the TCP and C2H2, while downregulated genes are potentially controlled by TFs of the AP2/ERF, bHLH, and bZIP families. MYB TFs regulate both induced and repressed responses. Notably, WRKY is the only family with both modulated TFs and target genes exclusively within the groups of upregulated transcripts. Results support the modulation of stress responses upon CiLV-C infection, being consistent with the specific induction of cell growth and HR/SA-mediated defenses, and repression of both ET and ABA branches of the JA pathway.

Genes Related to HR Are Induced at the Presymptomatic Stage

Due to the overrepresentation of HR-related processes within the upregulated GO network and the HR-like phenotype induced by CiLV-C infection (Arena et al., 2016), the DEGs associated with either HR or the mechanisms underlying the development of such resistance response were thoroughly reviewed. Hierarchical clusters were generated with data from all *A. thaliana* genes included in the categories "response to SA", "response to ROS", "cell death", and "plant-type HR" (**Figure 6A**). Within the sets of DEGs assigned to those categories ($\alpha \leq 0.05$, $|\log2FC| \geq 0.5$), 63, 34, 36, and 23 genes, respectively, were upregulated at 6 dai (**Supplementary Table 6**). The induction of several genes directly or indirectly related to HR at the presymptomatic stage provides additional evidence supporting the hypothesis that symptoms of CiLV-C infection result from an HR-like resistance.

To validate the RNA-Seq data and support the involvement of the SA pathway and HR in response to CiLV-C infection, the expression of selected DEGs was assessed by RT-qPCR (**Figure 6B**). Six SA- and HR-related genes upregulated at the presymptomatic stage were selected: the signaling component *EDS1* (*enhanced disease susceptibility 1*), the SA biosynthetic enzyme *ICS1* (*isochorismate synthase 1*), the regulator *GRX480* (*glutaredoxin 480*), the receptor-like kinase (RLK) *CRK9* (*cysteine-rich RLK 9*), the transcription factor *WRKY70* (*WRKY DNA-binding protein 70*), and the defense protein *PR1* (*pathogenesis-related protein 1*). Expression profiles of those genes were assessed in a new, independent, experiment with plants infested with nonviruliferous (mock) and CiLV-C viruliferous mites at 6 hai, 2 and 6 dai (**Figure 6B**). All the evaluated genes were induced at 6 dai, in line with the RNA-Seq data, supporting the results described in this work.

Genes Involved in ER Stress and Unfolded Protein Response (UPR) Are Upregulated in CiLV-C-Infected Plants

Subcellular localization of CiLV-C proteins revealed that P15, P61, and P24 accumulate in association with the endoplasmic reticulum (ER) membranes, inducing disruption of the ER network (Leastro et al., 2018). To investigate whether the CiLV-C infection triggers ER stress, which could potentially induce HR-like response (Williams et al., 2014), we verified the expression levels of genes related to the ER stress and UPR activity. Genes assigned with the GO categories "endoplasmic reticulum unfolded protein response" and "response to endoplasmic reticulum stress" were reviewed. Nine and 17 genes included in each category, respectively, were differentially expressed, including the transcription factor *bZIP60* and the chaperones *ER luminal binding proteins* (BiP) and *calreticulins* (CRT). Without exception, all DEGs from both categories were upregulated at 6 dai (**Supplementary Table 7**). Even though the number of DEGs related to ER stress and UPR was not large enough to classify both GO terms as enriched ones, the upregulation of all DEGs at the presymptomatic stage suggests the involvement of these processes in the transcriptional changes that culminate in the development of the disease symptoms.

Expression of CiLV-C P61 Protein Triggers a Hypersensitive-Like Response and Mimics Plant Responses to Viral Infection

The transcriptome analysis revealed that plant response to CiLV-C infection is spearheaded by the activation of the plant immune

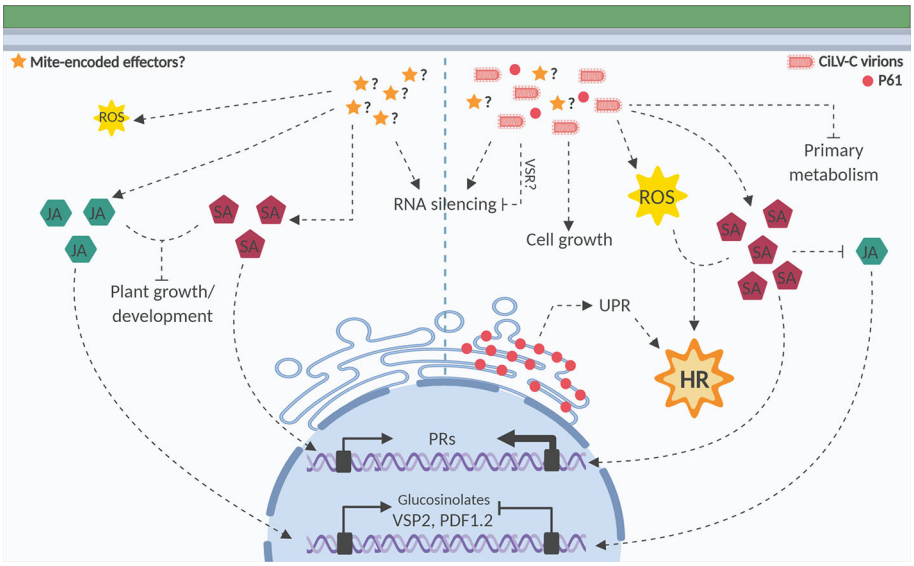


FIGURE 8 | Model representing the interaction of *Arabidopsis thaliana* plants/*Brevipalpus* mites/CiLV-C. The current model integrates the experimental facts obtained in this work with those revealed from previous assays of interaction between *A. thaliana* plants and *B. yothersi* mites (Arena et al., 2016; Arena et al., 2018), and CiLV-C (Arena et al., 2016). Left and right halves of the diagram show hallmarks during the plant interaction with nonviruliferous and viruliferous mites, respectively. Left: During feeding, *Brevipalpus* mites use their stylets to pierce mesophyll cells and inject saliva that might contain mite-encoded effectors. In response to the interaction with mites, salicylic acid (SA) is accumulated and induces SA-dependent genes (*e.g.*: *pathogenesis-related proteins*, PRs), which benefit mite performance. In parallel, jasmonic acid (JA) is increased, triggering the expression of JA-dependent genes such as the ones involved in the synthesis of glucosinolates and the markers from the JA pathway VSP2 (*vegetative storage protein 2*) and PDF1.2 (*plant defensin 1.2*). RNA silencing and a reactive oxygen species (ROS) burst restricted to the few cells affected by mite feeding are also triggered. Plant growth and developmental processes are repressed (Arena et al., 2018). Right: Upon plant interaction with viruliferous mites, CiLV-C reaches mesophyll cells with mite salivary flow and moves locally to cells surrounding inoculation sites. Virus presence intensifies ROS production and the SA-mediated response and activates several genes related to cell death and hypersensitive response (HR). Those responses are also triggered by the transient expression of P61, placing this protein in the epicenter of processes triggering typical HR lesions; P61 protein accumulates in the ER membrane, which might cause an ER stress that triggers the upregulation of genes related to unfolded protein response (UPR) and contribute to the induction of cell death. CiLV-C infection also triggers the induction of genes associated with cell growth, which might be involved in hyperplasia and hypertrophy processes. JA response is repressed in virus-infected plants, probably as a consequence of the antagonism exerted by the increased SA production. Viral infection also inhibits the plant primary metabolism, *e.g.* photosynthesis, likely contributing to the development of the chlorosis symptoms. A putative virus suppressor of RNA silencing (VSR) may target and inactivate the first antiviral defense line, leading to the upregulation of a second line that enhances the RNA silencing activity (Arena et al., 2016). Ultimately, as a result of the activation of several defense mechanisms, CiLV-C remains restricted to cells around the mite feeding sites, where chlorotic/necrotic lesions develop as a probable consequence of an HR-like resulting from an incompatible interaction.

system, with marked induction of genes related to cell death, ROS production, SA pathway, and HR. We hypothesize that the induction of those defense processes leads to the HR-like response that characterizes the phenotype of the viral infection. Next, we investigated the phenomenon of the activation of the plant defenses under the viral perspective, by searching for CiLV-C components capable of triggering such responses in the infected plants. To unravel the role of CiLV-C-encoded proteins in triggering the plant responses identified in the transcriptome analysis, the six viral ORFs were cloned in expression vectors and individually expressed into *N. benthamiana* leaves by agroinfiltration. The putative elicitor activity of the viral protein was assessed by: i) visual inspection of leaf phenotypic characteristics, ii) histochemical detection of H_2O_2, the main ROS detected during plant–pathogen interactions, iii) evaluation of the expression profile of SA- and HR-related genes *pathogenesis-related 1* (PR1), PR2, *hairpin-induced 1* (HIN1) and *hypersensitive-related 203J* (HSR203J), and iv) quantification of the SA and JA hormonal contents.

While the other CiLV-C proteins did not produce any altered phenotype (**Supplementary Figure 2**), the *Agrobacterium*-mediated transient expression of P61 consistently induced cell death on the infiltrated areas at 3 days after infiltration, which contrasted with the healthy phenotype observed in leaves infiltrated with both the *A. tumefaciens* carrying the empty vector and the infiltration buffer (**Figure 7A**). Histochemical analysis of the P61 infiltrated leaves revealed the production and accumulation of large amounts of H_2O_2 (**Figure 7B**). RT-qPCR assays showed that, although plants reacted to the infection by *A. tumefaciens* containing the empty vector, the presence of P61 clearly upregulated the expression of all the evaluated plant genes (**Figure 7C**). LC-MS/MS analyses revealed that SA levels were almost threefold higher on plants expressing this viral protein relative to those with the negative control, while JA levels were more than 42-fold lower (**Figure 7D**). Altogether, the results showed that the ectopic expression of the P61 protein triggers the same HR-related processes reported during plant infection with CiLV-C *i.e.* cell death, ROS production, SA pathway, and induction of HR-related genes. Mimicry of

responses typically observed during CiLV-C-plant interaction indicated P61 as a viral effector that elicits plant defenses that likely culminate in the HR-like phenotype characteristic of the viral infection (**Figures 1A** and **7G**)

To verify whether the HR-like is triggered by either the P61 protein or its RNA sequence, a frameshift mutant preventing the production of the protein was produced (**Figure 7E**). With the insertion of two nucleotides (T and A) downstream the start codon of the *p61* ORF, the construction resulted in a modification of the ORF generating an *amber* stop codon immediately after the first codon. The clones carrying the genes for the *p61* frameshift mutant, the *p61* wild type, and that encoding GFP were agroinfiltrated in different spots throughout the same *N. benthamiana* leaf (**Figure 7E**). Cell death was observed in areas infiltrated with the *p61* wild type 3 days after infiltration. HR was observed in neither the areas infiltrated with *gfp* nor the *p61* frameshift mutant. Since the frameshift mutation did not affect the synthesis of the *p61* RNA, the result demonstrates that the HR-like phenotype is triggered only in the presence of the P61 protein.

DISCUSSION

In this work, we have dissected the interaction between plants and the kitavirid CiLV-C, a virus atypically unable to accomplish systemic infection in any of its known plant hosts, at the molecular level. The accumulation of CiLV-C genomic and subgenomic RNA molecules was coordinately studied with the transcriptome profile of infected plants to reveal the major mechanisms underlying the global plant response to the infection. Further, the transient expression of individual viral proteins allowed us to identify P61 as a putative viral effector causing the HR-like symptoms associated with the viral infection.

The analysis of the kinetics of viral RNA accumulation in *A. thaliana* plants infected by CiLV-C revealed the main features of an asynchronous process that we classified in three major steps (**Figures 1B, C**). The earliest stage occurs from the viral inoculation by viruliferous mites until 24 hai, which is characterized by a low number of viral molecules, low replication rate, or more likely, replication restricted to a few cells (**Figures 1B, C**). Accordingly, during the first 6 h of this stage, the plant transcriptional response was undetectable, at least with the experimental approach used here (**Figures 2E, F**). From 2 to 4 dai, the lowest difference between subgenomic and genomic RNA takes place (**Figures 1B, C**), suggesting a high replication rate, and a moderate number of plant transcriptional responses to viral infection is detected (**Figures 2E, F**). The third stage that ranges from 6 to 10 dai is marked by an increased accumulation of both subgenomic and genomic RNAs, reaching the highest detection levels (**Figures 1B, C**). Over the first hours of this step, which precedes the appearance of symptoms in *A. thaliana* leaves, a massive reprogramming of the plant transcript profile is observed (**Figures 2E, F**).

Early transcriptome changes in response to the infection by CiLV-C involve the upregulation of cell growth-related processes (**Figure 3C**). Induction of genes related to cell growth might be related to the development of hyperplasia or hypertrophy, histological changes typically promoted by phytopathogens, including several plant viruses (Hull, 2009). Likewise, symptomatic areas of sweet orange leaves with citrus leprosis show a higher division activity and larger size of the parenchyma cells (Marques et al., 2007). Hyperplasia and hypertrophy are also detected in citrus plant tissues infected by *Brevipalpus*-transmitted dichorhaviruses (Marques et al., 2010), suggesting that upregulation of cell growth responses might be a common pattern during the infection by *Brevipalpus* transmitted viruses (BTVs) in plants.

CiLV-C infection downregulated the interconnected ET and JA pathways at 2 and 6 dai (**Figures 3C** and **4B**), respectively. Similarly, downregulation of JA-responsive genes involved in the biosynthesis of glucosinolates, compounds acting against herbivores (Jander, 2014), was also detected at the last time point (**Figure 4B**). The concomitance of the induction of SA and reduction of JA during the plant infection by CiLV-C is likely a consequence of the SA-JA antagonism (Arena et al., 2016). The repression of antiherbivory defenses upon virus infection is used as a viral strategy to increase vector fitness or attraction and encourage virus transmission (Mauck et al., 2018; Donnelly et al., 2019; Carr et al., 2020). For instance, the polerovirus potato leafroll virus attenuates the induction of JA and ET by aphids, affecting vector fecundity and settling (Patton et al., 2020). By exploiting the natural SA–JA antagonism, the tospovirus tomato spotted wilt virus triggers the SA- to reduce JA-mediated defenses, rendering a more attractive host to the thrips vector *Frankliniella occidentalis* (Abe et al., 2012). Likewise, the induction of SA and reduction of JA pathways and related defenses upon CiLV-C infection might account for an improvement in *Brevipalpus* vector performance. In agreement with this hypothesis, CiLV-C infected *A. thaliana* leaves are preferred for mite colonization and oviposition (Arena et al., 2016), and experimental evidence suggests that mite density in CiLV-C infected sweet orange trees is higher than in the healthy ones (Andrade et al., 2013). Furthermore, the oviposition of *B. yothersi* mites in *A. thaliana* mutant plants compromised in SA signaling is lower (Arena et al., 2018), pointing out the role of SA response on the improvement of the mite performance. On this basis, we speculate that CiLV-C/*Brevipalpus* interaction is a mutualistic relationship, in which the better performance of mites on CiLV-C plants, showing a boosted SA response and likely a suppressed antiherbivore defense, improves the transmission rate of the virus.

Along with the reduction of the JA pathway and glucosinolate production, the presymptomatic stage during the CiLV-C infection was marked by a repression of the primary metabolism (**Figure 4B**). Detected downregulated processes in *A. thaliana* infected by CiLV-C comprised the metabolism of lipids, amino acids, and carbohydrates, including photosynthesis (**Figure 4B**). Inhibition of photosynthesis is a general rule across plant–virus interactions and is usually associated with changes in the chloroplasts and the development of chlorosis and necrosis (Goulart et al., 2019), some of the most common viral symptoms and a key feature in cilevirus infections. Chloroplasts are prime targets for plant viruses to help fulfill essential stages of viral infections such as replication and movement, undergoing massive structural and functional disturbance (Zhao et al., 2016; Bhattacharyya and Chakraborty, 2018). Conversely,

chloroplasts play active roles in defense against viruses because they are the sites for the biosynthesis of SA and HR-related ROS, which might compromise photosynthesis (Zhao et al., 2016; Bhattacharyya and Chakraborty, 2018). Whatever the cause, plant viruses commonly damage chloroplasts, leading to reduced photosynthetic activity and development of chlorotic or even necrotic symptoms as a result of cell death (Zhao et al., 2016; Bhattacharyya and Chakraborty, 2018; Goulart et al., 2019). In this sense, the infection by CiLV-C seems to follow the commonly observed trend.

Even though the profile of some plant transcripts changed in the early stage of the CiLV-C infection, strikingly, the majority of the DEGs were identified at 6 dai (**Figure 3A**). As expected, the outcome of several biological processes identified at the presymptomatic stage might contribute to the development of the disease symptoms. Plant response at this phase of the viral infection is typified by the upregulation of the plant immune system (**Figure 4A**). A large number of upregulated genes are related to the SA-mediated pathway, response to ROS, cell death, and HR (**Figures 4A** and **6A**). In agreement with molecular responses detected in this work, histochemical analyses of tissues affected by CiLV-C revealed the accumulation of ROS and the presence of dead cells (Arena et al., 2016). In all its plant hosts, CiLV-C is restricted to cells around the vector inoculation sites where symptoms of viral infection arise (Freitas-Astúa et al., 2018; **Figures 1A** and **7G**). Phenotypically, these symptoms resemble the outcome of an HR, a cell death resistance process accompanied by pathogen restriction at the inoculation site during an incompatible interaction. Transcriptome changes associated with the induction of HR-like phenotype support the hypothesis that the lesions caused by citrus leprosis may be a consequence of an incompatible rather than a compatible interaction (Arena et al., 2016). Commonly, the development of an HR resistance is associated with the recognition of the viral protein by corresponding plant resistance (R) proteins in a host-specific manner (Garcia-Ruiz, 2019). It is noteworthy, however, that HR-like associated phenotypes (cell death and virus restriction) developed during CiLV-C infection occur over a large spectrum of CiLV-C hosts rather than in a specific plant species. Under these circumstances, alternative mechanisms leading to an HR-like phenotype cannot be ruled out.

Despite the fact that ER stress and UPR were not within the enriched GO terms, we identified the upregulation of genes related to both processes at the presymptomatic stage (**Supplementary Table 7**). Upon stress conditions, the accumulation of unfolded/misfolded proteins in the ER triggers the UPR, a protective response that improves protein folding activity and removes proteins from the ER (Afrin et al., 2020). When the relief of the ER stress fails, the programmed cell death can be activated (Eichmann and Schäfer, 2012; Williams et al., 2014). Virus-infected plants can upregulate UPR-related genes, and some virus-encoded proteins targeting ER induce UPR (Zhang and Wang, 2016). For instance, the TuMV protein 6K2 can induce UPR (Zhang et al., 2015), probably through its physical interaction with and remodeling of the ER (Laliberté and Sanfaçon, 2010; Zhang and Wang, 2016). Furthermore, evidence suggests that SA induces UPR in plants (Poór et al., 2019), and virus-induced ER stress triggers ROS production and Ca^{2+} influx that alert host defense systems (Zhang and Wang, 2016). Ultimately, ER stress may culminate in HR cell death during plant–virus infections (Williams et al., 2014). This is the case, for example, of the potato virus X movement protein, which activates the transcription factor bZIP60 to initiate the UPR and elicit programmed cell death (Ye et al., 2013). Similarly, the induction of UPR-related genes during CiLV-C infection raises the possibility that ER stress might take part in the processes triggering the plant immune system and HR-like cell death. Further studies are underway aiming to assess the contribution of the ER–UPR pathways during CiLV-C infection.

Finally, we showed that the transient expression of the CiLV-C P61 protein reproduces processes observed during plant interaction with CiLV-C, *i.e.* necrotic lesions in *A. thaliana* and *N. benthamiana*, and increased expression of HR-related genes, modulation of SA and JA pathways, and ROS burst in *A. thaliana* (Arena et al., 2016) (**Figure 7**). Besides, we demonstrated that all the tested processes are triggered by the expression of P61, but not by its mRNA (**Figure 7**). Several viral proteins that mimic responses of virus infection have been identified. For instance, the polerovirus P0 protein triggers HR necrotic lesions in *N. glutinosa* (Wang et al., 2015), the CP from cucumber mosaic virus interacts with a chloroplast ferredoxin protein causing chlorotic symptoms in tobacco (Qiu et al., 2018), and the βC1 protein from the beta satellite of tomato yellow leaf curl China virus interacts with the *MYC2* transcription factor decreasing levels of JA-responsive genes and enhancing the performance of its vector (Li et al., 2014). Induction of HR-like phenotype and related defenses brings CiLV-C P61 to the epicenter of processes triggering typical HR lesions. P61 causes structural remodeling of the ER membranes (Leastro et al., 2018), suggesting that the cell death caused by P61 expression might result from an unmitigated ER stress. Since ER stress signaling and UPR can be controlled by a nonfully known mechanism involving SA (Poór et al., 2019; Pastor-Cantizano et al., 2020), the interaction of P61 with *A. thaliana* or *N. benthamiana* plants might represent an interesting model for plant UPR regulation studies. Furthermore, due to the increased SA and reduced JA levels upon P61 expression (**Figure 7D**), it is suggested that P61 might trigger the cross-talk between hormonal pathways and modify the CiLV-C vector performance. New experimental approaches have been scheduled to verify the relevancy of these findings.

Results obtained in this work have enlarged and strengthened the previously proposed model depicting the plant response to components of the citrus leprosis pathosystem *i.e.* CiLV-C, and *Brevipalpus* mites (Arena et al., 2016; **Figure 8**). Comprehensively, we have provided host transcriptome data and viral protein expression evidence supporting the current understanding that the symptoms of CiLV-C infection arise from an HR-like resistance. CiLV-C is likely unable to overcome the myriad of plant defenses activated upon plant/virus interaction, which might prevent viral spread, thus restraining the viral infection in patches of tissues around the mite's feeding/inoculation foci, where programmed cell death is further triggered. Although offbeat within the universe of known plant-infecting viruses, local lesions and the lack of systemic movement are common across members of the family *Kitaviridae* (Quito-Avila et al., 2020). Likely derived from a common ancestor with nege and nege-like viruses (Ramos-

González et al., 2020), lack of systemic infection of kitavirids in plants might be a consequence of their unfitness in plants, *i.e.* failure to circumvent the plant defenses, which indirectly suggests a relatively short coevolutionary history of the biosystem kitavirus–plant. While the specific involvement of the innate immunity, including the putative existence of an R gene, UPR mechanism, and gene silencing on the kitavirus infection has been pointed out, the role of abiotic factors, such as temperature, on the pathosystem cannot be ruled out. Plant rhabdoviruses of the genus *Dichorhavirus*, which are also transmitted by *Brevipalpus* mites and produce nonsystemic infection under natural conditions, show a striking ability to infect plants systemically when incubated above 28°C (Dietzgen et al., 2018). Alternatively, since the rewire of the plant hormone metabolism seems to be relevant for *Brevipalpus* mite performance on CiLV-C infected plants, and on the other hand, the resistance breaking may affect virus fitness *i.e.* transmission and survival (García-Arenal and Fraile, 2013), a valid question would be what is the actual contribution of HR-like resistance to the fitness of this pathosystem. Ongoing experiments to clarify these and other aspects of CiLV-C plant and vector interplays, *e.g.* whether cilevirus multiply in their mite vector, will help to shed light on the forces and boundaries shaping the kitavirus evolutionary history.

DATA AVAILABILITY STATEMENT

The RNA-Seq raw data are available at sequence read archive (SRA) with the ID PRJNA454529.

AUTHOR CONTRIBUTIONS

GA and PR-G conceptualized and wrote the original draft. GA and PR-G worked on the formal analysis. JF-A, BF, CC, and MM supervised the study. JF-A, BF, CC, and MM were responsible for funding acquisition. GA, PR-G, CC, and JF-A worked on the investigation and methodology. All authors contributed to the article and approved the submitted version.

FUNDING

The authors are grateful to CAPES and FAPESP for scholarships and research grants associated with this work (CAPES PNPD20132154–33141010001P4–PNPD–IBSP, 88882.157041/2017-01; FAPESP 2014/50880-0, 2014/00366-8, 2016/23870-9, 2017/50222-0, and 2019/02137-0). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brazil (CAPES) Finance Code 001.

ACKNOWLEDGMENTS

The authors thank Dr. Valdenice M. Novelli (Centro APTA Citros Sylvio Moreira-IAC) for providing the mite population and Rafaela Gasparotto (Centro APTA Citros Sylvio Moreira-IAC) for technical support. This manuscript partly overlaps the dissertation of one of its authors (Arena, 2018).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.01188/full#supplementary-material

SUPPLEMENTARY FIGURE 1 | Accumulation of CiLV-C genes through the course of viral infection in *Arabidopsis thaliana* plants. Normalized relative quantities (NRQs) of molecules were determined by RT-qPCR at eight time points after infestation with viruliferous *Brevipalpus yothersi* mites. Data are presented as $\log_{10}$NRQ values in comparison with the time point with the lowest quantity (with $\log_{10}$NRQ set to zero). Different letters correspond to different quantities between the time points assessed (ANOVA and Student's t-test, $\alpha < 0.05$). *RdRp*, RNA-dependent RNA polymerase; *MP*, movement protein; *IR*, intergenic region; h, hours after infestation; d, days after infestation.

SUPPLEMENTARY FIGURE 2 | Phenotype resulted from the expression of CiLV-C proteins in *Nicotiana benthamiana* plants. The viral proteins were transiently expressed in *N. benthamiana* leaves using *Agrobacterium*-mediated infiltration. Left half (I) of each leaf was infiltrated with *A. tumefaciens* GV3101 containing the empty vector, while right half (II) was infiltrated with the bacteria containing the construct for the expression of the corresponding CiLV-C protein. A: *p61*, B: *RdRp* motif, C: *p24*, D: *p15*, E: *mp*, F: *p29*, and G: *methyltransferase* motif.

REFERENCES

Abe, H., Tomitaka, Y., Shimoda, T., Seo, S., Sakurai, T., Kugimiya, S., et al. (2012). Antagonistic plant defense system regulated by phytohormones assists interactions among vector insect, thrips and a tospovirus. *Plant Cell Physiol.* 53, 204–212. doi: 10.1093/pcp/pcr173

Afrin, T., Diwan, D., Sahawneh, K., and Pajerowska-Mukhtar, K. (2020). Multilevel Regulation of Endoplasmic Reticulum Stress Responses in Plants: Where Old Roads and New Paths Meet. *J. Exp. Bot.* 71, 1659–1667. doi: 10.1093/jxb/erz487

Aguilar, E., Brosseau, C., Moffett, P., and Tenllado, F. (2019). Cell death triggered by the P25 protein in Potato virus X -associated synergisms results from ER stress in Nicotiana benthamiana. *Mol. Plant Pathol.* 20, 194–210. doi: 10.1111/mpp.12748

Anders, S., McCarthy, D. J., Chen, Y., Okoniewski, M., Smyth, G. K., Huber, W., et al. (2013). Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat. Protoc.* 8, 1765–1786. doi: 10.1038/nprot.2013.099

Andrade, D. J., Falconi, R. S., Siqueira, D. S., Barbosa, C. L., Ferraudo, A. S., and Oliveira, C. A. L. (2013). The influence of citrus rootstocks on the relationship between the mite Brevipalpus phoenicis and citrus leprosis disease. *Pest Manage. Sci.* 69, 81–87. doi: 10.1002/ps.3365

Aoyama, T., and Chua, N. (1997). A glucocorticoid-mediated transcriptional induction system in transgenic plants. *Plant J.* 11, 605–612. doi: 10.1046/j.1365-313x.1997.11030605.x

Arena, G. D., Bergamini, M. P., Tassi, A. D., Kitajima, E. W., Kubo, K. S., and Freitas-Astúa, J. (2013). Citrus leprosis virus C infects Arabidopsis thaliana, the model for plant-pathogen interactions. *J. Plant Pathol.* 95, 448. doi: 10.4454/JPP.V95I2.003

Arena, G. D., Ramos-González, P. L., Nunes, M. A., Alves, M. R., Camargo, L. E. A., Kitajima, E. W., et al. (2016). Citrus leprosis virus C infection results in hypersensitive-like response, suppression of the JA/ET plant defense pathway and promotion of the colonization of its mite vector. *Front. Plant Sci.* 7, 1757. doi: 10.3389/FPLS.2016.01757

Arena, G. D., Ramos-González, P. L., Rogerio, L. A., Ribeiro-Alves, M., Casteel, C. L., Freitas-Astúa, J., et al. (2018). Making a Better Home: Modulation of Plant Defensive Response by Brevipalpus Mites. *Front. Plant Sci.* 9, 1147. doi: 10.3389/fpls.2018.01147

Arena, G. D. (2018). *Leprose dos citros: um modelo para o estudo da interação planta/ácaro/vírus. [dissertation].* (Campinas (SP): Universidade Estadual de Campinas.).

Bassanezi, R. B., Czermainski, A. B. C., Laranjeira, F. F., Moreira, A. S., Ribeiro Júnior, P. J., Krainski, E. T., et al. (2019). Spatial patterns of the Citrus leprosis virus and its associated mite vector in systems without intervention. *Plant Pathol.* 68, 85–93. doi: 10.1111/ppa.12930

Bastianel, M., Novelli, V. M., Kitajima, E. W., Kubo, K. S., Bassanezi, R. B., Machado, M. A., et al. (2010). Citrus Leprosis: Centennial of an unusual mite–virus pathosystem. *Plant Dis.* 94, 284–292. doi: 10.1094/PDIS-94-3-0284

Beard, J. J., Ochoa, R., Braswell, W. E., and Bauchan, G. R. (2015). Brevipalpus phoenicis (Geijskes) species complex (Acari: Tenuipalpidae)-a closer look. *Zootaxa* 3944, 167. doi: 10.11646/zootaxa.3944.1.1

Bhattacharyya, D., and Chakraborty, S. (2018). Chloroplast: the Trojan horse in plant-virus interaction. *Mol. Plant Pathol.* 19, 504–518. doi: 10.1111/mpp.12533

Carr, J. P., Tungadi, T., Donnelly, R., Bravo-Cazar, A., Rhee, S. J., Watt, L. G., et al. (2020). Modelling and manipulation of aphid-mediated spread of non-persistently transmitted viruses. *Virus Res.* 277, 197845. doi: 10.1016/j.virusres.2019.197845

Casteel, C. L., and Falk, B. W. (2016). "Plant Virus-Vector Interactions: More Than Just for Virus Transmission," in *Current Research Topics in Plant Virology*. Eds. A. Wang and X. Zhou (Cham: Springer International Publishing), 217–240. doi: 10.1007/978-3-319-32919-2_9

Casteel, C. L., De Alwis, M., Bak, A., Dong, H., Whitham, S. A., and Jander, G. (2015). Disruption of Ethylene Responses by Turnip mosaic virus Mediates Suppression of Plant Defense against the Green Peach Aphid Vector. *Plant Physiol.* 169, 209–218. doi: 10.1104/pp.15.00332

Cooley, M. B., Pathirana, S., Wu, H. J., Kachroo, P., and Klessig, D. F. (2000). Members of the Arabidopsis HRT/RPP8 family of resistance genes confer resistance to both viral and oomycete pathogens. *Plant Cell* 12, 663–676. doi: 10.1105/tpc.12.5.663

Couto, D., and Zipfel, C. (2016). Regulation of pattern recognition receptor signalling in plants. *Nat. Publ. Gr.* 16, 537–552. doi: 10.1038/nri.2016.77

Cui, H., Tsuda, K., and Parker, J. E. (2015). Effector-triggered immunity: from pathogen perception to robust defense. *Annu. Rev. Plant Biol.* 66, 487–511. doi: 10.1146/annurev-arplant-050213-040012

Dietzgen, R. G., Freitas-Astúa, J., Chabi-Jesus, C., Ramos-González, P. L., Goodin, M. M., Kondo, H., et al. (2018). Dichorhaviruses in Their Host Plants and Mite Vectors. *Adv. Virus Res.* 102, 119–148. doi: 10.1016/BS.AIVIR.2018.06.001

Donnelly, R., Cunniffe, N. J., Carr, J. P., and Gilligan, C. A. (2019). Pathogenic modification of plants enhances long-distance dispersal of nonpersistently transmitted viruses to new hosts. *Ecology* 100, e02725. doi: 10.1002/ecy.2725

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. *Trends Plant Sci.* 15, 573–581. doi: 10.1016/j.tplants.2010.06.005

Eichmann, R., and Schäfer, P. (2012). The endoplasmic reticulum in plant immunity and cell death. *Front. Plant Sci.* 3, 200. doi: 10.3389/fpls.2012.00200

Freitas-Astúa, J., Ramos-González, P. L., Arena, G. D., Tassi, A. D., and Kitajima, E. W. (2018). Brevipalpus-transmitted viruses: parallelism beyond a common vector or convergent evolution of distantly related pathogens? *Curr. Opin. Virol.* 33, 66–73. doi: 10.1016/J.COVIRO.2018.07.010

García-Arenal, F., and Fraile, A. (2013). Trade-offs in host range evolution of plant viruses. *Plant Pathol.* 62, 2–9. doi: 10.1111/ppa.12104

Garcia-Ruiz, H. (2019). Host factors against plant viruses. *Mol. Plant Pathol.* 20, 1588–1601. doi: 10.1111/mpp.12851

Garita, L. C., Tassi, A. D., Calegario, R. F., Freitas-Astúa, J., Salaroli, R. B., Romão, G. O., et al. (2014). Experimental host range of Citrus leprosis virus C (CiLV-C). *Trop. Plant Pathol.* 39, 43–55. doi: 10.1590/S1982-56762014005000004

Goulart, L., Gilza, Z., De Souza, B., and Siqueira, M. (2019). Transcriptomics of plant – virus interactions : a review. *Theor. Exp. Plant Physiol.* 31, 103–125. doi: 10.1007/s40626-019-00143-z

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., and Vandesompele, J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* 8, R19. doi: 10.1186/gb-2007-8-2-r19

Hull, R. (2009). *Comparative Plant Virology*. *2nd ed* (London: Elsevier). doi: 10.1017/CBO9781107415324.004

Ilarduya, O. M., Xie, Q., and Kaloshian, I. (2003). Aphid-Induced Defense Responses in Mi-1 -Mediated Compatible and Incompatible Tomato Interactions. *Mol. Plant-Microbe Interact.* 16, 699–708. doi: 10.1094/MPMI.2003.16.8.699

Jander, G. (2014). Revisiting Plant-Herbivore Co-Evolution in the Molecular Biology Era. *Annu. Plant Rev. Insect-Plant Interact.* 47, 361–384. doi: 10.1002/9781118829783.ch11

Jin, J., Tian, F., Yang, D.-C., Meng, Y.-Q., Kong, L., Luo, J., et al. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982

Kiełbowicz-Matuk, A. (2012). Involvement of plant C2H2-type zinc finger transcription factors in stress responses. *Plant Sci.* 185-186, 78–85. doi: 10.1016/j.plantsci.2011.11.015

Kuchibhatla, D. B., Sherman, W., Chung, B. Y. W., Cook, S., Schneider, G., Eisenhaber, B., et al. (2014). Powerful sequence similarity search methods and in-depth manual analyses can identify remote homologs in many apparently "orphan" viral proteins. *J. Virol.* 88, 10–20. doi: 10.1128/JVI.02595-13

Laliberté, J.-F., and Sanfaçon, H. (2010). Cellular Remodeling During Plant Virus Infection. *Annu. Rev. Phytopathol.* 48, 69–91. doi: 10.1146/annurev-phyto-073009-114239

Leastro, M. O., Kitajima, E. W., Silva, M. S., Resende, R. O., and Freitas-Astúa, J. (2018). Dissecting the Subcellular Localization, Intracellular Trafficking, Interactions, Membrane Association, and Topology of Citrus Leprosis Virus C Proteins. *Front. Plant Sci.* 9, 1299. doi: 10.3389/fpls.2018.01299

Li, W., Xu, Y.-P., Zhang, Z.-X., Cao, W.-Y., Li, F., Zhou, X., et al. (2012). Identification of genes required for nonhost resistance to Xanthomonas oryzae pv. oryzae reveals novel signaling components. *PLoS One* 7, e42796. doi: 10.1371/journal.pone.0042796

Li, R., Weldegergis, B. T., Li, J., Jung, C., Qu, J., Sun, Y., et al. (2014). Virulence Factors of Geminivirus Interact with MYC2 to Subvert Plant Resistance and Promote Vector Performance. *Plant Cell Online* 26, 4991–5008. doi: 10.1105/tpc.114.133181

Li, W., Yadeta, K. A., Elmore, J. M., and Coaker, G. (2013). The Pseudomonas syringae effector HopQ1 promotes bacterial virulence and interacts with tomato 14-3-3 proteins in a phosphorylation-dependent manner. *Plant Physiol.* 161, 2062–2074. doi: 10.1104/pp.112.211748

Locali, E. C., Freitas-astua, J., De Souza, A. A., Takita, M. A., Astua-monge, G., Antonioli, R., et al. (2003). Development of a Molecular Tool for the Diagnosis of Leprosis , a Major Threat to Citrus Production in the Americas. *Plant Dis.* 87, 1317–1321. doi: 10.1094/PDIS.2003.87.11.1317

Locali-Fabris, E. C., Freitas-Astúa, J., Souza, A. A., Takita, M. A., Astúa-Monge, G., Antonioli-Luizon, R., et al. (2006). Complete nucleotide sequence, genomic organization and phylogenetic analysis of Citrus leprosis virus cytoplasmic type. *J. Gen. Virol.* 87, 2721–2729. doi: 10.1099/vir.0.82038-0

Love, M.II, Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics* 21, 3448–3449. doi: 10.1093/bioinformatics/bti551

Manassero, N. G. U., Viola, I. L., Welchen, E., and Gonzalez, D. H. (2013). TCP transcription factors: architectures of plant form. *Gruyter* 4, 111–127. doi: 10.1515/bmc-2012-0051

Marques, J. P. R., De Freitas-Astúa, J., Kitajima, E. W., and Appezzato-da-Glória, B. (2007). Lesões foliares e de ramos de laranjeira-doce causadas pela leprose-dos-citros. *Pesqui. Agropecuária Bras.* 42, 1531–1536. doi: 10.1590/S0100-204X2007001100003

Marques, J. P., Kitajima, E., Freitas-Astúa, J., and Apezzato-da-Glória, B. (2010). Comparative morpho-anatomical studies of the lesions causedby citrus leprosis virus on sweet orange. *Ann. Braz. Acad. Sci.* 82, 501–511. doi: 10.1590/S0001-37652010000200025

Mauck, K. E., Chesnais, Q., and Shapiro, L. R. (2018). Evolutionary Determinants of Host and Vector Manipulation by Plant Viruses. *Adv. Virus Res.* 101, 189–250. doi: 10.1016/BS.AIVIR.2018.02.007

Nunes, M. A., Bergamini, M. P., Coerini, L. F., Bastianel, M., Novelli, V. M., Kitajima, E. W., et al. (2012). Citrus leprosis virus C naturally infecting Commelina benghalensis , a prevalent monocot weed of citrus orchards in Brazil. *Plant Dis.* 96, 770–770. doi: 10.1094/PDIS-11-11-0925-PDN

Nuruzzaman, M., Sharoni, A. M., and Kikuchi, S. (2013). Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. *Front. Microbiol.* 4, 248. doi: 10.3389/fmicb.2013.00248

Olsen, A. N., Ernst, H. A., Leggio, L., and Skriver, K. (2005). NAC transcription factors: structurally distinct, functionally diverse. *Trends Plant Sci.* 10, 79–87. doi: 10.1016/j.tplants.2004.12.010

Pascon, R. C., Kitajima, J. P., Breton, M. C., Assumpção, L., Greggio, C., Zanca, A. S., et al. (2006). The Complete Nucleotide Sequence and Genomic Organization of Citrus Leprosis Associated Virus, Cytoplasmatic type (CiLV-C). *Virus Genes* 32, 289–298. doi: 10.1007/s11262-005-6913-1

Pastor-Cantizano, N., Ko, D. K., Angelos, E., Pu, Y., and Brandizzi, F. (2020). Functional Diversification of ER Stress Responses in Arabidopsis. *Trends Biochem. Sci.* 45, 123–136. doi: 10.1016/j.tibs.2019.10.008

Patton, M. F., Bak, A., Sayre, J. M., Heck, M. L., and Casteel, C. L. (2020). A polerovirus, Potato leafroll virus, alters plant–vector interactions using three viral proteins. *Plant Cell Environ.* 43, 387–399. doi: 10.1111/pce.13684

Phukan, U. J., Jeena, G. S., and Shukla, R. K. (2016). WRKY Transcription Factors: Molecular Regulation and Stress Responses in Plants. *Front. Plant Sci.* 7, 760. doi: 10.3389/fpls.2016.00760

Pieterse, C. M. J., Van der Does, D., Zamioudis, C., Leon-Reyes, A., and Van Wees, S. C. M. (2012). Hormonal modulation of plant immunity. *Annu. Rev. Cell Dev. Biol.* 28, 489–521. doi: 10.1146/annurev-cellbio-092910-154055

Poór, P., Czékus, Z., Tari, I., and Ördög, A. (2019). The Multifaceted Roles of Plant Hormone Salicylic Acid in Endoplasmic Reticulum Stress and Unfolded Protein Response. *Int. J. Mol. Sci.* 20, 5842. doi: 10.3390/ijms20235842

Pumplin, N., and Voinnet, O. (2013). RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. *Nat. Rev. Microbiol.* 11, 745–760. doi: 10.1038/nrmicro3120

Qiu, Y., Zhang, Y., Wang, C., Lei, R., Wu, Y., Li, X., et al. (2018). Cucumber mosaic virus coat protein induces the development of chlorotic symptoms through interacting with the chloroplast ferredoxin I protein. *Sci. Rep.* 8, 1205. doi: 10.1038/s41598-018-19525-5

Quito-Avila, D. F., Freitas-Astúa, J., and Melzer, M. J. (2020). "Bluner-, Cile-, and Higreviruses (Kitaviridae).," in *Reference Module in Life Sciences* (Amsterdam: Elsevier). doi: 10.1016/B978-0-12-809633-8.21248-X

Ramos-González, P., Chabi-Jesus, C., Guerra-Peraza, O., Breton, M., Arena, G., Nunes, M., et al. (2016). Phylogenetic and Molecular Variability Studies Reveal a New Genetic Clade of Citrus leprosis virus C. *Viruses* 8, 153. doi: 10.3390/v8060153

Ramos-González, P. L., Chabi-Jesus, C., Arena, G. D., Tassi, A. D., Kitajima, E. W., and Freitas-Astúa, J. (2018). Citrus leprosis: a unique multietiologic disease. *Citrus Am.* 1, 4–19.

Ramos-González, P. L., dos Santos, G. F., Chabi-Jesus, C., Harakava, R., Kitajima, E. W., and Freitas-Astúa, J. (2020). Passion Fruit Green Spot Virus Genome Harbors a New Orphan ORF and Highlights the Flexibility of the 5′-End of the RNA2 Segment Across Cileviruses. *Front. Microbiol.* 11, 206. doi: 10.3389/fmicb.2020.00206

Rodrigues, V., Bastianel, M., Kubo, K., Fadel, A. L., Nicolini, F., Novelli, V. M., et al. (2007). Desenvolvimento de um método para a otimização da transmissão experimental do vírus da leprose dos citros. *Laranja* 28, 29–38.

Tautz, D., and Domazet-Lošo, T. (2011). The evolutionary origin of orphan genes. *Nat. Rev. Genet.* 12, 692–702. doi: 10.1038/nrg3053

Wang, K.-D., Empleo, R., Nguyen, T. T. V., Moffett, P., and Sacco, M. A. (2015). Elicitation of hypersensitive responses in *Nicotiana glutinosa* by the suppressor of RNA silencing protein P0 from poleroviruses. *Mol. Plant Pathol.* 16, 435–448. doi: 10.1111/mpp.12201

Williams, B., Verchot, J., and Dickman, M. B. (2014). When supply does not meet demand-ER stress and plant programmed cell death. *Front. Plant Sci.* 5, 211. doi: 10.3389/fpls.2014.00211

Xia, X.-J., Zhou, Y.-H., Shi, K., Zhou, J., Foyer, C. H., and Yu, J.-Q. (2015). Interplay between reactive oxygen species and hormones in the control of plant development and stress tolerance. *J. Exp. Bot.* 66, 2839–2856. doi: 10.1093/jxb/erv089

Ye, C., Chen, S., Payton, M., Dickman, M. B., and Verchot, J. (2013). TGBp3 triggers the unfolded protein response and SKP1-dependent programmed cell death. *Mol. Plant Pathol.* 14, 241–255. doi: 10.1111/mpp.12000

Zhang, L., and Wang, A. (2016). "ER Stress , UPR and Virus Infections in Plants," in *Current Research Topics in Plant Virology*. Eds. A. Wang and X. Zhou (Cham: Springer), 173–195. doi: 10.1007/978-3-319-32919-2

Zhang, L., Chen, H., Brandizzi, F., Verchot, J., and Wang, A. (2015). The UPR Branch IRE1- bZIP60 in Plants Plays an Essential Role in Viral Infection and Is Complementary to the Only UPR Pathway in Yeast. *PLoS Genet.* 11, e1005164. doi: 10.1371/journal.pgen.1005164

Zhao, S., and Fernald, R. D. (2005). Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J. Comput. Biol.* 12, 1047–1064. doi: 10.1089/cmb.2005.12.1047

Zhao, J., Zhang, X., Hong, Y., and Liu, Y. (2016). Chloroplast in Plant-Virus Interaction. *Front. Microbiol.* 7, 1565. doi: 10.3389/fmicb.2016.01565

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.

frontiers in Microbiology

OPINION
published: 29 October 2020
doi: 10.3389/fmicb.2020.561092

Sequence Relationships of RNA Helicases and Other Proteins Encoded by Blunervirus RNAs Highlight Recombinant Evolutionary Origin of Kitaviral Genomes

Sergey Y. Morozov[1,2]*, *Ekaterina A. Lazareva*[2] and *Andrey G. Solovyev*[1,2,3,4]

[1] A. N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow, Russia, [2] Department of Virology, Biological Faculty, Moscow State University, Moscow, Russia, [3] Institute of Molecular Medicine, Sechenov First Moscow State Medical University, Moscow, Russia, [4] Institute of Agricultural Biotechnology, Moscow, Russia

Keywords: kitaviruses, blunerviruses, negeviruses, evolution, RNA helicase, RNA genome, insects, plants

OPEN ACCESS

Edited by:
Sead Sabanadzovic, Mississippi State University, United States

Reviewed by:
Michael Melzer, University of Hawaii, United States
Mart Krupovic, Institut Pasteur, France

***Correspondence:**
Sergey Y. Morozov
morozov@belozersky.msu.ru

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *11 May 2020*
Accepted: *28 September 2020*
Published: *29 October 2020*

Citation:
Morozov SY, Lazareva EA and Solovyev AG (2020) Sequence Relationships of RNA Helicases and Other Proteins Encoded by Blunervirus RNAs Highlight Recombinant Evolutionary Origin of Kitaviral Genomes. Front. Microbiol. 11:561092. doi: 10.3389/fmicb.2020.561092

INDICATIONS OF INTER-KINGDOM RNA VIRUS TRANSFER BETWEEN ARTHROPODS AND PLANTS

The family *Kitaviridae* includes a small number of positive-stranded RNA plant viruses. Interestingly, this family name is not derived from the name of the type virus member, as usually, but from the family name of Dr. Elliot Watanabe Kitajima, a prominent virologist participating in kitavirus studies (Ramos-González et al., 2020; Quito-Avila et al., in press). Members of the kitavirus genera *Blunervirus*, *Cilevirus*, and *Higrevirus* produce *in planta* bacilliform (or near spherical) virus particles and have multipartite genomes composed of two, three, or four RNAs for cile-, higre- and bluner-viruses, respectively. In addition to common particle morphology, most viruses share some important biological peculiarities, such as lack of systemic movement (Ramos-González et al., 2020; Quito-Avila et al., in press).

Molecular phylogenetic analyses of the RNA-dependent RNA polymerase (RdRp) domain or concatenated sequences of methyl transferase (MT), replicative RNA helicase (HEL), and RdRp unambiguously showed that kitaviruses are most closely related to members of family *Virgaviridae* [this family name is also not derived from the name of the type virus, like *Bromoviridae*, but from Latin word virga (rod), as all viruses in this family are rod-shaped] followed by those of the family *Bromoviridae* (Quito-Avila et al., 2013; Nunes et al., 2017; Hao et al., 2018; Ramos-González et al., 2020). Accordingly, cileviruses and blunerviruses encode a putative movement protein (MP) with sequence similarity to the MP of *Bromoviridae* (Hao et al., 2018; Quito-Avila et al., in press). Currently, all these viruses are assigned to the recently established positive-stranded RNA virus order *Martellivirales* [this family has been named to honor Prof. Giovanni Paolo Martelli for his contribution to virus taxonomy] (Koonin et al., 2020). However, kitavirus replicative proteins show closest similarity to arthropod-infecting negeviruses (proposed genera Nelorpivirus and Sandewavirus) and nege-like viruses (Kallies et al., 2014; Shi et al., 2016; Nunes et al., 2017; Ramos-González et al., 2020). Recent detailed phylogenetic analysis clearly showed that higrevirus RdRp are closer to cilevirus RdRp than to corresponding blunervirus protein, and these viral RdRp sequences form two separated branches in a phylogenetic dendrogram (Kondo et al., 2020; Ramos-González et al., 2020).

SP24 is the most conserved protein among non-replicative polypeptides encoded by all kitaviruses (Nunes et al., 2017; Ramos-González et al., 2020). Moreover, SP24 was found in negeviruses, many nege-like viruses (Kondo et al., 2020), and in unrelated arthropod-infecting

viruses, such as chroparaviruses (Chronic Bee Paralysis Virus) (Kuchibhatla et al., 2014), which, unlike kitaviruses and negeviruses that represent order *Martellivirales*, show genome similarity to arthropod tombus-like viruses (Shi et al., 2016) and belong to separate virus order *Tolivirales* (Koonin et al., 2020). SP24-like proteins (Pfam: 16504) have several trans-membrane domains and a less conserved N-terminal region enriched with positively charged residues. It has been proposed that SP24 molecules could directly interact in virus particles not only with lipids but also with viral RNA due to its positively charged N-terminal region (Kuchibhatla et al., 2014; Solovyev and Morozov, 2017).

Interestingly, two related arthropod viruses with largest genomic RNAs (ca. 16Kb) among nege-like viruses, *Pyrrhocoris apterus* virus 1 (accession MK024711) and Wuhan heteroptera virus 1 (NC_033461) each encode in their monopartite genomes three proteins distantly related to SP24 among seven non-replicative proteins (**Supplementary Table 1**) (Shi et al., 2016; Koloniuk and Vinokurov, 2019).

The available data on the RdRp and SP24 evolutionary lineages strongly suggest the evolutionary connections between insect nege-like viruses and plant kitaviruses and occurrence of cross-kingdom transfer of positive-stranded RNA viruses between insects and plants during evolution (Kondo et al., 2019, 2020; Ramos-González et al., 2020; Quito-Avila et al., in press). In general, the current views suggest a close temporal parallelism between the evolutionary development of land plants and terrestrial arthropods that started over 450 million years ago and contributed significantly to rapid insect radiation (Kenrick et al., 2012; Labandeira, 2013; Morris et al., 2018). Remarkably, many modern negative-stranded bunyaviruses and rhabdoviruses as well as double-stranded reoviruses using arthropods as vectors have developed an ability to replicate in species from two kingdoms (Whitfield et al., 2018; Chen et al., 2019). These observations agree with the evolutionary inter-kingdom transfer from arthropods to plants recently observed for retrotransposons (Lin et al., 2016; Gao et al., 2018).

It is currently assumed that multiple inter-kingdom jumps between hosts created the modern plant and insect positive-stranded RNA virus diversity, which started its expansion from arthropod hosts as a primary reservoir (Dolja and Koonin, 2018; Shi et al., 2018). In line with this, some present-day insect positive-stranded RNA viruses can replicate in isolated plant cells (like most kitaviruses) or even move systemically in whole plants like tea plant necrotic ring blotch blunervirus (Dasgupta et al., 2001; Annamalai et al., 2008; Hao et al., 2018; Jiwaji et al., 2019). Nege-like plant viruses give us an additional example of inter-kingdom jump and adaptation of typical arthropod-infecting virus to plant host. Indeed, Fragaria vesca-associated virus 1 isolated from plants displaying yellow spot and mosaic symptoms has a single genomic RNA with gene arrangement typical for some nege-like viruses and encodes a replication protein and four non-replicative proteins, two of which are moderately similar to proteins (including SP24) of nege-like viruses, particularly, Hubei Wuhan insect virus 9, barley aphid RNA virus 3, Aphis glycines virus 3 and barley aphid RNA virus 2 (Lenz et al., 2020). Importantly, our previous TBLASTn search of NCBI plant Transcriptome Shotgun Assemblies (TSA) (see Supplementary Figure in Solovyev and Morozov, 2017) revealed several plant-specific "Virus-like RNA assemblies" (VLRAs) encoding replication-like domains moderately similar to nege-like viruses and homologs of SP24 protein (particularly, *Amaranthus tuberculatus* VLRA GGGT01091955; *Humulus lupulus* VLRA GAAW01021049; *Triticum polonicum* VLRA GEDQ01066052).

Outside nege-kitavirus representatives, identification of a unique plant virus in gentian (*Gentiana* sp., asterids; Gentianales; Gentianaceae) strongly argues in favor of quite recent horizontal virus transfer between plant and arthropod hosts. This RNA virus—Gentian Kobu-sho-associated virus (GKaV)—is discovered in a hyperplastic (tumorous disorder-affected) Japanese gentians, possesses a very large genomic RNA (22 Kb) and codes for a polyprotein of more than 810 kDa with prominent sequence similarity to proteins encoded by flaviviruses (Atsumi et al., 2013; Kobayashi et al., 2013). Particularly, phylogeny of viral replicative SF2 (superfamily 2) helicase and RdRp domains revealed most close relation of GKaV to arthropod-infecting flavi-like viruses (Matsumura et al., 2016). Our BLAST analysis of replicative helicase alone also showed that GKaV was most similar to Hermitage virus (accession AMO03217), Apis flavivirus (YP_009388303), Takaungu virus (AMO03219), Lampyris noctiluca flavivirus 1 (QBP37018), Diaphorina citri flavi-like virus (YP_009259672), and Wuhan centipede virus (YP_009254745). It is likely that GKaV is not unique among plant viruses because TBLASTn search of NCBI plant TSA revealed GKaV-related RdRp-coding sequences in *Croton tiglium* (rosids; Malpighiales; Euphorbiaceae) (accession GGDV01007611) (data not shown).

NOVEL PUTATIVE BLUNERVIRUS-LIKE RNAs IN THE TRANSCRIPTOME OF *PAULOWNIA TOMENTOSA*

Our TBLASTn search of very recent NCBI plant TSA data collection resulted in the identification of a new quadripartite set of VLRAs in *Paulownia tomentosa* (dicots, order Lamiales) having a similarity to blunervirus RNAs. Three of these *P. tomentosa* VLRAs exhibit a gene arrangement quite similar to that in RNAs 1, 2, and 3 of blunerviruses (**Figures 1A,B**). Pairwise BLASTP alignment of the replicative proteins encoded by RNA1 and RNA2 showed their obvious similarity to kitavirus polypeptides (**Supplementary Table 1**).

In contrast to replicative proteins, most *Paulownia tomentosa* VLRA 3-encoded polypeptides in have only marginal similarity to kitavirus non-replicative proteins. Only SP24 is moderately similar to analogs from kita- and nege-like viruses (**Figure 1B** and **Supplementary Table 1**). Other proteins show only rather short amino sequence motifs allowing motif-based sequence comparisons with some kitaviral polypeptides (**Supplementary Table 2**). Particularly, p31 protein (**Figure 1B**) contains a rather long C-terminal hydrophobic segment similar to that of cilevirus p61, which exhibits some features of glycoprotein (Leastro et al., 2018), whereas p30 protein includes

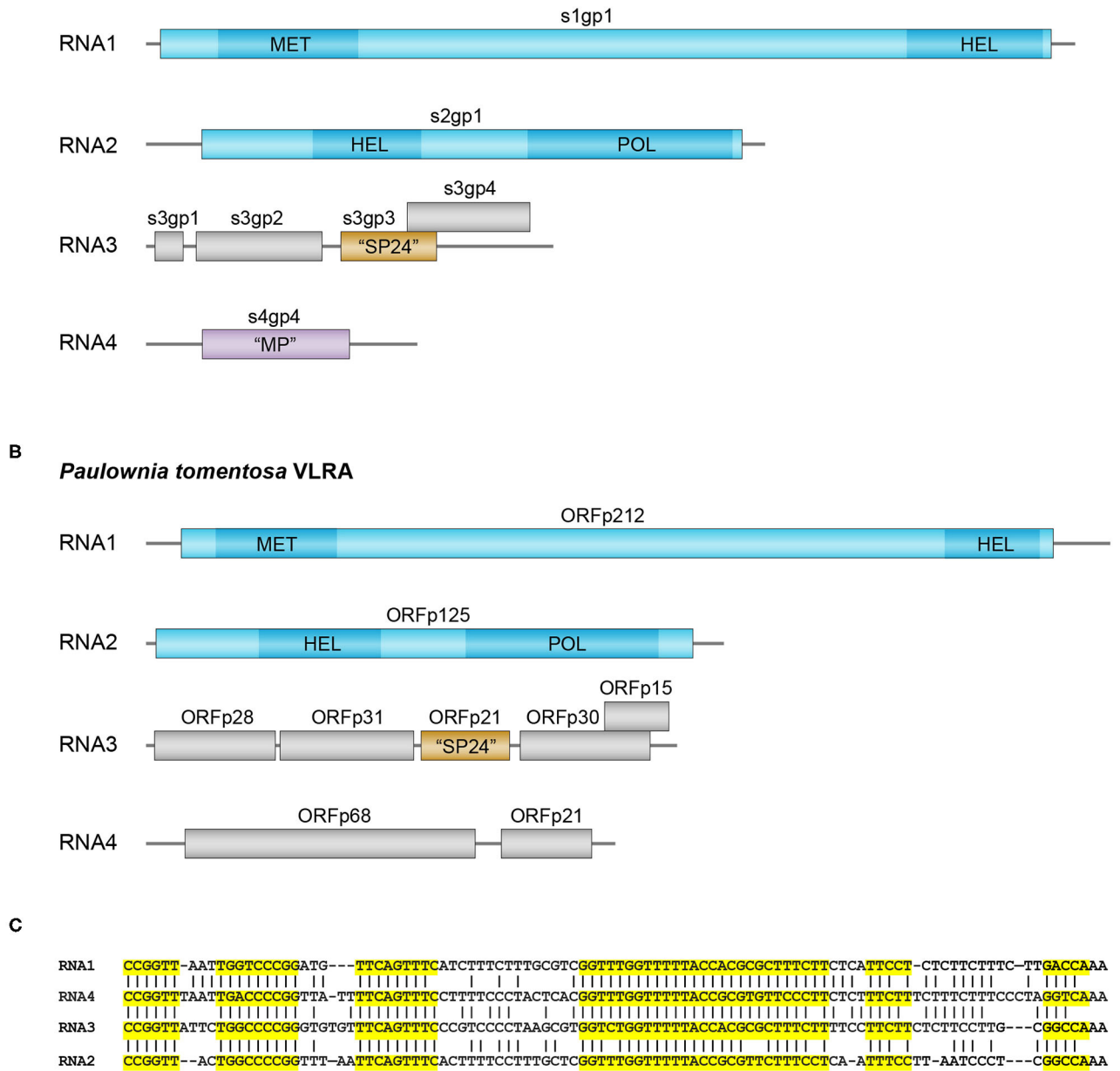


FIGURE 1 | Genomes and sequence analysis of blunerviruses. **(A)** Genome structure of the blueberry necrotic ring blotch blunervirus. Genomic segments (left end corresponds to 5′ end, whereas right end corresponds to 3′ end) encoding putative replication-related proteins are shown in blue, the MET, SF1HEL, and RdRp domains of the putative replication proteins are shown in dark blue. Predicted SP24 ORF protein in segment 3 is in brown. **(B)** Genome structure of the predicted blunervirus from *Paulownia tomentosa*. Putative replication-related genomic segments are shown in blue, the MET, SF1 HEL, and RdRp domains of the putative replication proteins are shown in dark blue. Predicted SP24 ORF protein in segment 3 is in brown. **(C)** Nucleotide sequence alignment of the 3′-terminal regions in the predicted VLRA RNAs from *Paulownia tomentosa*. Highly conserved RNA blocks are highlighted by yellow background.

a common motif with RNA3-encoded p22 of tea plant necrotic ring blotch virus (**Supplementary Table 2**).

The forth bicistronic *P. tomentosa* VLRA 4 encodes no proteins with significant similarity to other kitavirus non-replicative polypeptides, although pairwise BLASTN analysis of the 3′-untranslated regions from four *P. tomentosa* VLRAs indicated a high degree of sequence conservation among them and strongly suggested that the bicistronic VLRA 4 is indeed the component of the single virus-like multipartite genome (**Figure 1C**). VLRA 4 proteins could be hardly attributed to viral MPs, because these proteins show no similarity to the blunervirus-encoded orhtologs of well-known MPs found in

ilarvirus genomes (*Bromoviridae*) (Quito-Avila et al., 2013, in press; Hao et al., 2018; Bujarski, in press). Instead, VLRA 4 protein p68 (**Figure 1B**) contains a sequence motif quite similar to the nuclear localization signal (NLS) of cucumoviral 2b proteins known to have silencing suppressor functions (Du et al., 2014; Bujarski, in press) (**Supplementary Table 2**). Thus it can be speculated that VLRAs discovered by *in silico* analyses of *Paulownia tomoentosa* transcriptome represent a genome of a putative kitavirus that encodes silencing suppressors (Leastro et al., 2020).

Sequence analyses of blunervirus genomes reveal their unusual organization, as two of their four genomic RNAs (RNA1 and RNA2) encode SF1 helicases. Most probably, at least one of the blunervirus SF1 helicases represents a replicative protein. Particularly, RNA1-encoded helicase is a part of protein combining MT and HEL domains similarly to the replicative RNA1-encoded protein in the members of *Bromoviridae* (Hao et al., 2018). According to the phylogenetic trees generated using RdRp sequences, these viruses show relatedness to kitaviruses. Phylogenetic tree showed that all cilevirus and higrevirus RdRp domains are grouped as a single cluster, and RdRp domains of blunervirus replicative proteins also form a separate branch (Kondo et al., 2020; Ramos-González et al., 2020). Accordingly, pairwise sequence comparisons revealed that SF1 helicase domains of Met-HEL and HEL-RdRp proteins encoded by *P. tomentosa* VLRAs like BNRBV Met-HEL helicase are most similar to other blunervirus helicases and nege-like virus helicases (**Supplementary Table 1**). Strikingly, among six most similar relatives of BNRBV HEL-RdRp helicase SF1 helicases of grapevine leafroll-associated virus 7 (belongs to family *Closteroviridae*), barley aphid RNA negev-like virus 4 and "jiviviruses" (**Supplementary Table 1**), which is a recently discovered group of positive-stranded plant RNA viruses represented by capsid-less members with three genomic RNAs and named after similar jingmen-like viruses and virga-like viruses (Matsumura et al., 2017; Chiapello et al., 2020). "Jivivirus" RNA1 (around 4 Kb in length) codes for a Met-HEL protein similar in pairwise sequence comparisons to the protein encoded by RNA1 in members of *Bromoviridae* and *Kitaviridae*, whereas RNA2 (around 3 Kb) encodes an RdRp domain also showing relationship to closteroviruses and negev-like viruses (**Supplementary Table 1**). However, phylogenetic trees of "jivivirus" Met-HEL and RdRp domains revealed obvious similarity to virga-like viruses (Chiapello et al., 2020). Most strikingly, jivivirus RNA3 (around 2 Kb) encodes an SF2 RNA helicase with prominent sequence similarity to flavi-like insect jingmen viruses which contain the respective gene in RNA3 of a quadripartite genome (Chiapello et al., 2020; Garry and Garry, 2020) (**Supplementary Figure 1** and **Supplementary Table 1**). Thus, the structure of "jivivirus" genomes formally corresponds to a reassortant between virga-like and flavi-like viruses and represents a second example among plant viruses when a single genome encodes both SF1 and SF2 RNA helicases. The first such recombinant viral genome has been found for monopartite positive-stranded RNA viruses of the alga genus *Chara*, where SF1 replicative helicase is related to benyvirus proteins, and SF2 helicase is similar to flavi-like viruses (Gibbs et al., 2011; Vlok et al., 2019). Our TBLASTn search of NCBI plant TSA database revealed VLRAs similar to jivivirus genomes in a broad variety of seed plants including members of classes Magnoliopsida, Pinopsida and Ginkgoopsida (**Supplementary Table 1** and Legend to **Supplementary Figure 1**). These results suggest that a recombination/reassortment event between virga-like and flavi-like viruses might have occurred at rather early stages of seed plant evolution.

Based on these results and those presented in our previous papers (Morozov and Solovyev, 2012, 2015; Lazareva et al., 2017), we conclude that it seems logical to consider blunerviruses, and kitaviruses in general, as natural genetic chimeric systems combining in their multipartite genomes viral RNAs of different origin and showing traces of recombination even within individual genomic components.

AUTHOR CONTRIBUTIONS

SM collected and analyzed the data and authored drafts of the paper. EL prepared figure and reviewed the final draft. AS authored drafts of the paper, prepared figure, and reviewed the final draft. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Russian Foundation for Basic Research (grant 18-54-15003).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.561092/full#supplementary-material

REFERENCES

Annamalai, P., Rofail, F., Demason, D. A., and Rao, A. L. (2008). Replication-coupled packaging mechanism in positive-strand RNA viruses: synchronized coexpression of functional multigenome RNA components of an animal and a plant virus in Nicotiana benthamiana cells by agroinfiltration. *J. Virol.* 82, 1484–1495. doi: 10.1128/JVI.01540-07

Atsumi, G., Tomita, R., Kobayashi, K., and Sekine, K. T. (2013). Prevalence and genetic diversity of an unusual virus associated with Kobu-sho disease of gentian in Japan. *J. Gen. Virol.* 94, 2360–2365. doi: 10.1099/vir.0.053637-0

Bujarski, J. J. (in press). Bromoviruses (*Bromoviridae*). *Reference Module in Life Sciences*. doi: 10.1016/B978-0-12-809633-8.21563-X

Chen, Y., Dessau, M., Rotenberg, D., Rasmussen, D. A., and Whitfield, A. E. (2019). Entry of bunyaviruses into plants and vectors. *Adv. Virus Res.* 104, 65–96. doi: 10.1016/bs.aivir.2019.07.001

Chiapello, M., Rodríguez-Romero, J., Nerva, L., Forgia, M., Chitarra, W., Ayllón, M. A., et al. (2020). Putative new plant viruses associated with *Plasmopara viticola* infected grapevine samples. *Ann. Appl. Biol.* 176, 180–191. doi: 10.1111/aab.12563

Dasgupta, R., Garcia, B. H. 2nd, and Goodman, R. M. (2001). Systemic spread of an RNA insect virus in plants expressing plant viral movement protein genes. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4910–4915. doi: 10.1073/pnas.081288198

Dolja, V. V., and Koonin, E. V. (2018). Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus Res.* 244, 36–52. doi: 10.1016/j.virusres.2017.10.020

Du, Z., Chen, A., Chen, W., Liao, Q., Zhang, H., Bao, Y., et al. (2014). Nuclear-cytoplasmic partitioning of cucumber mosaic virus protein 2b determines the balance between its roles as a virulence determinant and an RNA-silencing suppressor. *J. Virol.* 88, 5228–5541. doi: 10.1128/JVI.00284-14

Gao, D., Chu, Y., Xia, H., Xu, C., Heyduk, H., Abernathy, B., et al. (2018). Horizontal transfer of non-LTR retrotransposons from arthropods to flowering plants. *Mol. Biol. Evol.* 35, 354–364. doi: 10.1093/molbev/msx275

Garry, C. E., and Garry, R. F. (2020). Proteomics computational analyses suggest that the envelope glycoproteins of segmented jingmen flavi-like viruses are class II viral fusion proteins (b-Penetrenes) with mucin-like domains. *Viruses* 12:260. doi: 10.3390/v12030260

Gibbs, A. J., Torronen, M., Mackenzie, A. M., Wood, J. T., Armstrong, J. S., Kondo, H., et al. (2011). The enigmatic genome of chara australis virus. *J. Gen. Virol.* 92, 2679–2690. doi: 10.1099/vir.0.033852-0

Hao, X., Zhang, W., Zhao, F., Liu, Y., Qian, W., Wang, Y., et al. (2018). Discovery of plant viruses from tea plant (*Camellia sinensis* (L.) *O. kuntze*) by metagenomic sequencing. *Front. Microbiol.* 9:2175. doi: 10.3389/fmicb.2018.02175

Jiwaji, M., Matcher, G. F., de Bruyn, M. M., Awando, J. A., Moodley, H., Waterworth, D., et al. (2019). Providence virus: an animal virus that replicates in plants or a plant virus that infects and replicates in animal cells? *PLoS ONE* 14:e0217494. doi: 10.1371/journal.pone.0217494

Kallies, R., Kopp, A., Zirkel, F., Estrada, A., Gillespie, T. R., Drosten, C., et al. (2014). Genetic characterization of goutanap virus, a novel virus related to negeviruses, cileviruses and higreviruses. *Viruses* 6, 4346–4357. doi: 10.3390/v6114346

Kenrick, P., Wellman, C. H., Schneider, H., and Edgecombe, G. D. (2012). A timeline for terrestrialization: consequences for the carbon cycle in the palaeozoic. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 519–536. doi: 10.1098/rstb.2011.0271

Kobayashi, K., Atsumi, G., Iwadate, Y., Tomita, R., Chiba, K.-i., Akasaka, S., et al. (2013). Gentian Kobu-sho-associated virus: a tentative, novel double-stranded RNA virus that is relevant to gentian kobu-sho syndrome. *J. Gen. Plant Pathol.* 79, 56–63. doi: 10.1007/s10327-012-0423-5

Koloniuk, K. S., and Vinokurov, I. (2019). Discovery and characterization of a novel alphavirus-like RNA virus from the red firebug *Pyrrhocoris apterus* L. (Heteroptera). *J. Invertebr. Pathol.* 166:107213. doi: 10.1016/j.jip.2019.107213

Kondo, H., Chiba, S., Maruyama, K., Andika, I. B., and Suzuki, N. (2019). A novel insect-infecting virga/nege-like virus group and its pervasive endogenization into insect genomes. *Virus Res.* 262, 37–47. doi: 10.1016/j.virusres.2017.11.020

Kondo, H., Fujita, M., Hisano, H., Hyodo, K., Andika, I. B., and Suzuki, N. (2020). Virome analysis of aphid populations that infest the barley field: the discovery of two novel groups of nege/kita-like viruses and other novel RNA viruses. *Front. Microbiol.* 11:509. doi: 10.3389/fmicb.2020.00509

Koonin, E. V., Dolja, V. V., Krupovic, M., Varsani, A., Wolf, Y. I., Yutin, N., et al. (2020). Global organization and proposed megataxonomy of the virus world. *Microbiol. Mol. Biol. Rev.* 84, e00061–19. doi: 10.1128/MMBR.00061-19

Kuchibhatla, D. B., Sherman, W. A., Chung, B. Y., Cook, S., Schneider, G., Eisenhaber, B., et al. (2014). Powerful sequence similarity search methods and in-depth manual analyses can identify remote homologs in many apparently "orphan" viral proteins. *J. Virol.* 88, 10–20. doi: 10.1128/JVI.02595-13

Labandeira, C. C. (2013). A paleobiologic perspective on plant-insect interactions. *Curr. Opin. Plant Biol.* 16, 414–421. doi: 10.1016/j.pbi.2013.06.003

Lazareva, E. A., Lezzhov, A. A., Komarova, T. V., Morozov, S. Y., Heinlein, M., and Solovyev, A. G. (2017). A novel block of plant virus movement genes. *Mol. Plant Pathol.* 18, 611–624. doi: 10.1111/mpp.12418

Leastro, M. O., Castro, D. Y. O., Freitas-Astúa, J., Kitajima, E. W., Pallás, V., and Sánchez-Navarro, J. Á. (2020). Citrus leprosis virus C encodes three proteins with gene silencing suppression activity. *Front. Microbiol.* 11:1231. doi: 10.3389/fmicb.2020.01231

Leastro, M. O., Kitajima, E. W., Silva, M. S., Resende, R. O., and Freitas-Astua, J. (2018). Dissecting the subcellular localization, intracellular trafficking, interactions, membrane association, and topology of citrus leprosis virus C proteins. *Front. Plant Sci.* 9:1299. doi: 10.3389/fpls.2018.01299

Lenz, O., Pribylová, J., Fránová, J., and Koloniuk, I. (2020). Fragaria vesca-associated virus 1: a new virus related to negeviruses. *Arch. Virol.* 165, 1249–1252. doi: 10.1007/s00705-020-04603-9

Lin, X., Faridi, N., and Casola, C. (2016). An ancient transkingdom horizontal transfer of penelope-like retroelements from arthropods to conifers. *Genome Biol. Evol.* 8, 1252–1266. doi: 10.1093/gbe/evw076

Matsumura, E. E., Coletta-Filho, H. D., de Oliveira-Dorta, S., Nouri, S., and Machado, M. A. (2016). Genetic structure and molecular variability analysis of citrus sudden death-associated virus isolates from infected plants grown in Brazil. *Viruses* 8:330. doi: 10.3390/v8120330

Matsumura, E. E., Coletta-Filho, H. D., Nouri, S., Falk, B. W., Nerva, L., Oliveira, T. S., et al. (2017). Deep sequencing analysis of RNAs from citrus plants grown in a citrus sudden death-affected area reveals diverse known and putative novel viruses. *Viruses* 9:92. doi: 10.3390/v9040092

Morozov, S. Y., and Solovyev, A. G. (2012). Did silencing suppression counter-defensive strategy contribute to origin and evolution of the triple gene block coding for plant virus movement proteins? *Front. Plant Sci.* 3:136. doi: 10.3389/fpls.2012.00136

Morozov, S. Y., and Solovyev, A. G. (2015). Phylogenetic relationship of some 'accessory' helicases of plant positive-stranded RNAviruses: toward understanding the evolution of triple gene block. *Front. Microbiol.* 6:508. doi: 10.3389/fmicb.2015.00508

Morris, J. L., Puttick, M. N., Clark, J. W., Edwards, D., Kenrick, P., Pressel, S., et al. (2018). The timescale of early land plant evolution. *Proc. Natl. Acad. Sci. U.S.A.* 115, E2274–E2283. doi: 10.1073/pnas.1719588115

Nunes, M. R., Contreras-Gutierrez, M. A., Guzman, H., Martins, L. C., Barbirato, M. F., Savit, C., et al. (2017). Genetic characterization, molecular epidemiology, and phylogenetic relationships of insect-specific viruses in the taxon negevirus. *Virology* 504, 152–167. doi: 10.1016/j.virol.2017.01.022

Quito-Avila, D. F., Brannen, P. M., Cline, W. O., Harmon, P. F., and Martin, R. R. (2013). Genetic characterization of blueberry necrotic ring blotch virus, a novel RNA virus with unique genetic features. *J. Gen. Virol.* 94, 1426–1434. doi: 10.1099/vir.0.050393-0

Quito-Avila, D. F., Freitas-Astúa, J., and Melzer, M. J. (in press). "Bluner-, cile-, and higreviruses (*Kitaviridae*)," in *Reference Module in Life Sciences* (Elsevier). doi: 10.1016/B978-0-12-809633-8.21248-X

Ramos-González, P. L., Dos Santos, G. F., Chabi-Jesus, C., Harakava, R., Kitajima, E. W., and Freitas-Astúa, J. (2020). Passion fruit green spot virus genome harbors a new orphan ORF and highlights the flexibility of the 5'-End of the RNA2 segment across cileviruses. *Front. Microbiol.* 11:206. doi: 10.3389/fmicb.2020.00206

Shi, M., Lin, X. D., Chen, X., Tian, J. H., Chen, L. J., Li, K., et al. (2018). The evolutionary history of vertebrate RNA viruses. *Nature* 556, 197–202. doi: 10.1038/s41586-018-0012-7

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

Solovyev, A. G., and Morozov, S. Y. (2017). Non-replicative integral membrane proteins encoded by plant alpha-like viruses: emergence of diverse orphan ORFs and movement protein genes. *Front. Plant. Sci.* 8:1820. doi: 10.3389/fpls.2017.01820

Vlok, M., Gibbs, A. J., and Suttle, C. A. (2019). Metagenomes of a freshwater charavirus from british columbia provide a window into ancient lineages of viruses. *Viruses* 11:299. doi: 10.3390/v11030299

Whitfield, A. E., Huot, O. B., Martin, K. M., Kondo, H., and Dietzgen, R. G. (2018). Plant rhabdoviruses – their origins and vector interactions. *Curr. Opin. Virol.* 33, 98–207. doi: 10.1016/j.coviro.2018.11.002

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 29 April 2021
doi: 10.3389/fmicb.2021.641252

Molecular Epidemiology of Citrus Leprosis Virus C: A New Viral Lineage and Phylodynamic of the Main Viral Subpopulations in the Americas

Camila Chabi-Jesus[1,2], Pedro L. Ramos-González[2], Matheus Postclam-Barro[2], Rafaela Salgado Fontenele[3], Ricardo Harakava[2], Renato B. Bassanezi[4], Alecio S. Moreira[4,5], Elliot W. Kitajima[2], Arvind Varsani[3,6] and Juliana Freitas-Astúa[2,5]**

[1] Escola Superior de Agricultura "Luiz de Queiroz", University of São Paulo, São Paulo, Brazil, [2] Instituto Biológico/IB, São Paulo, Brazil, [3] The Biodesign Center for Fundamental and Applied Microbiomics, Center for Evolution and Medicine, School of Life Sciences, Arizona State University, Tempe, AZ, United States, [4] Fundo de Defesa da Citricultura, Araraquara, Brazil, [5] Embrapa Mandioca e Fruticultura, Cruz das Almas, Brazil, [6] Structural Biology Research Unit, Department of Integrative Biomedical Sciences, University of Cape Town, Observatory, Cape Town, South Africa

Despite the importance of viral strains/variants as agents of emerging diseases, genetic and evolutionary processes affecting their ecology are not fully understood. To get insight into this topic, we assessed the population and spatial dynamic parameters of citrus leprosis virus C (CiLV-C, genus *Cilevirus*, family *Kitaviridae*). CiLV-C is the etiological agent of citrus leprosis disease, a non-systemic infection considered the main viral disorder affecting citrus orchards in Brazil. Overall, we obtained 18 complete or near-complete viral genomes, 123 complete nucleotide sequences of the open reading frame (ORF) encoding the putative coat protein, and 204 partial nucleotide sequences of the ORF encoding the movement protein, from 430 infected *Citrus* spp. samples collected between 1932 and 2020. A thorough examination of the collected dataset suggested that the CiLV-C population consists of the major lineages CRD and SJP, unevenly distributed, plus a third one called ASU identified in this work, which is represented by a single isolate found in an herbarium sample collected in Asuncion, Paraguay, in 1937. Viruses from the three lineages share about 85% nucleotide sequence identity and show signs of inter-clade recombination events. Members of the lineage CRD were identified both in commercial and non-commercial citrus orchards. However, those of the lineages SJP were exclusively detected in samples collected in the citrus belt of São Paulo and Minas Gerais, the leading Brazilian citrus production region, after 2015. The most recent common ancestor of viruses of the three lineages dates back to, at least, ∼1500 years ago. Since citrus plants were introduced in the Americas by the Portuguese around the 1520s, the Bayesian phylodynamic analysis suggested that the ancestors of the main CiLV-C lineages likely originated in contact with native vegetation of South America. The intensive expansion of CRD and SJP lineages in Brazil started probably linked to the beginning of the local citrus industry. The high prevalence of CiLV-C in the citrus belt of Brazil likely ensues from the intensive connectivity between orchards, which represents a potential risk toward pathogen saturation across the region.

Keywords: citrus leprosis disease, *Cilevirus*, *Kitaviridae*, *Brevipalpus* mites, virus evolution

OPEN ACCESS

Edited by:
Akio Adachi,
Kansai Medical University, Japan

Reviewed by:
Xuefeng Wang,
Chinese Academy of Agricultural Sciences, China
Lauro Velazquez-Salinas,
Plum Island Animal Disease Center, Agricultural Research Service, United States Department of Agriculture, United States
Avijit Roy,
United States Department of Agriculture (USDA), United States

***Correspondence:**
Pedro L. Ramos-González
plrg1970@gmail.com
Juliana Freitas-Astúa
juliana.astua@embrapa.br

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *13 December 2020*
Accepted: *06 April 2021*
Published: *29 April 2021*

Citation:
Chabi-Jesus C, Ramos-González PL, Postclam-Barro M, Fontenele RS, Harakava R, Bassanezi RB, Moreira AS, Kitajima EW, Varsani A and Freitas-Astúa J (2021) Molecular Epidemiology of Citrus Leprosis Virus C: A New Viral Lineage and Phylodynamic of the Main Viral Subpopulations in the Americas. Front. Microbiol. 12:641252. doi: 10.3389/fmicb.2021.641252

INTRODUCTION

Brazil is the leading sweet orange producer in the world. With almost 197.7 million sweet orange (*Citrus x sinensis (L.) Osbeck*) trees[1], the citrus belt São Paulo (SP)—Minas Gerais (MG) is the largest citrus cultivation area in South America and accounts for more than 80% of the Brazilian sweet orange production (Bassanezi et al., 2019). Citrus orchard yields may be impacted by citrus leprosis (CL) disease, ranked first among the viral diseases affecting this crop in Brazil (Ramos-González et al., 2018). Control of CL reaches up to US$ 54 million/year, a value representing about 5% of the management cost of orchards in the main Brazilian citrus belt (Bassanezi et al., 2019).

Despite the multi-etiological character of CL, citrus leprosis virus C (CiLV-C) is, by far, the prevalent causal agent in Brazil (Ramos-González et al., 2016, 2017, 2018; Chabi-Jesus et al., 2018). The virus infects several species within the genus *Citrus* and their hybrids, although with different degrees of severity. While sweet oranges show high susceptibility, mandarins (*C. reshni*, *C. reticulata,* and *C. deliciosa*) are moderately resistant, and lemons (*C. limon*) and limes (*C. aurantifolia*) are considered resistant (Bastianel et al., 2018). CiLV-C also naturally infects *Commelina benghalensis* and *Swinglea glutinosa* and can be experimentally transmitted to plants of 28 families (León et al., 2008; Nunes et al., 2012; Garita et al., 2014; Arena et al., 2017).

Citrus leprosis virus C is the type species of the genus *Cilevirus,* family *Kitaviridae* (Locali-Fabris et al., 2006, 2012; Freitas-Astúa et al., 2018). In addition to cileviruses, the family also includes members of the genera *Higrevirus* and *Blunervirus* (Melzer et al., 2012; Quito-Avila et al., 2013). Kitaviruses have bacilliform or spherical virions, divided positive-sense single-stranded RNA genomes, and likely share common ancestors with arthropod-infecting viruses of the group negevirus and nege/kita-like viruses (Roy et al., 2015; Kondo et al., 2020; Quito-Avila et al., 2020; Ramos-González et al., 2020).

Aside from CiLV-C, the genus *Cilevirus* also includes citrus leprosis virus C2 and passion fruit green spot virus (PfGSV) (Roy et al., 2013; Ramos-González et al., 2020). The canonical cilevirus genome comprises six open reading frames (ORFs) split into two molecules, RNA1 and RNA2. RNA1 is ∼9.0 kb in length and includes two ORFs encoding the RNA-dependent RNA polymerase (*RdRp*) and the putative coat protein (*p29*). RNA2 is ∼5.0 kb in length and has four ORFs (*p15*, *p61*, *p32,* and *p24*). In CiLV-C, the RNA2 also contains an intergenic region (IR) of ∼1 kb located between the ORFs *p15* and *p61*. P15, P61, and P24 are proteins without definitively associated functions, although the first two seem to be involved in the suppression of the RNA silencing mechanism (Leastro et al., 2020) and the latter one is conserved among cileviruses, higreviruses, and an increasing number of arthropod-infecting viruses (Kuchibhatla et al., 2014; Kondo et al., 2020; Ramos-González et al., 2020). The *p32* encodes a movement protein (MP) of the 30K superfamily (Mushegian and Elena, 2015; Leastro et al., 2021).

[1]https://www.fundecitrus.com.br/

CiLV-C does not systemically infect its host plants, it only causes local chlorotic and/or necrotic lesions in leaves, fruits, and branches (**Figure 1**), which may result from an incompatible interaction led by a hypersensitivity-like response (Arena et al., 2016, 2020). The viral spread, even to different points in an infected plant, is exclusively mediated by viruliferous mites of, mainly, the species *Brevipalpus yothersi* (Ramos-González et al., 2016). *B. papayensis* is also able to transmit the virus under experimental conditions (Nunes et al., 2018). Nonetheless, the CiLV-C/*Brevipalpus* spp. interaction has not been fully characterized yet. While biological and electron microscopy data suggest a circulative transmission (Kitajima et al., 2003, 2008; Tassi et al., 2017), non-conclusive molecular assays suggested the viral multiplication in the *Brevipalpus* cells. Negative-sense CiLV-C viral genomes detected in mite extracts (Roy et al., 2015) could have been remnants from the infected plant cells after mite feeding (Tassi et al., 2017). Moreover, the specific detection of the viral negative-strand RNA by reverse transcription-polymerase chain reaction (RT-PCR) could have been the result of false amplification due to either self-priming of the positive-strand RNA or the primer activity of other cellular nucleic acids (Haddad et al., 2007; Boncristiani et al., 2009; Haist et al., 2015; Strydom and Pietersen, 2018).

Preliminary studies revealed that the CiLV-C population has a low genetic variability ($\pi < 0.01$) and is subdivided into the clades CRD and SJP (Ramos-González et al., 2016). Type viruses of each lineage share ∼85% genome nucleotide identity, except the 5′-ends of their RNA2 molecules. With ∼98% nucleotide sequence identity, the high uniformity of the genomic segments compressing the *p15-IR* regions is likely a consequence of a natural recombination process (Ramos-González et al., 2016). The lineage CRD is prevalent throughout Latin America,

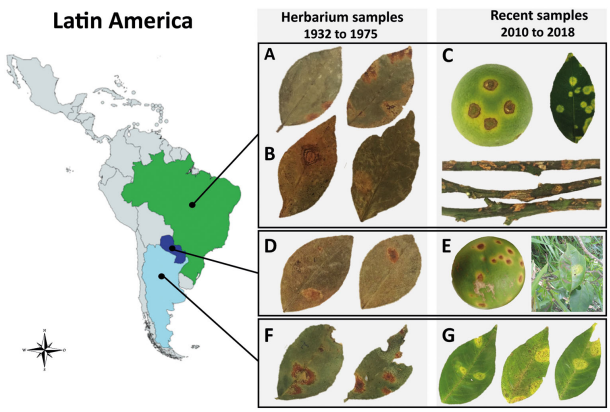


FIGURE 1 | Symptoms of citrus leprosis (CL) disease and place of collection of the infected *Citrus* spp. samples across Latin America. **(A,B,D,F)** Samples conserved at the Herbarium of Instituto Biológico, São Paulo, Brazil. **(C,E,G)** Fresh citrus samples. **(A)** CiLV-C isolate Jbt02, Jaboticabal, São Paulo, Brazil, 1975. **(B)** CiLV-C isolate Urg01, Uruguaiana, Rio Grande do Sul, Brazil, 1937. **(C)** CiLV-C isolate Bar25, São Paulo, Brazil, 2018. **(D)** CiLV-C isolate Asu02 from Asunción, Paraguay, 1937. **(E)** CiLV-C isolate PY03, Paraguay, 2010 (fruit) and 2015 (leaf). **(F)** CiLV-C isolate Ar06, Misiones, Argentina, 1937. **(G)** CiLV-C isolate AR04, Corrientes, Argentina, 2017.

whereas, until 2015, the lineage SJP was only detected in three counties in the northwestern region of the state of São Paulo, Brazil (Ramos-González et al., 2016). However, most aspects concerning the diversity, distribution, transmissibility, and virulence of these strains remain largely unknown.

In this study, we investigated the distribution, dynamic and evolutionary parameters of the CiLV-C population through the analysis of 430 fresh or herbarium samples of CL-affected *Citrus* spp. tissues collected from commercial or non-commercial citrus orchards between 1932 and 2020. We also reconstructed the evolutionary history of CiLV-C and contextualized it with the origin and expansion of citrus crops in the Americas.

MATERIALS AND METHODS

Citrus Samples

A total of 430 individual or mixed lesions from leaves, fruits, or branches were collected from 304 citrus plants showing typical chlorotic and/or necrotic symptoms of CL (**Table 1** and **Supplementary Table 1**). The RNA extracts were obtained from: (*i*) eight sweet orange (*Citrus sinensis*) samples stored at the Herbarium of Instituto Biológico, São Paulo, collected from 1932 to 1975 in Brazil ($n = 6$), Argentina ($n = 1$), and Paraguay ($n = 1$); *ii*) 41 leaf samples of *Citrus* spp. stored in −80°C freezer collected from 2003 to 2015 in Brazil ($n = 31$), Argentina ($n = 6$), Bolivia ($n = 1$), Paraguay ($n = 1$), and Colombia ($n = 1$); (*iii*) 37 sweet orange fruit samples collected from commercial citrus orchards in the citrus belt SP-MG in the period 2015–2016; (*iv*) 18 leaf samples of *Citrus* spp. collected from non-commercial citrus orchards in Brazil ($n = 10$), Argentina ($n = 7$), and Paraguay ($n = 1$) in the period 2015–2019; (*v*) 325 fruit lesions from 199 sweet orange trees collected in 196 commercial citrus orchards in the citrus belt SP-MG in the period 2017–2020; and (*vi*) one lesion from a sweet orange fruit collected in a commercial organic orchard, State of Pará, Brazil, in 2020.

RNA Isolation

RNA extraction was performed either from fresh or herbarium plant tissues. For fresh samples, about 100 mg of leaf lesions were ground in liquid nitrogen and the total RNA was extracted using Trizol® according to the manufacturer's recommendation (Thermo Fisher Scientific, Waltham, MA, United States). For the herbarium samples, in addition to the treatment with 0.01% diethylpyrocarbonate (DEPC) solution and 120°C sterilization, mortars and pestles were kept in an oven at 200°C for 48 h before the extractions. Approximately 600 mg of dry symptomatic tissues were ground in liquid nitrogen and processed following the Trizols® procedure modified as previously described (Ramos-González et al., 2016). Regardless of the origin of samples, final RNA solutions were precipitated using 0.1 volume of sodium acetate 3 M and 2.5 volume of isopropanol, kept at −80°C for 12 h, and centrifuged at 10,000 × g for 10 min at 4°C. The concentration and quality of the RNA extracts were assessed by NanoDrop ND8000 spectrophotometer (Thermo Fisher Scientific), and 1.2% agarose gel stained with ethidium bromide (10 mg/mL) or Bioanalyzer 2100 (Agilent Technologies, Santa Clara, United States), respectively. For samples collected in commercial citrus orchards from 2017 to 2020, the total RNA extracts were obtained from a single lesion found on the affected fruits. With this, we aimed to reduce the interference resulting from a putative intra-host virus variability and to detect whether

TABLE 1 | Summary of the set of *Citrus* spp. samples gathered in this study.

Place of collection	Orchard type	Year of collection	Number of analyzed samples (individual or mixed lesions)					
			CiLV-C strains					Total
			CRD	SJP	ASU[b]	CDR+SJP same tree[c]	CDR+SJP same lesion[d]	
Brazil								
SP and MG[a]	Non-commercial	1932–2020	18	2	–	0	0	20
	Commercial	2003–2014	12	0	–	0	0	12
		2015–2016	7	33	–	3	–	43
		2017–2020	42	207	–	2	74	325
Other states	Non-commercial	1937–2018	10	0	–	0	0	10
	Commercial	2020	1	0	–	0	0	1
Other countries								
Argentina	Non-commercial	1937–2019	14	0	–	0	0	14
Colombia	Non-commercial	2008	1	0	–	0	0	1
Bolivia	Non-commercial	2003	1	0	–	0	0	1
Paraguay	Non-commercial	1937–2019	2	0	1	0	0	3
Total			108	242	1	5	74	430

CiLV-C strains were detected by reverse transcription-polymerase chain reactions and/or high-throughput sequencing.
[a] *Brazilian states abbreviations: São Paulo (SP) and Minas Gerais (MG).*
[b] *Identification by HTS, giving that all available primers were not able to identify this strain.*
[c] *Members of the clades SJP and CRD detected in samples from the same tree.*
[d] *Members of the clades SJP and CRD detected in the same CL lesion.*

viruses belonging to more than a clade could be infecting a single lesion (detailed in **Supplementary Table 1**).

Detection of CiLV-C and Other Citrus Leprosis Symptom Producing Viruses by RT-PCR

Five hundred nanograms of total RNA were used for cDNA synthesis in a final reaction volume of 20 μL using the RevertAid H Minus First-Strand cDNA Synthesis Kit (Thermo Fisher Scientific). The presence of CL-associated viruses was assessed by PCR using cDNA as template (3 μL), specific primer pairs (**Table 2**), and GoTaq G2 Master Mix Green kit (Promega, Madison, WI, United States). For CiLV-C, in addition to primers for the detection of *p29* (Ramos-González et al., 2016) and *p32* (Locali et al., 2003), a set of strain-specific and degenerate primers were developed based on available GenBank *p24* sequences (CiLV-C RNA2 of the isolates SJP01 and Crd01: KP336747 and NC008170, respectively). To do this, sequences were aligned using MUSCLE implemented in MEGA version 7.0.21 (Kumar et al., 2016), and primers were designed using Geneious software platform version 11.1.4 (Kearse et al., 2012) (**Table 2**). The thermal cycles were as follows: 94°C, 3 min; 35 cycles of 94°C, 30 s; 54°C, 30 s; 72°C, 30 s; and a final extension at 72°C for 5 min. To confirm the specificity of primers, the amplicons were resolved on a 1% agarose gel, excised, purified, and Sanger sequenced. The putative presence of the cilevirus citrus leprosis virus C2 (CiLV-C2) (Roy et al., 2013) and the dichorhaviruses citrus leprosis virus N (CiLV-N) (Ramos-González et al., 2017), citrus chlorotic spot virus (CiCSV) (Chabi-Jesus et al., 2018), and orchid fleck virus (OFV) (Kubo et al., 2009) were screened by PCR using previously described primers (**Table 2**).

Partial Sequencing of CiLV-C Isolates

The complete sequence of *p29* (795 nts) and partial sequence of *p32* (288 nts) in the RNA1 and RNA2, respectively, of CiLV-C isolates, were obtained using described primers (Locali et al., 2003; Ramos-González et al., 2016; **Table 2**). Amplicons were obtained from 26 samples collected in non-commercial citrus regions in Brazil and Argentina, from 2006 to 2019, and 31 from commercial citrus orchards inside the citrus belt SP-MG, in the period 2017 to 2019 (**Supplementary Table 1**). After RT-PCR, amplicons were purified using Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, United States), and cloned into pGEM-T-Easy (Promega, Madison, WI, United States). Plasmids were transformed into *Escherichia coli* DH10β competent cells by electroporation, and 5–10 recombinant clones derived from each sample were sequenced by the Sanger method (Instituto Biológico, SP, Brazil). PCR products from some samples collected in non-commercial orchards (**Supplementary Table 1**) were directly sequenced after the purification using the Wizard SV Gel and PCR Clean-Up System (Promega).

High-Throughput Sequencing of CiLV-C Genomes

Small RNA (sRNA) from the herbarium samples were sequenced on an Illumina HiSeq 2500 system (Illumina, San Diego, United States) either at Genewiz (South Plainfield, NJ, United States) or BGI (Shenzhen, Guangdong, China) (**Table 3**). For most of the non-herbarium samples, total RNA extracts

TABLE 2 | Primer list used for the detection of viruses associated with citrus leprosis disease by reverse transcription-polymerase chain reaction.

Virus[a]	Target	Primer sequence (5′–3′)	Ta[b] (°C)	Amplicon size (bp)	References
CiLV-C	*p32*	F: GCGTATTGGCGTT GGATTTCTGAC	56	339	Locali et al., 2003
		R: TGTATACCAAGCC GCCTGTGAACT			
	p29	F: ACCGTGAATTTGT ATTTTGTCA		1,000	Ramos-González et al., 2016
		R: CAGCTGGAAGAGA CTAGAAA			
	p15	F: GTCAAGTGATATCC ATTTTGCTTG		667	Ramos-González et al., 2016
		R: TCATCGTCTTTTC TGTAACCG			
	p24	F: CGCAGTTTCCTAA TAACACC	54	322	This study
		R: GCTTTATGCTGAA CTCCC			
CiLV-C CRD	*p29*	F: CAGAAGGCCGAGG TTGTAAAG	56	330	Ramos-González et al., 2016
		R: GTAGTGATCACT GAACTCGAATACC			
	p24	F: ATGTTGGCAACG GAAAGTT	54	522	This study
		R: GTGAACAGGGTTG AAAAAGTT			
CiLV-C SJP	*p29*	F: GTAARCAAAAGG TCGAGGTTGTCC	56	456	Ramos-González et al., 2016
		R: TCTGTTGTCTAGC AGCRAGTAATG			
	p24	F: CTCATGATATCCTTG ATGACC	54	393	This study
		R: GACTAATAAGGTT GAGAAGGTTG			
CiLV-C2	*p29*	F: ATGAGTAACATTG TGTCGTTTTCGTTGT	56	795	Roy et al., 2013
		R: TCACTCTTCCTGTT CATCAACCTGTT			
OFV +CiLV-N	*L*	F: CAASTGTCATGCC TGCATGG	54	362	Ramos-González et al., 2017
		R: TTGATRCATGATG CRAGRCTGTATG			
CiCSV	*G*	F: CTGTTTTGCCCAT GCTAC		500	Chabi-Jesus et al., 2018
		R: CCTCCTCTTCTAG CGTCAT			

[a] *Virus or clade-specific viruses identified by each primer pair. Virus acronyms: CiLV-C, citrus leprosis virus C; CiLV-C CRD, lineage CRD; CiLV-C SJP, lineage SJP; CiLV-C2, citrus leprosis virus C2; OFV, orchid fleck virus; CiLV-N, citrus leprosis virus N; CiCSV, citrus chlorotic spot virus.*

[b] *Ta: PCR annealing temperature. F and R indicate forward and reverse primers, respectively.*

TABLE 3 | Brief description of the 18 high-throughput sequencing (HTS) libraries obtained in this study.

Host species/ variety	Local of collection city/state/ country	Year of collection	Sequencing company	HTS library	Number of reads	RNA molecule	CiLV-C-derived reads and contigs			
							% of viral-derived reads in the library	Number of assembled contigs	Contig length range (nts)	Viral coverage (%)
Herbarium samples										
Citrus sinensis	Jacarei, SP, BR[2]	1932	Genewiz, United States	siRNA	59,435,658	1	0.01	15	150–1423	99.3
						2	0.01	9	148–2116	99.6
C. sinensis (Washington Navel)	Piracicaba, SP, BR	1932	Genewiz, United States	siRNA	71,037,882	1	0.01	6	121–4168	100
						2	0.01	3	1478–1916	100
C. sinensis	Uruguaiana, RS, BR	1937	BGI, China	siRNA	63,116,147	1	2.1	3	235–7026	100
						2	1.1	2	24499–2317	100
C. sinensis (Washington Navel)	Asuncion, PY	1937	BGI, China	siRNA	61,507,832	1	20.2	6	403–3709	100
						2	14.6	4	1085–3781	100
C. sinensis	Santa'Ana, Misiones, AR	1937	BGI, China	siRNA	62,116,588	1	6.5	2	427–8144	100
						2	4.3	1	4785	100
C. sinensis (Washington Navel)	Limeira, SP, BR	1939	Genewiz, United States	siRNA	74,407,978	1	0.02	6	187–7577	100
						2	0.01	4	4850–4858	100
C. sinensis (Pera)	Sao Paulo, SP, BR	1941	BGI, China	siRNA	59,404,138	1	0.05	19	147–988	100
						2	0.05	9	144–707	100
Citrus sp.	Jaboticabal, SP, BR	1975	Genewiz, United States	siRNA	56,790,128	1	16.4	22	117–4750	100
						2	7.7	19	125–4428	100
Sample stored at −80°C										
C. sinensis	AR	2006	Esalq, USP, Brazil	mRNA	14,872,149	1	14.2	23	154–8724	100
						2	2.5	5	162–5276	100
Samples from fresh tissues										
C. sinensis	Piracicaba, SP, BR	2016	Esalq, USP, Brazil	mRNA	15,553,431	1	5.5	2	630–8147	100
						2	1.2	1	4963	100
C. sinensis	Corrientes, AR	2017	Esalq, USP, Brazil	mRNA	14,410,186	1	1.6	1	8747	100
						2	1.4	2	292–4742	100
C. reticulata	Piracicaba, SP, BR	2018	Esalq, USP, Brazil	mRNA	15,165,394	1	0.9	3	416–8755	100
						2	0.2	1	4975	100
C. sinensis	Sud Mennucci, SP, BR	2018	Esalq, USP, Brazil	mRNA	16,377,921	1	20.6	32	93–821	100
						2	6.8	30	58–892	100

(Continued)

TABLE 3 | Continued

Host species/ variety	Local of collection city/state/ country	Year of collection	Sequencing company	HTS library	Number of reads	RNA molecule	CiLV-C-derived reads and contigs			
							% of viral-derived reads in the library	Number of assembled contigs	Contig length range (nts)	Viral coverage (%)
C. sinensis	Vitoria, ES, BR	2018	Esalq, USP, Brazil	mRNA	13,111,075	1	6.7	8	181–7800	100
						2	1.6	1	4968	100
C. sinensis	Capitão Poço, PA, BR	2020	Esalq, USP, Brazil	mRNA	15,822,801	1	3.8	6	151–7781	100
						2	2.2	7	218–2345	100
C. sinensis	Limeira, SP, BR	2020	Esalq, USP, Brazil	mRNA	15,390,982	1	26.9	23	69–4377	100
						2	11.5	34	104–1031	100
C. reticulata	Jumirim, SP, BR	2020	Instituto Biológico, SP, Brazil	siRNA	10,557,136	1	13.3	8	250–3340	100
						2	6.7	4	470–1428	100
C. sinensis	Santo Antônio da Posse, SP, BR	2020	Esalq, USP, Brazil	mRNA	16,571,118	1	9.9	27	82–7820	100
						2	6.5	24	58–892	100

[a]*Country: BR, Brazil; AR, Argentina; PY, Paraguay.*

were processed at the Laboratory of Animal Biotechnology of the University of São Paulo (Piracicaba, SP, Brazil). Poly(A) enrichment of the RNA extracts and cDNA libraries were prepared with Illumina TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, United States). Sequencing was performed in an Illumina HiSeq 2500 system using HiSeq SBS v4 High Output Kit (Illumina, San Diego, United States). Paired-end reads of 2 × 125 bp were generated. The viral sequence in the sample BR_SP_Jmr01 was obtained using the Ion GeneStudio^{TM} S5 System (Thermo Fisher Scientific) at the Instituto Biológico, São Paulo, Brazil. The sRNA library from this sample was obtained using Ion Total RNA-Seq Kit v2 (Thermo Fisher Scientific). The quality of reads obtained by all the methodologies was checked using FastQC (Andrews, 2010) and the adaptor sequences were removed using the Trimmomatic (Bolger et al., 2014). Reads from all types of libraries were assembled with SPAdes (Bankevich et al., 2012) although using different *k*-mer sizes: 15, 17, 19 for sRNA libraries, and 33, 43, 55 for poly(A)-enriched RNA libraries. Viral contigs were identified using the Basic Local Alignment Search Tool (BLASTx and/or BLASTn) implemented in Geneious using a local database including viral reference genomes retrieved from the NCBI virus database[2] (Hatcher et al., 2017). After the identification and when necessary, reads were mapped to the reference genomes in an iterative mapping approach (Tsai et al., 2010) using Bowtie2 or BBMap to fill gaps and extend the end sequences of viral genomes (Langmead and Salzberg, 2012; Bushnell, 2014).

Recombination and Reassortment Analyses

Recombination events were assessed using seven methods (RDP, GENECONV, Bootscan, Maxchi, Chimaera, SiScan, and Topal) implemented in RDP version 5.5 (Martin et al., 2015) and GARD (Kosakovsky Pond et al., 2006). Sequences were aligned using the MUSCLE, MAFFT, and Clustal software, implemented in Geneious. Recombination events detected by more than three programs ($p \leq 0.05$) implemented in the RDP vs. 5.5 were considered as recombinants. Due to the length heterogeneity of sequences available, four independent analyses were carried out: (*i*) complete sequences of each CiLV-C genome (8,984 nts of RNA1 and 5,077 nts of RNA2; $n = 23$); (*ii*) *p29* (RNA1) (795 nts; $n = 190$); (*iii*) partial *p32* (RNA2) (288 nts; $n = 270$); (*iv*) and a partial RNA2 concatenated sequences [complete *p15* ORF (393 nts)—intergenic region (934 nts, upstream the *p61* ORF)—partial *p32* (288 nts); $n = 56$]. The reassortment events were analyzed based on the topology of the phylogenetic trees. All CiLV-C sequences available at the GenBank were retrieved and incorporated into this and further *in silico* analyses (**Supplementary Table 1**).

Phylogenetic Analyses Based on Complete Genomes and CiLV-C ORFs

Nucleotide sequence alignments were performed using MUSCLE implemented in MEGA version 7.0.21 (Kumar et al., 2016). Best-fit models for nucleotide substitutions were determined with

[2]https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/

Bayesian Information Criterion (BIC) implemented in MEGA version 7.0.21 (Kumar et al., 2016). They were as follow: model GTR+G for RNA1 (n = 23), HKY+G+I for both the RNA2 (n = 23) and the concatenated partial RNA2 sequences (*p15*-IR-*p32*, n = 56), and HKY+G for the nucleotide sequences alignments of *p29* (n = 190) and *p32* (n = 270). Phylogenetic trees were generated by Bayesian inference using a variant of Markov chain Monte Carlo (MCMC) with MrBayes, implemented in Geneious (Huelsenbeck and Ronquist, 2001; Kearse et al., 2012), with 6,000,000 generations and cognate sequences from CiLV-C2_Colombia (NC038848 and NC038849) as outgroup. Genomic regions involved in the recombination events were excluded before the phylogenetic tree building to minimize their influence on tree topologies. Trees were viewed and edited using iTOL version 4 (Letunic and Bork, 2019). Nucleotide distances within and between clades were calculated using MEGA version 7.0.21.

Temporal Phylogenetic Analyses of the CiLV-C Population

Assessment of the time to the most recent common ancestor (tMRCA) of CiLV-C isolates was carried out using BEAST software version 1.10.04 (Suchard et al., 2018). Two datasets were evaluated. They comprised (*i*) the concatenated sequences of all CiLV-C ORFs (except the *p15* of RNA2 because its putative recombinant origin) (11,473 nts; n = 23 isolates) and (*ii*) concatenated sequences of the complete *p29* and the partial *p32* ORFs (1,083 nts, n = 132 isolates). In samples from which more than one haplotype was sequenced, only those showing divergent sequences were included in the analyses (**Supplementary Table 2**), but nucleotide diversity inside a given sample was always lower than 0.007. "Non-clock" maximum likelihood phylogenetic trees were reconstructed with the best evolutionary model (TN93+G) using IQtree software version 1.5.5 (Nguyen et al., 2015), and the temporal signal was evaluated by TempEst.

To assess the evolutionary history of the CiLV-C population, the Bayesian Markov Chain Monte Carlo (MCMC) was estimated using the BEAST version 1.10.4 The best model of nucleotide substitution for the two analyzed datasets was TN93+G. The Bayesian skygrid model (number of parameters = 20; time of last transition point = 88) was selected as the tree coalescent model and using the strict clock. The MCMC analyses were performed with 100 million generations, sampling a tree every 1,000 steps. MCMC convergence was assessed by estimating the effective sample sizes (ESS) using Tracer version 1.7 (Rambaut et al., 2018). ESS > 100 are moderate values whilst values > 200 are considered better, according to the instruction manual of the software[3]. The maximum clade credibility (MCC) tree was created by discarding the initial 10% of the chains and summarized in TreeAnnotator version 1.10.4. The phylogenetic tree was viewed and edited using IcyTree (Vaughan, 2017).

Population Genetics and Selection Tests

Population genetic parameters, i.e., diversity of nucleotide (π) and haplotype (Hd), the number of polymorphic sites (s), nucleotide differences (k), average mutation rates (θ), and haplotypes (H); and the ratio of non-synonymous (dN) to synonymous (dS) nucleotide substitutions (ω = dN/dS) were calculated using DnaSP v. 6.12.03 (Rozas et al., 2017). Selection in polymorphic sites of the CiLV-C ORFs was calculated using Fast Unconstrained Bayesian AppRoximation for Inferring Selection (FUBAR), Fixed Effects Likelihood (FEL), and Mixed Effects Model of Evolution (MEME) methods with the GTR model, implemented in Datamonkey 2.0 (Weaver et al., 2018). For the identification of the amino acids under selection and their involvement in the protein structure, predicted secondary structures of deduced amino acid sequences of the P29 and MP proteins from definitive and tentative members of the genus *Cilevirus*: CiLV-C Crd01 (NC008169 and NC008170), CiLV-C_SJP01 (KP336746 and KP336747), CiLV-C2_Colombia (NC038848 and NC038849), hibiscus strain of CiLV-C2_Hawaii (MG253805 and MG253804) and PfGSV_Snp1 (MK804171 and MK804172) were obtained using PROMALS (PROfile Multiple Alignment with Local Structure) (Pei and Grishin, 2007).

Tests of Selective Neutrality and Differentiation in the CiLV-C Population

CiLV-C population expansion was evaluated by the statistical tests Tajima's D (Tajima, 1989), Fu and Li's F and D (Fu and Li, 1993), and Fu's FS (Fu, 1997), implemented in DnaSP v. 6.12.03 package (Rozas et al., 2017). They estimated the difference between two measures of genetic diversity, i.e., the mean number of pairwise differences and the number of segregating sites. In these tests, negative values denote populations in expansion or after a recent bottleneck, whereas positive values mean a decrease in population size and/or balancing selection.

Demographic expansions of CiLV-C subpopulations assessed by mismatch distributions (distribution of pairwise nucleotide differences) were performed based on the sum of squared deviation (SSD) and Harpending's Raggedness index (HRI) using Arlequin v. 3.5.2.2 (Excoffier and Lischer, 2010). The HRI test determines whether an observed mismatch distribution is drawn from an expanded (small raggedness index or non-significative) or a stationary population (large raggedness index), while the SSD quantifies the smoothness of the observed mismatch distribution and a non-significant result indicates an expanding population (Rogers and Harpending, 1992; Harpend, 1994).

Genetic subdivision of the CiLV-C population was assessed using the following tests implemented in DnaSP and Arlequin v. 3.5.2.2: the nearest-neighbor statistic (*Snn*) (Hudson, 2000), Hudson's test statistics [*Hst* (haplotype-based statistics), *Kst* (nucleotide-based statistics)], Wright's fixation index (*Fst*), and gene flow (*Nm*) (Hudson et al., 1992). The *Snn* is a measure of how often the nearest neighbors of sequences are found in the same locality (Hudson, 2000). *Snn* values range from 0.5 to 1, being the lowest indexes a sign that isolates from both locations are part of the same population and the highest ones that the populations in the two locations are highly differentiated. *Hst* and *Kst* statistics calculate the level of differentiation based on haplotypes and nucleotides, respectively, and values close to zero mean no differentiation. On the other hand, based on

[3]https://beast2.blogs.auckland.ac.nz/increasing-esss/

the proportion of the total genetic variance contained in a subpopulation, the *Fst* test provides insights into the evolutionary processes that influence the structure of genetic variation within and among populations (Hudson et al., 1992). *Nm*, the number of migrants successfully entering a population per generation, was used to measure gene flow (migration) between populations [FST≠1/(4*Nm*+1)]. Besides, the partitioning of variation at different levels was calculated by Analysis of Molecular Variance (AMOVA) in Arlequin using 1,000 permutations.

RESULTS

The presence of CiLV-C was confirmed in all the 430 symptomatic samples collected from 304 plants of sweet orange, six of mandarin, and in other five citrus plants whose species could not be determined (**Supplementary Table 1**). RT-PCR tests for the specific detection of citrus-infecting *Brevipalpus*-transmitted viruses other than CiLV-C indicated the absence of the cilevirus CiLV-C2 and the dichorhaviruses CiLV-N, CiCSV, and OFV (**Table 1**). Overall, this study included CiLV-C isolates collected from Argentina, Brazil, Bolivia, Colombia, and Paraguay during the period 1932–2020. Based on the high production volume and the large size of the farming area, ∼92% of the samples were collected from the citrus belt SP-MG, Brazil. All the analyzed samples showed typical symptoms of citrus leprosis disease, i.e., chlorotic and/or necrotic lesions on leaves and fruits and necrotic lesions on branches (**Figure 1**).

Near-Complete Genome Sequencing of New CiLV-C Isolates Reveals a Novel Divergent Strain

Total RNA extracts of leaves from eight herbarium, nine fresh, and one −80°C frozen-conserved samples were obtained (**Table 3**). RNA integrity number (RIN) of the extracts prepared from the herbarium samples was low, i.e., 1.6–2.1. Despite this, *de novo* assembling of the raw reads using the SPAdes enabled the recovery of more than 80–90% of the CiLV-C genomes from the high-throughput sequencing (HTS) libraries. Particularly, from the sample of sweet orange collected in 1941, in SP, Brazil, few and shorter contigs were obtained and only ∼60% of the CiLV-C genome could be determined. In this case, gaps between contiguous contigs were filled after a new round of assembling using BBMap and the CiLV-C genome as a reference. In sum, approaches combining *de novo* assemblies and the iterative mapping increased the genome coverages by about 100% in several samples. The lower genome coverage observed during the initial assembly steps of some herbarium samples seemed independent of the collection year and the laboratory where the HTS libraries were processed. Rather, it appeared to be intrinsically associated with the conservation procedure of every single sample, as previously observed (Hartung et al., 2015). Recovery rates of viral genomes higher than 98% were obtained from all fresh or −80°C conserved samples using the same wet lab and *in silico* procedures and tools. Overall, the complete or near-complete genomes of 18 studied isolates of CiLV-C were obtained.

The pairwise comparison of the genome sequences of the studied HTS CiLV-C isolates with those of the type viruses of the clades CRD (isolate Crd01, NC008169 and NC008170) and SJP (isolate SJP01, KP336746, and KP336747) showed values of nucleotide sequence identity that ranged from 84.0 to 99.8% and separates them into three groups (**Figure 2** and **Supplementary Table 3**). Fourteen isolates showed the highest identity values with the reference sequence of the clade CRD, three with that of the clade SJP, while the isolate CiLV-C_PY_Asu02, collected in Asunción, Paraguay in 1937, typified a novel diversity of CiLV-C and shared less than 86% nucleotide sequence identity with the reference genomes.

CiLV-C_PY_Asu02 shows the same genomic organization of the type member of the genus *Cilevirus*. The profile of nucleotide sequence identity across the genomes of the isolates PY_Asu02 and Crd01 follows the same pattern observed in the comparison between the type viruses of the clades SJP and CRD (**Figure 2**). Deduced amino acid sequences of ORFs from CiLV-C_PY_Asu02 and CiLV-C Crd01 show pairwise identity values ranging from 81% for P61, to 100% for P15, which resembles what is observed in the comparison among proteins from viruses of the clades SJP and CRD (**Table 4**). Notably, the stretch of nucleotide sequences at the 5′-end of the RNA2 in CiLV-C_PY_Asu02, which includes the ORF *p15* and part of the IR, is highly conserved between the isolates of the three clades, suggesting a common origin. CiLV-C_PY_Asu02 shows percentages of nucleotide sequence identity lower than 50% with the cileviruses CiLV-C2 and PfGSV (**Table 4**), confirming that the isolate PY_Asu02 represents a viral diversity previously unknown among members of the species *Citrus leprosis virus C*.

RNA2 of CiLV-C Strains Harbor Signals of Recombination

Putative signals of recombination events were detected in the RNA2 of CiLV-C using two datasets comprising the complete (n = 23) and partial (n = 56) sequences of the molecule. Partial RNA2 molecules contained the concatenate sequences of *p15*-IR-*p32*.

In the analyses of the complete RNA2 molecules, at least four out of seven programs implemented in the RDP version 5.5 identified recombinant events involving *p15* and the intergenic region (IR) (**Figure 3** and **Supplementary Table 4**). For instance, in the isolates BR_SP_SJP01, BR_SP_SJP05, BR_SP_Lim09, BR_SP_SAP03, and BR_SP_SdM15 of the clade SJP, as well as in the isolate CiLV-C_PY_Asu02, the seven programs suggested two breakpoints, one inside the IR and the second one closer to the 3′-end of their molecules. In both cases, isolates from the clade CRD (BR_SP_SPa11 or BR_RS_Urg01) were detected as the minor parents while the major parents were indeterminates. The third event, in the 3′-end of the RNA2 of CiLV-C_BR_SP_SJP01, involves CiLV-C_BR_SP_Lim09 as the major parent and an unknown minor parent.

When the alignments of the concatenated sequences were used, despite the artificial organization of the sequences, four events could be detected. The first two were identified in the *p15* and the IR of thirteen isolates belonging to the clade SJP, with

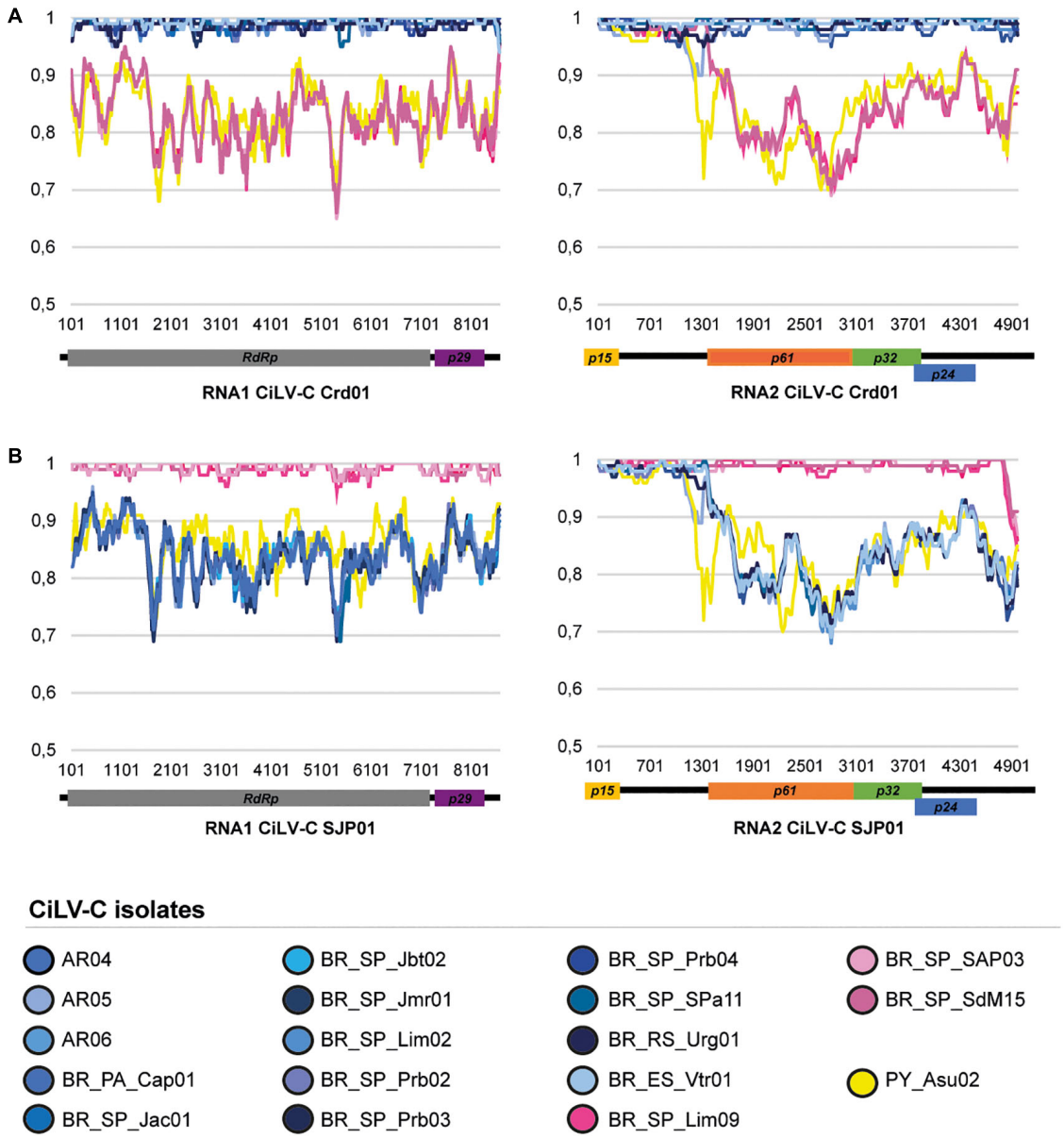


FIGURE 2 | Similarity plots of the nucleotide sequences of the RNA1 and RNA2 of the CiLV-C isolates. The full-length genome sequences of the 18 studied CiLV-C isolates were compared to those of the type-member viruses of the clades CRD **(A)** and SJP **(B)**. Curves depict the comparison between the analyzed and a reference genome. Each plotted point is the percent identity (vertical axis) within a sliding window 200 bp wide centered on the position plotted, with a step size between points of 20 bp. The horizontal axis indicates the nucleotide positions across the RNA1 and RNA2 molecules of the reference genomes. The horizontal bars above the curves are a cartoon of the ORFs of the CiLV-C genome. Plots were generated using SimPlot (Lole et al., 1999).

CiLV-C_BR_PA_Bel01 as the minor parent and unknown major parents; the third event was detected within the IR of the isolate CiLV-C_BR_PR_Mgf01, with the isolate CiLV-C_AR05 as the minor parent and an unknown major parent. Finally, the fourth event, detected in the isolate PY_Asu02, has breakpoints in the IR and the beginning of *p32*, and the isolate BR_PR_Ldb01 was indicated as the minor parent while the major parent could not be identified (**Supplementary Table 4**). Recombination events using separately either the partial sequences of the ORF *p32* ($n = 270$) and those of the complete ORF *p29* ($n = 190$) were not detected.

Phylogenetic and Genetic Analyses Support the Existence of Three Distinct Clades of CiLV-C: CRD, SJP, and ASU

Datasets grouping CiLV-C sequences generated in this work and those retrieved from GenBank were used for phylogenetic analyses. They comprised, in total, 23 complete or near-complete genomes, 167 complete *p29* sequences, and 247 partial *p32* sequences from isolates collected in six countries across Latin America in the period 1932–2020.

TABLE 4 | Nucleotide (nt) and deduced amino acid (aa) sequence identities between CiLV-C_PY_Asu02 and other members of the genus *Cilevirus*.

CiLV-C PY_Asu02[a]	CiLV-C Crd01		CiLV-C SJP01		CiLV-C2_Co		CiLV-C2_Hw		PfGSV_Snp1	
	nt	aa	nt	aa	nt	aa	nt	aa	nt	aa
RNA1	**86.2**	–	85.1	–	57.2	–	58.0	–	57.1	–
RdRp	86.3	93.4	**88.2**	**94.9**	59.6	58.86	59.8	58.7	59.2	57.9
p29	**86.0**	**94.0**	85.6	89.8	43.7	34.54	45.5	33.9	44.4	32.6
RNA2	**85.8**	–	85.1	–	44.3	–	43.3	–	43.4	–
p15	**98.5**	**100**	**98.5**	**100**	28.5	19.61	34.9	20.4	27.7	19.2
IR	**85.0**	–	84.0	–	29.1	–	31.4	–	26.9	–
p61	82.1	81.1	**82.9**	**83.8**	44.7	31.72	43.4	33.9	45.3	32.6
p32	**89.3**	**95.0**	86.5	94.0	55.3	51.16	54.9	50.5	52.5	53.9
p24	88.6	94.0	**89.2**	**94.4**	60.9	60.63	60.3	59.7	60.9	60.4

The highest values are highlighted in bold.

[a]*GenBank accession numbers of each isolate: CiLV-C_PY_Asu02: MT554532 (RNA1) and MT554546 (RNA2); CiLV-C Crd01: NC008169 and (RNA1) and NC008170 (RNA2); CiLV-C SJP01: KP336746 (RNA1) and KP336747 (RNA2); CiLV-C2_Co: NC038848 (RNA1) and NC038849 (RNA2); CiLV-C2_Hw: MG253805 (RNA1) and MG253804 (RNA2); and PfGSV_Snp1: MK804171 (RNA1) and MK804172 (RNA2).*

Bayesian phylogenetic reconstructions using the data sets of *p29* and *p32* showed three major branches, where the two largest ones encompassed the isolates of the previously identified lineages CRD and SJP (**Figure 4A** and **Supplementary Figure 1**). The third branch, supported with a high value of posterior probability (0.88) and hereafter called the lineage ASU, included only the isolate CiLV-C_PY_Asu02, collected in Asunción, Paraguay, in 1937. The subdivision in three clades of the CiLV-C population was also supported by the analysis of intra- and inter- clades genetic distances. Using both *p29* and *p32*, inter-clade genetic distances ($0.106 \leq d_{inter-clade} \leq 0.173$) were, generally, 11-fold higher than the intraclade distances ($d_{intra-clade} \leq 0.01$) (**Figure 4B**). Moreover, trees constructed with the complete genome sequences showed the same clade topology observed in the analyses using the independent ORFs (**Figure 4C**). It is noteworthy that for the construction of the RNA2 tree, the first 1,434 nts of the 5′-end of each molecule were removed considering the putative origin by recombination of this genomic region. Reassortment events between isolates from either different or the same clade were not observed.

Nucleotide (π) and haplotype diversity (Hd) values intrinsic to the CiLV-C population were calculated based on ORFs *p29* and *p32* (partial sequence) (**Table 5**). Although the nucleotide diversity of the whole population (SJP+CRD+ASU) was relatively very low ($\pi \sim 0.07$), its value was roughly 10-fold higher than those observed for the independent clades SJP and CRD ($\pi \sim 0.006$–0.01), and the pattern was similar regardless of the analyzed ORF, i.e., *p29* or *p32*. The haplotype diversity was close to 1 in any of the analyzed groups.

Viruses of the Clade CRD Are Spread Across the Continent Whereas Those of the Clade SJP Are Restricted to the Brazilian Citrus Belt SP-MG

Spatial and temporal distribution of isolates of the clades CRD and SJP were assessed by RT-PCR using clade-specific primers for the detection of *p29* (RNA1) and *p24* (RNA2) sequences. Results were screened for the spatial and temporal distribution of members of each clade (**Table 1**, **Figure 5**, and **Supplementary Table 1**). Unfortunately, the genomic characterization of the isolate CiLV-C_PY_Asu02 was obtained when most of the samples had been already evaluated. Therefore, information on the current distribution of ASU clade viruses, if they are still circulating, is not available. An *in silico* analysis indicated that in case viruses of the clade ASU might be present in the evaluated samples, primers used in the detection of members of the clades CRD and SJP would not be able to detect them, at least under the thermal cycling conditions performed in this study.

A broad-based analysis of the data indicated the presence of members of the clade CRD distributed between commercial (141/186; 75.8%) and non-commercial (45/186, 24.2%) orchards all over the period 1932–2020. In contrast, members of the clade SJP were almost exclusively found in commercial orchards in the citrus belt SP-MG (**Table 1** and **Figure 5**). In commercial orchards, from the samples collected during the period 2003–2020, 64% of lesions (240/375) were infected with viruses of the clade SJP, 16.2% with CRD (61/375), and 19.8% (75/375) exhibited mixed infections involving viruses of the two clades (**Figure 5**). Interestingly, out of commercial orchards, isolates of the clade SJP were detected in only two samples collected in a backyard tree in São José do Rio Preto, SP, in the neighborhood of the citrus belt SP-MG.

CiLV-C SJP and CRD Subpopulations Are Mainly Under Purifying Selection

Analysis based on the nucleotide sequences of *p29* and *p32* suggested purifying selection on the CiLV-C subpopulations CRD and SJP (**Table 5**). The strength of negative selection is weaker on *p32* ($\omega = 0.137$) than on *p29* ($\omega = 0.278$). When subpopulations were compared by each ORF, the highest values corresponded to *p29* of the clade CRD ($\omega_{p29} = 0.448$). Consistent with the low ω values detected, a large number of amino acids under purifying selection were identified across the amino acid sequences of P29 and MP, as supported by the FUBAR, FEL, and MEME programs implemented in Datamonkey 2.0 (Weaver et al., 2018;

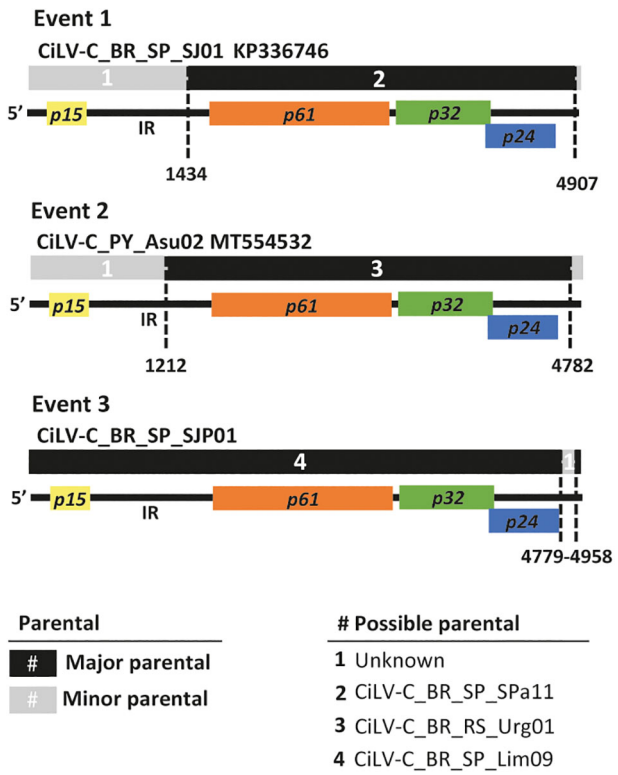


FIGURE 3 | Evidence of recombination in the RNA2 of CiLV-C. The diagram depicts three recombination events detected using the alignment of the completed RNA2 molecules of 23 isolates. Multicolored horizontal bars represent the genome of the recombinant molecules. Major and minor parentals are indicated by black and gray bars and numbers (#) as described in the legend. Dashed vertical lines indicate the recombination breakpoints using a cartoon representing the RNA2 molecule of CiLV-C as reference. Event 1 shows the recombination found in all isolates of the clade SJP, event 2 was detected in the isolate PY_Asu02 (clade ASU), and event 3 in BR_SP_SJP01 (clade SJP). Recombination events were detected by more than four programs implemented in the software RDP. Details of these and other recombination events are shown in **Supplementary Table 4**.

Supplementary Table 5). Much of the positions detected under negative selection using the FEL method matched with those described in a previous report (Ramos-González et al., 2016).

CiLV-C SJP and CRD Are Two Genetically Distinct and Expanding Subpopulations

Snn, *Fst*, *Kst,* and *Hst* tests were implemented to evaluate the genetic differentiation between the CiLV-C subpopulations. *Snn* and *Fst* values $\geq$ 1 showed a highly structured population with significant genetic differentiation between the clades CRD and SJP (**Table 6**). *Kst* and *Hst* values indicated a higher level of genetic differentiation considering the *p29* than the *p32* sequences. Gene flow (*Nm*) values calculated for both ORF sequences were smaller than 0.03 indicating that the gene flow among populations was infrequent or almost non-existent (**Table 6**).

Moreover, the molecular variance (AMOVA) test carried out for the detection of genetic differentiation between the CiLV-C subpopulations revealed the largest variance between the subpopulations CRD and SJP (∼93%), whereas it reached only 7% within each subpopulation (**Table 7**). With AMOVA statistic values close to 1, the *Fst* results based on *p29* and *p32* allowed us to refute the Null hypothesis of the non-differentiation between CiLV-C subpopulations.

Neutrality tests of the CiLV-C subpopulations were estimated using three statistic tests (Fu and Li's D and F, Fu's Fs, and Tajima's D) (**Table 8**). All values for *p29* and *p32* were negative or non-significant. These results suggested that the CiLV-C subpopulations are not neutral, but possibly expanding. To further address this question, we determined whether the data fit the sudden expansion model using the sum of square deviations (SSD) and Harpending's Raggedness index (HRI). Both SSD and HRI values were non-significant (**Table 8**), supporting the hypothesis of population expansion.

The Most Recent Common Ancestor of CiLV-C Lineages Dates Back, at Least, Approximately 1,500 Years Ago

To investigate the evolutionary dynamics and the time to the most recent common ancestor (tMRCA) of CiLV-C lineages, two MCC trees were constructed using two datasets. The topology of both trees was similar (**Figure 6** and **Supplementary Figure 2**). However, the tree using the concatenated *RdRd-p29-p61-p32-p24* reached only moderated values of ESS ($100 \leq$ ESS ≤ 200) for some statistic parameters, likely as a consequence of the lower number of sequences (n = 23). The MCC tree generated with the concatenated *p29-p32*, with a larger dataset (n = 132), showed a moderate temporal signal (correlation coefficient = 0.3 and R^2 = 0.26, **Supplementary Figure 3**), and ESS values always > 300. This MCC tree included sequences of isolates collected in Argentina (n = 4), Panama (n = 1), Paraguay (n = 1), and Brazil (n = 126) during the period 1932–2020. Therefore, we selected the tree generated with the concatenated *p29–p32* as the best representative of the evolutionary history of CiLV-C in the Americas.

Based on the concatenated *p29–p32*, the most recent common ancestor (MRCA) of the three CiLV-C lineages dates back to 500 A.D. [supported by a posterior probability (PP) of 1 and 95% highest probability density (HPD) of 115–875 years A.D.] (node A in **Figure 6**). This ancestral virus diverged into viruses that gave rise to two lineages: SJP and another one, intermediary, that subsequently diverged into the lineages ASU and CRD (node B). The MRCA of the lineages ASU and CRD dates back to 740 A.D. (PP = 0.87 and HPD 95% = 400–1061 years A.D.). Hence, the diversification events represented in nodes A and B overlap in the posterior probability distributions (**Figure 6**). Diversification of clade CRD happened around 160 years ago, ∼1860 A.D. (node C, PP = 1 and HPD 95% = 1812–1897 years A.D.), whereas the diversification of clade SJP occurred ∼1940 A.D., less than 100 years ago (node D, PP = 1 and HPD 95% = 1856–1918 years A.D.) (**Figure 6**). Despite the moderate ESS of the MCC based on the concatenated *RdRd-p29-p61-p32-p24*, for the sake of a better

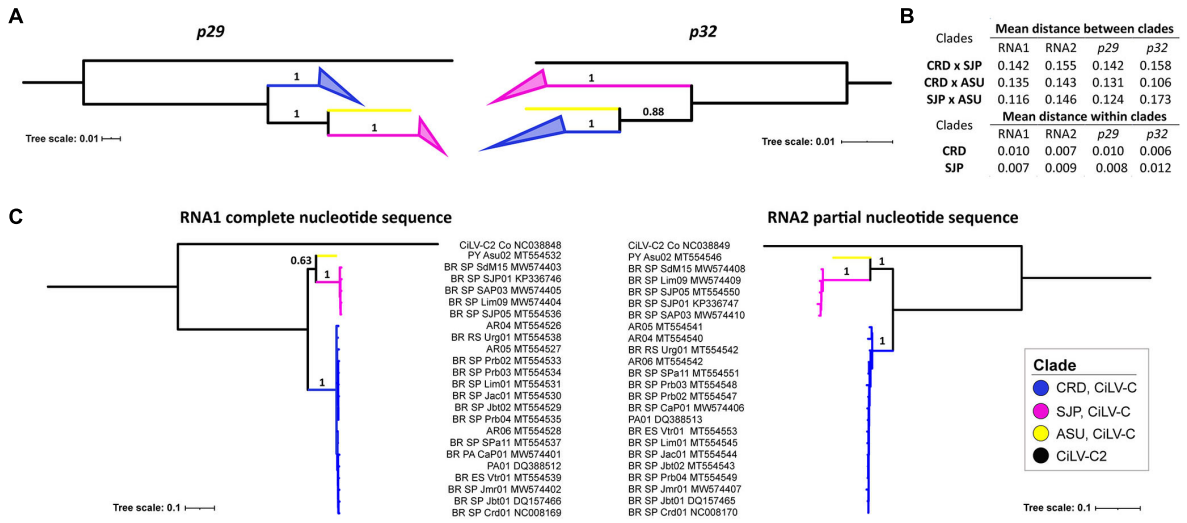


FIGURE 4 | Phylogenetic relationships and genetic distances among CiLV-C strains. Phylogenetic trees were generated by Bayesian inference using MrBayes and based on *p29* (190 isolates, 795 nts) and *p32* (270 isolates, 288 nts, partial) **(A)**, and the complete sequences of the RNA1 molecules and partial sequences of the RNA2 molecules (*p61*, *p32,* and *p24*) of 23 isolates of CiLV-C **(B)**. Inter- and intra-clade nucleotide distances were calculated considering independent ORFs and complete genomic molecules **(C)**. The maximum clade credibility trees were generated with 6,000,000 generations and using the appropriate sequences of CiLV-C2_Colombia (NC038848 and NC038849) as an outgroup. Tree branches were identified according to CiLV-C lineages.

understanding of the evolutionary viral process, the tMRCA of the three CiLV-C lineages was also assessed considering that dataset. The virus representing the MRCA of the main lineages dates back to ∼4000 B.C. (**Supplementary Figure 2**).

DISCUSSION

Citrus leprosis (CL) disease is a serious multi-etiological viral pathology affecting citrus in Latin America. In Brazil, the form of CL caused by the cilevirus CiLV-C is prevalent and, economically, the most detrimental virus-induced disease affecting the citrus orchards (Ramos-González et al., 2018). A former study reported the existence of two viral clades inside the population of CiLV-C (Ramos-González et al., 2016), but the ecology of these subpopulations could not be deeply assessed due to the limited number of viral isolate sequences available. In the present approach, which includes both fresh and herbarium citrus samples infected by CiLV-C, we obtained the complete and partial nucleotide sequences of the viral ORFs *p29* and *p32*, respectively, from 268 viral isolates. Among them, we revealed the near-complete genomes of 18 isolates, which were mostly collected during the first half of the last century.

Global analyses of the generated dataset by phylogenetic and populational statistical tools confirmed the existence of the two previously identified clades CRD and SJP (Ramos-González et al., 2016), whereas negative results of the Fu and Li's D and Tajima's D neutrality tests and the non-significative pairwise mismatch distribution indicated the expansion of the viral population or more likely a population that frequently undergoes bottlenecks (**Table 8**). Analyses also showed that CRD and SJP subpopulations are genetically well-differentiated ($Fst \geq 0.92$), have very low genetic diversity ($\pi \sim 0.01$) where almost each haplotype is unique ($0.8 < Hd < 0.99$), and, as a whole, are under purifying selection ($\omega < 0.5$). Such values are the quantitative expression ensuing from the sum of biological factors underlying the CL pathosystem, which, in general, may lead the CiLV-C subpopulations to continuous bottlenecks (Ramos-González et al., 2016). CiLV-C has a very limited natural known host range, infects a reduced number of cells around the inoculation sites, and, in nature, is exclusively transmitted by *Brevipalpus* mites, the only mean by which the virus can reach new infection foci even within a single leaf due to the absence of systemic movement capacity (Freitas-Astúa et al., 2018). Interestingly, the population of coffee ringspot virus, a dichorhavirus with some

TABLE 5 | Population genetics parameters and assessment of selection pressure for ORF *p29* (795 nts) and *p32* (288 nts, partial sequence) of CiLV-C.

ORF	Dataset (number of isolates)	H	Hd	π	ω (dN/dS)
p29	SJP+CRD+ASU (190)	159	0.9975	0.07519	0.278
	SJP+CRD (189)	158	0.9975	0.07501	0.294
	SJP (106)	87	0.995	0.00803	0.359
	CRD (83)	71	0.995	0.00994	0.448
p32	SJP+CRD+ASU (270)	70	0.930	0.07284	0.137
	SJP+CRD (269)	69	0.929	0.07224	0.165
	SJP (190)	49	0.917	0.01010	0.362
	CRD (80)	27	0.777	0.00631	0.145

Independent analyses were carried out with different combinations of viral isolates according to phylogenetic clades.
H, number of haplotypes; Hd, haplotype diversity; π, nucleotide diversity; and ω = dN/dS, ratio of non-synonymous (dN) to synonymous (dS) substitutions.

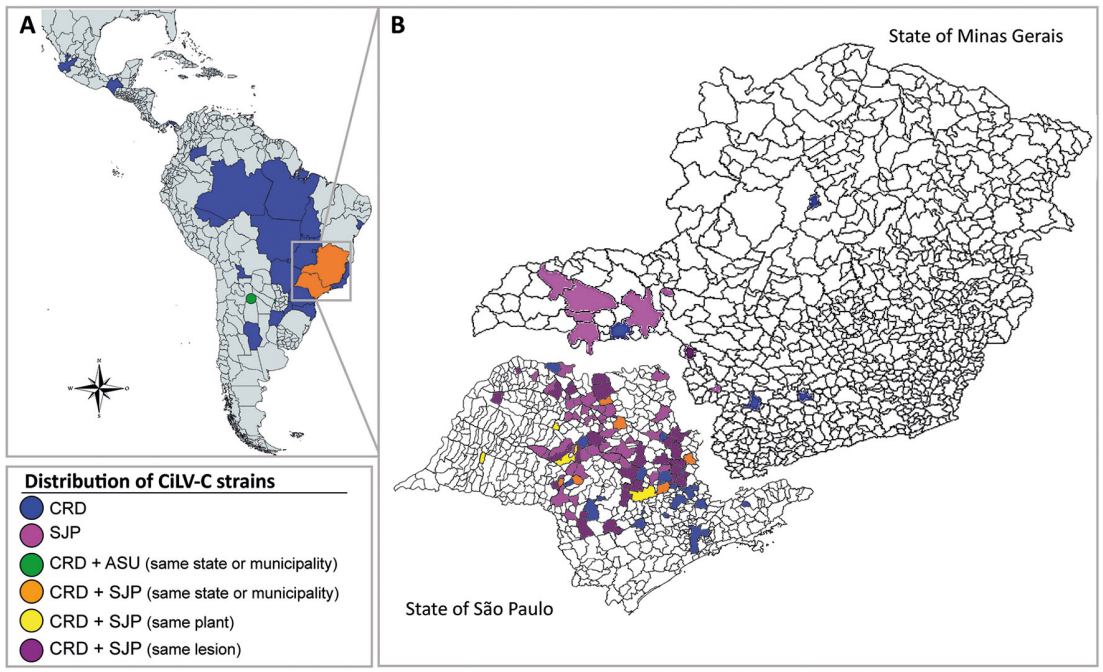


FIGURE 5 | Distribution of CiLV-C strains in Latin America **(A)** and across municipalities in the Brazilian states of São Paulo and Minas Gerais **(B)** in the period from 1932 to 2020. Viral detection was based on RT-PCR assays for specific identification of *p29* and *p24* from isolates of the CRD and SJP clade. CiLV-C_ASU was identified by high-throughput sequencing. The color-coded legend indicates the presence of a viral isolate, concomitantly or at a different time in the same plant, in single or mixed infections. The map also depicts the distribution of CiLV-C isolates whose sequences were already available in the GenBank database.

biological features comparable to those of CiLV-C, e.g., absence of systemic infection, few natural hosts, and *Brevipalpus* mite transmission, also displays low variability and purifying selection (Ramalho et al., 2016).

While the vast majority of the samples evaluated in this study were infected by CiLV-C isolates of the clades CRD and SJP, the herbarium sample collected in Paraguay, in 1937, revealed to be unique (**Figures 4**, **6**). In the phylogenetic analyses using either individual ORFs or complete genomic sequences, the isolate PY_Asu02 was separated into a third branch, called clade ASU (**Figure 4**). The genome of CiLV-C_PY_Asu02 has the typical organization of cileviruses, with roughly 86% nucleotide sequence identity with the reference genomes of the clades CRD and SJP. Notably, the stretch encompassing the 5′-end of its RNA2 shows high conservation with members of the other two clades and likely originated by recombination (**Figure 3**), as also observed in other CiLV-C strains (Ramos-González et al., 2016). Recombinant strains can result in viruses with improved virulence, best adaptability to a changing environment, or expansion in the host ranges and vector species (García-Arenal et al., 2001). Even though the role of the 5′-end of the CiLV-C RNA2 is not well understood yet, the highly conserved nucleotide sequence across viruses of the three CiLV-C clades (**Figure 2**) highlights the participation of this region in viral biology and, particularly, points out the involvement of recombination events shaping the genome of cileviruses.

The set of CiLV-C infected samples gathered in this work is the largest ever undertaken. Its composition is heterogeneous, showing temporal and geographical biases following the trend of the relative importance of citrus crop and the incidence of CL across Latin America. Despite this, the holistic analysis of the dataset discloses ecological relationships between members of the clades CRD and SJP whose extension and authenticity still need to be proven, e.g., (*i*) inside the citrus belt SP-MG, viruses of both lineages are unevenly distributed, (*ii*) members of the clade SJP are more frequently found in single infection (63%) than those of the CRD (16%), (*iii*) mixed infections in the same lesion were detected in 20% of the samples, whereas (*iv*) simultaneous infection of viruses of the clades CRD-SJP in the same orchard or tree accounts for 1% of samples (**Figure 5**), (*v*) viruses of the clade SJP were detected neither in areas outside the citrus belt SP-MG nor in samples collected before 2015, and (*vi*) except for the isolate PY_Asu02, which belongs to clade ASU, all other samples collected out of the citrus belt SP-MG were infected by viruses of the clade CRD, including a sample from an organic commercial

TABLE 6 | Population genetic differentiation analysis based on ORFs *p29* (795 nts) and *p32* (288 nts, partial) of CiLV-C.

ORF	Subpopulation	*Snn*	K_{st}[a]	H_{st}[a]	F_{st}	*Nm*
p29	CRD and SJP	1.00000	0.40534	0.00246	0.93675	0.03
p32	CRD and SJP	1.00000	0.50932	0.07276	0.94339	0.03

[a]*p-value: 0.01 < p < 0.05.*
Snn, nearest-neighbor statistic; Hst, Hudson's test statistics; Kst, nucleotide-based statistics; Fst, Wright's fixation index, and Nm, gene flow.

TABLE 7 | Analysis of molecular variance (AMOVA) for CiLV-C sub-populations of the clades CRD and SJP, based on ORFs *p29* (795 nts) and *p32* (288 nts, partial).

Source of variance		d.f.[a]	Sum of squares	Variance components	Percentage of variation	AMOVA statistics	*p*-value
p29	Among sub-populations	1	5185.937	55.66483 Va	94.07	0.94067	0.00
	Within sub-populations	187	656.593	3.51119 Vb	5.93		
	Total	188	5842.529	59.17602			
p32	Among sub-populations	1	2258.226	20.22056 Va	92.49	0.92491	0.00
	Within sub-populations	267	438.320	1.64165 Vb	7.51		
	Total	268	2696.546	21.86221			

[a]*Degrees of freedom.*

TABLE 8 | Genetic diversity indices, pairwise mismatch distributions, and neutrality tests (Fu and Li's, Fu's Fs, and Tajima's D) based on ORFs *p29* (795 nts) and *p32* (288 nts, partial) of CiLV-C.

ORF	Sub-population	*s*	*k*	θ	Mismatch		Neutrality tests			
					SSD[a]	HRI[a]	Fu and Li's D[b]	Fu and Li's F[b]	Fu's Fs[c]	Tajima's D
p29	SJP	138	6.383	144	0.0069	0.0075	−5.59231	−5.12367	−127.897	−2.54236[d]
	CRD	154	7.841	158	0.0048	0.0033	−6.13159	−5.57164	−86.537	−2.55671[d]
p32	SJP	50	2.908	51	0.0235	0.0700	−5.86479	−5.01720	−33.099	−1.99819[e]
	CRD	27	1.494	28	0.0718	0.2135	−4.67051	−4.50782	−29.520	−2.28804[f]

[a]*Not significant p-values.*
[b]$p < 0.02$.
[c]*None of the statistics gave significant p-values.*
[d]$p < 0.001$.
[e]$p < 0.05$.
[f]$P < 0.01$.
s, number of segregating sites; k, mean number of nucleotide differences; θ, mean mutation rate per site; SSD, sum of squared deviation; HRI, Harpending's Raggedness index.

orchard collected in Pará, Brazil. The higher prevalence of viruses of the clade SJP in commercial orchards of the citrus belt SP-MG suggests this subpopulation could have some adaptive advantages over those in the CRD one. Preliminary studies on the diversity of *Brevipalpus* mites in Brazil revealed a large genetic variability of the *B. yothersi* population, however, the association, if any, between a given mite haplotype and any host plant or geographical origin remains to be addressed (Sánchez-Velázquez et al., 2015; Salinas-Vargas et al., 2016). Similarly, the nature of the *Brevipalpus* mite interaction with clade-specific CiLV-C strains and its significance on virus ecology are also lacking. On the other hand, the existence of a large number of CL lesions with mixed infections gives support to the recombination events detected *in silico* and suggests the simultaneous transmission of viruses from different clades by a single mite, and/or the sequential arrival of mites bearing a single viral strain to the same lesion. Conceivably all possibilities may happen in nature, but whatever is the case, the increased attractiveness of *Brevipalpus* mites by CiLV-C-infected leaves (Arena et al., 2016) might be a factor contributing to the occurrence of mixed infections.

Natives of Southwest Asia, plants of the genus *Citrus* have spread worldwide and their introduction into America, in the early sixteenth century, is intrinsically linked to the advent of the European colonization. Commercial exploration of the citrus crop in Brazil began in the seventeenth century (Carvalho et al., 2019; Passos et al., 2019), but expanded and became one of the most important Brazilian commodities in the middle of the twentieth century (Mattos et al., 2005; Neves et al., 2011; Carvalho et al., 2019). According to the phylodynamic analysis based on *p29-p32*, the most recent common ancestor (tMRCA) of CiLV-C lineages was dated in 500 A.D., whereas intensive but less deep diversification processes inside the clades CRD and SJP have been occurring from the nineteenth century. Based on a second dataset that included the sequences *RdRp-p29-p61-p32-p24* of 23 isolates of the three lineages, the MRCA of the main lineages dates back more than 3000 years before. However, since the MCC quality parameters with the second dataset only reached moderate ESS values, the resultant tMRCA must be only considered as a framework for further evaluations. It could be expected that the analysis of new sequences will allow higher accuracy in the determination of tMRCA providing a better description of the evolutionary history of CiLV-C lineages. Regardless of this, bringing together the timelines described by the two datasets we can conclude that (*i*) ancestors of the three viral clades might have been originated in contact with native ecosystems of South America and (*ii*), the expansion of the citriculture in Brazil and other countries in the region, has contributed, although in a low rate, with the intra-clade diversifications of CiLV-C.

CL is considered an emerging disease affecting a crop that, at least in Brazil, encompasses a relatively low genetic diversity and high-density plantings. Increasing pieces of evidence strongly hypothesized that wild ecosystems are a major source of diversity of plant viruses, which have co-evolved with their wild hosts long before they were domesticated (Pagan and Holmes, 2010; Rodríguez-Nevado et al., 2017). In addition to some *Citrus* spp. and their hybrids, only plants of the species *Commelina*

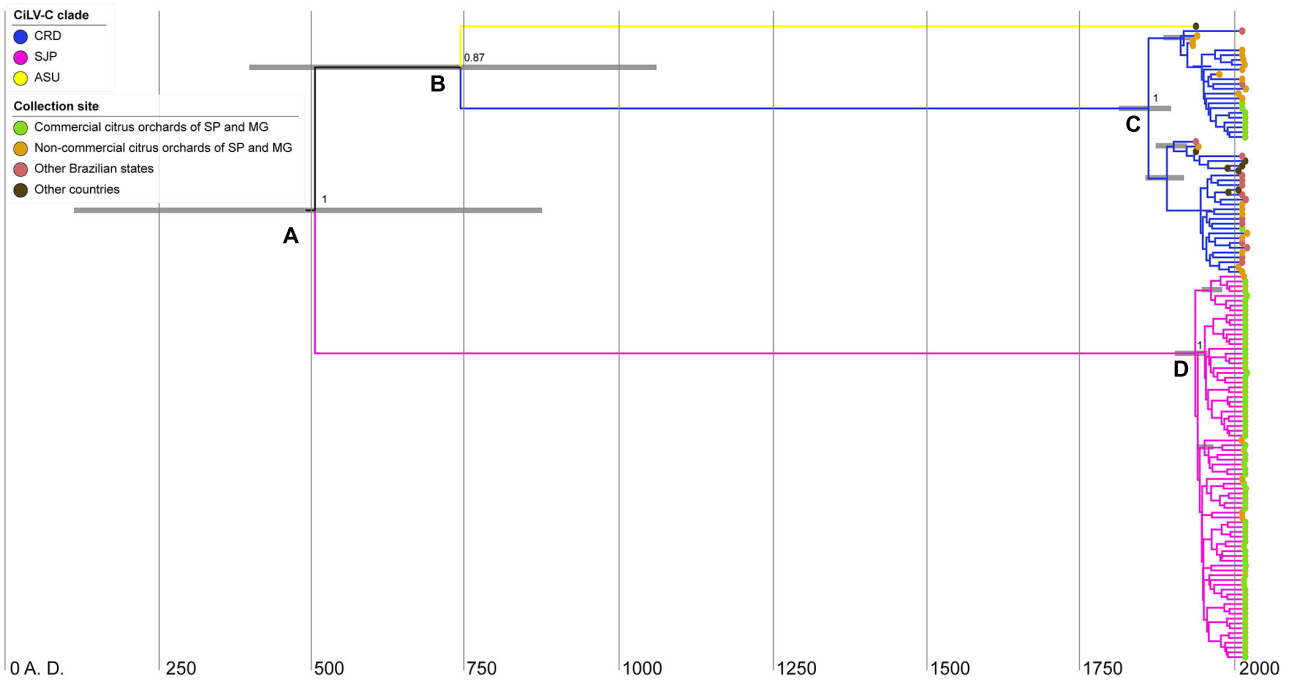


FIGURE 6 | Bayesian maximum-clade-credibility time-scaled phylogenetic tree using the concatenated sequences of the *p29* (795 nts) and *p32* (288 nts, partial) ORFs from 132 CiLV-C isolates collected in the period 1932–2020 in South America. Horizontal gray bars on nodes **(A–D)** indicate the uncertainty for the date of each node (95% highest posterior density—HPD—intervals). Figures near the main nodes represent the posterior probability values. The phylogenetic tree was edited using IcyTree.

benghalensis and *Swinglea glutinosa* are reported as natural hosts of CiLV-C (León et al., 2008; Nunes et al., 2012; Garita et al., 2014; Freitas-Astúa et al., 2018), all of them are exotic to the Americas (Weniger et al., 2001; Webster et al., 2005). Therefore, the natural host range of CiLV-C is likely not yet fully known or the interaction with its wild hosts might no longer exist in nature. The distribution and dispersion of cileviruses are mediated by the polyphagous *Brevipalpus* mites, capable of feeding on more than 150 genera of plants, including several crops, ornamentals, and forest plants (Childers et al., 2003). Consequently, the CiLV-C vector is likely the main path between native and exotic plants in the Americas (Freitas-Astúa et al., 2018). Similar to CiLV-C, the cilevirus CiLV-C2 also infects citrus plants and seems to have a narrow range of known natural hosts, e.g., *Citrus* spp., *S. glutinosa, Hibiscus* sp., and *Dieffenbachia* sp. (Melzer et al., 2013; Roy et al., 2015, 2018). In contrast, PfGSV is the only known cilevirus infecting native plants of the Americas, e.g., *Passiflora* spp. (Kitajima et al., 1997; Ramos-González et al., 2020). This virus can be found naturally infecting more than twenty plant species and, notably, symptoms are not lesions as locally restricted as those observed in citrus plants infected by CiLV-C. Since the interaction of CiLV-C with citrus plants is dominated by a hypersensitive-like response (Arena et al., 2016, 2020), the expansion of the host range to *Citrus* spp. might have resulted in CiLV-C fitness reduction. Alternatively, based on the low variability of viruses inside each CiLV-C clade, a thought-provoking question is whether CiLV-C can be considered a specialist virus, whose interaction with citrus is carefully selected to act as a helper (effector-like) factor of the mite infestation to suppress the plant defenses (Arena et al., 2016, 2018, 2020).

Altogether, this study provides the most complete snapshot of the CiLV-C population to date. Throughout molecular epidemiology analyses, we have revealed the structure, sources of the genetic variability, and forces involved in the recent evolution of this viral population. The evolutionary history of CiLV-C may be strongly influenced by interaction with its main known host, *Citrus* spp. during a relatively short period, which at most includes the last 500 years. Maximum values of variability inside the population are typified by its subdivision into the clades ASU, first identified in this study, CRD, and SJP. These three clades are the outcome of diversification processes that occurred before the viral contact with the citrus host. Moreover, besides the highly frequent bottlenecks as a result of mite transmission, the incompatible host-virus interaction with an intensive crop with a relatively low genetic variability, likely prevents the expansion and diversification of the CiLV-C subpopulations. In practical terms, our results confirm the possibility of recovering viral sequences present in herbarium citrus leaf samples despite the low RIN values of the RNA extracts, as previously successfully described using dried fruit peels (Hartung et al., 2015). Moreover, the current study reinforces the prevalence and wide distribution of CiLV-C in the largest citrus commercial area of Brazil and reveals the urgency for updating detection systems able to identify the presence of CiLV-C variants whose epidemiological profiles are currently unknown.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in the NCBI Genbank Database. The accession numbers can be found in **Supplementary Table 1**.

AUTHOR CONTRIBUTIONS

CC-J and PR-G conceptualized and wrote the manuscript. CC-J, PR-G, MP-B, and RH worked in laboratory analyses. CC-J, PR-G, and RF worked on the formal analysis. PR-G, AV, and JF-A supervised the study. JF-A and EK were responsible for funding acquisition. AM and RB assisted in the collection of fresh citrus samples. All authors reviewed the final manuscript.

FUNDING

The authors are appreciative of the São Paulo Research Foundation (FAPESP) for financial support in projects (2017/50222-0, 2017/50334-3, and 2019/25078-9) and scholarships (2016/01960-6, 2019/08186-2, and 2020/06378-9), Coordination for the Improvement of Higher Education Personnel (CAPES) financial support PNPD20132154-33141010001P4-PNPD—IBSP. This study was financed in part by CAPES Finance Code 001. EK, RH, and JF-A were CNPq productivity research fellows.

ACKNOWLEDGMENTS

We acknowledge Dr. Alexander Banguela-Castillo for his help with the bioinformatics analyses.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.641252/full#supplementary-material

Supplementary Figure 1 | Bayesian maximum clade credibility tree inferred using a Markov Chain Monte Carlo (MCMC) of the CiLV-C of the lineages CRD (blue branches), SJP (pink branches), and ASU (yellow branches) based on *p29* **(A)** and *p32* **(B)** ORFs. CiLV-C2 sequences (black branches) were used as an outgroup and trees were generated with 6,000,000 generations. The color-coded based collection sites show whether the sample was obtained from commercial orchards in the citrus belt of São Paulo (SP) and Minas Gerais (MG) states (in green), from non-commercial citrus trees of SP and MG (in gold), from other Brazilian states (in mauve), or other countries, i.e., Argentina, Colombia, Mexico, Panama, and Paraguay (in brown).

Supplementary Figure 2 | Bayesian maximum-clade-credibility (MCC) time-scaled phylogenetic tree (**A**) and analyses of the temporal signal by linear regression approach (**B**) using the complete sequence of *RdRp–p29–p61–p32–p24* ORFs from 23 CiLV-C isolates collected in the period 1932–2020 in South America. **(A)** Horizontal gray bars on nodes A, B, C, and D indicate the uncertainty for the date of each node (95% highest posterior density—HPD—intervals). Figures near the main nodes represent the posterior probability values. The color-coded based collection sites show whether the sample was obtained from commercial orchards in the citrus belt of São Paulo (SP) and Minas Gerais (MG) states (in green), from non-commercial citrus trees of SP and MG (in gold), from other Brazilian states (in mauve), or other countries, i.e., Argentina, Colombia, Mexico, Panama, and Paraguay (in brown). The phylogenetic tree was edited using IcyTree; **(B)** The temporal signal was calculated using TempEst v. 1.5.3 software. Root-to-tip divergence as a function of sampling time for Maximum likelihood non-clock tree generated by IQtree v. 1.5.5 software.

Supplementary Figure 3 | Analyses of the temporal signal by linear regression approach from the concatenate of the *p29* (795 nt) and partial *p32* (288 nt) of 132 isolates of CiLV-C using TempEst v. 1.5.3 software. Root-to-tip divergence as a function of sampling time for Maximum likelihood non-clock tree generated by IQtree v. 1.5.5 software.

Supplementary Table 1 | Complete list of plant samples and citrus leprosis virus C (CiLV-C) sequences analyzed in this study.

Supplementary Table 2 | Nucleotide and haplotypic diversities of ORFs *p29* and *p32* of CiLV-C. Nucleotide sequences were amplified and cloned from samples of infected *Citrus sinensis* trees collected in commercial orchards of the citrus belt São Paulo—Minas Gerais, Brazil, in 2017.

Supplementary Table 3 | Nucleotide and deduced amino acid identities (%) among CiLV-C isolates described in this study and the type members of the clades CRD (CiLV-C_BR_SP_Crd01, GenBank accession numbers NC008169 and NC008170) and SJP (CiLV-C_BR_SP_SJP01, GB acc. numbers KP336746 and KP336747).

Supplementary Table 4 | Recombination events detected in CiLV-C sequences determined by RDP software version 5.5.

Supplementary Table 5 | Summary of selection analysis of the *p29* and *p32* in CiLV-C.

REFERENCES

Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]*. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed October 25, 2020).

Arena, G. D., Ramos-González, P. L., Falk, B. W., Casteel, C. L., Freitas-Astúa, J., and Machado, M. A. (2020). Plant immune system activation upon citrus leprosis virus C infection is mimicked by the ectopic expression of the P61 viral protein. *Front. Plant Sci.* 11:1188. doi: 10.3389/fpls.2020.01188

Arena, G. D., Ramos-González, P. L., Nunes, M. A., Alves, M. R., Camargo, L. E. A., Kitajima, E. W., et al. (2016). Citrus leprosis virus C infection results in hypersensitive-like response, suppression of the JA/ET plant defense pathway and promotion of the colonization of its mite vector. *Front. Plant Sci.* 7:1757. doi: 10.3389/FPLS.2016.01757

Arena, G. D., Ramos-González, P. L., Nunes, M. A., Jesus, C. C., Calegario, R. F., Kitajima, E. W., et al. (2017). Arabidopsis thaliana as a model host for *Brevipalpus* mite-transmitted viruses. *Sci. Agric.* 74, 85–89. doi: 10.1590/1678-992x-2015-0380

Arena, G. D., Ramos-González, P. L., Rogerio, L. A., Ribeiro-Alves, M., Casteel, C. L., Freitas-Astúa, J., et al. (2018). Making a better home: modulation of plant defensive response by *Brevipalpus* mites. *Front. Plant Sci.* 9:1147. doi: 10.3389/fpls.2018.01147

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477.

Bassanezi, R. B., Czermainski, A. B. C., Laranjeira, F. F., Moreira, A. S., Ribeiro, P. J., Krainski, E. T., et al. (2019). Spatial patterns of the *Citrus leprosis* virus and its associated mite vector in systems without intervention. *Plant Pathol.* 68, 85–93. doi: 10.1111/ppa.12930

Bastianel, M., Pereira-Martin, J. A., Novelli, V. M., Freitas-Astúa, J., and Nunes, M. A. (2018). *Citrus leprosis* resistance within the citrus group. *VirusDisease* 29, 491–498. doi: 10.1007/s13337-018-0489-6

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Genome analysis Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Boncristiani, H. F., Di Prisco, G., Pettis, J. S., Hamilton, M., and Chen, Y. P. (2009). Molecular approaches to the analysis of deformed wing virus replication and pathogenesis in the honey bee, *Apis mellifera*. *Virol. J.* 6:221. doi: 10.1186/1743-422X-6-221

Bushnell, B. (2014). *BBMap: A Fast, Accurate, Splice-Aware Aligner*. Berkeley, CA: Lawrence Berkeley National Lab, 3–5.

Carvalho, S. A., Girardi, E. A., Filho, F., de, A. A. M., Ferrarezi, R. S., and Filho, H. D. C. (2019). Advances in citrus propagation in Brazil. *Rev. Bras. Frutic.* 41, 1–36. doi: 10.1590/0100-29452019422

Chabi-Jesus, C., Ramos-González, P. L., Tassi, A. D., Guerra-Peraza, O., Kitajima, E. W., Harakava, R., et al. (2018). Identification and characterization of citrus chlorotic spot virus, a new dichorhavirus associated with *Citrus leprosis*-like symptoms. *Plant Dis.* 102, 1588–1598. doi: 10.1094/PDIS-09-17-1425-RE

Childers, C. C., French, J. V., and Rodrigues, J. C. V. (2003). *Brevipalpus californicus*, *B. obovatus*, *B. phoenicis*, and *B. lewisi* (Acari: Tenuipalpidae): a review of their biology, feeding injury and economic importance. *Exp. Appl. Acarol.* 30, 5–28. doi: 10.1023/B:APPA.0000006543.34042.b4

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Freitas-Astúa, J., Ramos-González, P. L., Arena, G. D., Tassi, A. D., and Kitajima, E. W. (2018). Brevipalpus-transmitted viruses: parallelism beyond a common vector or convergent evolution of distantly related pathogens? *Curr. Opin. Virol.* 33, 66–73. doi: 10.1016/J.COVIRO.2018.07.010

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.

Fu, Y. X., and Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics* 133, 693–709.

García-Arenal, F., Fraile, A., and Malpica, J. M. (2001). Variability and genetic structure of plant virus populations. *Annu. Rev. Phytopathol.* 39, 157–186. doi: 10.1146/annurev.phyto.39.1.157

Garita, L. C., Tassi, A. D., Calegario, R. F., Freitas-Astúa, J., Salaroli, R. B., Romão, G. O., et al. (2014). Experimental host range of citrus leprosis virus C (CiLV-C). *Trop. Plant Pathol.* 39, 43–55. doi: 10.1590/S1982-56762014005000004

Haddad, F., Qin, A. X., Giger, J. M., Guo, H., and Baldwin, K. M. (2007). Potential pitfalls in the accuracy of analysis of natural sense-antisense RNA pairs by reverse transcription-PCR. *BMC Biotechnol.* 7:21. doi: 10.1186/1472-6750-7-21

Haist, K., Ziegler, C., and Botten, J. (2015). Strand-specific quantitative reverse transcription-polymerase chain reaction assay for measurement of arenavirus genomic and antigenomic RNAs. *PLoS One* 10:e0120043. doi: 10.1371/journal.pone.0120043

Harpend, H. C. (1994). Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Hum. Biol.* 66, 591–600.

Hartung, J. S., Roy, A., Fu, S., Shao, J., Schneider, W. L., and Brlansky, R. H. (2015). History and diversity of citrus leprosis virus recorded in herbarium specimens. *Phytopathology* 105, 1277–1284. doi: 10.1094/PHYTO-03-15-0064-R

Hatcher, E. L., Zhdanov, S. A., Bao, Y., Blinkova, O., Nawrocki, E. P., Ostapchuck, Y., et al. (2017). Virus variation resource-improved response to emergent viral outbreaks. *Nucleic Acids Res.* 45, D482–D490. doi: 10.1093/nar/gkw1065

Hudson, R. R. (2000). A new statistic for detecting genetic differentiation. *Genetics* 155, 2011–2014.

Hudson, R. R., Slatkin, M., and Maddison, W. P. (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics* 132, 583–589.

Huelsenbeck, J. P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755. doi: 10.1093/bioinformatics/17.8.754

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kitajima, E. W., Chagas, C. M., and Rodrigues, J. C. V. (2003). *Brevipalpus*-transmitted plant virus and virus-like diseases: cytopathology and some recent cases. *Exp. Appl. Acarol.* 30, 135–160.

Kitajima, E. W., Kubo, K. S., Ferreira, P. T. O., de Alcântara, B. K., Boari, A. J., Gomes, R. T., et al. (2008). Chlorotic spots on Clerodendrum, a disease caused by a nuclear type of *Brevipalpus* (Acari:Tenuipalpidae) transmitted virus. *Sci. Agric.* 65, 36–49. doi: 10.1590/S0103-90162008000100006

Kitajima, E. W., Rezende, J. A. M., Rodrigues, J. C. V., Chiavegato, L. G., Piza Júnior, C. T., and Morozini, W. (1997). Green spot of passion fruit, a possible viral disease associated with infestation by the mite *Brevipalpus phoenicis*. *Fitopatol. Bras.* 22, 555–559.

Kondo, H., Fujita, M., Hisano, H., Hyodo, K., Andika, I. B., and Suzuki, N. (2020). Virome analysis of aphid populations that infest the barley field: the discovery of two novel groups of Nege/Kita-like viruses and other novel RNA viruses. *Front. Microbiol.* 11:509. doi: 10.3389/fmicb.2020.00509

Kosakovsky Pond, S. L., Posada, D., Gravenor, M. B., Woelk, C. H., and Frost, S. D. W. (2006). GARD: a genetic algorithm for recombination detection. *Bioinformatics* 22, 3096–3098. doi: 10.1093/bioinformatics/btl474

Kubo, K. S., Stuart, R. M., Freitas-Astúa, J., Antonioli-Luizon, R., Locali-Fabris, E. C., Coletta-Filho, H. D., et al. (2009). Evaluation of the genetic variability of orchid fleck virus by single-strand conformational polymorphism analysis and nucleotide sequencing of a fragment from the nucleocapsid gene. *Arch. Virol.* 154, 1009–1014. doi: 10.1007/s00705-009-0395-8

Kuchibhatla, D. B., Sherman, W. A., Chung, B. Y. W., Cook, S., Schneider, G., Eisenhaber, B., et al. (2014). Powerful sequence similarity search methods and in-depth manual analyses can identify remote homologs in many apparently "orphan" viral proteins. *J. Virol.* 88, 10–20. doi: 10.1128/JVI.02595-13

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

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Leastro, M. O., Castro, D. Y. O., Freitas-Astúa, J., Kitajima, E. W., Pallás, V., and Sánchez-Navarro, J. Á (2020). *Citrus leprosis* virus C encodes three proteins with gene silencing suppression activity. *Front. Microbiol.* 11:1231. doi: 10.3389/fmicb.2020.01231

Leastro, M. O., Freitas-Astúa, J., Kitajima, E. W., Pallás, V., and Sánchez-Navarro, J. A. (2021). Unravelling the involvement of cilevirus p32 protein in the viral transport. *Sci. Rep.* 11:2943. doi: 10.1038/s41598-021-82453-4

León, M. G., Becerra, C. H., Freitas-Astúa, J., Salaroli, R. B., and Kitajima, E. W. (2008). Natural infection of *Swinglea glutinosa* by the *Citrus leprosis* virus cytoplasmic type (CiLV-C) in Colombia. *Plant Dis.* 92, 1364–1364. doi: 10.1094/PDIS-92-9-1364C

Letunic, I., and Bork, P. (2019). Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239

Locali, E. C., Freitas-Astua, J., de Souza, A. A., Takita, M. A., Astua-Monge, G., Antonioli, R., et al. (2003). Development of a molecular tool for the diagnosis of leprosis, a major threat to citrus production in the Americas. *Plant Dis.* 87, 1317–1321. doi: 10.1094/PDIS.2003.87.11.1317

Locali-Fabris, E. C., Freitas-Astúa, J., and Machado, M. A. (2012). "Genus *Cilevirus*. international committee on taxonomy of viruses," in *Virus Taxonomy*, eds A. King, M. Adams, E. Carstens, and E. Lefkowitz (London: Academic Press), 1139–1142.

Locali-Fabris, E. C., Freitas-Astúa, J., Souza, A. A., Takita, M. A., Astúa-Monge, G., Antonioli-Luizon, R., et al. (2006). Complete nucleotide sequence, genomic organization and phylogenetic analysis of citrus leprosis virus cytoplasmic type. *J. Gen. Virol.* 87, 2721–2729. doi: 10.1099/vir.0.82038-0

Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., et al. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* 73, 152–160.

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., and Muhire, B. (2015). RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol.* 1:vev003. doi: 10.1093/ve/vev003

Mattos, D. Jr., de Negri, J. D., Pio, R. M., and Pompeu, J. Jr. (2005). *Citros*. New York, NY: IAC, doi: 10.1146/annurev-phyto-080417-050052

Melzer, M. J., Sether, D. M., Borth, W. B., and Hu, J. S. (2012). Characterization of a virus infecting *Citrus volkameriana* with citrus leprosis-like symptoms. *Phytopathology* 102, 122–127. doi: 10.1094/PHYTO-01-11-0013

Melzer, M. J., Simbajon, N., Carillo, J., Borth, W. B., Freitas-Astúa, J., Kitajima, E. W., et al. (2013). A cilevirus infects ornamental hibiscus in Hawaii. *Arch. Virol.* 158, 2421–2424. doi: 10.1007/s00705-013-1745-0

Mushegian, A. R., and Elena, S. F. (2015). Evolution of plant virus movement proteins from the 30K superfamily and of their homologs integrated in plant genomes. *Virology* 476C, 304–315. doi: 10.1016/j.virol.2014.12.012

Neves, M. F., Trombin, V. G., Lopes, F. F., Kalaki, R., and Milan, P. (2011). "The citrus sector in Brazil," in *The Orange Juice Business*, eds M. Fava Neves, V. G. Trombin, F. F. Lopes, R. Kalaki, and P. Milan (Wageningen: Wageningen Academic Publishers), 18–20. doi: 10.3920/978-90-8686-739-4_1

Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

Nunes, M. A., Bergamini, M. P., Coerini, L. F., Bastianel, M., Novelli, V. M., Kitajima, E. W., et al. (2012). Citrus leprosis virus C naturally infecting *Commelina benghalensis*, a prevalent monocot weed of citrus orchards in Brazil. *Plant Dis.* 96, 770–770. doi: 10.1094/PDIS-11-11-0925-PDN

Nunes, M. A., de Carvalho Mineiro, J. L., Rogero, L. A., Ferreira, L. M., Tassi, A. D., Novelli, V. M., et al. (2018). First report of *Brevipalpus papayensis* Baker (Acari: Tenuipalpidae) as vector of Coffee ringspot virus and *Citrus leprosis* virus C. *Plant Dis.* 102, 1046–1046. doi: 10.1094/PDIS-07-17-1000-PDN

Pagan, I., and Holmes, E. C. (2010). Long-term evolution of the luteoviridae: time scale and mode of virus speciation. *J. Virol.* 84, 6177–6187. doi: 10.1128/jvi.02160-09

Passos, O. S., Souza, J. S., Bastos, D. C., Girardi, E. A., Gurgel, F. L., Garcia, M. V. B., et al. (2019). "Citrus industry in Brazil with emphasis on tropical areas," in *Citrus – Health Benefits and Production Technology*, eds M. Sajid and Amanullah (London: IntechOpen), doi: 10.5772/intechopen.80213

Pei, J., and Grishin, N. V. (2007). PROMALS: towards accurate multiple sequence alignments of distantly related proteins. *Bioinformatics* 23, 802–808. doi: 10.1093/bioinformatics/btm017

Quito-Avila, D. F., Brannen, P. M., Cline, W. O., Harmon, P. F., and Martin, R. R. (2013). Genetic characterization of Blueberry necrotic ring blotch virus, a novel RNA virus with unique genetic features. *J. Gen. Virol.* 94, 1426–1434. doi: 10.1099/vir.0.050393-0

Quito-Avila, D. F., Freitas-Astúa, J., and Melzer, M. J. (2020). " Bluner-, Cile-, and Higreviruses (Kitaviridae)," in *Reference Module in Life Sciences*, (Amsterdam: Elsevier), doi: 10.1016/B978-0-12-809633-8.21248-X

Ramalho, T. O., Figueira, A. R., Wang, R., Jones, O., Harris, L. E., and Goodin, M. M. (2016). Detection and survey of coffee ringspot virus in Brazil. *Arch. Virol.* 161, 335–343. doi: 10.1007/s00705-015-2663-0

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Syst. Biol.* 67, 901–904. doi: 10.1093/sysbio/syy032

Ramos-González, P. L., Chabi-Jesus, C., Arena, G. D., Tassi, A. D., Kitajima, E. W., and Freitas-Astúa, J. (2018). Citrus leprosis: a unique multietiologic disease. *Citrus Am.* 1, 4–19. doi: 10.6084/m9.figshare.13513857.v2

Ramos-González, P. L., Chabi-Jesus, C., Guerra-Peraza, O., Breton, M. C., Arena, G. D., Nunes, M. A., et al. (2016). Phylogenetic and molecular variability studies reveal a new genetic clade of *Citrus leprosis* virus C. *Viruses* 8:153. doi: 10.3390/v8060153

Ramos-González, P. L., Chabi-Jesus, C., Guerra-Peraza, O., Tassi, A. D., Kitajima, E. W., Harakava, R., et al. (2017). *Citrus leprosis* virus N: a new dichorhavirus causing *Citrus leprosis* disease. *Phytopathology* 107, 963–976. doi: 10.1094/PHYTO-02-17-0042-R

Ramos-González, P. L., dos Santos, G. F., Chabi-Jesus, C., Harakava, R., Kitajima, E. W., and Freitas-Astúa, J. (2020). Passion fruit green spot virus genome harbors a new orphan ORF and highlights the flexibility of the 5'-end of the RNA2 segment across Cileviruses. *Front. Microbiol.* 11:206. doi: 10.3389/fmicb.2020.00206

Rodríguez-Nevado, C., Montes, N., and Pagán, I. (2017). Ecological factors affecting infection risk and population genetic diversity of a novel potyvirus in its native wild ecosystem. *Front. Plant Sci.* 8:1958. doi: 10.3389/fpls.2017.01958

Rogers, A. R., and Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9, 552–569. doi: 10.1093/oxfordjournals.molbev.a040727

Roy, A., Choudhary, N., Guillermo, L. M., Shao, J., Govindarajulu, A., Achor, D., et al. (2013). A novel virus of the genus *Cilevirus* causing symptoms similar to *Citrus leprosis*. *Phytopathology* 103, 488–500. doi: 10.1094/PHYTO-07-12-0177-R

Roy, A., Hartung, J. S., Schneider, W. L., Shao, J., Leon, M. G., Melzer, M. J., et al. (2015). Role bending: complex relationships between viruses, hosts, and vectors related to citrus leprosis, an emerging disease. *Phytopathology* 105, 872–884. doi: 10.1094/PHYTO-12-14-0375-FI

Roy, A., Stone, A. L., Melzer, M. J., Shao, J., Hartung, J. S., Mavrodieva, V., et al. (2018). Complete nucleotide sequence of a novel *Hibiscus*-infecting *Cilevirus* from Florida and its relationship with closely associated *Cileviruses*. *Genome Announc.* 6:e01521-17. doi: 10.1128/genomeA.01521-17

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302. doi: 10.1093/molbev/msx248

Salinas-Vargas, D., Santillán-Galicia, M. T., Guzmán-Franco, A. W., Hernández-López, A., Ortega-Arenas, L. D., and Mora-Aguilera, G. (2016). Analysis of genetic variation in *Brevipalpus yothersi* (Acari: Tenuipalpidae) populations from four species of *Citrus* host plants. *PLoS One* 11:e0164552. doi: 10.1371/journal.pone.0164552

Sánchez-Velázquez, E. J., Santillán-Galicia, M. T., Novelli, V. M., Nunes, M. A., Mora-Aguilera, G., Valdez-Carrasco, J. M., et al. (2015). Diversity and genetic variation among *Brevipalpus* populations from Brazil and Mexico. *PLoS One* 10:e0133861. doi: 10.1371/journal.pone.0133861

Strydom, E., and Pietersen, G. (2018). Development of a strand-specific RT-PCR to detect the positive sense replicative strand of Soybean blotchy mosaic virus. *J. Virol. Methods* 259, 39–44. doi: 10.1016/j.jviromet.2018.05.014

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., and Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* 4, 1–5. doi: 10.1093/ve/vey016

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.

Tassi, A. D., Garita-Salazar, L. C., Amorim, L., Novelli, V. M., Freitas-Astúa, J., Childers, C. C., et al. (2017). Virus-vector relationship in the *Citrus leprosis* pathosystem. *Exp. Appl. Acarol.* 71, 227–241. doi: 10.1007/s10493-017-0123-0

Tsai, I. J., Otto, T. D., and Berriman, M. (2010). Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps. *Genome Biol.* 11:R41. doi: 10.1186/gb-2010-11-4-r41

Vaughan, T. G. (2017). IcyTree: rapid browser-based visualization for phylogenetic trees and networks. *Bioinformatics* 33, 2392–2394. doi: 10.1093/bioinformatics/btx155

Weaver, S., Shank, S. D., Spielman, S. J., Li, M., Muse, S. V., and Kosakovsky Pond, S. L. (2018). Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. *Mol. Biol. Evol.* 35, 773–777. doi: 10.1093/molbev/msx335

Webster, T. M., Burton, M. G., Culpepper, A. S., York, A. C., and Prostko, E. P. (2005). Tropical Spiderwort (Commelina benghalensis): a tropical invader threatens agroecosystems of the Southern United States. *Weed Technol.* 19, 501–508. doi: 10.1614/wt-04-234r.1

Weniger, B., Um, B. H., Valentin, A., Estrada, A., Lobstein, A., Anton, R., et al. (2001). Bioactive acridone alkaloids from *Swinglea glutinosa*. *J. Nat. Prod.* 64, 1221–1223. doi: 10.1021/np0005762

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 03 May 2021
doi: 10.3389/fmicb.2021.660237

A Virus Infecting *Hibiscus rosa-sinensis* Represents an Evolutionary Link Between Cileviruses and Higreviruses

*Alejandro Olmedo-Velarde, John Hu and Michael J. Melzer**

Department of Plant and Environmental Protection Sciences, University of Hawaii, Honolulu, HI, United States

Hibiscus (*Hibiscus* spp.) are popular ornamental and landscape plants in Hawaii which are susceptible to foliar diseases caused by viruses belonging to the genera *Cilevirus* and *Higrevirus* (family *Kitaviridae*). In this study, a virus infecting *H. rosa-sinensis* plants displaying foliar symptoms consistent with infection by a kitavirus, including yellow chlorotic blotches with a green perimeter, was characterized. The genome consisted of two RNAs 8.4 and 4.4 kb in length, and was organized most similarly to cileviruses, but with important distinctions. These included the location of the p29 homolog as the 3′-terminal open reading frame (ORF) of RNA2 instead of its typical locus at the 3′-end of RNA1; the absence of a p15 homolog on RNA2 and the adjacent intergenic region which also harbors small putative ORFs of unknown function; and the presence of an ORF encoding a 10 kDa protein at the 3′-terminal end of RNA1 that was also found to be present in the hibiscus green spot virus 2 genome. Spherical particles approximately 55–65 nm in diameter were observed in infected leaf tissue, and viral RNA was detected by reverse-transcription PCR in individual mites collected from symptomatic plants tentatively identified as *Brevipalpus yothersi*. Although phylogenetic analyses placed this virus between the higrevirus and cilevirus clades, we propose the tentative taxonomic placement of this virus, designated hibiscus yellow blotch virus (HYBV), within the genus *Cilevirus*.

Keywords: kitavirid, *Kitaviridae*, hibiscus, yellow blotch, HTS

OPEN ACCESS

Edited by:
Sergey Morozov, Lomonosov Moscow State University, Russia

Reviewed by:
Jesús Navas-Castillo, Institute of Subtropical and Mediterranean Horticulture La Mayora, Spain
Pedro Luis Ramos-González, Biological Institute of São Paulo, Brazil

***Correspondence:**
Michael J. Melzer
melzer@hawaii.edu

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: *29 January 2021*
Accepted: *06 April 2021*
Published: *03 May 2021*

Citation:
Olmedo-Velarde A, Hu J and Melzer MJ (2021) A Virus Infecting Hibiscus rosa-sinensis Represents an Evolutionary Link Between Cileviruses and Higreviruses. Front. Microbiol. 12:660237. doi: 10.3389/fmicb.2021.660237

INTRODUCTION

The family *Kitaviridae* encompasses three genera of positive-sense ssRNA plant viruses: *Blunervirus*, *Cilevirus*, and *Higrevirus* (Melzer et al., 2018). Although related, there are considerable physical and genetic distinctions between members of the different genera. First, cile- and higreviruses are associated with a bacilliform virion, whereas a spherical virion has been observed for the lone blunervirus for which microscopy has been reported (Kitajima et al., 1974; Melzer et al., 2012; Hao et al., 2018). Second, cile-, higre-, and blunerviruses have bi-, tri-, and tetrapartite genomes, respectively (Quito-Avila et al., 2020). Third, different lineages of movement protein (MP) are present in the family: blunerviruses and cileviruses have a 3A/30K superfamily MP, whereas the lone higrevirus member possesses a triple gene block-like MP module (Quito-Avila et al., 2020). Finally, the replication-associated polyproteins are encoded by a single genomic RNA

for cile- and higreviruses, but are split between two genomic RNAs for blunerviruses (Quito-Avila et al., 2013, 2020). As such, higreviruses and cileviruses share a closer phylogenetic relationship when conserved protein sequences are analyzed.

In recent years, several unclassified insect-infecting viruses, namely nelorpi- and sandewaviruses, Nunes et al. (2017) and arthropod viruses (Shi et al., 2016) have been characterized that resemble kitaviruses and appear to reside within the kitavirus clade. The RNA-dependent RNA polymerase (RdRp) of kitaviruses and that of the unsegmented negeviruses (nelorpi- and sandewaviruses) have a common phylogenetic origin and homologs of p24, a predicted virion membrane protein, are currently found only in these plant- and arthropod-infecting viruses (Kuchibhatla et al., 2014). The evidence of a common ancestry between these viruses has led to the hypothesis that plant-infecting kitaviruses arose from these arthropod-infecting viruses, with the arthropod vector being a potential origin (Nunes et al., 2017; Ramos-Gonzalez et al., 2020). The arthropod vector has only been confirmed for cileviruses, with *Brevipalpus* spp. (Acari: Tenuipalpidae) mites responsible for transmission in a persistent circulative, and likely propagative manner (Roy et al., 2015; Freitas-Astúa et al., 2018).

The family *Kitaviridae* is currently composed of five recognized species among the three genera: *Blueberry necrotic ring blotch virus* and *Tea plant necrotic ring blotch virus* (genus *Blunervirus*); *Citrus leprosis virus C* and *Citrus leprosis virus C2* (genus *Cilevirus*); and *Hibiscus green spot virus 2* (genus *Higrevirus*). The genomes of tomato fruit blotch virus, a putative blunervirus, and passion fruit green spot virus, a putative cilevirus, have also recently been described (Cuiffo et al., 2020; Ramos-Gonzalez et al., 2020). To better understand the diversity and evolutionary history of this family, it is imperative that additional members be described. In this study, we use high-throughput sequencing (HTS) to identify and characterize a new kita-like virus in Hawaii that infects *Hibiscus rosa-sinensis* in a non-systemic manner. This virus has distinctive genomic characteristics and phylogenetic analyses indicate it is an intermediate of cileviruses and higreviruses. Its bipartite genome suggests a tentative and temporary placement in the genus *Cilevirus*, and the name hibiscus yellow blotch virus (HYBV) is proposed.

MATERIALS AND METHODS

Tissue Collection and Virus Indexing

In July 2019, *Hibiscus rosa-sinensis* (L.) leaves displaying viral-like symptoms consistent with those caused by a brevipalpus-transmitted virus (BTV) (Kitajima et al., 2003) were collected from a single tree in Pearl City, Hawaii. Symptoms were characterized by yellow blotches surrounded by a green halo on both green and senescing leaves (**Figure 1**). Additional samples from the same tree were collected in in November 2019 and March 2020. Leaf samples displaying green ringspot symptoms consistent with BTV infection in Hawaii (Melzer et al., 2013) were subsequently collected from six *Hibiscus* spp. trees growing within 1.6 km of the original *H. rosa-sinensis* tree.

To determine if the *H. rosa-sinensis* tree was infected with hibiscus green spot virus 2 (HGSV-2) and the hibiscus strain of citrus leprosis virus C2 (CiLV-C2H), which have been associated with similar, yet distinct, symptoms in Hawaii's hibiscus plants, virus-specific RT-PCR assays were performed using existing protocols (Melzer et al., 2012, 2013).

Genome Sequencing

Double stranded RNAs (dsRNAs) were extracted from ∼5 g of symptomatic leaf tissue collected from the *H. rosa-sinensis* tree in Pearl City using either CF-11 (Whatman, Maidstone, United Kingdom) or C6288 (Sigma, St. Louis, MO, United States) cellulose chromatography (Morris and Dodds, 1979) and resolved by 1X TBE-1% agarose gel electrophoresis. Using dsRNAs as template, randomly amplified cDNAs were generated (Melzer et al., 2010) and prepared for HTS using a Nextera XT DNA Library Prep kit (Illumina, San Diego, CA, United States). HTS was performed on an Illumina MiSeq 2 × 300 bp (V2) platform at the University of Hawaii Advanced Studies in Genomics, Proteomics, and Bioinformatics (ASGPB) Laboratory. Genome assembly and bioinformatic analyses were performed as described (Olmedo-Velarde et al., 2019). Briefly, paired-end reads were trimmed, and quality filtered using Trimmomatic 0.35.3 (Bolger et al., 2014). Trinity 2.2.0 (Grabherr et al., 2011) produced *de novo* assembled contigs that were annotated using BLASTX search (Altschul et al., 1997) against the viral genome database[1]. Contiguous sequences (contigs) with similarity to cilevirus sequences were then used as reference for an iterative mapping approach (Dey et al., 2019) using Geneious mapper plug-in implemented in Geneious v. 10.1.3 (Kearse et al., 2012) and raw reads. A group of overlapping primer sets (**Supplementary Table 1**) that were designed as detailed below and based on the contigs of HYBV were used to validate the HTS output and bridge gaps by RT-PCR. Termini were characterized by 5′ RACE as detailed by Navarro et al. (2018) on poly-A tailed dsRNAs and poly-dG-tailed cDNAs that were generated using *E. coli* Poly(A) Polymerase (New England Biolabs, Ipswich, MA, United States) and Terminal Deoxynucleotidyl Transferase (Thermo Fisher Scientific, Waltham, MA, United States), respectively. 3′ RACE was performed by RT-PCR using an oligo-dT primer to target the poly-A tracts at the 3′-end of both RNA 1 and 2. Primers employed for genome validation and RACE experiments are detailed in **Supplementary Table 1**. Amplicons were cloned into pGEM-T Easy (Promega, Madison, WI, United States) and three to five clones were sequenced.

Genomic and Proteomic Analyses

The NCBI ORFfinder program[2] was used to identify putative open reading frames (ORFs) *in silico*. Conserved domains were predicted using either the NCBI conserved domain search tool[3] or HMMSCAN[4] implemented in HMMER (Potter et al., 2018). HMMSCAN was also used for the prediction of transmembrane

[1]ftp://ftp.ncbi.nih.gov/genomes/Viruses/all.fna.tar.gz

[2]www.ncbi.nlm.nih.gov/orffinder

[3]www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi

[4]www.ebi.ac.uk/Tools/hmmer/search/hmmscan

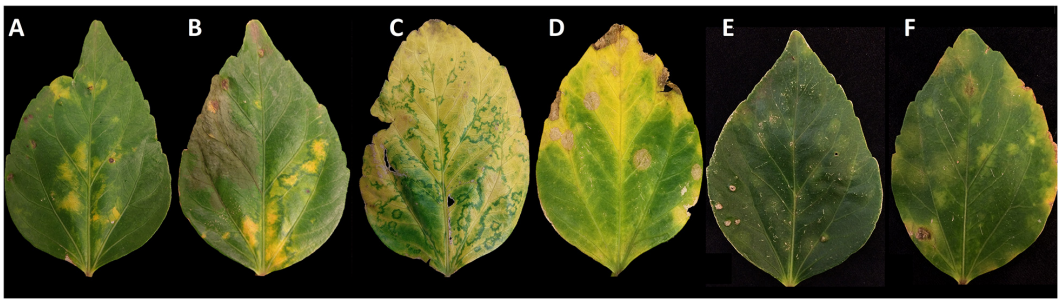


FIGURE 1 | Variation in lesion symptoms displayed by *Hibiscus rosa-sinensis* leaves infected by hibiscus yellow blotch virus. Yellow blotches developed on mature leaves, and a dark green perimeter became visible as the leaves senesced. Lesions observed in panels **(A,B)** were predominant in July 2019; lesions observed in panels **(C,D)** were predominant in November 2019; lesions in panels **(E,F)** were predominant in March 2020. The lesions observed in panel **(D)** present a necrotic center that was sometimes observed in senescent leaves.

helices, signal peptides, coiled coils and protein disorders. Protein transmembrane helices were also predicted and visualized using TMHMM (Krogh et al., 2001) implemented in Geneious 10.1.3. Signal peptides for protein cleavage were predicted using SignalP-5.0[5].

BLASTP searches were used to retrieve protein homologs and infer putative function. Furthermore, putative orphan proteins showing no homology to any protein in any database were aligned using their structural information and the Expresso algorithm (Armougom et al., 2006) implemented in T-Coffee[6] (Di Tommaso et al., 2011). Protein sequence alignment was evaluated using TCS (Chang et al., 2014) implemented in T-Coffee.

Pairwise protein sequence comparisons using orthologous sequences retrieved from GenBank were performed using LALIGN[7] (Huang and Miller, 1991). In addition, the percentage pairwise protein identities of multiple alignments of the replication-associated polyproteins, putative MPs, and the putative structural p23 orthologs were determined using sequence demarcation tool (SDT) 1.2 (Muhire et al., 2014) and the MUSCLE algorithm implemented in SDT 1.2.

The 5′ and 3′ untranslated regions (UTR) of all isolates of recognized and putative/newly described kitaviruses, including HYBV, were analyzed. Intra-species UTR sequences underwent multiple alignment using the ClustalW algorithm (Thompson et al., 1994) implemented in Geneious 10.1.3 (Kearse et al., 2012). Nucleotide composition such as A/T% and their length were determined for all the 5′ and 3′ UTR. Conserved 5′ and 3′ termini were identified, and their length, nucleotide identity as well as the best consensus sequence were manually determined based on these alignments.

Phylogenetic Analyses

Phylogenetic relationships between HYBV and members of the family *Kitaviridae*, *Virgaviridae*, *Bromoviridae*, *Closteroviridae*, and negeviruses were inferred using the amino acid sequences of multiple proteins/domains conserved among some or all of these viruses. Multiple protein alignment was performed with ClustalW (Thompson et al., 1994) implemented in MEGA 7.0.25 (Kumar et al., 2016). Ambiguous positions for each alignment were curated using Gblocks 0.91b[8] (Talavera and Castresana, 2007). The best model of protein evolution for each alignment was used to generate a maximum likelihood tree with 1,000 bootstrap repetitions. Bayesian phylogeny was inferred using BEAST 2.6.2 (Bouckaert et al., 2019) and the best model of protein evolution with three Markov chain Monte Carlo runs of 10,000,000 generations with sampling every 1,000 trees. The runs were combined using LogCombiner in BEAST and 10% of the sample trees were discarded as "burn-in." Tracer 1.7.1 (Rambaut et al., 2018) was used to confirm sample sizes were above 200 for all the parameters. Maximum clade credibility and posterior probabilities were annotated using TreeAnnotator in BEAST 2.6.2. The output trees were visualized in FigTree 1.4.4.

Transmission Electron Microscopy

Using leaf samples collected in March 2020 from the Pearl City *H. rosa-sinensis* tree, ultra-thin sections for transmission electron microscopy were prepared and observed as described (Melzer et al., 2012; Olmedo-Velarde et al., 2019). Briefly, 1 × 2 mm pieces excised from asymptomatic tissue and the margin of lesions from symptomatic leaves were fixed using 2% glutaraldehyde and 2% paraformaldehyde, and post-fixed using a 0.1 M sodium cacodylate solution containing 1% osmium tetroxide. Ultra-thin sections were embedded, stained with uranyl acetate and lead citrate.

Partially purified virion preparations were obtained as described by Colariccio et al. (2000). Briefly, ∼5 g of symptomatic *H. rosa-sinensis* leaf tissue was powdered with liquid nitrogen and mixed with 20 mL of extraction buffer (0.05 M phosphate buffer pH 7.0 containing sodium DIECA, 0.1% (w/v) ascorbic acid and 0.02 M sodium sulphite) for 30 min. After clarification, 0.5% (w/v) sodium chloride and 6% (w/v) 6000 polyethylene glycol were added to the suspension, and partially purified virions were pelleted by centrifugation at 8,000 × *g* for 10 min. The pellet was resuspended with 1 mL of extraction buffer

[5]www.cbs.dtu.dk/services/SignalP-5.0/

[6]http://tcoffee.crg.cat/apps/tcoffee/do:expresso

[7]www.ebi.ac.uk/Tools/psa/lalign

[8]https://ngphylogeny.fr

overnight. All steps were performed at ∼4°C. A 10-fold dilution of the partially purified virion preparations was negatively stained on formvar/carbon-coated grids using 1% uranyl acetate (UA) or 1% phosphotungstic acid (PTA). A density plot based on the diameter of observed particles was created with ggplot2 (Wickham, 2016).

Ultra-thin sections and negatively stained partially purified virion preparations were viewed with a HT7700 120 kV transmission electron microscope (Hitachi High Technologies America Inc., Dallas, TX, United States) at the University of Hawaii Biological Electron Microscope Facility.

Virus Detection

Total RNA was extracted from 100 to 200 mg of plant samples collected in July 2019 using NucleoSpin RNA Plus kit (Macherey-Nagel, Düren, Germany) or Spectrum Plant Total RNA kit (Sigma-Aldrich, St. Louis, MO, United States). These RNA extracts were reverse transcribed into cDNA using random primers and M-MLV reverse transcriptase (Promega) using the manufacturer's protocol. Two microliters of cDNAs were tested by endpoint PCR using GoTaq Green Master Mix (Promega). Virus-specific primer sets for HYBV were designed to target the RNA-dependent RNA polymerase (RdRp) domain and p10 in RNA 1, and p33 in RNA 2, respectively. Primer3 (Untergasser et al., 2012) was used for the primer design with consideration of thermodynamic primer features (Arif and Ochoa-Corona, 2013). HYBV-RdRp-F/R, HYBV-p10-F/R, and HYBV-p33-F/R (**Supplementary Table 1**) were used for specific detection of HYBV in endpoint RT-PCR assays using 0.5 μM as final primer concentration and 55°C as the annealing temperature. Furthermore, one-step quantitative reverse transcription (RT-qPCR) assays were implemented using HYBV-RdRp-F/R, and CiLV-C2-RdRP-F/R and HGSV-2-RdRp-F/R (**Supplementary Table 1**). The two latter primer sets were designed as described above and based on a consensus sequence of an alignment of the RdRp sequences of CiLV-C2 and HGSV-2 available in GenBank as well as sequences from additional isolates of CiLV-C2 and HGSV-2. One-step RT-qPCR assays were implemented using iScript One-Step RT-PCR Kit with SYBR Green (Bio-Rad, Hercules, CA, United States) and 0.125 μM of each primer. Each reaction was performed in three replicates. Cycling parameters for all RT-qPCR assays consisted of cDNA synthesis at 50°C for 40 min. Later, an initial denaturation was performed at 95°C for 1 min, followed by 35 cycles of denaturation at 95°C for 10 s and an annealing-extension step at 60°C for 1 min during which time data was collected. Melt curve analysis was performed as follows: 95°C for 1 min, pre-melting conditioning at 60°C for 1 min followed by a melting temperature cycle range from 60°C to 95°C. Positive cDNA controls, specific for each virus, and non-template controls (DEPC-treated water) were used in all of the assays.

Mite Barcoding and Virus Detection in Flat Mites

Four individual flat mites (*Brevipalpus* spp.), collected in 2020 from the symptomatic *H. rosa-sinensis* tree in Pearl City were used for direct reverse transcriptase (DRT)-PCR assays (Druciarek et al., 2019). Briefly, individual *Brevipalpus* mites were introduced into a PCR tube containing 10 μl of water and random hexamer primers and crushed using a needle under a dissecting microscope. Then, cDNA was synthesized using random primers and SuperScript III reverse transcription kit (Thermo Fisher Scientific) using the manufacturer's instructions. Two microliters of ten-fold diluted cDNA reactions were used in endpoint PCR for DNA barcoding and internal PCR control using the 28S rRNA primers, D1D2w2: 5′-ACAAGTACCDTRAGGGAAAGTTG-3′, 28Sr0990: 5′-CCTTGGTCCGTGTTTCAAGAC-3′ (Sonnenberg et al., 2007; Mironov et al., 2012; Druciarek et al., 2019) that produce a ∼700 bp expected amplicon. The cytochrome oxidase unit I (COI) gene was additionally amplified using the COI primers, DNF: 5′-TACAGCTCCTATAGATAAAAC-3′, DNR: 5′-TGATTTTTTGGTCACCCAGAAG-3′ (Navajas et al., 1996) that produce a ∼450 bp expected amplicon. All DNA barcoding PCR assays were performed using Q5 High Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, United States). Furthermore, endpoint RT-PCR assays using primers HYBV-RdRp-F/R, HYBV-p10-F/R and HYBV-p33-F/R were performed as detailed above to detect the presence of HYBV in the mite specimens. All amplicons were gel extracted, purified and bi-directionally sequenced or cloned into pGEM-T Easy with three clones sequenced per amplicon.

RESULTS

Symptom Monitoring and Virus Indexing

The expression of symptoms in a *H. rosa-sinensis* plant consistent with BTV infection was periodically observed from July 2019 to March 2020. The predominant symptomology varied over this 9-month period, ranging from faint, circular chlorotic blotches (March 2020), chlorotic blotches with a green perimeter (July 2019), to circular necrotic lesions (November 2019) that may represent an advanced stage of the faint, circular chlorotic blotches (**Figure 1**). Symptoms were often most dramatic in senescing leaves. Total RNA extracted from symptomatic leaf tissue (**Figure 1**) tested negative for CiLV-C2 and HGSV-2 in two-step RT-PCR assays. Positive and non-template (water) controls performed as expected (data not shown). Subsequent testing using the two-step RT-PCR assay and primers HYBV-RdRp-F/R revealed that all symptomatic tissues were positive for HYBV (data not shown).

Molecular Characterization of Hibiscus Yellow Blotch Virus

Agarose gel electrophoresis revealed the presence of two dsRNA bands of ∼8 and ∼4.5 Kbp isolated from symptomatic *H. rosa-sinensis tissue* (**Supplementary Figure 1**). HTS of a library generated from these dsRNAs produced ∼24 M paired-end reads that were *de novo* assembled into 2,178 contiguous sequences (contigs). Of these, 27 showed similarity to cileviruses and higreviruses. Iterative mapping of raw reads and reassembling of these contigs led to the generation of three larger contigs of

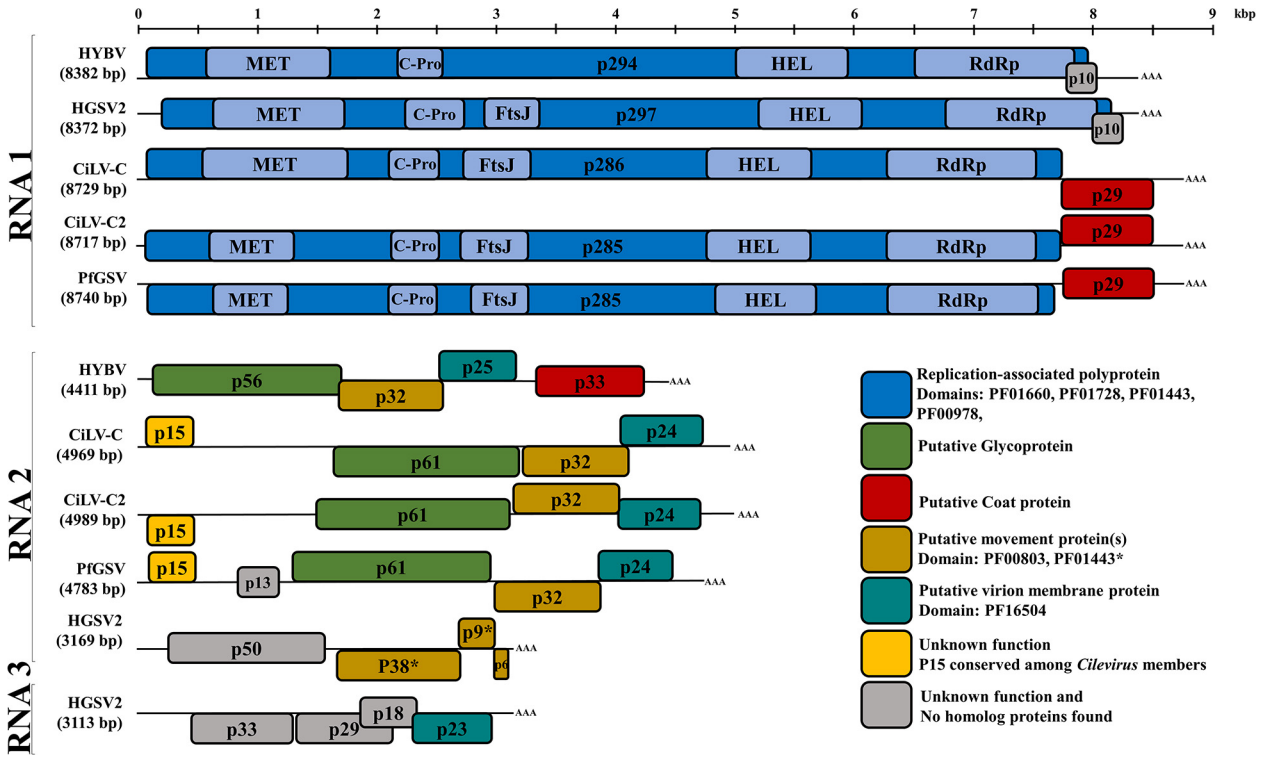


FIGURE 2 | Genome organization comparison between hibiscus yellow blotch virus (HYBV) and *Cilevirus* and *Higrevirus* members. Boxes represent open reading frames (ORFs), and similarly colored ORFs encode proteins of similar function as indicated. RNA 1 (8382 bp) contains two ORFs coding for the replication-associated polyprotein (p294) and a protein with unknown function (p10). RNA 2 (4411 bp) contains four ORFs coding for the putative glycoprotein (p56), putative movement protein (p32), putative virion membrane protein (p25) and putative coat protein (p33). Cileviruses represented: citrus leprosis virus C (CiLV-C) and CiLV-C2, and the proposed cilevirus, passion fruit green spot virus (PfGSV). Higreviruses represented: hibiscus green spot virus 2 (HGSV2). MET, methyltransferase; C-Pro, cysteine like protease; FtsJ, methyltransferase; HEL, helicase; RdRp, RNA-dependent RNA polymerase; AAA, poly(A) tail. Scale at top of figure indicates approximate size in kilobase pairs.

8,127, 3,676, and 507 bp in length. BLASTX searches showed the three contigs putatively coded proteins showing low to moderate identity to the replication-associated polyprotein coded by RNA 1 of cileviruses and HGSV-2, p61 through p24 coded by RNA 2 of cileviruses, and to p29 coded by RNA 1 of cileviruses, respectively. Using the sequence of the 8,127 bp contig, 5′ and 3′ RACE was performed to complete the RNA 1. Using the sequence of the 3,676 bp contig, 5′ RACE was used to determine the 5′ terminal sequence, and 3′ RACE using primer 909 (**Supplementary Table 1**) resulted in a ∼1,620 bp amplicon which included the sequence of the 507 bp contig. RT-PCR using primer 909 and HYBV-p33-R (**Supplementary Table 1**) validated the 3′ RACE result by bridging the ∼307 nt sequence gap between the 3,676 and 507 bp contigs.

Excluding the poly-A tails at their 3′ end, RNA 1 and RNA 2 of HYBV were 8,382 and 4,411 nt, respectively (**Figure 2**). The 8,382 bp RNA 1 molecule (GenBank accession MT472637) had a large ORF that putatively encoded a 294 kDa replication-associated polyprotein of 2,645 amino acids (aa). This putative polyprotein possessed viral methyltransferase (MET; PF01660, aa residues 144–495), cysteine protease (C-Pro; PF02338, aa residues 681–816), viral helicase 1 (HEL; PF01443, aa residues 1,640–1,932) and RdRp 2 (PF00978, aa residues 2,148–2,587) domains (**Figure 2**). This polyprotein was most similar to that of CiLV-C2H, with an identity of 37% (**Table 1** and **Supplementary Figure 2**). The individual MET, C-Pro, HEL and RdRp 2 domains were 49%, 39%, 44% and 61% identical to those of CiLV-C2H, respectively. No transmembrane helices were found in this protein, however, coiled coil and disorder regions were identified at aa residues 1620–1640 and 1461–1468, respectively. In the 3′-terminal region of RNA 1, an 85 aa ORF was identified that putatively encoded a 10 kDa protein harboring two transmembrane helices. This protein of unknown function, designated p10, showed no similarity to any protein in the current databases as determined by a BLASTP search. However, it did resemble the size, genomic location and secondary structure of a previously undescribed ORF of HGSV-2 (**Figure 2** and **Supplementary Figure 3A**). A structural alignment of the two putative proteins obtained using Expresso implemented in T-Coffee revealed an identity of 20% (**Table 1** and **Supplementary Figures 3B,C**).

The 4,411 bp RNA 2 molecule (MT472638) possessed four ORFs. From the 5′ end, the first ORF putatively encoded for a 485 aa protein with a molecular weight of 56 kDa. This protein possessed three transmembrane helices in the C-terminus and shared low homology (19–22% aa identity) to the cilevirus p61

TABLE 1 | Percent amino acid identities between orthologous proteins of hibiscus yellow blotch virus (HYBV) and *Kitaviridae* members: higreviruses include: hibiscus green spot virus 2 (HGSV-2); cileviruses include: citrus leprosis virus C (CiLV-C) and CiLV-C2 (with isolates indicated), and the proposed cilevirus, passion fruit green spot virus (PfGSV); blunerviruses include, blueberry necrotic ring blotch virus (BNRBV) and tea plant necrotic ring blotch virus (TPNRBV), and the putative blunervirus tomato fruit blotch virus (TFBV).

Protein	HGSV2[2]	CiLV-C (CRD)	CiLV-C (SJP)	CiLV-C2 (Citrus-Co)	CiLV-C2 (Hibiscus-HI)	PfGSV (Snp1)	BNRBV	TPNRBV	TFBV
RNA1-Polyprotein	33.9	35.9	36.1	36.3	**36.6**	36.4	32.1	26.3	26.0
RNA1-p10	19.7	–	–	–	–	–	–	–	–
RNA2-p56[1]	–	20.8	**22.3**	20.7	19.5	19.1	–	–	–
RNA2-p32	–[2]	40.6	39.6	**45.9**	44.4	45.3	36.6	32.7	34.1
RNA2-p25[1]	27.7	**40.0**	39.4	38.2	37.3	39.8	23.2	26.9	22.4
RNA2-p33[1]	–	20.3	23.1	25.6	**32.8**	28.1	–	–	–

Orthologous proteins showing the highest identity to the HYBV proteins are bold.
[1]*p56, p25, and p33 of HYBV shows resemblance to p61, p24 and p29 of cileviruses, respectively.*
[2]*HGSV-2 codes for a triple gene block-like movement proteins rather than 3A or 30K movement proteins (Melzer et al., 2018).*

(**Table 1**), which represents the putative glycoprotein. A signal peptide that has been predicted for the p61 of some cileviruses (Ramos-Gonzalez et al., 2020) was also identified in the HYBV homolog, with a predicted cleavage site at Arg_{20}/Val_{21}. The second ORF putatively encoded for a 301 aa protein (32 kDa) in which several disorder regions were identified with the largest located in the C-terminus, at aa residues 270–301. This protein has a conserved domain (PF00803, aa residues 6–227) of viral movement proteins (MP) and likely represents the MP of HYBV. It shared moderate (33–46%) identity with orthologs of kitaviruses (**Table 1**). The third ORF putatively encoded a 230 aa protein, p25, and harbors an SP24 conserved domain (PF16504, aa residues 33–156) which possesses four transmembrane helices. SP24 is present in kitavirids, negeviruses and chroparaviruses, with the latter two being taxons of insect-specific viruses (Kuchibhatla et al., 2014). Two disorder regions were found in the C-terminus of the HYBV p25, and the protein shared moderate (23–40%) identity with orthologs of recognized kitaviruses (**Table 1**). The 3′-terminal ORF putatively encoded a 312 aa protein orthologous (20–33% aa identity) to p29 encoded on RNA1 of cileviruses (**Table 1**), which is a predicted coat protein (Leastro et al., 2018). A long disorder region was identified in aa 52–163 of the p33.

Transmission Electron Microscopy

Ultra-thin sections prepared from foliar lesions revealed the presence of small congregations of spherical structures approximately 50–60 nm in diameter (**Figure 3**). These spherical structures, contained in cytosolic vesicles, were typically in close proximity to the endoplasmic reticulum. Although not common in cells from symptomatic tissue, these structures were not observed in ultra-thin sections obtained from asymptomatic leaves. Despite extensive examination of six thin sections from both symptomatic and asymptomatic leaves, neither electron-dense viroplasm or bacilliform virus-like particles were observed.

Abundant rod-shaped and spherical particles were observed in the partially purified, negatively stained virus preparations. Rod-shaped virions were approximately 20 nm in width and up to 370 nm in length, and were consistent with tobamovirus virion morphology (**Figure 3**). Spherical particles encompassing a wide range of diameters were observed. A density plot based on the measurement of 156 particles indicated diameters of 33 and 66 nm were most prevalent, and these particle sizes were most readily observed with PTA and UA staining, respectively (**Figure 3**).

Phylogenetic Placement of Hibiscus Yellow Blotch Virus

Phylogenies of HYBV were inferred with two character-based algorithms: Maximum likelihood and Bayesian inference. Both algorithms predicted a similar relationship between HYBV and other viruses for each protein sequence analyzed. For the RdRp and p24 proteins, HYBV formed a monotypic lineage between the cilevirus and higrevirus clades (**Figure 4** and **Supplementary Figure 4**). For the p61 and p29 proteins, which only have homologs in cileviruses, HYBV consistently formed a basal branch (**Supplementary Figure 4**). For the MP, which has homologs in cileviruses, blunerviruses, and other plant-infecting viruses, HYBV was placed between the cilevirus and blunervirus clades (**Supplementary Figure 4**). Furthermore, in the RdRp and p24 phylogenies, *Kitaviridae* members and negeviruses formed a monophyletic group sharing a common ancestor (**Figure 4** and **Supplementary Figure 4**).

Protein Comparisons Between Genera in the *Kitaviridae*

To evaluate genetic divergence between HYBV and cileviruses in the context of inter-genera divergence within the family *Kitaviridae*, protein sequence identity matrices of the replication-associated polyprotein, p24, and putative MP of kitavirids were obtained using SDT 1.2. The matrices show overall low to moderate protein identity levels, from 21 to 45%, for the three proteins of HYBV with their kitavirid homologs (**Table 1** and **Supplementary Figure 2**). Cileviruses, including the recently characterized passion fruit green spot virus (PfGSV), present high protein intra-genus identities ranging from 59 to 81, 62 to 89, and 51 to 75% for the replication-associated polyprotein, p24 and MP, respectively. However, cileviruses present a low protein identity, not surpassing 45%, with their homologs of HYBV, HGSV-2 and blunerviruses. The sole higrevirus, HGSV-2, has low protein identity ranges of 24–36 and 22–33% with other kitavirids for the replication-associated polyprotein and p24,

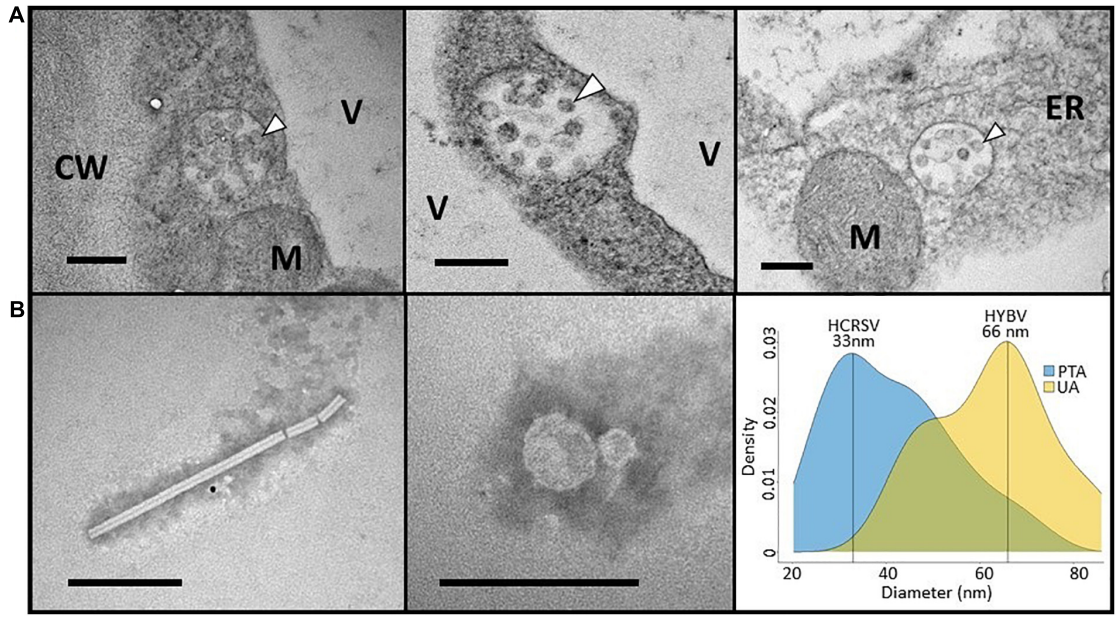


FIGURE 3 | (A) Electron micrographs of hibiscus yellow blotch virus-infected *Hibiscus rosa-sinensis* leaves containing aggregates of electron dense spherical structures (arrowheads) typically between 50 and 60 nm in diameter. Similar structures were not observed in healthy tissue. CW, cell wall; ER, endoplasmic reticulum; M, mitochondrion; V, vacuole. **(B)** Electron micrographs of a rod-shaped virion (left) and spherical particles (center) partially purified from *H. rosa-sinensis* leaves co-infected with hibiscus latent Fort Pierce virus (HLFPV), hibiscus chlorotic ringspot virus (HCRSV), and hibiscus yellow blotch virus (HYBV). The rod-shaped virion is consistent with the HLFPV virion, whereas the spherical particles could represent HCRSV and HYBV virions. A density plot was generated to determine the frequency of spherical particles stained with either phosphotungstic acid (PTA) or uranyl acetate (UA) based on their diameter (right). Particle diameters of 33 and 66 nm were most frequently recorded, which may correspond to HCRSV and HYBV virions, respectively. Bar = 200 nm for all micrographs.

respectively. Blunerviruses also show low protein identity ranges of 24–36, 20–31, and 29–38% with other kitavirids, including other blunerviruses, for the replication-associated polyprotein, p24 and MP, respectively. In general, kitavirids present broad protein identity ranges of 24–81%, 20–89%, and 29–75% for the replication-associated polyprotein, p24 and MP.

Analyses of Untranslated Regions of Kitavirids

Multiple nucleotide sequence alignments revealed that cileviruses, HGSV-2, and HYBV share common features in their UTR, including AT richness. They possess conserved 5′ and 3′ termini, and conserved long 3′ UTRs in all of their genomic segments (**Supplementary Table 2**). All their 5′ termini start with a G or C, while most of their 3′ termini end with a C. Most of their UTR length are uniform among intra-species genomic segments. Furthermore, cileviruses present a long conserved 3′ termini of ∼122 nucleotides and the last three conserved nucleotides are GAC. Similar to cileviruses, HYBV possesses a long conserved 3′ terminus of 171 nucleotides and the last three conserved nucleotides for each RNA segment are GCC. HGSV-2 possesses a shorter conserved 3′ terminus of 78 nucleotides with the last three not clearly conserved. Whereas, blunerviruses also presented AT-rich 5′ and 3′ UTRs, but shorter semi-conserved termini in their genomic segments. Their UTR lengths are variable among intra-species genomic segments.

Hibiscus Yellow Blotch Virus Detection in Flat Mites

Four individual flat mites (*Brevipalpus* spp.) were collected from a symptomatic *H. rosa-sinensis* infected with HYBV. HYBV RNA 1 and RNA 2 were amplified by DRT-PCR in two of these four mites using three different HYBV-specific primer sets (**Supplementary Figure 5**). Direct sequencing showed the three sequences share 100% nucleotide identity to the RdRp and p10 regions in RNA 1 and p33 in RNA 2 of HYBV isolate present in the from the symptomatic *H. rosa-sinensis* tree. Using 28s rRNA primers, a prominent ∼700 bp amplicon of expected size was produced from each of the four individual mites using DRT-PCR, as well as a faint ∼340 bp amplicon which was sequenced and found to be non-specific (**Supplementary Figure 5**). Direct sequencing of the ∼700 bp 28S rRNA amplicon for those two individual mites, and a pairwise alignment showed that both sequences shared 100% nucleotide identity. A BLASTN search using the consensus sequence (MT812697) showed that it shares 99.7% nucleotide identity to *B. yothersi* (MK293649) with 88% query coverage. Direct sequencing of the 28S rRNA region was performed for the other two mites in which HYBV was not detected. A pairwise alignment revealed that both sequences shared 100% nucleotide identity. A BLASTN search of a consensus of both sequences (MT812698) showed that it shares 98.4% and 98.1% nucleotide identity to *B. azores* (MK919272) and *B. feresi* (MK919273), respectively, both with 89% query coverage. Furthermore, direct sequencing of the

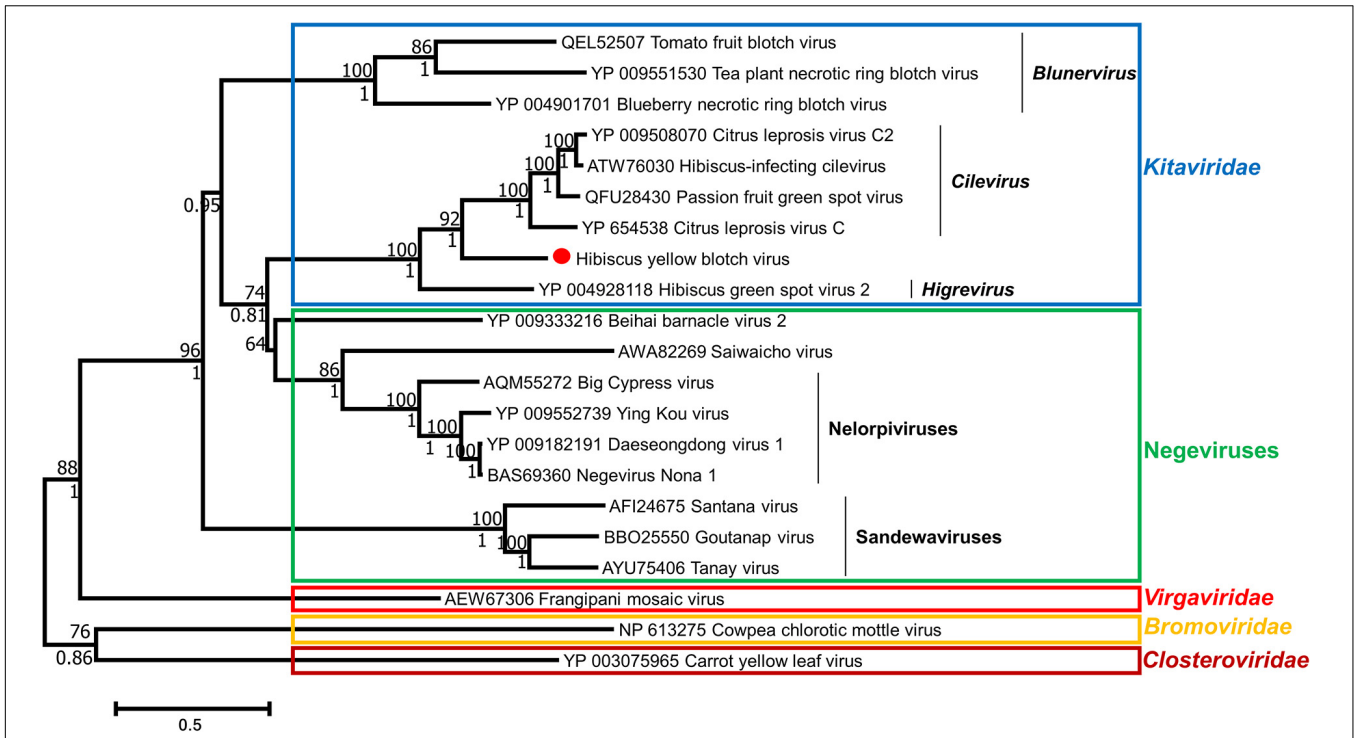


FIGURE 4 | Phylogenetic relationships among members of the families *Kitaviridae*, *Virgaviridae*, *Bromoviridae*, *Closteroviridae* and negeviruses based on an RNA-dependent RNA polymerase (RdRp) multiple protein alignment using CLUSTAL and inferred using maximum likelihood algorithm implemented in MEGA 7.0.25. Bootstrap values generated by maximum likelihood are shown above the branches after 1000 repetitions. Posterior probabilities that were calculated using Bayesian inference with three Markov Chain Monte Carlo runs of 10,000,000 generations and implemented in BEAST 2 are shown under the branches. Missing values indicate values below 50 (bootstrap) or 0.7 (posterior probability). GenBank accession numbers are provided prior to each virus name. Scale at bottom indicates the number of substitutions per given branch length.

∼450 bp amplicons generated using the COI primers (data not shown) further corroborated the identity of *B. yothersi* after a BLASTN search that show the consensus sequence (MT796740) shares 99.3% nucleotide identity to *B. yothersi* (KP180426) with 100% query coverage. A consensus sequence (MT796741) for the COI gene of the two other mites, in which HYBV was not detected, showed 100% identity to *B. obovatus* (DQ450495) after a BLASTN search with a 90% query coverage.

Natural Mixed Infections in Hibiscus

In addition to HYBV, HTS data from the dsRNA library also identified the complete genomic sequences of hibiscus chlorotic ringspot virus (HCRSV, *Betacarmovirus*) and hibiscus latent Fort Pierce virus (HLFPV, *Tobamovirus*) in the Pearl City sample using *de novo* assembly and an iterative mapping approach. The genome of HCRSV from the sample was 3,969 nt in length (MT512573) and assembled from 1,386,006 reads with and average depth of 33,785. The genome shared 95.4% nucleotide identity to an HCRSV isolate from Singapore (X86448). No dsRNAs associated with this virus were observed following agarose gel electrophoresis, however, any dsRNAs of this size would have been obscured by the loading dye (**Supplementary Figure 1**). The RdRp, p2, p3, and CP of HCRSV shared 95%, 95.7% and 92.3% protein identity, respectively, to those of the HCRSV isolate from Singapore, while the CP shared 99.4% protein identity to that of HCRSV isolate SB01 from Brazil (AZL87708). The genome of HLFPV was 6,408 nt in length (MT512572) assembled from 1,667,706 reads with an average depth of 35,941. The genome shared 99.4% nucleotide identity to HLFPV isolate J from Japan (AB917427). This genome size was consistent in size with a faint dsRNA observed following agarose gel electrophoresis (**Supplementary Figure 1**) Replication-associated polyprotein, MP and CP of HLFPV show 98.5%, 99.7%, and 97.5% protein identity, respectively, to those of HLFPV isolate J (AB917427).

Samples from six *Hibiscus* spp. plants designated A-F displaying symptoms typical of BTV infection (Kitajima et al., 2003) were collected within 1.6 km of the original HYBV-infected tree (**Supplementary Figure 6**). The presence of CiLV-C2, HGSV-2 and HYBV, was assessed by one-step RT-qPCR and two-step RT-PCR assays. Three out of six of the samples (A, B, and D) tested positive for HYBV in both PCR assays. Only one sample (A), which also tested positive for HYBV, tested positive for CiLV-C2 in both PCR assays. None of the samples tested positive for HGSV-2. In all of the assays, positive and non-template controls tested as expected (**Supplementary Figure 7**). The genetic diversity of HYBV in these samples was assessed for RNA1 and RNA2 by sequencing amplicons generated using primer sets HYBV-RdRp-F/HYBV-p10-R and 909/HYBV-p33-R,

respectively (**Supplementary Table 1**). These amplicons were 100% identical for RNA1 and >99.6% identical for RNA2.

DISCUSSION

The *Kitaviridae* family is a recently created taxon currently comprised of three genera of plant pathogenic viruses: *Blunervirus*, *Cilevirus*, and *Higrevirus* (Melzer et al., 2018; Quito-Avila et al., 2020). Kitavirids share an evolutionary history based on the three proteins they commonly possess: replication-associated polyprotein, p24 protein, and in the case of blunerviruses and cileviruses, the MP. Additionally, mites (eriophyid and flat or false spider) have been reported as the putative or confirmed vectors of these viruses (Burkle et al., 2012; Rodrigues et al., 2016). Interestingly, systemic infection of their plant host is rare; most kitavirid infections of plants are restricted to localized lesions (Quito-Avila et al., 2020). Several *Kitaviridae* members have been characterized through the use of HTS (Quito-Avila et al., 2013; Roy et al., 2013; Hao et al., 2018; Ramos-Gonzalez et al., 2020) which is an increasingly important and efficient technique for virus discovery and characterization, including in numerous studies in different agricultural systems (Villamor et al., 2019).

In this study, we have characterized a new kitavirid infecting *H. rosa-sinensis* in Hawaii using HTS of a dsRNA library. The symptoms displayed on leaves infected with this virus resemble those associated with infection by cileviruses and other kitavirids (**Figure 1**) (Melzer et al., 2013; Rodrigues et al., 2016). RT-PCR assays specific for CiLV-C2 and HGSV-2 indicated these two kitavirids, which commonly infect hibiscus in Hawaii, were absent, suggesting the presence of another pathogen. HTS and bioinformatic analyses allowed the complete bipartite genome of this new kitavirid to be characterized. The genomic RNAs, determined to be 8.3 and 4.4 kb in size, were consistent with the dsRNAs observed by agarose gel electrophoresis (**Supplementary Figure 1**). RNA 1 codes for a RdRp and a putative p10 protein with unknown function. RNA 2 codes for p56, p32, p25 and p33 proteins that likely represent the putative glycoprotein, putative movement protein, putative virion membrane protein and putative coat protein, respectively. This genomic organization resembles that of cileviruses, but with some key differences (**Figure 2**). First, p33 is located in the 3′ region of RNA 2, but the homolog of this protein in cileviruses (p29) is located in the 3′ region of RNA 1. These represent the putative viral coat proteins (Leastro et al., 2018), and this observation suggests a gene rearrangement event in an ancestor of these viruses. Second, p15 is the most variable conserved orphan protein present in the 5′ region of RNA 2 of the three classified and putative *Cilevirus* species (Ramos-Gonzalez et al., 2020). This new kitavirid lacks a p15 homolog in the 5′ region of its RNA 2. The absence of a p15 homolog was supported by an inability to identify this putative protein *in silico*, using 5′ RACE, and through degenerate RT-PCR assays targeting conserved amino acid sequences identified following the alignment of cilevirus p15 sequences. Based on sequence comparisons and recombination analyses of the 5′ end of the cilevirus RNA 2, it has been previously suggested that the high nucleotide variability present in that region is the result of continuous illegitimate (non-homologous) recombination processes that may occur at the inter-species level (Ramos-Gonzalez et al., 2016). Therefore, recombination processes may have contributed to the loss of a p15 homolog in an ancestor of HYBV. Finally, a 3′-terminal ORF encoding a putative 10 kDa protein is present in HYBV RNA 1. Although the predicted p10 protein shares no homology to viral proteins in in GenBank, a pairwise protein structural alignment of the HYBV p10 and the product of a similarly sized and positioned ORF encoded by HGSV-2 using EXPRESSO T-Coffee revealed they share statistically significant structural similarities. A global pairwise alignment showed they share 19.7 % protein identity. Furthermore, both p10 proteins of HYBV and HGSV-2 possess two transmembrane domains (**Supplementary Figure 3**). This suggests these putative orphan proteins of HYBV and HGSV-2 are either distant orthologs or represent structural convergence of two unrelated proteins. It has been suggested that some orphan proteins may help plant viruses to infect and colonize its arthropod host/vector (Kuchibhatla et al., 2014; Solovyev and Morozov, 2017). Based on this same hypothesis, it has been speculated that kitaviruses have two sets of movement genes, and specifically cileviruses have the MP gene in conjunction with the p24 (putative virion membrane protein) and p61 (putative glycoprotein) proteins involved in virus movement within the mite host (Solovyev and Morozov, 2017). Based on the genomic organization of HYBV, this new kitavirid represents a distinct member within the family *Kitaviridae* that shares genomic similarities with members of both *Cilevirus* and *Higrevirus* genera.

Phylogenetic analyses of all conserved protein products support the placement of HYBV within the *Kitaviridae* family, specifically in an intermediate position between the cilevirus and higrevirus clades (**Figure 4** and **Supplementary Figure 4**). The sequence identities of the three conserved kitavirid proteins, namely RdRp, p24 and MP, of HYBV with kitavirid homologs was found to be low to moderate (<45%). A similar scenario is observed for HGSV-2, the sole *Higrevirus* member and members of the genus *Blunervirus* (<38%). Conversely, the *Cilevirus* proteins have high identities (>54%) (**Table 1** and **Supplementary Figure 2**). The relatively high protein identity levels among the *Cilevirus* protein homologs and the phylogenetic relatedness may suggest a more recent divergence within the genus, whereas blunerviruses diverged much earlier or might have undergone a more rapid evolution due to host or vector selective pressures. Considering the low protein identity level in the genus *Blunervirus*, it is plausible that in the future when more viruses belonging to the blunervirus clade are characterized, the genus may be divided into additional genera. The phylogenetic placement of HYBV, coupled with its distinctive genome organization, suggests this virus represents a distinct lineage intermediate of cileviruses and higreviruses. However, until more kitavirids are characterized which will allow greater resolution of the family's taxonomy, it seems appropriate to consider HYBV a basal member of the genus *Cilevirus*, as it shares the most features with members of this taxon.

Multiple nucleotide sequence alignments of the 5′ and 3′ UTR revealed that HYBV possesses long conserved 5′ and 3′ UTRs rich in AT among its genomic segments (**Supplementary Table 2**), a feature shared by other cileviruses, most notably CiLV-C. Both 5′ and 3′ termini of HYBV RNA 1 and 2 were also highly conserved, with the latter being longer and sharing a higher nucleotide identity. Previously, it was found that the blunervirus BNRBV contains AT-rich UTRs and conserved termini (Quito-Avila et al., 2013). Also, the formation of stem-loop secondary structures in the 3′ UTR of the four RNAs was predicted and may be associated to the regulation of virus genome replication as well as protein synthesis (Quito-Avila et al., 2013). Interestingly, most of the cileviruses, HYBV, HGSV-2 and BNRBV have a G/C, and C as their first and last nucleotide, respectively (**Supplementary Table 2**). While sequencing clones for the 5′ RACE experiments for HYBV RNA 1 and 2, some indicated G (rather than the more common C) as the first nucleotide (data not shown). Considering the multiple plant hosts, limited *Brevipalpus* species associated to kitavirids, and conserved UTRs in some kitavirids, it is plausible that these regions may play a role in regulation of different virus infectious cycle processes within their hosts.

Flat mites are polyphagous and the presence of several *Brevipalpus* species on the same host plant has been reported previously (Salinas-Vargas et al., 2016). DNA barcoding using dRT-PCR (Druciarek et al., 2019) targeting the 28S rRNA and COI indicated mites tentatively identified as *B. yothersi* and an unidentified *Brevipalpus* sp. were present on the HYBV-infected *H. rosa-sinensis* plant. For this latter specimen, the COI sequence suggested its identity to be *B. obovatus*, however, the 28S rRNA sequence provided no clear identity, with resemblance to sequences attributed to *B. azores* and *B. feresi*. *B. yothersi*, *B. azores* and *B. feresi* are three of the seven recently created *Brevipalpus* species that were derived from the *B. phoenicis* species complex (Beard et al., 2015). Therefore, the identity of this *Brevipalpus* specimen needs further identification based on morphological keys using scanning electron microscopy (Beard et al., 2013). *B. yothersi* has been reported as the main vector of CiLV-C and CiLV-C2 (Roy et al., 2013; García-Escamilla et al., 2018; Ferreira et al., 2020). In this study we demonstrated the ingestion and potential acquisition of HYBV by *B. yothersi* using dRT-PCR assays targeting both RNA 1 and 2 of the virus. Considering the low number of individual *Brevipalpus* used in this study, the non-systemic nature of infection, and the ability of several *Brevipalpus* species to vector CiLV-C (Nunes et al., 2018), it is plausible that the other *Brevipalpus* sp. that could not be identified by DNA barcoding in this study may play a role in the transmission of HYBV. Additional transmission experiments are required to validate the transmission of HYBV by *B. yothersi* and other *Brevipalpus* species, and include confirmation of the mite species.

High-throughput sequencing data indicated that HCRSV and HLFPV were co-infecting the *H. rosa-sinensis* plant harboring HYBV. HCRSV infection has been associated with distinct mild ringspot symptoms and HLFPV with a latent infection (Kamenova and Adkins, 2004; Zhou et al., 2006). Neither of these are consistent with the observed symptoms of yellow chlorotic blotches with a dark green perimeter that are typical of a BTV infection (**Figure 1** and **Supplementary Figure 6**), making HYBV the most likely causal agent of the observed symptoms. However, the presence of HCRSV and HLFPV, as well as the presence of CiLV-C2, as co-infections with HYBV may impact host symptoms and other aspects of this pathosystem. The development of an infectious clone of HYBV would greatly help elucidate its role as a causal agent of the observed disease. Until such a clone is available, the conventional and quantitative RT-PCR assays developed in this study for the detection of HYBV will help to further determine any relationship between HYBV infection, co-infection with other pathogens, and symptom expression. Although the presence of non-viral pathogens or physiological disorders cannot be excluded, the observed symptoms were absent from other hibiscus plants in the immediate area that would presumably be exposed to the same pathogen inoculum and growing conditions as the symptomatic plant. Additional symptomatic plants were observed up to 1.6 km away, some of which tested negative for HYBV, CiLV-C2, and HGSV-2 (**Supplementary Figure 6**). This suggests additional related viruses may be present in Hawaii's hibiscus, adding to the complexity of this pathosystem.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/genbank/, MT472637 and MT472638.

AUTHOR CONTRIBUTIONS

AO-V and MM worked on the investigation, methodology, formal analyses, and supervised the study. AO-V and MM conceptualized and wrote the original draft. AO-V, JH, and MM wrote, reviewed, and edited the manuscript. MM was responsible for funding acquisition. MM and JH were responsible for the resources. All authors contributed to the article and approved the submitted version.

FUNDING

This project was funded, in part, by USDA-NIFA Hatch Project HAW09050-H awarded to MM, and managed by the College of Tropical Agriculture and Human Resources at the University of Hawaií at Mānoa.

ACKNOWLEDGMENTS

The authors thank Ms. Tina Carvalho (Biological Electron Microscope Facility, University of Hawaií at Mānoa) for providing technical assistance and expertise in the transmission electron microscope.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.660237/full#supplementary-material

Supplementary Figure 1 | Agarose gel electrophoresis of dsRNAs extracted from hibiscus yellow blotch virus (HYBV)-infected *Hibiscus rosa-sinensis* leaves (right lane). Blue arrows at ∼8 Kbp and ∼4.5 Kbp indicate the observed dsRNAs that likely represent RNA 1 and RNA 2 of HYBV, respectively. The black arrow at ∼6 Kbp likely represents the genome of hibiscus latent Fort Pierce virus. Numbers on left give size in kilobase pairs (Kbp) for select fragments of the Thermo Fisher 1 kb Plus ladder (left lane).

Supplementary Figure 2 | Pairwise protein sequence similarity matrix of the replication-associated polyprotein, p24 and putative movement protein of hibiscus yellow blotch virus (HYBV) with their homologs of *Cilevirus*, *Higrevirus*, and *Blunervirus* members using sequence demarcation tool (SDT) 1.2. Cilevirus species and isolates include: citrus leprosis virus C (CiLV-C) and CiLV-C2 (Citrus isolate from Colombia: Citrus-Col. Hibiscus isolate from Hawaii: HibiscusHI), and the proposed cilevirus, passion fruit green spot virus (PfGSV). Higrevirus member: hibiscus green spot virus 2 (HGSV2). Blunerviruses include: blueberry necrotic ring blotch virus (BNRBV) and tea plant necrotic ring blotch virus (TPNRBV), and the proposed blunervirus, tomato fruit blotch virus (TFBV).

Supplementary Figure 3 | Structural alignment results of hibiscus yellow blotch virus (HYBV) p10 and hibiscus green spot virus 2 (HGSV-2) p10 using the Expresso algorithm implemented in T-Coffee. **(A)** Presence of two transmembrane helices in both HYBV-p10 and HGSV-2-p10 proteins and visualization of the structural alignment. **(B)** Expresso T-Coffee matrix result showing a significant score for the alignment (>50) and at least three structurally conserved motifs are observed in the alignment (Highlighted in light red). **(C)** Transitive consistency score of the protein alignment showing congruent structurally conserved positions revealed by Expresso-T-Coffee.

Supplementary Figure 4 | Phylogenetic placement of hibiscus yellow blotch virus (HYBV) with other viruses possessing conserved proteins. **(A)** MP, putative movement protein (p32 of HYBV); **(B)** p61, putative glycoprotein (p56 of HYBV) and p29, putative coat protein (p33 of HYBV); and **(C)** p24, putative virion membrane protein (p25 of HYBV). All phylogenies were inferred by a multiple protein alignment using CLUSTAL and the maximum likelihood algorithm implemented in MEGA 7.0.25. Bootstrap values generated by maximum likelihood are shown above the branches after 1000 repetitions. Posterior probabilities that were calculated using Bayesian inference with three Markov Chain Monte Carlo runs of 10,000,000 generations and implemented in BEAST 2 are shown under the branches. Missing values indicate values below 50 (bootstrap) or 0.7 (posterior probability).

Supplementary Figure 5 | Agarose gel electrophoresis of 28S rRNA, hibiscus yellow blotch (HYBV) RNA 1 and RNA 2 amplicons using direct RT-PCR (dRT-PCR) assays described by Druciarek et al. (2019). dRT-PCR assays were performed on four individual *Brevipalpus* mites (B1–B4) collected from a HYBV-infected *Hibiscus rosa-sinensis* plant. Upper panel: 28S rRNA and HYBV-RNA 1 amplicons were amplified using primer sets D1D2w2/28Sr0990 and HYBV-RdRp-F/R with expected amplicon sizes of ∼700 and 223 bp, respectively. Bottom panel: HYBV-RNA 1 and RNA2 amplicons were amplified using primer sets HYBV-p10-F/R and HYBV-p33-F/R with expected amplicons of 210 and 438 bp, respectively. A non-template control (N) was included in all the assays. Sanger sequencing of the 28S rRNA amplicon later suggested mites B1 and B2 were *B. yothersi*, while the identity of mites B3 and B4 was ambiguous. DNA ladder is 100 bp (Thermo Fisher), and sizes of select fragments are provided in base pairs (bp).

Supplementary Figure 6 | Leaves collected from symptomatic *Hibiscus* spp. plants growing with 1.6 km of the hibiscus yellow blotch virus (HYBV)-infected *H. rosa-sinensis* plant. Samples A, B, and D tested positive for HYBV using RT-PCR, whereas Samples C, E, and F tested negative. Sample A also tested positive for citrus leprosis virus C2 by RT-PCR. None of the samples **(A–F)** tested positive for hibiscus green spot virus 2 by RT-PCR.

Supplementary Figure 7 | Detection of hibiscus yellow blotch virus (HYBV), citrus leprosis virus C2 (CiLV-C2), and hibiscus green spot virus 2 (HGSV2) in *Hibiscus* spp. samples by RT-qPCR and RT-PCR assays. Six samples (A–F) of *Hibiscus* spp. (**Supplementary Figure 3**) collected within 1.6 km of the HYBV-infected *H. rosa-sinensis* underwent both assays using HYBV-RdRp-F/R (main text), CiLV-C2-RdRp-F/R and HGSV2-RdRp-F/R (**Supplementary Table 1**) for virus detection. Samples A, B, and D tested positive for HYBV in both one-step RT-qPCR **(A)** and two-step RT-PCR **(B)** assays. Due to similar amplification curve of samples A (dark red), B (blue), and D (light green), individual amplification curves are provided for each sample and compared with a positive control (red). Sample A was the only CiLV-C2-positive sample in both one-step RT-qPCR **(A)** and two-step RT-PCR **(B)** assays. No samples tested positive for HGSV2. Expected size of amplicons for HYBV, CiLV-C2, and HGSV2 are 223 bp, 386 bp, and 227 bp, respectively. NTC and (N) represent non-template controls for **(A)** and **(B)**, respectively. The DNA ladder in panel **(B)** is the 100 bp Ladder from Thermo Fisher, and fragment sizes are provided at left in base pairs (bp).

Supplementary Table 1 | Primers used in this study for 5′ and 3′ RACE, genomic sequence validation, and detection of hibiscus yellow blotch virus (HYBV). Primers used for the detection of citrus leprosis virus C2 (CiLV-C2) and hibiscus green spot virus 2 (HGSV-2) in one-step RT-qPCR assays are also provided.

Supplementary Table 2 | Kitavirus intra-species nucleotide composition analysis of the 5′ and 3′ untranslated regions (UTR), and 5′ and 3′ conserved termini among the genomic segments of each virus species.

REFERENCES

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.

Arif, M., and Ochoa-Corona, F. (2013). Comparative assessment of 5' A/T-rich overhang sequences with optimal and sub-optimal primers to increase PCR yields and sensitivity. *Mol. Biotechnol.* 55, 17–26. doi: 10.1007/s12033-012-9617-5

Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., et al. (2006). Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res.* 34, W604—W608. doi: 10.1093/nar/gkl092

Beard, J. J., Ochoa, R., Bauchan, G. R., Trice, M. D., Redford, A. J., Walters, T. W., et al. (2013). *Flat Mites of the World*, 2 Edn. Fort Collins, CO: CPHST.

Beard, J. J., Ochoa, R., Braswell, W. E., and Bauchan, G. R. (2015). Brevipalpus phoenicis (Geijskes) species complex (Acari: Tenuipalpidae) a closer look. *Zootaxa* 3944, 1–67.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., et al. (2019). BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 15:e1006650. doi: 10.1371/journal.pcbi.1006650

Burkle, C., Olmstead, J. W., and Harmon, P. F. (2012). A potential vector of Blueberry necrotic ring blotch virus and symptoms on various host genotypes. *Phytopathology* 102:S4.17.

Chang, J. M., Di Tommaso, P., and Notredame, C. (2014). TCS: a new multiple sequence alignment reliability measure to estimate alignment accuracy and improve phylogenetic tree reconstruction. *Mol. Biol. Evol.* 31, 1625—1637. doi: 10.1093/molbev/msu117

Colariccio, A., Lovisolo, O., Boccardo, G., Chagas, C. M., d'Aquilio, M., and Rossetti, V. (2000). Preliminary purification and double stranded RNA analysis of citrus leprosis virus. *Proc. Conf. Int. Organ. Cit. Virol.* 14, 159–163.

Cuiffo, M., Kinoti, W. M., Tiberini, A., Forgia, M., Tomassoli, L., Constable F. E., et al. (2020). A new blunervirus infects tomato crops in Italy and Australia. *Arch. Virol.* 165, 2379–2384. doi: 10.1007/s00705-020-04760-x

Dey, K. K., Sugikawa, J., Kerr, C., and Melzer, M. J. (2019). Air potato (*Dioscorea bulbifera*) plants displaying virus-like symptoms are co-infected with a novel potyvirus and a novel ampelovirus. *Virus Genes* 55, 117–121. doi: 10.1007/s11262-018-1616-6

Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J. M., et al. (2011). T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res.* 39, W13—W17. doi: 10.1093/nar/gkr245

Druciarek, T., Lewandowski, M., and Tzanetakis, I. (2019). A new, sensitive and efficient method for taxonomic placement in the Eriophyoidea and virus detection in individual eriophyoids. *Exp. Appl. Acarol.* 78, 247–261. doi: 10.1007/s10493-019-00382-4

Ferreira, L. M., Nunes, M. A., Sinico, T. E., Soares, A. J., and Novelli, V. M. (2020). Brevipalpus species vectoring citrus leprosis virus (*Cilevirus* and *Dichorhavirus*). *J. Econ. Entomol.* 113, 1628–1634.

Freitas-Astúa, J., Ramos-González, P. L., Arena, G. D., Tassi, A. D., and Kitajima, E. W. (2018). Brevipalpus-transmitted viruses: parallelism beyond a common vector or convergent evolution of distantly related pathogens? *Curr. Opin. Virol.* 33, 66–73. doi: 10.1016/j.coviro.2018.07.010

García-Escamilla, P., Duran-Trujillo, Y., Otero-Colina, G., Valdovinos-Ponce, G., Santillán-Galicia, M. T., Ortiz-García, C. F., et al. (2018). Transmission of viruses associated with cytoplasmic and nuclear leprosis symptoms by Brevipalpus yothersi and B. californicus. *Trop. Plant Pathol.* 43, 69–77.

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.

Hao, X., Zhang, W., Zhao, F., Liu, Y., Qian, W., Wang, Y., et al. (2018). Discovery of plant viruses from tea plant (Camellia *sinensis* (L.) O. Kuntze) by metagenomic sequencing. *Front. Microbiol.* 9:2175. doi: 10.3389/fmicb.2018.02175

Huang, X., and Miller, W. (1991). A time-efficient, linear-space local similarity algorithm. *Adv. Appl. Math.* 12, 337–357.

Kamenova, I., and Adkins, S. (2004). Transmission, in planta distribution, and management of hibiscus latent fort pierce virus, a novel tobamovirus isolated from florida hibiscus. *Plant Dis.* 88, 674–679.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649.

Kitajima, E. W., Chagas, C. M., and Rodrigues, J. C. V. (2003). Brevipalpus-transmitted plant virus and virus-like diseases: cytopathology and some recent cases. *Exp. Appl. Acarol.* 30, 135–160.

Kitajima, E. W., Rosillo, M. A., Portillo, M. M., Müller, G. W., and Costa, A. S. (1974). Microscopia electronica de tecidos foliares de laranjeiras infectadas pela lepra explosive da Argentina. *Fitopatologia (Lima)* 9, 254–258.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580.

Kuchibhatla, D. B., Sherman, W. A., Chung, B. Y., Cook, S., Schneider, G., Eisenhaber, B., et al. (2014). Powerful sequence similarity search methods and in-depth manual analyses can identify remote homologs in many apparently "orphan" viral proteins. *J. Virol.* 88, 10—20. doi: 10.1128/JVI.02595-13

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA 7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874.

Leastro, M. O., Kitajima, E. W., Silva, M. S., Resende, R. O., and Freitas-Astúa, J. (2018). Dissecting the subcelular localization, intracelular trafficking, interactions, membrance association, and topology of Citrus leprosis virus C protein. *Front. Plant Sci.* 9:1299. doi: 10.3389/fpls.2018.01299

Melzer, M. J., Borth, W. B., Sether, D. M., Ferreira, S., Gonsalves, D., and Hu, J. S. (2010). Genetic diversity and evidence for recent modular recombination in Hawaiian Citrus tristeza virus. *Virus Genes* 40, 111–118. doi: 10.1007/s11262-009-0409-3

Melzer, M. J., Freitas-Astúa, J., Rodriguez, J. C. V., Roy, A., and Wei, G. (2018). *Create One New Family Kitaviridae Comprising Three Previously Unassigned Genera, Cilevirus, Blunervirus and Higrevirus 2018.002P. melzer@hawaii.edu*. Available online at: https://talk.ictvonline.org/ictv/proposals/2018.002P.A.Kitaviridae.zip. (accessed 08 Aug, 2020).

Melzer, M. J., Sether, D. M., Borth, W. B., and Hu, J. S. (2012). Characterization of a virus infecting Citrus volkameriana with citrus leprosis-like symptoms. *Phytopathol.* 102, 122–127. doi: 10.1094/PHYTO-01-11-0013

Melzer, M. J., Simbajon, N., Carillo, J., Borth, W. B., Freitas-Astúa, J., Kitajima, E. W., et al. (2013). A cilevirus infects ornamental hibiscus in Hawaii. *Arch. Virol.* 158, 2421–2424. doi: 10.1007/s00705-013-1745-0

Mironov, S. V., Dabert, J., and Dabert, M. (2012). A new feather mite species of the genus Proctophyllodes Robin, 1877 (Astigmata: *Proctophyllodidae*) from the long-tailed tit Aegithalos caudatus (Passeriformes Aegithalidae)-morphological description with DNA barcode data. *Zootaxa* 3253, 54–61.

Morris, T. J., and Dodds, J. A. (1979). Isolation and analysis of doubled-stranded RNA from virus-infected plant and fungal tissue. *Phytopathol.* 69, 854–858.

Muhire, B. M., Varsani, A., and Martin, D. P. (2014). SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One* 9:e108277. doi: 10.1371/journal.pone.0108277

Navajas, M., Fournier, D., Lagnel, J., Gutierrez, J., and Boursot, P. (1996). Mitochondrial COI sequences in mites: evidence for variations in base composition. *Insect Mol. Biol.* 5, 281–285. doi: 10.1111/j.1365-2583.1996.tb00102.x

Navarro, B., Minutolo, M., De Stradis, A., Palmisano, F., Alioto, D., and Di Serio, F. (2018). The first phlebo-like virus infecting plants: a case study on the adaptation of negative-stranded RNA viruses to new hosts. *Mol. Plant Pathol.* 19, 1075–1089. doi: 10.1111/mpp.12587

Nunes, M. A., de Carvalho Mineiro, J. L., Rogerio, L. A., Ferreira, L. M., Tassi, A., Novelli, V. M., et al. (2018). First reporto f Brevipalpus papayensis as vector of coffee ringspot virus and citrus leprosis virus C. *Plant Dis.* 102:1046. doi: 10.1094/PDIS-07-17-1000-PDN

Nunes, M. R. T., Contreras-Gutierrez, M. A., Guzman, H., Martins, L. C., Feitoza Barbirato, M., Savit, C., et al. (2017). Genetic characterization, molecular epidemiology and phylogenetic relationships of insect-specific viruses in the taxon Negevirus. *Virology* 504, 152–167. doi: 10.1016/j.virol.2017.01.022

Olmedo-Velarde, A., Park, A. C., Sugano, J., Uchida, J. Y., Kawate, M., Borth, W. B., et al. (2019). Characterization of Ti Ringspot-associated virus, a novel emaravirus associated with an emerging ringspot disease of cordyline fruticosa. *Plant Dis.* 103, 2345–2352.

Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., and Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Res.* 46, W200–W204. doi: 10.1093/nar/gky448

Quito-Avila, D. F., Brannen, P. M., Cline, O. W., Harmon, P. F., and Martin, R. R. (2013). Genetic characterization of Blueberry necrotic ring blotch virus, a novel RNA virus with unique genetic features. *J. Gen. Virol.* 94, 1426–1434. doi: 10.1099/vir.0.050393-0

Quito-Avila, D. F., Freitas-Astúa, J., and Melzer, M. J. (2020). " Bluner-, Cile-, and higreviruses (*kitaviridae*)," in *Reference Module in Life Sciences*, ed. B. D. Roitberg (Amsterdam: Elsevier), 1–5.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–904. doi: 10.1093/sysbio/syy032

Ramos-Gonzalez, P. L., Chabi-Jesus, C., Guerra-Peraza, O., Breton, M. C., Arena, G. D., Nunes, M. A., et al. (2016). Phylogenetic and molecular variability studies reveal a new genetic clade of Citrus leprosis virus C. *Viruses* 8:153. doi: 10.3390/v8060153

Ramos-Gonzalez, P. L., Santos, G. F., Chabi-Jesus, C., Harakava, R., Kitajima, E. W., and Freitas-Astúa, J. (2020). Passion fruit green spot virus genome harbors a new orphan ORF and highlights the flexibility of the 5'-End of the RNA2 segment across *Cileviruses*. *Front. Microbiol.* 11:206. doi: 10.3389/fmicb.2020.00206

Rodrigues, J. C. V., Childers, C. C., and Kitajima, E. W. (2016). "Brevipalpus spp. (Acari: Tenuipalpidae): vectors of cytoplasmic and nuclear viruses in plants," in *Vector-Mediated Transmission of Plant Pathogens*, ed. J. K. Brown (St. Paul, MN: APS Press), 309–318.

Roy, A., Choudhary, N., Leon, G. M., Shao, J., Govindarajulu, A., Achor, D., et al. (2013). A novel virus of the genus *Cilevirus* causing symptoms similar to citrus leprosis. *Phytopathology* 103, 488–500. doi: 10.1094/PHYTO-07-12-0177-R

Roy, A., Hartung, J. S., Schneider, W. L., Shao, J., Leon, G., Melzer, M. J., et al. (2015). Role bending: complex relationships between viruses, hosts, and vectors related to citrus leprosis, an emerging disease. *Phytopathology* 105, 1013–1025. doi: 10.1094/PHYTO-12-14-0375-FI

Salinas-Vargas, D., Santillan-Galicia, M. T., Guzman-Franco, A. W., Hernandez-Lopez, A., Ortega-Arenas, L. D., and Mora-Aguilera, G. (2016). Analysis of genetic variation in Brevipalpus yothersi (Acari: Tenuipalpidae) populations from four species of citrus host plants. *PLoS One* 11:e0164552. doi: 10.1371/journal.pone.0164552

Shi, M., Lin, X.-D., Tain, 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

Solovyev, A. G., and Morozov, S. Y. (2017). Non-replicative integral membrane proteins encoded by plant Alpha-like viruses: emergence of diverse orphan ORFs and movement protein genes. *Front. Plant Sci.* 8:1820. doi: 10.3389/fpls.2017.01820

Sonnenberg, R., Wolte, A. W., and Tautz, D. (2007). An evaluation of LSU rDNA D1-D2 sequences for their use in species identification. *Front. Zool.* 4:6. doi: 10.1186/1742-9994-4-6

Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56, 564–577. doi: 10.1080/10635150701472164

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., et al. (2012). Primer3 - new capabilities and interfaces. *Nucleic Acids Res.* 40:e115.

Villamor, D. E. V., Ho, T., Al Rwahnih, M., Martin, R. R., and Tzanetakis, I. E. (2019). High throughput sequencing for plant virus detection and Discovery. *Phytopathology* 109, 716–725. doi: 10.1094/PHYTO-07-18-0257-RVW

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag New York.

Zhou, T., Fan, Z. F., Li, H. F., and Wong, S. M. (2006). Hibiscus chlorotic ringspot virus p27 and Its isoforms affect symptom expression and potentiate virus movement in Kenaf (Hibiscus *cannabinus* L.). *Mol. Plant Microbe In.* 19, 948–957.

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.

frontiers in Plant Science

HYPOTHESIS AND THEORY
published: 05 November 2021
doi: 10.3389/fpls.2021.771983

Poorly Conserved P15 Proteins of Cileviruses Retain Elements of Common Ancestry and Putative Functionality: A Theoretical Assessment on the Evolution of Cilevirus Genomes

Pedro L. Ramos-González[1]†, Tirso Pons[2†], Camila Chabi-Jesus[1,3], Gabriella Dias Arena[1] and Juliana Freitas-Astua[1,4]**

[1] Laboratório de Biologia Molecular Aplicada, Instituto Biológico de São Paulo, São Paulo, Brazil, [2] National Centre for Biotechnology (CNB-CSIC), Madrid, Spain, [3] Escola Superior de Agricultura Luiz de Queiroz (ESALQ), Universidade de São Paulo, Piracicaba, Brazil, [4] Embrapa Mandioca e Fruticultura, Cruz das Almas, Brazil

OPEN ACCESS

Edited by:
Ken Komatsu,
Tokyo University of Agriculture and Technology, Japan

Reviewed by:
Ken-Taro Sekine,
University of the Ryukyus, Japan
Alejandro Olmedo-Velarde,
University of Hawaii, United States

***Correspondence:**
Pedro L. Ramos-González
plrg1970@gmail.com
Juliana Freitas-Astua
juliana.astua@embrapa.br

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Pathogen Interactions, a section of the journal Frontiers in Plant Science

Received: 07 September 2021
Accepted: 18 October 2021
Published: 05 November 2021

Citation:
Ramos-González PL, Pons T, Chabi-Jesus C, Arena GD and Freitas-Astua J (2021) Poorly Conserved P15 Proteins of Cileviruses Retain Elements of Common Ancestry and Putative Functionality: A Theoretical Assessment on the Evolution of Cilevirus Genomes. Front. Plant Sci. 12:771983. doi: 10.3389/fpls.2021.771983

The genus *Cilevirus* groups enveloped single-stranded (+) RNA virus members of the family *Kitaviridae*, order *Martellivirales*. Proteins P15, scarcely conserved polypeptides encoded by cileviruses, have no apparent homologs in public databases. Accordingly, the open reading frames (ORFs) *p15*, located at the 5′-end of the viral RNA2 molecules, are considered orphan genes (ORFans). In this study, we have delved into ORFs *p15* and the relatively poorly understood biochemical properties of the proteins P15 to posit their importance for viruses across the genus and theorize on their origin. We detected that the ORFs *p15* are under purifying selection and that, in some viral strains, the use of synonymous codons is biased, which might be a sign of adaptation to their plant hosts. Despite the high amino acid sequence divergence, proteins P15 show the conserved motif [FY]-L-x(3)-[FL]-H-x-x-[LIV]-S-C-x-C-x(2)-C-x-G-x-C, which occurs exclusively in members of this protein family. Proteins P15 also show a common predicted 3D structure that resembles the helical scaffold of the protein ORF49 encoded by radinoviruses and the phosphoprotein C-terminal domain of mononegavirids. Based on the 3D structural similarities of P15, we suggest elements of common ancestry, conserved functionality, and relevant amino acid residues. We conclude by postulating a plausible evolutionary trajectory of ORFans *p15* and the 5′-end of the RNA2 of cileviruses considering both protein fold superpositions and comparative genomic analyses with the closest kitaviruses, negeviruses, nege/kita-like viruses, and unrelated viruses that share the ecological niches of cileviruses.

Keywords: *Kitaviridae*, orphan ORF, structure-based phylogenetic analysis, synonymous codon usage bias, horizontal gene transfer, *Dichorhavirus*, miniproteins, small ORF

INTRODUCTION

Orphan genes (ORFans) code for proteins with unrecognized homologs in any other species (Fischer and Eisenberg, 1999; Arendsee et al., 2014). They are present in a variety of organisms, including prokaryotes, eukaryotes, and viruses, reaching up to one-third of the genome of some microorganisms and 50% to two-thirds in mimiviruses (Fischer and Eisenberg, 1999; Kaur Saini and Fischer, 2007; Yin and Fischer, 2008; Tautz and Domazet-Lošo, 2011; Colson et al., 2017). The quick origin of ORFans may provide adaptive mechanisms conferring rapid fitness to changing environments (Entwistle et al., 2019). Most of the known proteins encoded by ORFans interact with conserved proteins like transcription factors or receptors, acting as toxins, or as modulators of metabolic or regulatory networks (Singh and Wurtele, 2020).

Kitaviridae, order *Martellivirales,* is a family of heterogeneous plant-infecting viruses displaying two, three, or four segments of single-stranded positive-sense RNA molecules as genomes, which have been assigned into the genera *Cilevirus*, *Higrevirus,* and *Blunervirus*, respectively (Quito-Avila et al., 2021). Taken together, kitavirids have been detected in a relatively narrow range of natural hosts, where they normally produce local infections mainly characterized by chlorotic and/or necrotic lesions that in some cases resemble the outcome of a hypersensitive-like response (Arena et al., 2016, 2020). However, regardless of their failure to consummate the systemic movement through the plants, kitavirids pose serious threats to major crops such as citrus (Ramos-González et al., 2018). Likely most kitaviruses are transmitted by mites, but an effective transmission by *Brevipalpus* mites has been only confirmed for cileviruses (Kitajima and Alberti, 2014; Quito-Avila et al., 2021).

With several strains and isolates belonging to three species already described at the biological and molecular level, *Cilevirus* is the best-studied genus of kitavirids (Freitas-Astúa et al., 2018; Quito-Avila et al., 2021). Virions of citrus leprosis virus C (CiLV-C), citrus leprosis virus C2 (CiLV-C2), and passion fruit green spot virus (PfGSV) are enveloped, short bacilliform particles that encapsidate two poly-adenylated RNA molecules (Locali-Fabris et al., 2006; Roy et al., 2013; Kitajima and Alberti, 2014; Ramos-González et al., 2020). The canonical genome of cileviruses contains six open reading frames (ORF) distributed in two RNA segments. ORFs *RdRp* (RNA dependent-RNA polymerase) and *p29* (putative coat protein) are in RNA 1 (≈9.0 kb); whilst the ORFs *p32* (movement protein), *p61*, *p24,* and *p15* are contained in the RNA2 (≈5.0 kb) (**Figure 1**). RNA2 in certain cileviruses also includes some accessory ORFs, for instance, *p7* in CiLV-C2, and *p11-13* in PfGSV. ORFs *p61* and *p24* encode proteins that are likely involved in the virion structure (Solovyev and Morozov, 2017). P24 is a transmembrane (TM) protein that is also probably an integral component of the virus factory-like membranes, whereas P61 is a putative glycoprotein that in the case of CiLV-C acts as a viral effector in plants (Kuchibhatla et al., 2014; Arena et al., 2016, 2020; Leastro et al., 2018; Vinokurov and Koloniuk, 2019). ORF *p61* is taxonomically restricted to cileviruses, whereas, besides in the kitavirids, *p24* is also present in an increasingly discovered number of arthropod-infecting viruses including mosquito-specific viruses of the proposed group Negevirus (clades Nelorpivirus and Sandewavirus) (Vasilakis et al., 2013; Kuchibhatla et al., 2014; Kondo et al., 2020). P24 displays the highly conserved motif SP24 (TM-SP24, structure protein of 24 kDa, Pfam code PF16504). The glycoproteins P61 of cileviruses lack the motifs DiSA (Pfam code PF19226) and DiSB-ORF2_chro B (Pfam code PF16506) detected in the glycoproteins encoded by their taxonomically related arthropod-infecting viruses (Kuchibhatla et al., 2014). ORF *p15,* only present in cileviruses, encodes the smallest annotated polypeptide across the genomes of these viruses (Ramos-González et al., 2020).

In this study, based on an integrator process comprising genomics, population genetics, usage of synonymous codons, and tridimensional (3D) structures, we elaborated a hypothesis about the conserved putative functionality of the P15 proteins across the genus. Following a chain of logical thinking, we theorize on the origin and evolution of the ORF *p15* considering the genomic context in the 5′-end of the RNA2 in cileviruses. A meaningful part of the study has been founded on the comparisons of the protein folds using the principles of structure-based phylogenetic analyses.

THE OPEN READING FRAMES *p15* OF CILEVIRUSES ARE UNDER PURIFYING SELECTION AND ENCODE POORLY CONSERVED PROTEINS

Citrus leprosis virus C is the best-characterized cilevirus at both molecular and epidemiological levels. Its population is subdivided into three clades (Ramos-González et al., 2016; Chabi-Jesus et al., 2021). Average identity values of genomic nucleotide sequences between CiLV-C isolates belonging to different lineages range from 85 to 89%. Contrastingly, a highly conserved segment of approximately 1.5 kb, showing almost 100% of identity, is present at the 5′-end of their RNA2 molecules and comprises the ORF *p15* (393 nts) (**Figure 1**).

Based on the analysis of 58 nucleotide sequences of *p15* (**Supplementary Table 1**) using MEGA v. 10.1.8 (Kumar et al., 2018) and DnaSP v. 6.12.03 (Rozas et al., 2017) software suites, we confirmed that CiLV-C *p15* variability is very low with an overall genetic distance (*D*) and nucleotide diversity (π) $\leq$ 0.01 (**Figure 2**). According to population analyses, this low variability may result from constant purges of the non-synonymous substitutions by purifying selection ($\omega < 1$), and/or intra-species recombination processes, as already described (Ramos-González et al., 2016; Chabi-Jesus et al., 2021). P15 encoded by CiLV-C shows Cys residues that were earlier suggested as taking part in a putative Zn-finger structure (Ramos-González et al., 2016) (**Figure 3**). Transiently expressed P15 of CiLV-C of the strain CRD enters into the cell nucleus likely by passive diffusion and the formation of homodimers and heterodimers with P29 and the viral movement proteins can be detected in the cytosol of agroinfiltrated *Nicotiana benthamiana* plants (Leastro et al., 2018). When expressed from a viral vector, the infected

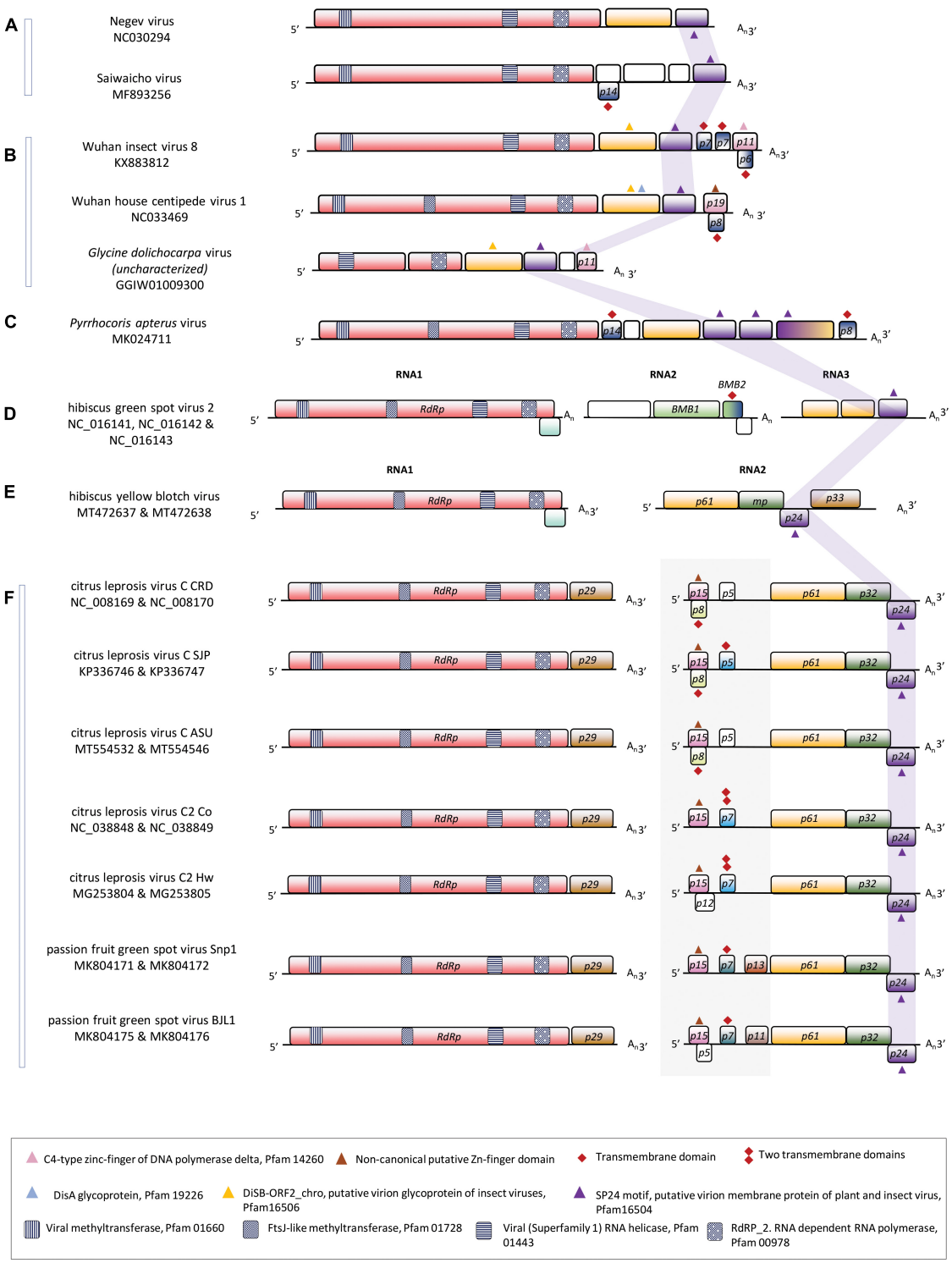


FIGURE 1 | Linear genomic maps of some members of the family *Kitaviridae* and some of their closest phylogenetically arthropod-infecting viruses. **(A)** Two members of the group Negevirus, **(B)** two members of the group centivirus, **(C)** unclassified kita/nege-like virus, **(D)** the sole member of the genus *Higrevirus*, **(E)** a cile-like virus, and **(F)** members of the genus *Cilevirus*. Open reading frames are represented by boxes and the fill colors indicate a conserved functional or structural relationship between them. White boxes indicate unknown features. The solid gray background box in **(F)** highlights the 5′-end of the RNA2 genomic segment of cileviruses. The purple ribbon points out the position of the ORFs containing the SP-24 motif (structural protein of 24 kDa, Pfam code PF16504). Small geometric symbols show the presence of relevant sequences as described in the legend box.

N. benthamiana plants show stunted growth and enhanced necrosis in their younger leaves suggesting pathophysiological disorders caused by its ectopic expression. A possible role as an RNA silencing suppressor was also suggested for P15 of CiLV-C strain CRD (Leastro et al., 2020).

Analyses conducted in this work indicated that despite the low number of isolates available for CiLV-C2 and PfGSV (**Supplementary Table 1**), intra-species variability of their *p15* is almost 10-fold higher than among the CiLV-C isolates. Yet, as well as for CiLV-C, ORFs *p15* from CiLV-C2 and PfGSV show signatures of purifying selection ($\omega < 1$). P15 proteins of these two viruses are phylogenetically closer to each other and more distant from that in CiLV-C (**Figure 2**).

Pairwise comparisons of the P15 amino acid (aa) sequences reveal very low identity values, approximately 14 or 20%, between the proteins from CiLV-C and PfGSV or CiLV-C2, respectively. Higher values, still moderate, approximately 55% aa sequence identity, are detected in the comparisons between P15 from CiLV-C2 and PfGSV (Ramos-González et al., 2016, 2020). Remarkably, the aa pairwise sequence identity values among P15 of CiLV-C and those from CiLV-C2 and PfGSV meaningfully deviate from those commonly observed within the cilevirus proteomes. The identity values between the P15 of PfGSV and CiLV-C are less than half of those shown by the next less conserved proteins among cileviruses, i.e., P29 and P61 (31–32%) (Ramos-González et al., 2020).

OPEN READING FRAME *p15* DISPLAYS SYNONYMOUS CODON USAGE BIAS THAT MIGHT REFLECT ADAPTION TO HOST-SPECIFIC CODON USAGE PATTERNS

Codon usage bias (CUB) refers to differences in the frequency of occurrence of synonymous codons in coding sequences (Behura and Severson, 2013). In viruses, CUB reflects changes in gene expression that occur as a consequence of, for instance, the interplay and co-evolution with their hosts, which may increase the viral fitness (Biswas et al., 2019; He et al., 2019; Khandia et al., 2019). Several indices, such as the Effective Number of codons (ENc), the Relative Synonymous Codon Usage (RSCU), COdon Usage Similarity Index (COUSIN), and Relative Codon Deoptimization Index (RCDI) are used to quantify the CUB (Sharp and Li, 1987; Wright, 1990; Bourret et al., 2019). These indices enable the comparison of genes within a genome and across different genotypes and species revealing elements about their evolution.

Effective Number of codons values range from 20 to 61, with a value of 20 indicating extreme bias, whereas, inversely, 61 indicates non-biased codon usage (Wright, 1990). Meanwhile, an RSCU value higher or lower than 1 denotes that the codon has a positive or negative CUB, respectively, and those with an RSCU value equal to the unit are randomly chosen. In this study, ENc and RSCU values were calculated using the web servers COUSIN[1] and CAIcal[2], respectively. Analyses of both ENc and RSCU of the cilevirus ORFs revealed codon usage bias. Even though the means of the ENc values of ORFs both at virus and genus levels were generally higher than 50, some viral isolates and ORFs, e.g., PfGSV_Snp1 and ORF *p15*, had mean values lower than 50 which are considered an indicator of skewed codon usage. Within this group, the ORF *p15* of CiLV-C2_Hw displayed a distinctly low ENc value, ≈35 (**Figure 4A** and **Supplementary Table 2**), which is generally accepted as a mark of genes with significant codon usage bias (Comeron and Aguadé, 1998; Butt et al., 2016). The values of RSCU were computed for every codon of each ORF by the respective genomes (**Supplementary Table 3**). When the RSCU patterns were hierarchically clustered, most of the ORFs were grouped in a large branch of the generated dendrogram. ORFs *p15*, however, were generally distributed in two small branches shared also with the ORFs *p32* and *p24* of a few isolates (**Figure 4B**). Differently, the ORF *p15* of CiLV-C2_Hw was clustered together with the ORFs *RdRp* and *p61* of that isolate, suggesting that they could have undergone a comparable mutational selection.

To get further insight into the codon preferences of *p15*, we assessed and compared the CUB of each ORF in each cilevirus strain using COUSIN[1] and the ratio RCDI/eRCDI[2] (RCDI/expected RCDI) (Puigbò et al., 2010). It should be noted that while ENc and RSCU values exclusively rely on the coding sequence of the analyzed gene, COUSIN and the ratio RCDI/eRCDI assess the similarity of the codon usage patterns among the coding sequences of at least two genes. In this study, the COUSIN score (Bourret et al., 2019) estimates the codon usage preferences of a viral ORF compared with those of a reference, the viral host, normalized over a null hypothesis of equal usage of synonymous codons. Sequences with scores ≥ 1 display adaptation to the host cellular machinery and higher values might reflect an increasing level of expression. RCDI reflects the similarity of the codon usage between a given coding sequence and a reference genome (Mueller et al., 2006). A lower RCDI value shows the best rate of viral gene translation in the host as well as being indicative of the possible co-evolution of virus and host genomes. The eRCDI value is the RCDI corresponding to a random sequence generated with similar G + C content and amino acid composition to the analyzed sequence and provides a threshold that allows for statistical analysis of the calculated RCDI values. A lower ratio RCDI/eRCDI means a better viral ORF adaptation to its host.

The values of COUSIN and the ratio RCDI/eRCDI were calculated using the set of host plants whose codon usage data could be directly retrieved from, or calculated using available genomic information in, public databases. Codon usage frequencies of some natural and experimental hosts of cileviruses, i.e., citrus (*Citrus* × *sinensis*), passion fruit (*Passiflora* spp.), orchid (*Oncidium* spp.) (Kitajima et al., 2010), *Arabidopsis thaliana* (Arena et al., 2017), and *Phaseolus vulgaris*

[1]http://cousin.ird.fr/index.php

[2]https://ppuigbo.me/programs/CAIcal/

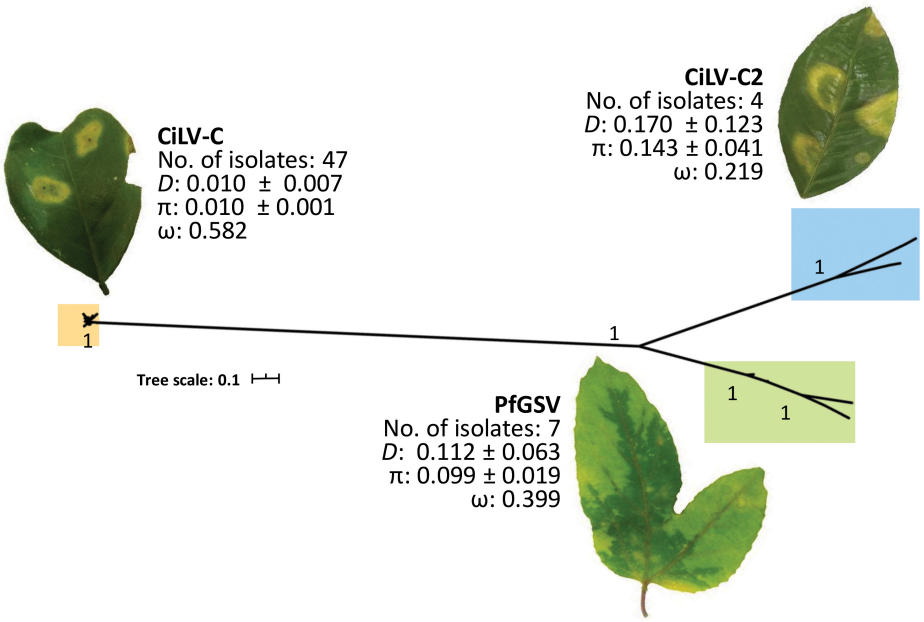


FIGURE 2 | Phylogenetic analysis of P15 proteins and intra-species variability of ORF *p15* from cileviruses. Nucleotide sequence alignments were performed using the MAFFT algorithm (Katoh and Standley, 2013). The best-fit nucleotide substitution model was estimated using MEGA X v. 10. 1. 8 (Kumar et al., 2018). Model HKY + G (Hasegawa et al., 1985), with the lowest Bayesian information criterion (BIC), was used for the estimation of the phylogenetic relationships. The phylogenetic tree was inferred using the Maximum Likelihood method implemented in the Galaxy platform (Afgan et al., 2018). The reliability of the inferred evolutionary relationships was assessed by 1,000 bootstrapped replications. The tree was edited and visualized using Interactive Tree Of Life (iTOL) v4 (Letunic and Bork, 2019). ORF *p15* sequences from 47, 4, and 7 isolates of CiLV-C, CiLV-C2, and PfGSV, respectively, were retrieved from the GenBank database accessed on 17th March 2020 (**Supplementary Table 1**). Intraspecific genetic distance (*D*) and Non-synonymous (dN) and synonymous (dS) rates of substitutions (ω) were calculated in MEGA X, and nucleotide diversity (π) and the confidence intervals for π values were estimated using DnaSP software v6.12.03 (Rozas et al., 2017). Photo of the CiLV-C2-infected orange leaf is courtesy of Dr. Guillermo Adolfo Leon Martinez, AGROSAVIA, 2021.

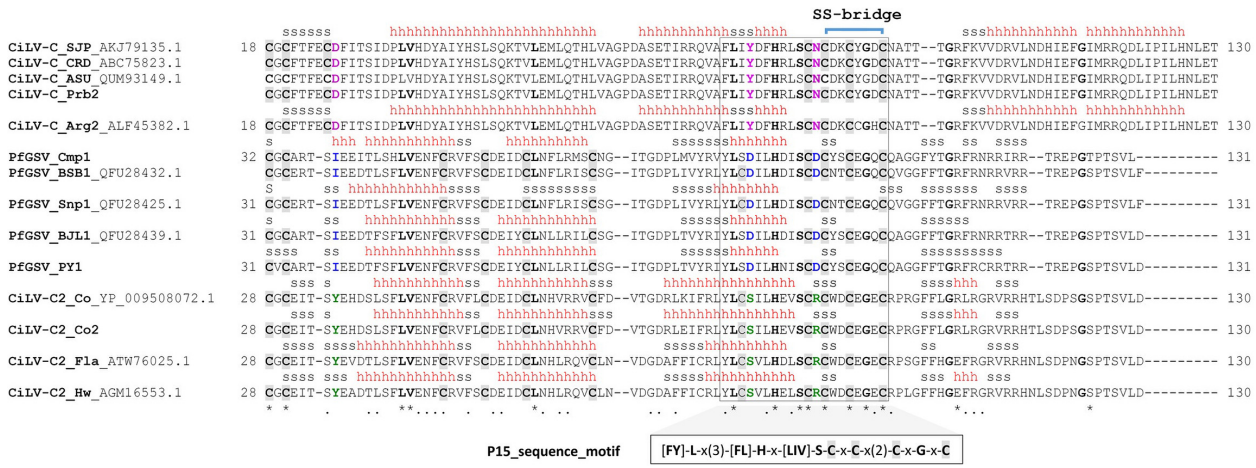


FIGURE 3 | Consensus sequence obtained from a multiple sequence alignment (MSA) of the cysteine-rich domain of the P15 protein family. MSA was built using ClustalO (Sievers and Higgins, 2014). The alignment was manually curated by analyzing gaps, conserved amino acid positions, and the predicted secondary structure. The non-aligning segments at the N-terminus of P15 proteins, presumably representing non-core regions, were removed. Conserved cysteine residues are highlighted by gray background and boldface letters. Secondary structure elements, as predicted by I-TASSER, are depicted on top. Helix and β-strand conformations are indicated by h and s, respectively. A conserved sequential motif {[FY]-L-x(3)-[FL]-H-x-[LIV]-S-C-x-C-x(2)-C-x-G-x-C} is displayed below the sequence alignment. Strictly conserved residues are in boldface and putative functional residues specific from each group are in magenta, blue, and green colors. The putative disulfide bridge between the cysteine residues (positions 85 and 92 in CiLV-C) is represented by a blue bracket. Asterisk and point symbols beneath the alignment indicate strictly conserved residues and conservative substitutions, respectively.

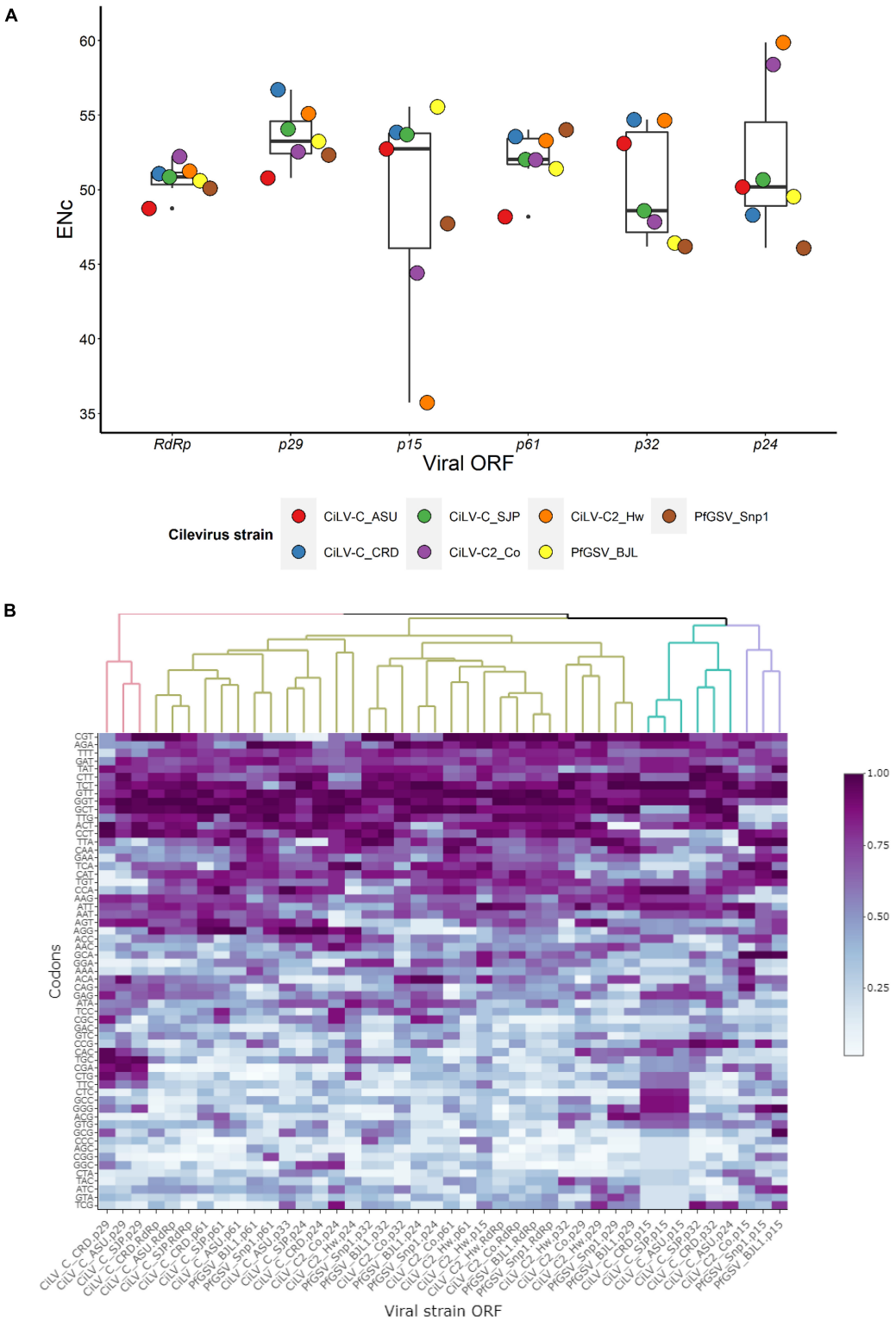


FIGURE 4 | Codon usage bias of cilevirus ORFs. **(A)** Boxplot of Effective Number of codons (ENc). ENc values were calculated using COUSIN webserver (http://cousin.ird.fr/index.php) and plotted in Rstudio (Rstudio_team, 2021). **(B)** RSCU profiles across ORF of seven different isolates of three species of cileviruses. RSCU values were obtained using the CAIcal webserver (https://ppuigbo.me/programs/CAIcal/). The heat map was drawn in Rstudio using the "heatmapply" function (Galili et al., 2018). RSCU values were transformed using the "percentize" method and dist_method = "Euclidean" and hclust_method = "complete" were selected. Darker purple represents higher RSCU values.

(Garita et al., 2013) were retrieved from Codon Usage Database[3] or calculated using countcodon v4[4] (**Supplementary Table 4**). Coding sequence information of *Passiflora edulis,* approximately 300 ORFs, was retrieved from the CoGe platform[5]. Orchids were included as cilevirus natural hosts since the infection by PfGSV of plants from two orchids species collected in Brazil was molecularly confirmed (unpublished results). For the sake of the comparisons, the analyses were performed considering both the sum of natural and experimental host plants and two subsets of plants, each comprising natural or experimental hosts. Altogether, the analysis included plants belonging to five families, including perennial woody (citrus) and perennial herbaceous (orchids) plants, short-lived evergreen perennial vine (passion fruit), and herbaceous annual (common beans and Arabidopsis) plants. Values of COUSIN and the ratio RCDI/eRCDI were visualized and compared using hierarchical clustering. The accuracy of the assessment could be potentially increased as new annotated genomes of cilevirus host plants are available.

Regardless of the viral species, the analysis by ORFs using the complete set of host plants revealed a >1 bias in the mean COUSIN values of *p61* (1.28 ± 0.44), *RdRp* (1.19 ± 0.36), *p32* (1.15 ± 0.37), and *p24* (1.09 ± 0.38), whereas mean values below or near to 1 were observed for the ORFs *p15* (0.89 ± 0.53) and *p29* (0.93 ± 0.35) (**Supplementary Table 5A**). COUSIN values of *p15*, however, showed the highest variability with a coefficient of variation surpassing 50%. Some CiLV-C2 ORFs reached COUSIN values higher than 1.5, whereas the minimum values, e.g., −0.15, corresponded to PfGSV isolates. The same trend of COUSIN values was obtained when calculations were performed with the two subgroups of plants conveniently separated into natural and experimental hosts (**Supplementary Tables 5B,C**). In a global analysis across the genus, the COUSIN profile of *p15* resembled that showed by *p32*, which encodes the viral movement protein (**Figure 5A**).

The study of the ratio RCDI/eRCDI of the cilevirus ORFs indicated that *RdRp* and *p61* bear the lowest mean values, followed in ascending order by *p24*, *p29*, and *p32*, and finally *p15* (**Supplementary Table 6A**). In the particular comparisons of *p15*, values from CiLV-C2 were collectively the lowest ones reached. In the evaluations using *P. vulgaris*, the ratios RCDI/eRCDI of *p15* genes from CiLV-C2 isolates were below the threshold of 0.9, similar to those held by other viral ORFs, e.g., *p29*, *p24*, and *p32*. The analyses suggest that the CUB of all those ORFs might be the outcome of equivalent processes of strain/host adaptation.

Regardless of the subgroup of studied plants, i.e., the subsets of natural, experimental, or the array containing all host plants, the mean values of the ratio RCDI/eRCDI of *p15* were higher than those shown by the remaining viral ORFs (**Supplementary Tables 6B,C**). At the genus level, when analyzed through hierarchical clustering, the RCDI/eRCDI profile of *p15* is unique among those shown by any other ORF of cileviruses (**Figure 5B**).

P15 PROTEINS DISPLAY A UNIQUE CONSERVED SEQUENCE MOTIF

Even though a low global sequence identity of less than 25% was observed for P15 proteins, the consensus sequence obtained from a multiple sequence alignment of the cysteine-rich domain reveals a stretch of highly conserved residues organized in the sequence motif {[FY]-L-x(3)-[FL]-H-x-[LIV]-S-C-x-C-x(2)-C-x-G-x-C} (**Figure 3** and **Supplementary Table 7**). Besides four full-conserved residues of cysteines and one of histidine, the consensus sequence motif also contains three other invariable amino acids: leucine, serine, and glycine. A ScanProsite[6] (de Castro et al., 2006) search using the conserved sequence motif in the UniProtKB/Swiss-Prot[7] (release 07-Apr-2021: 564,638 entries) and UniProtKB/TrEMBL[8] (release 07-Apr-2021: 214,406,399 entries) databases retrieved 30 hits spanned in 30 sequences, all of them exclusively corresponding to the P15 proteins from CiLV-C, CiLV-C2, and PfGSV. Additionally, no hits were retrieved from a similar search but against a randomized UniProtKB/Swiss-Prot database, i.e., reversed option. According to ScanProsite, the conserved sequence motif has an approximate number of expected random matches of 6.97×10^{-08} in 100,000 sequences (50,000,000 residues). Altogether, these results suggested a low probability of P15 matches by chance and, thus, highlighted the singular occurrence of the identified cysteine-rich domain in the P15 protein family.

THREE-DIMENSIONAL STRUCTURE PREDICTION SUGGESTS A HELICAL BUNDLE-LIKE SCAFFOLD FOR THE P15 PROTEINS

The existence of a conserved sequence motif among the P15 proteins dropped a hint that all of them might exhibit a similar fold. Interestingly, this motif overlaps a predicted α-turn-α motif in the secondary structure of P15 in CiLV-C and CiLV-C2, whereas in PfGSV the second α-helix is missing (**Figure 3**). The predictions by evaluation of sequence-structure fitness (Ouzounis et al., 1993) with proteins of known 3D structures of a representative group of ten P15 proteins from the three species of cileviruses were carried out by the iterative threading assembly refinement (I-TASSER) server[9] (Yang and Zhang, 2015). For the computational structural analyses, experimentally determined 3D structures were downloaded from the RCSB PDB database[10] and the protein 3D structure alignments were done using the

[3]http://www.kazusa.or.jp/codon/

[4]http://www.kazusa.or.jp/codon/countcodon.html

[5]https://genomevolution.org/coge/

[6]https://prosite.expasy.org/scanprosite/

[7]https://www.uniprot.org/statistics/Swiss-Prot

[8]https://www.uniprot.org/statistics/TrEMBL

[9]https://zhanglab.ccmb.med.umich.edu/I-TASSER/

[10]https://www.rcsb.org

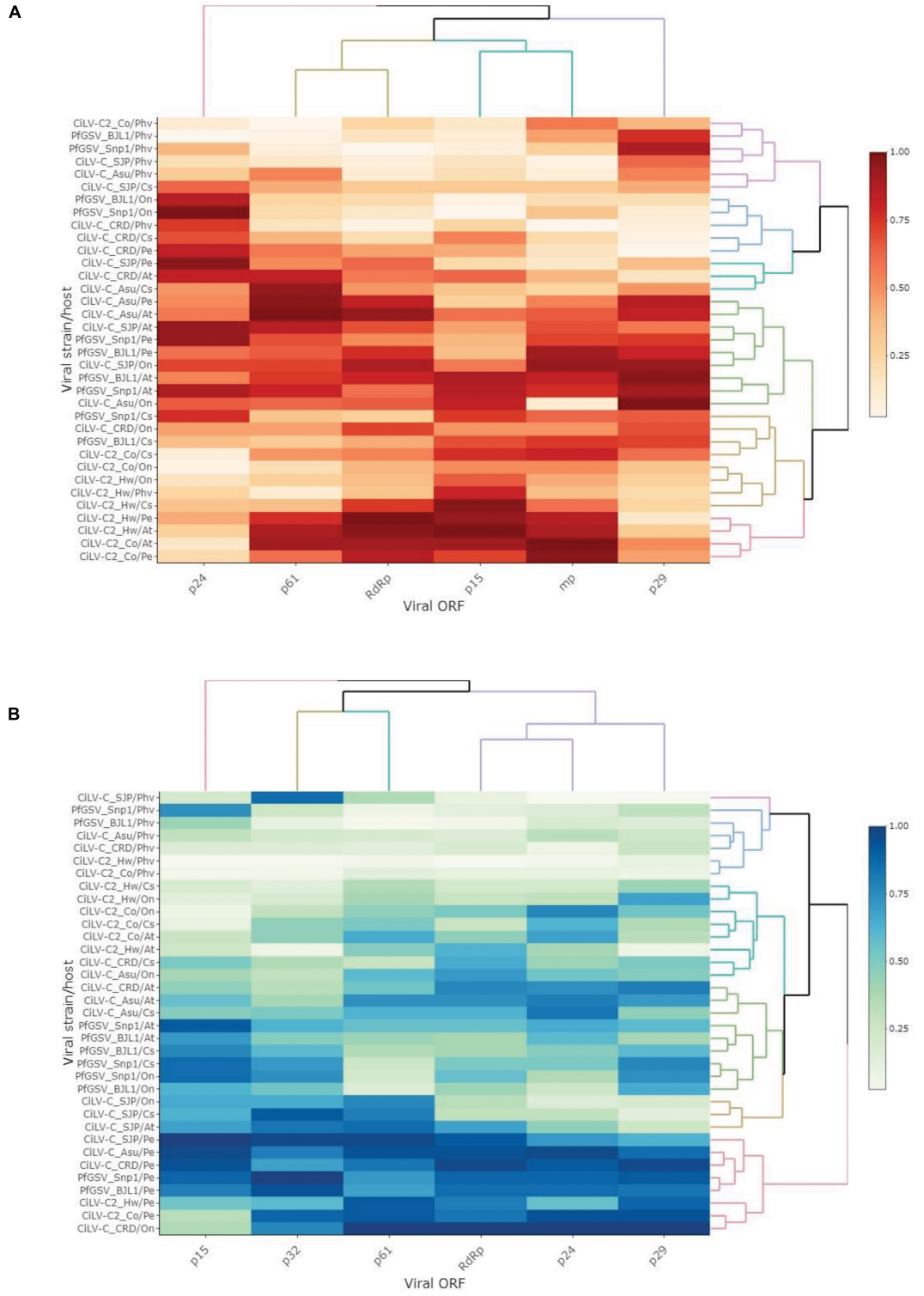


FIGURE 5 | Codon adaption of cilevirus ORFs. Heatmaps representing the values of COUSIN **(A)** and the RCDI/eRCDI **(B)**, respectively, of every ORF across cilevirus isolates. For the sake of comparative analyses, COUSIN and the RCDI/eRCDI values have been obtained simulating the interaction of all viruses with the same hypothetical host range. Codon usage database of host plants and values of ENc, COUSIN, and RCDI are shown in **Supplementary Tables 3**, **4**. The heat map was drawn in Rstudio (Rstudio_team, 2021) using the "heatmapply" function (Galili et al., 2018). Both COUSIN and RCDI/eRCDI values were transformed using the "percentize" method and dist_method = "Euclidean" and hclust_method = "complete" were selected. Darker red and light green to white represent the best values in each heat map, respectively. At, *Arabidopsis thaliana*; Cs, *Citrus* × *sinensis*; Pe, *Passiflora edulis*; On, *Oncidium* sp.; and Phv, *Phaseolus vulgaris*.

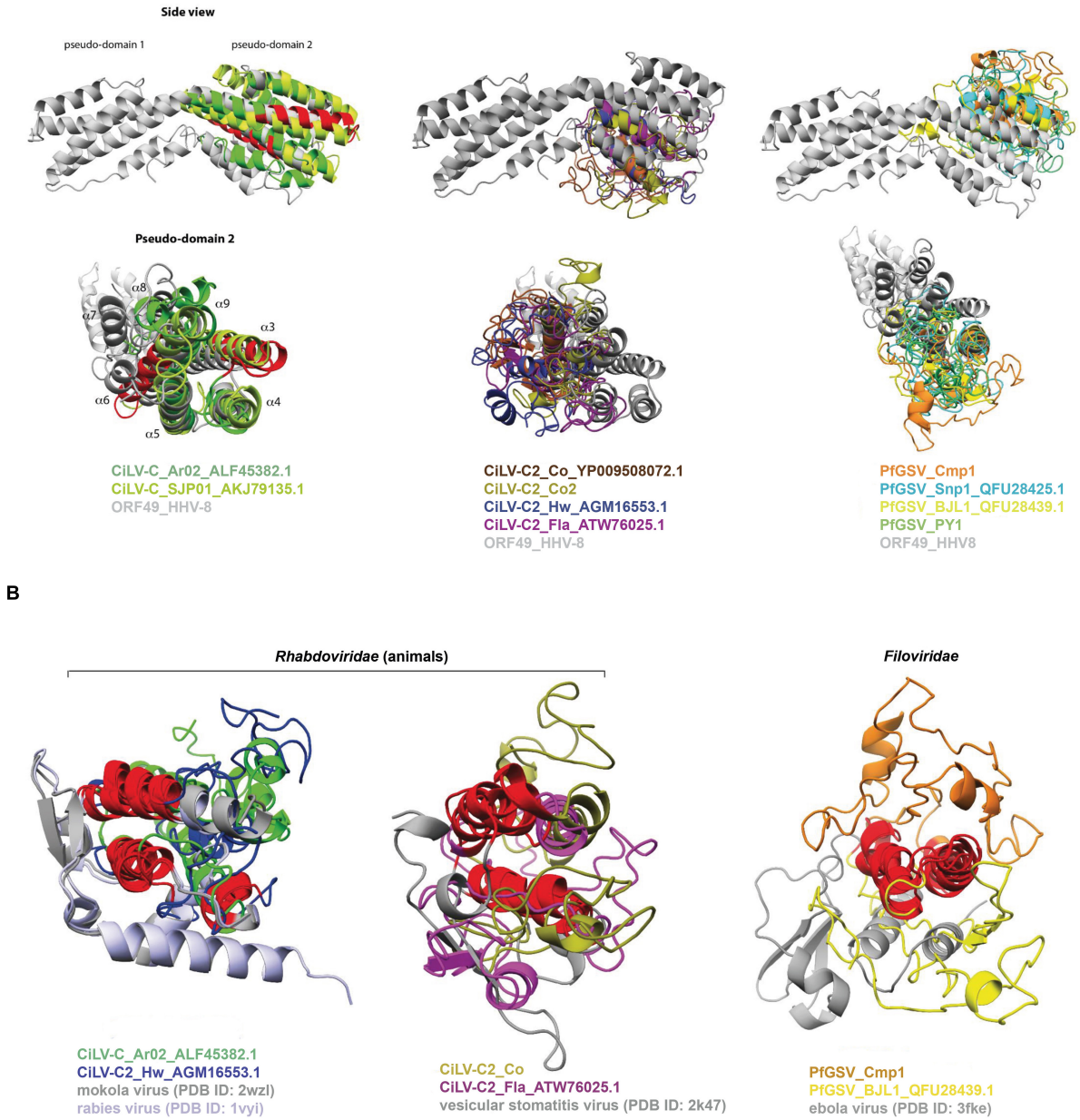


FIGURE 6 | 3D structural similarities between P15 and **(A)** the protein ORF49 from the human gammaherpesvirus 8 (ORF49_HHV-8, PDB ID: 5ipx) and **(B)** the rhabdoviruses rabies virus (PDB ID: 1vyi) and vesicular stomatitis Indiana virus (PDB ID: 2k47), and the filovirus Zaire ebolavirus (PDB ID: 3fke). Multiple structural alignments were done with the PDBeFold method. The conserved sequential motif {[FY]-L-x(3)-[FL]-H-x [LIV]-S-C-x-C-x(2)-C-x-G-x-C} is displayed in red. Structure figures were prepared using PyMOL (Delano, 2002).

PDBeFold server[11] (Krissinel and Henrick, 2004). The quality of protein structure templates and predicted full-length models were assessed using the *C*-score and TM-score implemented in I-TASSER[9] (Zhang and Skolnick, 2004). *C*-score is typically in the range [−5,2], where a *C*-score of higher value signifies a model with high confidence, and vice-versa. TM-score values range from [0,1], where 1 indicates a perfect match between two structures. A TM-score below 0.17 corresponds to randomly chosen unrelated proteins whereas structures with a score higher than 0.5 assume generally the same fold.

The best generated P15 3D model corresponded to that for CiLV-C_Ar02 (*C*-score −2.82, TM-score 0.39 ± 0.13, **Supplementary Table 8**). The model from this protein suggested the presence of a four-helical bundle-like scaffold with structural

[11]https://www.ebi.ac.uk/msd-srv/ssm/

similarity to the 12 α-helices bundled of protein ORF49 from the human gammaherpesvirus 8 (ORF49_HHV-8, PDB ID: 5ipx). The multiple structural alignments of 3D models of P15 proteins with the crystal structure of ORF49_ HHV-8 by using PDBeFOLD[11] revealed the "plasticity" of the helix-bundle protein folds (**Figure 6A**).

PDBeFold provides average root mean square deviation (RMSD) and *Q*-score values to estimate the quality of matching structures. RMSD distance between corresponding residues is calculated after an optimal rotation of one structure to another. An RMSD of less than about 3–4 Å would generally be considered very close, but there is no absolute rule. The *Q*-score assesses the number of residues in the matched secondary structural elements and their positions in space. *Q*-scores reach the value of 1 for identical structures and decline to zero with the lack of similarity, thus, the higher *Q*, the better.

Among the P15 from CiLV-C isolates, an overall RMSD = 2.47 Å and *Q*-score = 0.31 over 94 aligned residues were obtained, whereas in those from CiLV-C2 the values were RMSD = 3.83 Å and *Q*-score = 0.01 over 22 aligned residues. In P15 from the PfGSV isolates, the values were an overall RMSD = 4.1 Å and *Q*-score = 0.02 over 30 aligned residues. Detailed analysis of the 3D model for the CiLV-C_Ar02 P15 indicated that cysteine residues at positions 85 (Cys85) and 92 (Cys92) could form a disulfide bridge and consequently stabilize the four-helical bundle. Remarkably, this two-cysteine-residues pattern is also present in the conserved sequence motif across the P15 proteins. However, since the spatial localization of the two cysteines is not close enough in some 3D models, the formation of a disulfide bridge is not completely evident. Such discrepancies could ensue from the analysis of low-resolution 3D models.

The resolved crystal structures of ORF49_ HHV-8 (PDB ID: 5ipx) and ORF49 from the murine herpesvirus 68 (ORF49_MHV68, PDB ID: 6a4v) (Hew et al., 2017; Chung et al., 2018) show a fold consisting of 12 α-helices bundled into two pseudo-domains (left top and bottom panels in **Figure 6A**). By mapping the identical and similar amino acids onto the helical scaffold, the potentially functional residues in P15 from CiLV-C were predicted. Conserved leucine and valine residues were localized in hydrophobic regions along the interacting surfaces of the α-helices and have the potential to play a structural role. Polar and charged residues were localized on the surface and could be associated with protein fold stability or protein interactions (**Table 1**).

α-helical folds have also been detected in the C-terminal domain of the phosphoprotein (P_{CTD}) of mononegaviruses (Assenberg et al., 2010; Ivanov et al., 2010; Ng et al., 2020). It is of note that cileviruses and mononegaviruses of the genus *Dichorhavirus*, family *Rhabdoviridae* (Dietzgen et al., 2018), share several commonalities, including hosts and vectors, that have led to postulate the likely existence of evolutionary convergence among viruses of these two genera (Freitas-Astúa et al., 2018). The peculiar architecture observed in P_{CTD} and P15 prompted us to compare the 3D structure of these proteins.

TABLE 1 | Summary of the predicted functional residues for the P15 family of proteins.

Amino acid[a]	ORF49[b]	3D[c]	Description
Cys 18	His 14	α1	Polar, exposed, protein interaction
Cys 20	Phe 16	α1	Polar, exposed, protein interaction
Asp 26	Ser 22	Loop	Charged, exposed, protein interaction
Leu 34	Leu 30	α2	Hydrophobic, buried, helix packing
Val 35	Asp 31	α2	Hydrophobic, buried, helix packing
Leu 50	Ala 46	α3	Hydrophobic, buried, helix packing
Leu 74	Pro 70	α3	Hydrophobic, buried, helix packing
Tyr 76	Leu 72	α3	Polar, exposed, protein interaction
His 79	Ala 75	α3	Charged, exposed, protein interaction
Ser 82	Ala 78	α3	Polar, exposed, protein interaction
Cys 83	Asn 79	α3	Polar, exposed, protein interaction
Asn 84	Leu 80	α3	Polar, exposed, protein interaction
Cys 85	Ala 81	α3	Polar, exposed, disulfide bridge, protein fold stability
Cys 88	Leu 84	α3	Polar, exposed, protein interaction
Gly 90	Gln 86	α3	Polar, exposed, protein fold stability
Cys 92	Tyr 88	α3	Polar, exposed, disulfide bridge, protein fold stability
Gly 98	Lys 94	α4	Polar, exposed, protein fold stability
Gly 114	Phe 113	α5	Polar, exposed, protein fold stability

[a]*Amino acid position in the CiLV-C_Ar02_ALF45382.1 sequence.*
[b]*Equivalent position in the ORF49 sequence according to the structure superposition between ORF49_HHV-8 (PDB ID: 5ipx) and CiLV-C_Ar02_ALF45382.1 (I-TASSER 3D-model).*
[c]*Nomenclature of alpha-helices is based on the 3D structure of ORF49_ HHV-8. The structure superposition between ORF49_HHV-8 and CiLV-C 3D-models by PDBeFold is provided in* ***Supplementary File 1****.*

The P_{CTD} of the cytorhabdovirus lettuce necrotic yellows virus (LNYV) comprises five α-helices that have an overall topology that although with different structural features, seems to be conserved with those in other rhabdoviruses, filoviruses, and paramyxoviruses (Martinez et al., 2013). P_{CTD} crystal structure from plant-infecting rhabdoviruses other than LNYV has not been described. Hence, the predicted 3D models of P15 were aligned with the P_{CDT} of mononegaviruses, regardless of their range hosts. P15 of CiLV-C_Ar02 and CiLV-C2_Hw are structurally related to P_{CTD} of mokola virus (PDB ID: 2wzl) (Delmas et al., 2010) and rabies virus (PDB ID: 1vyi) (Mavrakis et al., 2004) with overall RMSD = 3.1 Å and *Q*-score = 0.06 over 42 aligned residues, whereas those of CiLV-C2_Co2 and CiLV-C2_Fla are structurally related to vesicular stomatitis Indiana virus (PDB ID: 2k47) with overall RMSD = 4.0 Å and *Q*-score = 0.04 over 32 aligned residues (**Figure 6B**). In the case of PfGSV_Cmp1 and PfGSV_BJL1, their modeled P15 are structurally related to Zaire ebolavirus (PDB ID: 3fke) with an overall RMSD = 3.1 Å and *Q*-score = 0.04 over 36 aligned residues. Based on structural similarities, it could be hypothesized that P15 proteins can perform functions related to those accomplished by ORF49 and P_{CTD} proteins. ORF49 interacts and upregulates the transcriptional activity

of a protein known as RTA (replication and transcription activator) in HHV-8, which drives the expression of all lytic genes (González et al., 2006). Present in a large number of viruses, P protein orthologs are involved in a wide range of functions, e.g., acting as chaperons, regulation of viral transcription, RNA binding, and suppressor of RNA silencing in plants (Masters and Banerjee, 1988; Peluso and Moyer, 1988; Martinez et al., 2013; Bejerman et al., 2016; Mann et al., 2016; Ng et al., 2020).

Zn-FINGER-LIKE PROTEINS IN ARTHROPOD-INFECTING VIRUSES CLOSELY RELATED TO KITAVIRUSES ARE NOT RELATED WITH THE CILEVIRUS P15 PROTEINS

Based on the predicted isoelectric point and the presence of a cysteine-rich pattern compatible with a potential Zn-finger domain, an acidic cilevirus P15-like protein was previously recognized in the genome of Wuhan insect virus 8 (WhIV-8) (Vinokurov and Koloniuk, 2019). The protein is encoded by the putative ORF *p11* (**Figure 1**) that was neither identified during the virus genome description (Shi et al., 2016) nor annotated at its corresponding GenBank locus KX883812.

Our search for supposed *p15* homologs based on the presence of a zinc-finger-like motif was expanded to nege/kita-like viruses and other related arthropod-infecting viruses. It also included previously ignored or poorly characterized putative ORFs. In members of the proposed subgroup Centivirus of the nege/kita-like viruses, a homolog of WhIV-8 *p11* was identified in the uncharacterized *Glycine dolichocarpa* virus (GDV, Transcriptomic shotgun assembly TSA GGIW01009300) (Kondo et al., 2020). Proteins encoded by the ORF *p11* in GDV and WhIV-8 have 41.4% amino acid sequence identity. The region comprising the cysteine-rich pattern in these two proteins shows a marginal identity (*E*-value < 1 according to MOTIF Search[12]) with the Pfam motif C4-type zinc-finger of DNA polymerase delta (PF14260). Similarly, the ORF p*19* in the putative centivirus Wuhan house centipede virus 1 (WHCV-1) and ORF *p9* of Hubei virga-like virus 9 (HVLV-9) encode proteins with signatures of a putative zinc-finger domain (Vinokurov and Koloniuk, 2019), which in the case of HVLV-9 also displays a subtle identity [*E*-value < 1 according to MOTIF Search[11]] with the Pfam motif SLBP (stem-loop binding protein) of the RNA binding superfamily (PF15247). The N-terminal of the ORF3-encoded protein of *Pyrrhocoris apterus* virus slightly resembles the canonical Pfam DSRM (double-stranded RNA binding motif, PF00035) architecture, whereas, although with low sequence similarity, the C-terminal end of this protein displays a conservative basic RNA-interacting motif. However, neither nucleotide sequences of these ORFs nor their predicted polypeptides show significant sequence identity to any element across cileviruses.

[12]https://www.genome.jp/tools/motif/

To get further insight into the putative relationships between P15 and P11 proteins of WhIV-8 and GDV, their 3D structures were predicted using the I-TASSER server[9]. Although P11 proteins show α helix scaffolds, the arrangement of helixes is different from that observed in P15. Multiple structural alignments of 3D models by PDBeFold[11] indicated that only one α-helix in P11 proteins shows structural similarity to α-helices in the P15 3D models (overall RMSD = 2.68 Å and *Q*-score = 0.01 over 17 aligned residues). The lack of overall sequence or 3D structure similarity between both groups of proteins drastically reduces the probability of a common origin.

THE 5′-END OF THE RNA2 IN CILEVIRUSES, BESIDES THE ORF *p15*, HARBORS ORFs ENCODING SMALL HYDROPHOBIC PROTEINS

Downstream the ORF *p15*, the genomic stretch extended up to *p61* in the 5′-end of the RNA2 of cileviruses is known as the intergenic region (IR). However, this is not a proper description because rather than lacking ORFs, the IR lacks ORFs greater than 300 nucleotides. *In silico* tools have traditionally considered a minimum of 100 amino acids for an ORF to be annotated as a putative coding protein sequence, which has led to the inaccurate annotation of certain genomes. However, despite the detection of those small ORFs, the IRs of cileviruses show a very low density of predicted coding sequences, only 20–30%, with large non-coding stretches (**Figure 1**).

In congruence with ORF *p15*, the interspecies comparisons of the IRs of cileviruses show low levels of nucleotide sequence identity (≤45%) (Ramos-González et al., 2020). With variable length, the IR is larger in CiLV-C isolates (1,095 ± 56 nucleotides) and shorter in PfGSV isolates (850 ± 21 nucleotides). The region harbors small ORFs encoding hydrophobic proteins with predicted TM domains, but which are unrelated to the SP24 motif (Melzer et al., 2013; Roy et al., 2013; Ramos-González et al., 2016, 2020). ORF *p13* in PfGSV_Snp1, the largest detected so far (333 nucleotides), presents four regularly distributed cysteine residues and low sequence identity with motifs identified in the cytokine-induced anti-apoptosis inhibitor 1 (CIAPIN1, PF05093), and PGC7/Stella/Dppa3-like domain (PGC7_Stella, PF15549) (Ramos-González et al., 2020).

Further inspection in the 5′-end of the RNA2 in cileviruses, allowed us to detect the ORF *p8* overlapping the ORF *p15* of CiLV-C. ORF *p8* is present in every sequenced virus of the three lineages of CiLV-C (Chabi-Jesus et al., 2021), but it is absent in CiLV-C2 and PfGSV. Previously overlooked, the ORF *p8* spans 54% of *p15* and has 213 nucleotides in the frame +2 (**Figure 1**). ORF *p8* is significantly longer than any other ORF expected in all reading frames within the ORF *p15*, which is a fair indicator of its real existence according to an *in silico* method for detecting overlapping genes (Schlub et al., 2018) performed in this study (**Supplementary Table 9**). The presence of the overlapping ORF *p8* in CiLV-C may limit the variability of *p15* and suggests a

different evolutionary trajectory in the formation of this genomic region among cileviruses.

The predicted P8 protein seems to be hydrophobic with a TM domain according to MEMSAT-SVM[13] and TMHMM Server v. 2.0[14]. This TM domain resembles the predicted properties of the small P7 proteins encoded downstream the ORF *p15*, in the IR of CiLV-C2 (Roy et al., 2013), CiLV-C_SJP, and PfGSV (**Figure 1**). Small hydrophobic proteins by themselves are outside the scope of this work but what is particularly pertinent is their constant presence in the 5′-end of the RNA2 of cileviruses, and the association with the small Zn-finger-like or nucleic acid binding-like proteins in some nege/kita-like viruses. The overlapping gene array observed in the pair P15/P8 in CiLV-C was also observed in the pairs P11/P6 and P19/P8 of the centiviruses WhIV-8 and WHCV-1, respectively (**Figure 1**). Overlapping genes *p6* and *p8* of centiviruses were detected following the same procedure described for the identification of *p8* in CiLV-C (**Supplementary Table 9**). Small hydrophobic proteins have been described in several taxonomic groups of plant RNA viruses where they can be involved in coupling viral replication and cell-to-cell movement showing properties similar to those of reticulons (Solovyev and Morozov, 2017; Morozov and Solovyev, 2020; Lazareva et al., 2021).

DISCUSSION

This study was driven by the search for commonalities between the ORFs *p15*, their encoded proteins, and the questioning of whether these polypeptides can be considered orthologs. Finally, these analyses led us to a critical appraisal of the 5′-end of the RNA2 of cileviruses under an evolutionary perspective.

Absent in kitavirids beyond the genus *Cilevirus* and with no apparent homology with any sequence available in public databases, ORF *p15* was considered an orphan gene (Arena et al., 2016). P15 proteins from different species show a remarkably reduced amino-acid sequence identity between them, with values as low as those shown by the global alignment of unrelated proteins (≈16%) (Konaté et al., 2019; Ramos-González et al., 2020). As well as P13 from PfGSV, it has been speculated that P15 proteins are viral auxiliary factors that act as regulators of host physiology during the plant-virus interplay (Ramos-González et al., 2020). Particularly, the putative role of P15 from CiLV-C of the strain CRD as a viral suppressor of the plant gene silencing antiviral defense was analyzed (Leastro et al., 2020), but further studies are needed to reach sound conclusions.

Together with the hallmark ORF *RdRp* (Koonin et al., 2006), the genomes of the typical cileviruses comprise the ORFs *p24, p32,* and the taxonomically restricted *p61*. Their RNA2 molecules also accommodate at their 5′-ends the *p15* and the IR, which in addition to *p11* and *p13* in PfGSV strains, harbor some putative small ORFs encoding predicted TM small polypeptides. Interestingly, hibiscus yellow blotch virus (HYBV) (Olmedo-Velarde et al., 2021), a specimen considered a cilevirus with an unusual genomic organization and phylogenetically intermediate among typical cileviruses and hibiscus green spot virus 2 (HGSV-2), genus *Higrevirus* (Melzer et al., 2012), poses an additional puzzle on the composition of the RNA2 segments of typical cileviruses and the role of the ORFs in their 5′-ends, e.g., *p15*. While the RNA1 in all these viruses encodes two proteins, the RdRp and P29 or P10 in typical cileviruses and HYBV, respectively, the RNA2 of HYBV resembles a truncated version of the RNA2 of typical cileviruses, lacking the 5′- end stretch upstream the homologs of the threesome *p61-p32-p24* (**Figure 1**). RNA2 of HYBV also includes an additional ORF, homolog to the cilevirus RNA1 *p29*, at its 3′-end. The very conserved synteny and composition among the genomes of HYBV and typical cileviruses suggest, under a reductionist approach, the rather dispensable role of the 5′-end of the RNA2 of cileviruses. Perhaps, this region only provides proteins and/or unknown regulatory non-coding RNAs that albeit non-essential could expand the viral fitness (Frías-Lasserre, 2012; Miller et al., 2016). Alternatively, we could speculate, for instance, that the putative functions encoded by this genomic region in typical cileviruses are taken over by other ORFs in HYBV, likely involving a compensatory evolutionary process (Rokyta et al., 2002; Harcombe et al., 2009). However, while the lack of function of the 5′ end of RNA2 in typical cileviruses does not appear to be a rational choice, in the absence of compelling evidence suggesting explicit roles, the analysis of *p15* across the genus, intrinsically, may shed some light on the 5′-end of cilevirus RNA2.

In this study, we described that *p15* in each of the three cilevirus species is under purifying selection, suggesting an ongoing action to conserve its nucleotide sequence, most likely precluding the change of amino acid residues and preserving function at directly selected sites (Cvijović et al., 2018). Furthermore, we detected that in some viral strains the synonymous codon usage of this ORF is biased showing the same codon patterns adopted by other viral genes and displaying COUSIN and RCDI values that suggest a certain level of adaptation to its plant hosts. We also revealed that although with extremely divergent amino acid sequences, P15 proteins show a conserved sequence motif, and their 3D models have preserved a backbone topology that resembles a helical bundle-like scaffold. Altogether, these two aspects suggest elements of common ancestry and conserved functionality of the P15 proteins (Arendsee et al., 2019; Ng et al., 2020; Ravantti et al., 2020).

The superimposition of the 3D models of P15 and the resolved crystal structures of ORF49_ HHV-8 and ORF49_MHV68 (Hew et al., 2017; Chung et al., 2018) and the P_{CTD} of negative-stranded RNA viruses of the order *Mononegavirales* (Martinez et al., 2013) indicated the presence of an α-helical bundle-like scaffold in the cilevirus protein. Bundles of four α-helices are frequently found in proteins with a wide variety of functions (Rhys et al., 2019). ORF49 shows a fold consisting of 12 helices bundled into two pseudo-domains, with charged patches on its surface likely acting in protein-protein interaction sites. The protein is stabilized by double-stranded oligonucleotides suggesting its interaction and

[13]http://bioinf.cs.ucl.ac.uk/psipred/

[14]http://www.cbs.dtu.dk/services/TMHMM/

binding with DNA molecules (Hew et al., 2017). The P_{CTD} of negative-stranded RNA viruses of the order *Mononegavirales* (Martinez et al., 2013) has been shown to regulate viral transcription, RNA binding, and immunomodulatory functions (Ng et al., 2020). The P protein of cytorhabdoviruses has weak local RNA silencing suppressor activity, strongly suppresses systemic RNA silencing, and the P_{CTD} is essential for both local RNA silencing suppression and the interaction with AGO1, AGO2, AGO4, RDR6, and SGS3, proteins of the main core of the gene silencing mechanism (Bejerman et al., 2016; Mann et al., 2016). Therefore, the presence of an α-helical bundle-like scaffold in P15 could provide a structural base by mean it could interact with other proteins and likely mediate the plant-virus interplay. P15 from CiLV-C of the strain CRD forms homo- and heterodimers with other viral proteins (Leastro et al., 2018), and considering the diversity of its interactions, it appears that the protein could act modulating a wide spectrum of viral and cellular pathways, as observed in other viral proteins (Deblasio et al., 2018; Dao et al., 2020; Li et al., 2020).

The Origin and Evolution of the 5′-End of Cileviruses: A Theory

Phylogenetic inferences suggest a common ancestor between kitaviruses and a variety of arthropod-infecting viruses including negeviruses (Vasilakis et al., 2013; Kondo et al., 2019; Ramos-González et al., 2020; Quito-Avila et al., 2021). Having regards to the presence of the orthologs *RdRp* and *p24* in these viruses (Kuchibhatla et al., 2014), a synteny-based analysis indicates a parsimonious evolutionary scenario in which cileviruses, and kitaviruses in general, derived from a monopartite ss(+) RNA through genome segmentation. In brief, the process could have started amid assortments of defective molecules resulting from error-prone replication and recombination events (Bangham and Kirkwood, 1993; García-Arriaza et al., 2004; Vignuzzi and López, 2019) and sub-genomic RNA transcripts derived from the ancestor. Those molecules with self-replication activity could have recruited segments supplying viral complementary functions. *Cis*- and *trans*-acting sequence motifs likely evolved providing these novel genomic segments with the capacity to coordinate essential processes for the persistence of split viral genomes, e.g., multiplication and packaging (Newburn and White, 2015). Further modifications of kitavirus genomes likely included the acquisition of auxiliary genes, as seems to be a general pathway in the formation of plant virome (Dolja et al., 2020). The clearest examples of horizontal gene transfer in the genome of current kitaviruses are epitomized by the ORF *p32* in cileviruses, and ORFs *BMB1* and *2*, which encode the binary movement block proteins in HGSV-2 (Quito-Avila et al., 2013; Solovyev and Morozov, 2017; Hao et al., 2018).

Multiplication of unnecessary viral sequences appears to have fitness costs and the size of virus genomes shows a natural trend to shrinkage (Belshaw et al., 2007; Zwart et al., 2014). The occurrence of IRs as long as observed in the RNA2 of cileviruses is atypical. Whether large IRs existed in the ancestors of kita/nege-like viruses, they were most likely purged by selection. The genomic organization of known negeviruses and kita/nege-like viruses are rather compact, with short stretches of intergenic sequences. Therefore, a relevant interrogation in this context is whether the IRs of cileviruses deviate from the selection rules. In a sense, most likely not, and some features might suggest it. The IR of PfGSV isolates are shorter and contain more predicted coding regions than those in CiLV-C isolates, whereas the presence of the overlapping ORF *p8* in CiLV-C may also indicate genomic compression.

As defined in this study, the 5′-end of the RNA2 genomic segments in typical cileviruses comprises the ORF *p15* and the IR. There is no obvious evidence allowing for a direct connection between the 5′-end of the cilevirus RNA2 molecules and any genetic element in the genome of their closely related negeviruses and nege/kita-like viruses. Yet, the relationship between the 5′-end of the RNA2 molecules from cileviruses would be difficult to establish, except for the presence of a conserved amino acid motif across the P15 family and the 3D structure similarities among these proteins. These two features, moreover, may be pertinent clues suggesting the common ancestry of the ORF *p15* present in the cileviruses CiLV-C, CiLV-C2, and PfGSV, giving support to a theory about their common origin.

The origin of the 5′-end of the RNA2 in typical cileviruses could have involved the following plausible scenarios: (*i*) a reminiscent ancient ORF from the monopartite ancestor or (*ii*) its acquisition from a heterologous source by recombination, i.e., horizontal gene transfer. Although we lack evidence to exclude any of these possibilities, our approach using prediction and computational analysis of 3D protein structures has provided information on a possible external origin of P15 proteins. The scaffold of α-helixes detected in P15 shows structural resemblance with those detected in well-characterized proteins from herpesvirids and mononegavirids. Importantly, the odds of the 3D structure of P15 is the result of convergence are very low considering the small number of protein folds available (Malik et al., 2017).

The external origin of the ORF *p15* could be considered part of the horizontal gene transfer events that likely happened in the course of kitavirus evolution, but what is peculiar in this case is the participation of a mononegavirid, probably an ancestor of the current plant rhabdoviruses, as its putative donor. Cileviruses and plant-infecting rhabdoviruses have a partially shared plant host range, and, in the case of dichorhaviruses, they also share mite vectors (Dietzgen et al., 2018, 2020; Freitas-Astúa et al., 2018). Likely, *Brevipalpus* mites, where dichorhavirus also multiplicate, have underpinned a primordial and essential role in the segmentation of both cileviruses and dichorhaviruses (Kondo et al., 2017).

However, whatever the scenario for the origin of *p15* in cileviruses, the very conservation of a sequence motif and fold in the encoded proteins indicate that the "ancestor" ORF was present before the radiation of, at least, CiLV-C, CiLV-C2, and PfGSV. Since then, it has been likely undergone a transformation process under selective pressure, acquiring or losing protein secondary structure elements and involving host-specific patterns to provide evolutionarily competitiveness to these viruses under new ecological conditions, e.g., the change of natural host range from arthropods to plants. Indeed, if this would be the case,

p15 likely became a fitness factor whose fixation might have indirectly contributed to the conservation of the IR. This large chunk of genomic sequence with apparently an unclear role is preserved between functional ORFs, that away of inactivity, it may be a zone of gain and loss of start and stop codons gradually incorporating novel ORFs.

Two mechanisms explaining the origin of orphan genes are currently accepted: (*i*) divergence from the coding sequence of a preexisting gene (Tautz and Domazet-Lošo, 2011; Vakirlis et al., 2020), and (*ii*) *de novo* appearance from intergenic or non-genic regions (Carvunis et al., 2012; Van Oss and Carvunis, 2019). The irregular organization and lack of identity among the small ORFs other than *p15* found in the 5′-end of the RNA 2 of cileviruses support the idea that this genomic region might act as "the cradle of new genes" (Forterre and Gaïa, 2016). Several of the predicted ORFs found in the IR are small, of ≈200 nucleotides, and they encode putative proteins where TM domains are detected by *in silico* analyses. Viral miniproteins with TM domains are encoded by a wide range of DNA and RNA viruses and they can bind and modulate cellular transmembrane proteins, providing flexible mechanisms to regulate several cellular activities (Dimaio, 2014; Finkel et al., 2018). Moreover, small ORFs can act as seeds for larger coding sequences (CDS) through a mechanism known as "CDS elongation" when their stop codons undergo mutations (Couso and Patraquim, 2017). Thus, small ORFs could provide an extra coding capacity that in the course of evolution could eventually become an additional source of fitness (Guerra-Almeida and Nunes-da-Fonseca, 2020).

CONCLUDING REMARKS

In this study, we have addressed one of the most intriguing features of the molecular biology of cileviruses: the 5′ end of the RNA2 molecule and its two most salient elements, the ORF *p15*, and the IR. We presented pieces of evidence on the putative biochemical activity of the protein P15 over the genus and elucubrate on its putative contribution to viral fitness. The results are supported by the analyses of 58 nucleotide sequences of isolates of the three known species whose P15 amino acid sequences cover a spectrum of identity ranging from 14 to 100%. The main predictions have been obtained by integrating the results of different computational methods because of the premise that each of these elements taken in isolation does not allow to draw solid conclusions about the activity and function of a given protein (Omer et al., 2017; Ponting, 2017; Ravantti et al., 2020). Taken them together, we posit a theory that partially describes the evolution of the cilevirus genomes that involves elements denoting the likely close interplay during the coevolutionary history of *Brevipalpus*-transmitted viruses. In light of the plausible scenario described by the theory, two aspects of the biology of cileviruses gain new distinctions: (*i*) since *p15* across cileviruses can be considered alleles descendant from a common ancestor likely co-opted from another virus, this ORF should not be longer considered an orphan gene, and (*ii*) the markedly chimeric origin of the RNA2 genome segment, which adds further support to what appears to be a common feature of members of the family *Kitaviridae* (Morozov et al., 2020).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

PR-G and TP: conceptualization. PR-G, TP, CC-J, and GA: formal analysis, investigation, and methodology. JF-A: funding acquisition. TP and JF-A: resources. PR-G and TP: writing—original draft. PR-G, TP, CC-J, GA, and JF-A: writing—review and editing. All authors approved the submitted version.

ACKNOWLEDGMENTS

The authors are grateful to CAPES, CNPq, and FAPESP for scholarships and research grants associated with this work (CAPES PNPD20132154 – 33141010001P4 – PNPD – IBSP, 88882.157041/2017-01; CNPq 312528/2020-5; FAPESP 2017/50222-0, 2019/25078-9, 2019/02137-0, and 2020/15413-2). The authors also thank the reviewers and editor for their useful and constructive comments on our work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.771983/full#supplementary-material

REFERENCES

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46, W537–W544. doi: 10.1093/nar/gky379

Arena, G. D., Ramos-González, P. L., Falk, B. W., Casteel, C. L., Freitas-Astúa, J., and Machado, M. A. (2020). Plant Immune System Activation Upon Citrus Leprosis Virus C Infection Is Mimicked by the Ectopic Expression of the P61 Viral Protein. *Front. Plant Sci.* 11:1188. doi: 10.3389/fpls.2020.01188

Arena, G. D., Ramos-González, P. L., Nunes, M. A., Alves, M. R., Camargo, L. E. A., Kitajima, E. W., et al. (2016). Citrus leprosis virus C infection results in hypersensitive-like response, suppression of the JA/ET plant defense pathway and promotion of the colonization of its mite vector. *Front. Plant Sci.* 7:1757. doi: 10.3389/FPLS.2016.01757

Arena, G. D., Ramos-González, P. L., Nunes, M. A., Jesus, C. C., Calegario, R. F., Kitajima, E. W., et al. (2017). *Arabidopsis thaliana* as a model host for *Brevipalpus* mite-transmitted viruses. *Sci. Agric.* 74, 85–89. doi: 10.1590/1678-992x-2015-0380

Arendsee, Z. W., Li, L., and Wurtele, E. S. (2014). Coming of age: Orphan genes in plants. *Trends Plant Sci.* 19, 698–708. doi: 10.1016/j.tplants.2014.07.003

Arendsee, Z., Li, J., Singh, U., Bhandary, P., Seetharam, A., and Wurtele, E. S. (2019). Fagin: Synteny-based phylostratigraphy and finer classification of young genes. *BMC Bioinformatics* 20:440. doi: 10.1186/s12859-019-3023-y

Assenberg, R., Delmas, O., Ren, J., Vidalain, P.-O., Verma, A., Larrous, F., et al. (2010). Structure of the Nucleoprotein Binding Domain of Mokola Virus Phosphoprotein. *J. Virol.* 84, 1089–1096. doi: 10.1128/jvi.01520-09

Bangham, C. R. M., and Kirkwood, T. B. L. (1993). Defective interfering particles and virus evolution. *Trends Microbiol.* 1, 260–264. doi: 10.1016/0966-842X(93)90048-V

Behura, S. K., and Severson, D. W. (2013). Codon usage bias: causative factors, quantification methods and genome-wide patterns: with emphasis on insect genomes. *Biol. Rev.* 88, 49–61. doi: 10.1111/J.1469-185X.2012.00242.X

Bejerman, N., Mann, K. S., and Dietzgen, R. G. (2016). Alfalfa dwarf cytorhabdovirus P protein is a local and systemic RNA silencing supressor which inhibits programmed RISC activity and prevents transitive amplification of RNA silencing. *Virus Res.* 224, 19–28. doi: 10.1016/j.virusres.2016.08.008

Belshaw, R., Pybus, O. G., and Rambaut, A. (2007). The evolution of genome compression and genomic novelty in RNA viruses. *Genome Res.* 17, 1496–1504. doi: 10.1101/gr.6305707

Biswas, K. K., Palchoudhury, S., Chakraborty, P., Bhattacharyya, U. K., Ghosh, D. K., Debnath, P., et al. (2019). Codon usage bias analysis of citrus tristeza virus: Higher codon adaptation to citrus reticulata host. *Viruses* 11:v11040331. doi: 10.3390/v11040331

Bourret, J., Alizon, S., and Bravo, I. G. (2019). COUSIN (COdon Usage Similarity INdex): A Normalized Measure of Codon Usage Preferences. *Genome Biol. Evol.* 11, 3523–3528. doi: 10.1093/gbe/evz262

Butt, A. M., Nasrullah, I., Qamar, R., and Tong, Y. (2016). Evolution of codon usage in Zika virus genomes is host and vector specific. *Emerg. Microbes Infect.* 5, 1–14. doi: 10.1038/emi.2016.106

Carvunis, A. R., Rolland, T., Wapinski, I., Calderwood, M. A., Yildirim, M. A., Simonis, N., et al. (2012). Proto-genes and de novo gene birth. *Nature* 487, 370–374. doi: 10.1038/nature11184

Chabi-Jesus, C., Ramos-González, P. L., Postclam-Barro, M., Fontenele, R. S., Harakava, R., Bassanezi, R. B., et al. (2021). Molecular Epidemiology of Citrus Leprosis Virus C: A New Viral Lineage and Phylodynamic of the Main Viral Subpopulations in the Americas. *Front. Microbiol.* 12:986. doi: 10.3389/fmicb.2021.641252

Chung, W. C., Kim, J., Kim, B. C., Kang, H. R., Son, J., Ki, H., et al. (2018). Structure-based mechanism of action of a viral poly(ADP-ribose) polymerase 1-interacting protein facilitating virus replication. *IUCrJ* 5, 866–879. doi: 10.1107/S2052252518013854

Colson, P., La Scola, B., Levasseur, A., Caetano-Anollés, G., and Raoult, D. (2017). Mimivirus: leading the way in the discovery of giant viruses of amoebae. *Nat. Rev. Microbiol.* 15, 243–254. doi: 10.1038/nrmicro.2016.197

Comeron, J. M., and Aguadé, M. (1998). An evaluation of measures of synonymous codon usage bias. *J. Mol. Evol.* 47, 268–274. doi: 10.1007/PL00006384

Couso, J. P., and Patraquim, P. (2017). Classification and function of small open reading frames. *Nat. Rev. Mol. Cell Biol.* 18, 575–589. doi: 10.1038/nrm.2017.58

Cvijović, I., Good, B. H., and Desai, M. M. (2018). The effect of strong purifying selection on genetic diversity. *Genetics* 209, 1235–1278. doi: 10.1534/genetics.118.301058

Dao, T. N. M., Kang, S.-H., Bak, A., and Folimonova, S. Y. (2020). A Non-Conserved p33 Protein of Citrus Tristeza Virus Interacts with Multiple Viral Partners. *Mol. Plant Microbe Interact.* 33, 859–870. doi: 10.1094/MPMI-11-19-0328-FI

de Castro, E., Sigrist, C., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P., Gasteiger, E., et al. (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.* 34, 362–365. doi: 10.1093/nar/gkl124

Deblasio, S. L., Xu, Y., Johnson, R. S., Rebelo, A. R., Maccoss, M. J., Gray, S. M., et al. (2018). The interaction dynamics of two potato leafroll virus movement proteins affects their localization to the outer membranes of mitochondria and plastids. *Viruses* 10:v10110585. doi: 10.3390/v10110585

Delano, W. (2002). PyMOL: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr.* 40, 82–92.

Delmas, O., Assenberg, R., Grimes, J. M., and Bourhy, H. (2010). The structure of the nucleoprotein binding domain of lyssavirus phosphoprotein reveals a structural relationship between the N-RNA binding domains of Rhabdoviridae and Paramyxoviridae. *RNA Biol.* 7, 322–327. doi: 10.4161/rna.7.3.11931

Dietzgen, R. G., Bejerman, N. E., Goodin, M. M., Higgins, C. M., Huot, O. B., Kondo, H., et al. (2020). Diversity and epidemiology of plant rhabdoviruses. *Virus Res.* 281:197942. doi: 10.1016/j.virusres.2020.197942

Dietzgen, R. G., Freitas-Astúa, J., Chabi-Jesus, C., Ramos-González, P. L., Goodin, M. M., Kondo, H., et al. (2018). "Dichorhaviruses in their Host Plants and Mite Vectors," in *Advances in Virus Research*, (Cambridge, MA: Academic Press Inc), 119–148. doi: 10.1016/bs.aivir.2018.06.001

Dimaio, D. (2014). Viral miniproteins. *Annu. Rev. Microbiol.* 68, 21–43. doi: 10.1146/annurev-micro-091313-103727

Dolja, V. V., Krupovic, M., and Koonin, E. V. (2020). Deep Roots and Splendid Boughs of the Global Plant Virome. *Annu. Rev. Phytopathol.* 58, 23–53. doi: 10.1146/ANNUREV-PHYTO-030320-041346

Entwistle, S., Li, X., and Yin, Y. (2019). Orphan Genes Shared by Pathogenic Genomes Are More Associated with Bacterial Pathogenicity. *mSystems* 4:18. doi: 10.1128/msystems.00290-18

Finkel, Y., Stern-Ginossar, N., and Schwartz, M. (2018). Viral Short ORFs and Their Possible Functions. *Proteomics* 18:201700255. doi: 10.1002/pmic.201700255

Fischer, D., and Eisenberg, D. (1999). Finding families for genomic ORFans. *Bioinformatics* 15, 759–762.

Forterre, P., and Gaïa, M. (2016). Giant viruses and the origin of modern eukaryotes. *Curr. Opin. Microbiol.* 31, 44–49. doi: 10.1016/j.mib.2016.02.001

Freitas-Astúa, J., Ramos-González, P. L., Arena, G. D., Tassi, A. D., and Kitajima, E. W. (2018). Brevipalpus-transmitted viruses: parallelism beyond a common vector or convergent evolution of distantly related pathogens? *Curr. Opin. Virol.* 33, 66–73. doi: 10.1016/J.COVIRO.2018.07.010

Frías-Lasserre, D. (2012). Non coding RNAs and viruses in the framework of the phylogeny of the genes, epigenesis and heredity. *Int. J. Mol. Sci.* 13, 477–490. doi: 10.3390/ijms13010477

Galili, T., O'Callaghan, A., Sidi, J., and Sievert, C. (2018). Heatmaply: An R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics* 34, 1600–1602. doi: 10.1093/bioinformatics/btx657

García-Arriaza, J., Manrubia, S. C., Toja, M., Domingo, E., and Escarmís, C. (2004). Evolutionary Transition toward Defective RNAs That Are Infectious by Complementation. *J. Virol.* 78, 11678–11685. doi: 10.1128/jvi.78.21.11678-11685.2004

Garita, L., Tassi, A. D., Calegario, R. F., Kitajima, E. W., Carbonell, S. A. M., and Freitas-Astúa, J. (2013). Common bean (Phaseolus vulgaris L.): Experimental local lesion host for citrus leprosis virus C (CiLV-C) and some other cytoplasmic-type Brevipalpus-transmitted viruses. *Plant Dis.* 97, 1346–1351. doi: 10.1094/PDIS-12-12-1143-RE

González, C. M., Wong, E. L., Bowser, B. S., Hong, G. K., Kenney, S., and Damania, B. (2006). Identification and Characterization of the Orf49 Protein of Kaposi's Sarcoma-Associated Herpesvirus. *J. Virol.* 80, 3062–3070. doi: 10.1128/jvi.80.6.3062-3070.2006

Guerra-Almeida, D., and Nunes-da-Fonseca, R. (2020). Small Open Reading Frames: How Important Are They for Molecular Evolution? *Front. Genet.* 11:574737. doi: 10.3389/fgene.2020.574737

Hao, X., Zhang, W., Zhao, F., Liu, Y., Qian, W., Wang, Y., et al. (2018). Discovery of Plant Viruses From Tea Plant (Camellia sinensis (L.) O. Kuntze) by Metagenomic Sequencing. *Front. Microbiol.* 9:2175. doi: 10.3389/fmicb.2018.02175

Harcombe, W. R., Springman, R., and Bull, J. J. (2009). Compensatory evolution for a gene deletion is not limited to its immediate functional network. *BMC Evol. Biol.* 9:106. doi: 10.1186/1471-2148-9-106

Hasegawa, M., Kishino, H., and Yano aki, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22, 160–174. doi: 10.1007/BF02101694

He, Z., Gan, H., and Liang, X. (2019). Analysis of synonymous codon usage bias in potato virus M and its adaption to hosts. *Viruses* 11:v11080752. doi: 10.3390/v11080752

Hew, K., Veerappan, S., Sim, D., Cornvik, T., Nordlund, P., and Dahlroth, S.-L. (2017). Structure of the Open Reading Frame 49 Protein Encoded by Kaposi's Sarcoma-Associated Herpesvirus. *J. Virol.* 91:16. doi: 10.1128/jvi.01947-16

Ivanov, I., Crépin, T., Jamin, M., and Ruigrok, R. W. H. (2010). Structure of the Dimerization Domain of the Rabies Virus Phosphoprotein. *J. Virol.* 84, 3707–3710. doi: 10.1128/jvi.02557-09

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kaur Saini, H., and Fischer, D. (2007). Structural and functional insights into Mimivirus ORFans. *BMC Genomics* 8:115. doi: 10.1186/1471-2164-8-115

Khandia, R., Singhal, S., Kumar, U., Ansari, A., Tiwari, R., Dhama, K., et al. (2019). Analysis of nipah virus codon usage and adaptation to hosts. *Front. Microbiol.* 10:886. doi: 10.3389/fmicb.2019.00886

Kitajima, E. W., and Alberti, G. (2014). Anatomy and fine structure of *Bevipalpus* mites (Tenuipalpidae) – Economically important plant virus vectors. Part 7. Ultrastructural detection of cytoplasmic and nuclear types of *Brevipalpus* transmitted viruses. *Zoologica* 160, 174–192.

Kitajima, E. W., Rodrigues, J. C. V., and Freitas-Astua, J. (2010). An annotated list of ornamentals naturally found infected by *Brevipalpus* mite-transmitted viruses. *Sci. Agric.* 67, 348–371. doi: 10.1590/S0103-90162010000300014

Konaté, M. M., Plata, G., Park, J., Usmanova, D. R., Wang, H., and Vitkup, D. (2019). Molecular function limits divergent protein evolution on planetary timescales. *Elife* 8:39705. doi: 10.7554/eLife.39705

Kondo, H., Chiba, S., Maruyama, K., Andika, I. B., and Suzuki, N. (2019). A novel insect-infecting virga/nege-like virus group and its pervasive endogenization into insect genomes. *Virus Res.* 262, 37–47. doi: 10.1016/j.virusres.2017.11.020

Kondo, H., Fujita, M., Hisano, H., Hyodo, K., Andika, I. B., and Suzuki, N. (2020). Virome Analysis of Aphid Populations That Infest the Barley Field: The Discovery of Two Novel Groups of Nege/Kita-Like Viruses and Other Novel RNA Viruses. *Front. Microbiol.* 11:509. doi: 10.3389/fmicb.2020.00509

Kondo, H., Hirota, K., Maruyama, K., Andika, I. B., and Suzuki, N. (2017). A possible occurrence of genome reassortment among bipartite rhabdoviruses. *Virology* 508, 18–25. doi: 10.1016/j.virol.2017.04.027

Koonin, E. V., Senkevich, T. G., and Dolja, V. V. (2006). The ancient virus world and evolution of cells. *Biol. Direct* 1:29. doi: 10.1186/1745-6150-1-29

Krissinel, E., and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 60, 2256–2268. doi: 10.1107/S0907444904026460

Kuchibhatla, D. B., Sherman, W. A., Chung, B. Y. W., Cook, S., Schneider, G., Eisenhaber, B., et al. (2014). Powerful sequence similarity search methods and in-depth manual analyses can identify remote homologs in many apparently "orphan" viral proteins. *J. Virol.* 88, 10–20. doi: 10.1128/JVI.02595-13

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Lazareva, E. A., Lezzhov, A. A., Chergintsev, D. A., Golyshev, S. A., Dolja, V. V., Morozov, S. Y., et al. (2021). Reticulon-like properties of a plant virus-encoded movement protein. *New Phytol.* 229, 1052–1066. doi: 10.1111/nph.16905

Leastro, M. O., Castro, D. Y. O., Freitas-Astúa, J., Kitajima, E. W., Pallás, V., and Sánchez-Navarro, J. Á (2020). Citrus Leprosis Virus C Encodes Three Proteins With Gene Silencing Suppression Activity. *Front. Microbiol.* 11:1231. doi: 10.3389/fmicb.2020.01231

Leastro, M. O., Kitajima, E. W., Silva, M. S., Resende, R. O., and Freitas-Astúa, J. (2018). Dissecting the Subcellular Localization, Intracellular Trafficking, Interactions, Membrane Association, and Topology of Citrus Leprosis Virus C Proteins. *Front. Plant Sci.* 9:1299. doi: 10.3389/fpls.2018.01299

Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239

Li, H., Li, F., Zhang, M., Gong, P., and Zhou, X. (2020). Dynamic Subcellular Localization, Accumulation, and Interactions of Proteins From Tomato Yellow Leaf Curl China Virus and Its Associated Betasatellite. *Front. Plant Sci.* 11:840. doi: 10.3389/fpls.2020.00840

Locali-Fabris, E. C., Freitas-Astúa, J., Souza, A. A., Takita, M. A., Astúa-Monge, G., Antonioli-Luizon, R., et al. (2006). Complete nucleotide sequence, genomic organization and phylogenetic analysis of citrus leprosis virus cytoplasmic type. *J. Gen. Virol.* 87, 2721–2729. doi: 10.1099/vir.0.82038-0

Malik, S. S., Azem-e-Zahra, S., Kim, K. M., Caetano-Anollés, G., and Nasir, A. (2017). Do viruses exchange genes across superkingdoms of life? *Front. Microbiol.* 8:02110. doi: 10.3389/fmicb.2017.02110

Mann, K. S., Johnson, K. N., Carroll, B. J., and Dietzgen, R. G. (2016). Cytorhabdovirus P protein suppresses RISC-mediated cleavage and RNA silencing amplification in planta. *Virology* 490, 27–40. doi: 10.1016/j.virol.2016.01.003

Martinez, N., Ribeiro, E. A., Leyrat, C., Tarbouriech, N., Ruigrok, R. W. H., and Jamin, M. (2013). Structure of the C-Terminal Domain of Lettuce Necrotic Yellows Virus Phosphoprotein. *J. Virol.* 87, 9569–9578. doi: 10.1128/jvi.00999-13

Masters, P. S., and Banerjee, A. K. (1988). Resolution of multiple complexes of phosphoprotein NS with nucleocapsid protein N of vesicular stomatitis virus. *J. Virol.* 62, 2651–2657. doi: 10.1128/JVI.62.8.2651-2657.1988

Mavrakis, M., McCarthy, A. A., Roche, S., Blondel, D., and Ruigrok, R. W. H. (2004). Structure and function of the C-terminal domain of the polymerase cofactor of rabies virus. *J. Mol. Biol.* 343, 819–831. doi: 10.1016/j.jmb.2004.08.071

Melzer, M. J., Sether, D. M., Borth, W. B., and Hu, J. S. (2012). Characterization of a virus infecting Citrus volkameriana with citrus leprosis-like symptoms. *Phytopathology* 102, 122–127. doi: 10.1094/PHYTO-01-11-0013

Melzer, M. J., Simbajon, N., Carillo, J., Borth, W. B., Freitas-Astúa, J., Kitajima, E. W., et al. (2013). A cilevirus infects ornamental hibiscus in Hawaii. *Arch. Virol.* 2013, 1745–1740. doi: 10.1007/s00705-013-1745-0

Miller, W. A., Shen, R., Staplin, W., and Kanodia, P. (2016). Noncoding RNAs of plant viruses and viroids: Sponges of host translation and RNA interference machinery. *Mol. Plant Microbe Interact.* 29, 156–164. doi: 10.1094/MPMI-10-15-0226-FI

Morozov, S. Y., and Solovyev, A. G. (2020). Small hydrophobic viral proteins involved in intercellular movement of diverse plant virus genomes. *AIMS Microbiol.* 6, 305–329. doi: 10.3934/microbiol.2020019

Morozov, S. Y., Lazareva, E. A., and Solovyev, A. G. (2020). Sequence Relationships of RNA Helicases and Other Proteins Encoded by Blunervirus RNAs Highlight Recombinant Evolutionary Origin of Kitaviral Genomes. *Front. Microbiol.* 11:561092. doi: 10.3389/fmicb.2020.561092

Mueller, S., Papamichail, D., Coleman, J. R., Skiena, S., and Wimmer, E. (2006). Reduction of the Rate of Poliovirus Protein Synthesis through Large-Scale Codon Deoptimization Causes Attenuation of Viral Virulence by Lowering Specific Infectivity. *J. Virol.* 80, 9687–9696. doi: 10.1128/jvi.00738-06

Newburn, L. R., and White, K. A. (2015). Cis-acting RNA elements in positive-strand RNA plant virus genomes. *Virology* 479–480, 434–443. doi: 10.1016/j.virol.2015.02.032

Ng, W. M., Stelfox, A. J., and Bowden, T. A. (2020). Unraveling virus relationships by structure-based phylogenetic classification. *Virus Evol.* 6:veaa003. doi: 10.1093/ve/veaa003

Olmedo-Velarde, A., Hu, J., and Melzer, M. J. (2021). A Virus Infecting Hibiscus rosa-sinensis Represents an Evolutionary Link Between Cileviruses and Higreviruses. *Front. Microbiol.* 12:988. doi: 10.3389/fmicb.2021.660237

Omer, S., Harlow, T. J., and Gogarten, J. P. (2017). Does Sequence Conservation Provide Evidence for Biological Function? *Trends Microbiol.* 25, 11–18. doi: 10.1016/j.tim.2016.09.010

Ouzounis, C., Sander, C., Scharf, M., and Schneider, R. (1993). Prediction of protein structure by evaluation of sequence-structure fitness: Aligning sequences to contact profiles derived from three-dimensional structures. *J. Mol. Biol.* 232, 805–825. doi: 10.1006/jmbi.1993.1433

Peluso, R., and Moyer, S. (1988). Viral proteins required for the *in vitro* replication of vesicular stomatitis virus defective interfering particle genome RNA. *Virology* 162, 369–376. doi: 10.1016/0042-6822(88)90477-1

Ponting, C. P. (2017). Biological function in the twilight zone of sequence conservation. *BMC Biol.* 15:71. doi: 10.1186/s12915-017-0411-5

Puigbò, P., Aragonès, L., and Garcia-Vallvé, S. (2010). RCDI/eRCDI: A web-server to estimate codon usage deoptimization. *BMC Res. Notes* 3:87. doi: 10.1186/1756-0500-3-87

Quito-Avila, D. F., Brannen, P. M., Cline, W. O., Harmon, P. F., and Martin, R. R. (2013). Genetic characterization of Blueberry necrotic ring blotch virus, a novel RNA virus with unique genetic features. *J. Gen. Virol.* 94, 1426–1434. doi: 10.1099/vir.0.050393-0

Quito-Avila, D. F., Freitas-Astúa, J., and Melzer, M. J. (2021). Bluner-, Cile-, and Higreviruses (Kitaviridae). *Encycl. Virol.* 2021, 247–251. doi: 10.1016/B978-0-12-809633-8.21248-X

Ramos-González, P. L., Chabi-Jesus, C., Arena, G. D., Tassi, A. D., Kitajima, E. W., and Freitas-Astúa, J. (2018). Citrus leprosis: a unique multietiologic disease. *Citrus Am.* 1, 4–19. doi: 10.6084/m9.figshare.13513857.v2

Ramos-González, P. L., Chabi-Jesus, C., Guerra-Peraza, O., Breton, M. C., Arena, G. D., Nunes, M. A., et al. (2016). Phylogenetic and molecular variability studies reveal a new genetic clade of Citrus leprosis virus C. *Viruses* 8:153. doi: 10.3390/v8060153

Ramos-González, P. L., Santos, G. F., dos, Chabi-Jesus, C., Harakava, R., Kitajima, E. W., et al. (2020). Passion Fruit Green Spot Virus Genome Harbors a New Orphan ORF and Highlights the Flexibility of the 5′-End of the RNA2 Segment Across Cileviruses. *Front. Microbiol.* 11:00206. doi: 10.3389/fmicb.2020.00206

Ravantti, J. J., Martinez-Castillo, A., and Abrescia, N. G. A. (2020). Superimposition of viral protein structures: A means to decipher the phylogenies of viruses. *Viruses* 12:v12101146. doi: 10.3390/v12101146

Rhys, G. G., Wood, C. W., Beesley, J. L., Zaccai, N. R., Burton, A. J., Brady, R. L., et al. (2019). Navigating the Structural Landscape of de Novo α-Helical Bundles. *J. Am. Chem. Soc.* 141, 8787–8797. doi: 10.1021/jacs.8b13354

Rokyta, D., Badgett, M. R., Molineux, I. J., and Bull, J. J. (2002). Experimental Genomic Evolution: Extensive Compensation for Loss of DNA Ligase Activity in a Virus. *Mol. Biol. Evol.* 19, 230–238. doi: 10.1093/oxfordjournals.molbev.a004076

Roy, A., Choudhary, N., Guillermo, L. M., Shao, J., Govindarajulu, A., Achor, D., et al. (2013). A novel virus of the genus *Cilevirus* causing symptoms similar to citrus leprosis. *Phytopathology* 103, 488–500. doi: 10.1094/PHYTO-07-12-0177-R

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol. Biol. Evol.* 34, 3299–3302. doi: 10.1093/molbev/msx248

Rstudio_team, (2021). *RStudio | Open source & professional software for data science teams - RStudio*. Available online at: https://www.rstudio.com/ (accessed June 9, 2021)

Schlub, T. E., Buchmann, J. P., and Holmes, E. C. (2018). A simple method to detect candidate overlapping genes in viruses using single genome sequences. *Mol. Biol. Evol.* 35, 2572–2581. doi: 10.1093/molbev/msy155

Sharp, P., and Li, W. (1987). The codon Adaptation Index–a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 15, 1281–1295. doi: 10.1093/NAR/15.3.1281

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

Sievers, F., and Higgins, D. G. (2014). Clustal omega, accurate alignment of very large numbers of sequences. *Methods Mol. Biol.* 1079, 105–116. doi: 10.1007/978-1-62703-646-7_6

Singh, U., and Wurtele, E. S. (2020). Genetic novelty: How new genes are born. *Elife* 9:55136. doi: 10.7554/eLife.55136

Solovyev, A. G., and Morozov, S. Y. (2017). Non-replicative Integral Membrane Proteins Encoded by Plant Alpha-Like Viruses: Emergence of Diverse Orphan ORFs and Movement Protein Genes. *Front. Plant Sci.* 8:1820. doi: 10.3389/fpls.2017.01820

Tautz, D., and Domazet-Lošo, T. (2011). The evolutionary origin of orphan genes. *Nat. Rev. Genet.* 12, 692–702. doi: 10.1038/nrg3053

Vakirlis, N., Carvunis, A.-R., and McLysaght, A. (2020). Synteny-based analyses indicate that sequence divergence is not the main source of orphan genes. *Elife* 9:53500. doi: 10.7554/eLife.53500

Van Oss, S. B., and Carvunis, A. R. (2019). De novo gene birth. *PLoS Genet.* 15:1008160. doi: 10.1371/journal.pgen.1008160

Vasilakis, N., Forrester, N. L., Palacios, G., Nasar, F., Savji, N., Rossi, S. L., et al. (2013). Negevirus: a Proposed New Taxon of Insect-Specific Viruses with Wide Geographic Distribution. *J. Virol.* 87, 2475–2488. doi: 10.1128/JVI.00776-12

Vignuzzi, M., and López, C. B. (2019). Defective viral genomes are key drivers of the virus–host interaction. *Nat. Microbiol.* 4, 1075–1087. doi: 10.1038/s41564-019-0465-y

Vinokurov, K. S., and Koloniuk, I. (2019). Discovery and characterization of a novel alphavirus-like RNA virus from the red firebug Pyrrhocoris apterus L. (Heteroptera). *J. Invertebr. Pathol.* 166:107213. doi: 10.1016/j.jip.2019.107213

Wright, F. (1990). The "effective number of codons" used in a gene. *Gene* 87, 23–29. doi: 10.1016/0378-1119(90)90491-9

Yang, J., and Zhang, Y. (2015). I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res.* 43, W174–W181. doi: 10.1093/nar/gkv342

Yin, Y., and Fischer, D. (2008). Identification and investigation of ORFans in the viral world. *BMC Genomics* 9:24. doi: 10.1186/1471-2164-9-24

Zhang, Y., and Skolnick, J. (2004). Scoring function for automated assessment of protein structure template quality. *Proteins Struct. Funct. Bioinforma.* 57, 702–710. doi: 10.1002/prot.20264

Zwart, M. P., Willemsen, A., Daròs, J. A., and Elena, S. F. (2014). Experimental evolution of pseudogenization and gene loss in a plant RNA virus. *Mol. Biol. Evol.* 31, 121–134. doi: 10.1093/molbev/mst175

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 28 March 2022
doi: 10.3389/fmicb.2022.836076

A Novel Lineage of Cile-Like Viruses Discloses the Phylogenetic Continuum Across the Family *Kitaviridae*

Pedro L. Ramos-González[1†], Camila Chabi-Jesus[1,2†], Aline D. Tassi[1,2†], Renata Faier Calegario[3], Ricardo Harakava[1], Claudia F. Nome[4], Elliot W. Kitajima[2] and Juliana Freitas-Astua[5*]*

[1] Laboratório de Biologia Molecular Aplicada, Instituto Biológico de São Paulo, São Paulo, Brazil, [2] Escola Superior de Agricultura Luiz de Queiroz (ESALQ), Universidade de São Paulo, Piracicaba, Brazil, [3] Departamento de Fitotecnia e Fitossanidade, Universidade Federal do Paraná, Curitiba, Brazil, [4] Instituto de Patologia Vegetal, Centro de Investigaciones Agropecuarias, Instituto Nacional de Tecnología Agropecuaria (INTA), Córdoba, Argentina, [5] Embrapa Mandioca e Fruticultura, Cruz das Almas, Brazil

OPEN ACCESS

Edited by:
Elvira Fiallo-Olivé,
La Mayora Experimental Station (CSIC), Spain

Reviewed by:
Michael Melzer,
University of Hawaii, United States
Massimo Turina,
National Research Council (CNR), Italy

***Correspondence:**
Pedro L. Ramos-González
plrg1970@gmail.com
Juliana Freitas-Astua
juliana.astua@embrapa.br

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Microbe and Virus Interactions with Plants, a section of the journal Frontiers in Microbiology

Received: 15 December 2021
Accepted: 15 February 2022
Published: 28 March 2022

Citation:
Ramos-González PL, Chabi-Jesus C, Tassi AD, Calegario RF, Harakava R, Nome CF, Kitajima EW and Freitas-Astua J (2022) A Novel Lineage of Cile-Like Viruses Discloses the Phylogenetic Continuum Across the Family Kitaviridae. Front. Microbiol. 13:836076. doi: 10.3389/fmicb.2022.836076

An increasing number of plant species have been recognized or considered likely reservoirs of viruses transmitted by *Brevipalpus* mites. A tiny fraction of these viruses, primarily those causing severe economic burden to prominent crops, have been fully characterized. In this study, based on high-throughput sequencing, transmission electron microscopy analyses of virions in plant-infected tissues, viral transmission experiments, and the morphoanatomical identification of the involved *Brevipalpus* mites, we describe molecular and biological features of viruses representing three new tentative species of the family *Kitaviridae*. The genomes of Solanum violifolium ringspot virus (SvRSV, previously partially characterized), Ligustrum chlorotic spot virus (LigCSV), and Ligustrum leprosis virus (LigLV) have five open reading frames (ORFs) > 500 nts, two distributed in RNA1 and three in RNA2. RNA1 of these three viruses display the same genomic organization found in RNA1 of typical cileviruses, while their RNA2 are shorter, possessing only orthologs of genes *p61*, *p32,* and *p24.* LigCSV and LigLV are more closely related to each other than to SvRSV, but the identities between their genomic RNAs were lower than 70%. In gene-by-gene comparisons, ORFs from LigCSV and LigLV had the highest sequence identity values (nt sequences: 70–76% and deduced amino acid sequences: 74–83%). The next higher identity values were with ORFs from typical cileviruses, with values below 66%. Virions of LigLV (≈ 40 nm × 55 nm) and LigCSV (≈ 54 nm × 66 nm) appear almost spherical, contrasting with the bacilliform shape of SvRSV virions (≈ 47 nm × 101 nm). Mites collected from the virus-infected plants were identified as *Brevipalpus papayensis, B. tucuman*, and *B. obovatus.* Viruliferous *B. papayensis* mites successfully transmitted LigCSV to *Arabidopsis thaliana*. SvRSV, LigCSV, and LigLV seem to represent novel sub-lineages of kitaviruses that descent on parallel evolutionary branches from a common ancestor

shared with the tentative cile-like virus hibiscus yellow blotch virus and typical cileviruses. Biological and molecular data, notably, the phylogenetic reconstruction based on the RdRp proteins in which strong support for monophyly of the family *Kitaviridae* is observed, mark an advance in the understanding of kitavirids.

Keywords: *Brevipalpus*-transmitted viruses, high-throughput sequencing, ornamental plants, *Brevipalpus* mites, virion morphology, viral mixed infections

INTRODUCTION

The family *Kitaviridae,* order *Martellivirales,* groups plant-infecting viruses with linear single-stranded (ss) positive (+) split RNA genomes which are assigned to the genera *Cilevirus*, *Higrevirus,* or *Blunervirus* (Quito-Avila et al., 2021). Except for the blunervirus tea plant necrotic ring blotch virus (TPNRBV) (Hao et al., 2018), kitavirids stand out by the production of non-systemic diseases wherein only the locally infected tissues typically develop chlorotic and/or necrotic lesions (Freitas-Astúa et al., 2018; Quito-Avila et al., 2021). In the case of cileviruses, whose vectorial transmission by mites of the genus *Brevipalpus* has been long-established (Alberti and Kitajima, 2014; de Lillo et al., 2021), the viral spread remains restricted to tissues around the mite feeding sites. Cyto- and physiopathological effects ensuing from the infection by the cilevirus citrus leprosis C (CiLV-C) resemble a hypersensitive like-response (Marques et al., 2010; Arena et al., 2016, 2020).

Based on the sequences of the RNA-dependent RNA polymerase (RdRp) and P24 proteins, kitavirids display a close phylogenetic relationship with an increasing number of unclassified arthropod-infecting viruses including those belonging to the groups centivirus, aphiglyvirus, and negevirus (Kondo et al., 2020; Morozov et al., 2020; Ramos-González et al., 2020). Kitaviruses have quasi-spherical or bacilliform virions that in the case of cileviruses are enveloped particles (Kitajima and Alberti, 2014), a feature shared with negeviruses (Vasilakis et al., 2013).

The genome of typical cileviruses is split into two molecules in which RNA1, ≈9 kb, comprises two open reading frames (ORFs), *RdRp* and *p29*, whereas RNA2, ≈5 kb, includes four canonical ORFs (*p15*, *p61*, *p32,* and *p24*) (Freitas-Astúa et al., 2018). The 5′-end of the cilevirus RNA2 is of variable length and organization (Ramos-González et al., 2021). Between *p15* and *p61*, or overlapping *p15*, cilevirus genomes harbor orphan ORFs which in some cases encode predicted small proteins with predicted transmembrane domains (TM) (Ramos-González et al., 2020, 2021). Hibiscus yellow blotch virus (HYBV), a cile-like virus infecting hibiscus plants in Hawaii, has a genomic organization that differs from those commonly observed in typical cileviruses i.e., CiLV-C, citrus leprosis virus C2 (CiLV-C2), and passion fruit green spot virus (PfGSV) (Locali-Fabris et al., 2006; Roy et al., 2013; Ramos-González et al., 2020; Olmedo-Velarde et al., 2021). Compared with typical cileviruses, HYBV lacks the genomic region upstream of the ORF *p61*, while its ORF ortholog of *p29* is located at the 3′-end of the RNA2 instead of its regular locus at the 3′-end of RNA1.

A disease called "*lepra explosiva de la ligustrina*" (LEL) that affects privet (*Ligustrum sinense*), an ornamental shrub, was first observed at the end of the 1930s, in Argentina (Vergani, 1942). LEL drew attention because of the similarity between its symptoms in leaves and twigs and those generated by "*lepra explosiva*," a serious disease of citrus concomitantly described in Argentina, Brazil, and Paraguay later referred to as citrus leprosis and caused by a range of cileviruses and dichorhaviruses (Bitancourt, 1940, 1955; Frezzi, 1940; Vergani, 1942; Bastianel et al., 2010; Ramos-González et al., 2018). Affected privet leaves showed isolated circular or irregular spots with 2–4 mm in diameter (sometimes reaching 6–8 mm in diameter) with yellowish-green halos, in which the center of the lesion was usually noticeably chlorotic. The transmission of LEL was first demonstrated to be carried out by *Tenuipalpus pseudocuneatus* Blanchard, a mites species later synonymized with *Brevipalpus obovatus* (Welbourn et al., 2003). The application of miticides successfully blocked the transmission of LEL to both new healthy plants and fresh shoots of the diseased plants (Vergani, 1942). The putative causal agent of the disease was tentatively named Ligustrum leprosis virus (Tassi, 2018).

In 1976, shrubs of *Ligustrum lucidum* in a hedgerow in the Agricultural Science campus of Federal University of Paraná, in Curitiba, State of Paraná (PR), Brazil, were found affected by ring spots in the leaves (Lima Neto et al., 1994). Abundant aggregated bacilliform particles of approximately 28 × 103 nm inside enlarged cisternae in the perinuclear region and/or large electronluscent vesicles, likely viroplasms, associated with the rough endoplasmic reticulum were detected in the parenchyma cells. Based on particle morphology, which was somehow similar to those of rhabdoviruses but of incongruent size, and, particularly, the difference in the cytopathological effects observed in the infected cells, it was suggested the presence of a new type of virus (Lima et al., 1991; Lima Neto et al., 1994).

A tentative cilevirus called Solanum violifolium ringspot virus (SvRSV) was first described in Piracicaba, State of São Paulo (SP), Brazil, in 2007 (Ferreira et al., 2007). SvRSV-infected leaves of the ornamental creeping plant *Solanum violifolium* show chlorotic spots which can turn into necrotic lesions, and occasionally, when senescent, they can display green islands. Experimentally, SvRSV can be successfully transmitted to *Arabidopsis thaliana* and other plants of the families Solanaceae, Amaranthaceae, and Malvaceae, mechanically or by using viruliferous mites of the species *B. yothersi* (previously known as *Brevipalpus phoenicis* sensu lato) (Ferreira et al., 2007; Arena et al., 2017). The virus is also transmitted by *B. obovatus*, which seems to be its natural vector (Ferreira et al., 2007).

In SvRSV-infected plants, abundant enveloped bacilliform particles with 50–60 nm in width and an extremely heterogeneous length, 100–1,000 nm, were commonly observed in the lumen of the endoplasmic reticulum and vesicles likely derived from

the endoplasmic reticulum (Ferreira et al., 2007). Cytopathic effects such as the presence of electron-dense areas were noticed in the cytoplasm of cells of the plant species *S. violifolium, Datura suaveolens*, *Salvia leucantha*, *Thumbergia erecta*, *Hibiscus cannabinus*, and *Capsicum annum*. A fragment of approximately 600 bp derived from the ORF encoding the RdRp of SvRSV was obtained from an infected *S. violifolium* plant collected in Piracicaba, SP, Brazil, in 2006 (GenBank accession number DQ514336). The fragment showed a relatively low nucleotide (nt) sequence identity (<62%) with the genome of CiLV-C (Ferreira et al., 2007). Using the same specific primer pair, SvRSV was detected in an infected *Gomesa bifolia* plant, a small, cool-temperature native orchid from South America collected in Córdoba, Argentina, in 2012 (GenBank accession number KT733671).

In this study, we describe the full-genomic sequences of SvRSV and two other related cile-like viruses found in several ornamental plants collected in Brazil and Argentina. Additionally, mites found on the leaf samples were identified by morphoanatomical evaluations and in some cases were used for viral transmission to arabidopsis plants. The size of virions detected in natural and experimental hosts was measured and compared with those of the typical cilevirus PfGSV. Finally, we mapped the distribution of two kitaviruses found infecting the same sample by detecting their RNA molecules in isolated leaf lesions.

MATERIALS AND METHODS

Plant Material

Leaf samples from *Solanum violifolium* Schott plants showing chlorotic and necrotic spots were collected in Piracicaba, SP, Brazil, in 2014 (sample Prb1). Branches of privet shrubs of the species *Ligustrum lucidum, L. japonicum,* and *L. sinense* with leaves mainly showing chlorotic spots were collected in urban public gardens of the cities of São Paulo, SP, Brazil, in 2018 (sample SPa1); Curitiba, PR, Brazil, in 2018 (samples Crb1 and Crb2); and La Falda, Córdoba, Argentina, in 2019 (sample Cdb1), respectively. Pictures of all samples were taken immediately after their collection (**Figure 1**).

Plant samples were examined under a stereo-microscope (Leica M125, Wetzlar, Germany) or a hand-held magnifying glass to detect the presence of *Brevipalpus* mites and collect them when found. Mites were kept in 90% ethanol solution until their taxonomic identification. Small pieces of leaf lesions were removed using new razor blades or sterile scissors. Fragments of these tissues were prefixed in a modified Karnovsky solution (2.5% glutaraldehyde and 2% paraformaldehyde in 0.05 M cacodylate buffer, pH 7.2) (Kitajima and Nome, 1999) for transmission electron microscopy (TEM) analyses or conserved in RNA stabilization solution (RNA*later*, Thermo Fisher Scientific, Waltham, MA, United States) until RNA extraction.

All of the preserved biological materials and fresh leaves and branches of plants collected in Brazil were sent to Instituto Biológico de São Paulo, SP, Brazil, and Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, SP, Brazil. Living mites were captured from fresh samples and used for virus transmission experiments. Symptomatic plant tissues from these samples were stored at -80°C until their use in molecular analyses. Healthy tissues from plants of each of the studied species were also collected and used as negative controls in reverse transcriptase-PCR viral detection tests.

Virion Particle Morphology and Cytopathology of the Infected Plant Tissues

The pretreatment of plant tissues for TEM analyses was carried out as previously described (Ramos-González et al., 2017). Copper grids with plant tissue sections embedded in Spurr's epoxy resin and stained with 3% uranyl acetate and Reynold's lead citrate were examined in a JEOL JEM 1011 (JEOL, Akishima, Japan) transmission electron microscope. For particle size measurements, all images were recorded at the same magnification (200,000x). Ten to fifteen tissue sections per virus/host pair from four independent plants were examined. Exceptionally, virions of the sample Cdb1 corresponded to a single plant. At least 50 particles per virus/host combination were measured using the program ImageJ (Collins, 2007). The measurements of length and width were carried out from different particles. For the sake of comparative analysis of the virion morphology, viral particles in the samples SPa1 and Prb1 were measured in tissues from both the naturally infected *Ligustrum* spp. and *Arabidopsis thaliana* plants which were experimentally infected using viruliferous mites found in field-collected branches from those samples. Virions of the sample Cdb1 were only measured in tissues of an *L. sinense* plant. In parallel, particles of the typical cilevirus PfGSV isolate BSB infecting both arabidopsis and *L. japonicum* plants were measured. Values of viral width and length were compared using the non-parametric tests Kruskal-Wallis and Wilcoxon signed-rank implemented in Rstudio (Rstudio_Team, 2021). Row data were represented using a violin plot using geom_violin (ggplot2, tidyverse) also implemented in Rstudio.

RNA Extraction, cDNA Synthesis, and Virus Detection

One to two grams of plant tissues were ground in mortars in presence of liquid N_2 and their RNA extracts were obtained using TRIzol™ reagent following the recommendation of the manufacturer (Life Technologies, Foster City, CA, United States). No more than 100 mg of plant tissue were processed with 1 mL of TRIzol™. The RNA concentration in the final extracts was determined at 260 nm using a Nanodrop ND-8000 micro-spectrophotometer (Thermo Scientific, Waltham, MA, United States). Approximately 500 ng of the RNA extracts were used to prepare the cDNA solutions using a mix of random hexamer primers and GoScript™ Reverse Transcriptase kit as described by the manufacturer (Promega, Madison, WI, United States).

For the preparation of RNA extracts from small pieces of tissues, i.e., isolated leaf lesions comprising no more than 10–30 mg, tissue fragments were ground using the FastPrep-24™

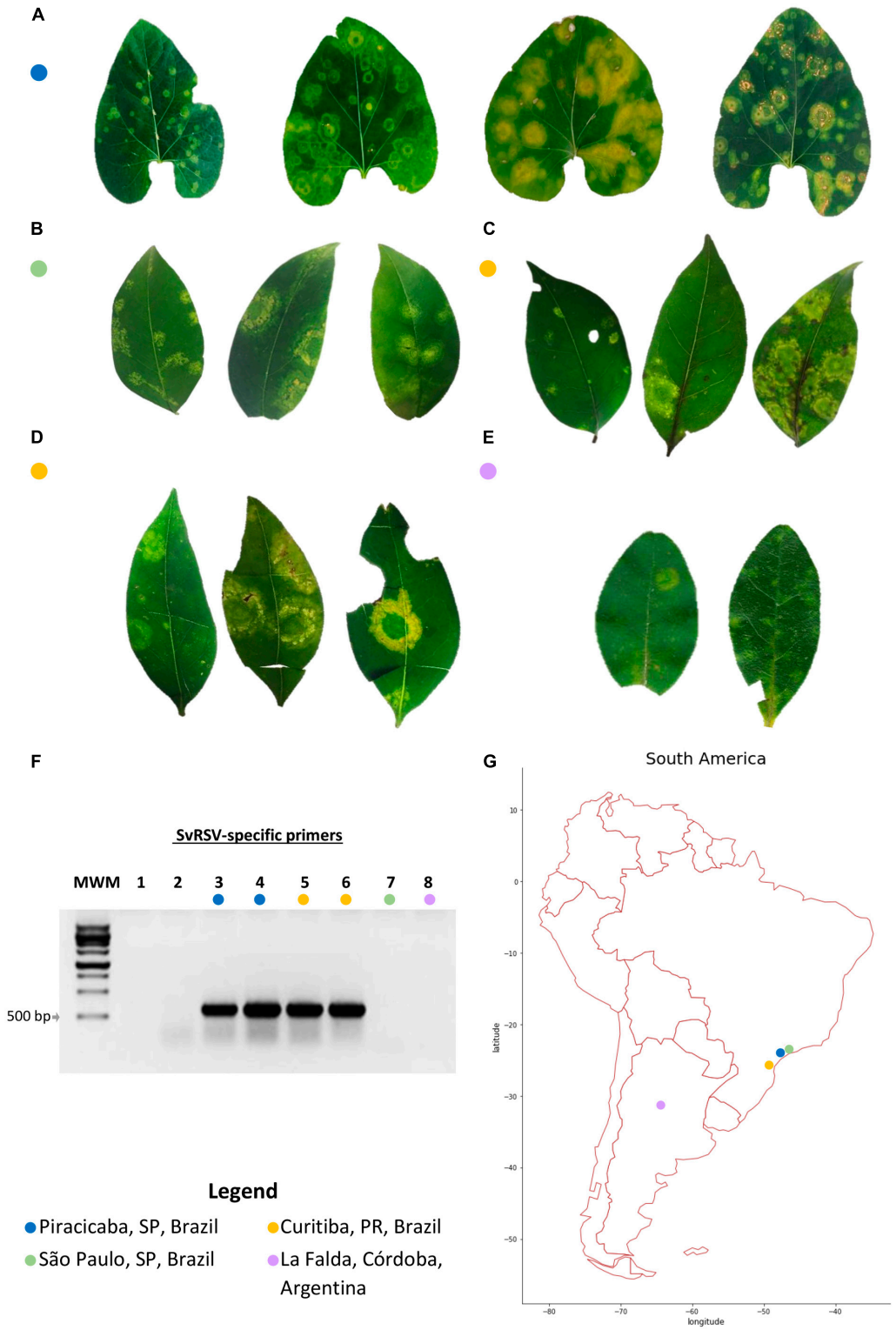


FIGURE 1 | Leaf symptoms, detection of the partially characterized Solanum violifolium ringspot virus (SvRSV) in the RNA extracts, and place of collection of the samples. **(A)** Sample Prb1. Leaves of the ornamental herbaceous *Solanum violifolium* plant were likely collected during different stages of disease progression. The leaf on the left shows small chlorotic ring spots and was likely collected a few days after the infection. In contrast, the three leaves on the right show large chlorotic spots that also include necrotic areas in form of concentric circles. **(B)** Sample SPa1. Leaves of the ornamental shrub *Ligustrum japonicum*, collected in a private garden, show chlorotic spots with yellow irregular lesions, and regular concentric rings. **(C)** Sample Crb1. Leaves of *L. lucidum* were collected in the public garden *Bosque do Papa* in the city of Curitiba, PR, showing yellow concentric spots of variable sizes. **(D)** Sample Crb2. Leaves of *L. japonicum* plants were collected in a

(Continued)

FIGURE 1 | hedgerow in the Agricultural Science campus of the Federal University of Paraná. Typical symptoms consisted of large yellowish concentric spots and necrotic areas were infrequently observed. **(E)** Sample Cdb1. Leaves of *L. sinense* shrubs showing small (<0.5 cm in diameter) to medium (≈ 1 cm in diameter) light yellowish to greenish spots. The largest lesions showed small rings of necrotic tissues. **(F)** 1% agarose gel electrophoresis of RT-PCR products for the detection of the tentative cilevirus SvRSV using specific primers. The expected size of the amplicons is ≈ 600 bp. MWM: molecular weight marker, M1181 Ladder (Sinapse Biotechnology, São Paulo, SP, BR). Lanes 1 and 2: Reverse-transcription reaction and PCR negative controls, respectively. Lanes 3–8 correspond with samples as indicated by the colored circles; lanes 3 and 4: two RNA extracts from sample Prb1; 5: sample Crb1; 6: sample Crb2; 7: sample SPa1; and 8: sample Cdb1. **(G)** South America map indicating the place of collection of samples evaluated in this study.

homogenizer (MP Biomedicals, Santa Ana, CA, United States). Samples were ground in the presence of three steel beads of 3 mm in diameter and 500 μL of TRIzol™ during two cycles of 30 s at 6.5 m s^{-1}. After the purification of the total RNA fraction, pellets were resuspended in 10 μL of Milli-Q water treated with DEPC. cDNA solutions were prepared using 2 μL of the RNA extracts and following the directions described in the GoScript™ Reverse Transcriptase kit (Promega).

The presence of viral genomes was tested by PCR using 2 μL of cDNA solutions as templates and GoTaq® G2 Green Master Mix (Promega). A set of specific and generic primers were used for the detection of the cileviruses CiLV-C, CiLV-C2, PfGSV, and the partially characterized kitavirus SvRSV (**Table 1**). Newly designed primer pairs obtained after genomic analyses carried out in this study were also used for viral screening (**Table 1**). Generated amplicons were separated and visualized in 1% agarose gel in 1X Tris-acetate-ethylenediaminetetraacetic acid (TAE) in the presence of ethidium bromide (0.5 μg/mL), under ultraviolet light.

High Throughput Sequencing and *in silico* Assembly of Viral Genomes

RNA extracts obtained by TRIzol™ method from the samples Prb1, Cdb1, SPa1, and Crb1 were further purified using RNeasy Mini Kit (QIAGEN, Venlo, The Netherlands). The concentration of RNA in the final solutions and the A260/A280 ratio were assessed using a NanoDrop ND-8000 micro-spectrophotometer. One to five hundred nanograms of each RNA extract were sent to the Animal Biotech Laboratory at Escola Superior de Agricultura Luiz de Queiroz, University of São Paulo (Piracicaba, SP, Brazil) for high-throughput sequencing (HTS) using HiSeq 2500 Technology (2 × 150 nt paired-end reads) (Illumina, San Diego, CA, United States). RNA extracts from each sample were independently processed resulting in four HTS libraries. The enrichment of poly-A RNA fractions, preparation of libraries, sequencing, quality assessments of reads, and the removal of the adaptor sequences, were carried out as previously described (Chabi-Jesus et al., 2021). Reads were *de novo* assembled using SPAdes (Bankevich et al., 2012) and Trinity (Haas et al., 2013) implemented on the Galaxy platform[1] v 21.09 (Afgan et al., 2018). Contigs were annotated with BlastX and BlastN implemented in Geneious software package v 11.1.4 (Kearse et al., 2012) using custom-organized plant viral genome databases retrieved from NCBI Virus[2] (Hatcher et al., 2017). The largest contigs producing the best *E-value* score with cileviruses sequences i.e., CiLV-C, CiLV-C2, and PfGSV, were selected for further detailed analyses. Nucleotide sequences corresponding to the genomic fragments of the partially characterized SvRSV isolates previously collected in Brazil and Argentina, i.e., GenBank # DQ514336 and KT733671, were also included in the nucleotide database during the Blast-based screenings of the contigs.

Three sets of primers corresponding to each of the kitaviruses identified in this study were designed using the Primer3 program (Untergasser et al., 2012). Primer target regions were conveniently selected to generate overlapping amplicons of approximately 0.7–1.0 kb (**Supplementary Table 1**). Generated amplicons were used for the validation of HTS results and, when required, to amplify the genome sequence of novel viral isolates. The 5′-ends of genomic molecules of new viruses were obtained using SMARTer® RACE 5′/3′ Kit (Clontech Laboratories, Mountain View, CA, United States) following the procedure previously described (Ramos-González et al., 2020).

The relative abundance of viral-specific reads and mean coverage of viral bases per generated HTS library were calculated using BBMap (Bushnell, 2014). Reads were mapped to each genomic segment of the virus detected in each sample.

In silico Viral Genome Analyses: Annotation, Pairwise Comparisons, and Phylogenetic Relationships

Viral ORFs were identified using the ORF finder.[3] The presence of signal peptides, conserved domain architecture, and transmembrane helices in predicted viral proteins was detected using SignalP 5.0[4] (Armenteros et al., 2019), MOTIF Search,[5] TMHMM Server 2.06[6] (Sonnhammer et al., 1998), and Deeploc v 1[7] (Armenteros et al., 2017), respectively. Nucleotide and predicted amino acid sequences were aligned and the identity values were assessed using MAFFT[8] (Katoh and Standley, 2013).

For phylogenetic analyses, sequences from kitaviruses, unclassified kita/nege like viruses including those of the groups negevirus, centivirus, and aphiglyvirus, were retrieved from GenBank after a BLAST search (cut-off *E-value* < e-10) using typical cilevirus sequences as the query. Some viruses of the family *Virgaviridae* were also included as an outgroup. Phylogenetic informative regions of the multiple sequence alignments (MSAs) were selected using BMGE software

[1]https://usegalaxy.org

[2]https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/

[3]https://www.ncbi.nlm.nih.gov/orffinder/

[4]http://www.cbs.dtu.dk/services/SignalP/

[5]https://www.genome.jp/tools/motif/

[6]http://www.cbs.dtu.dk/services/TMHMM/

[7]http://www.cbs.dtu.dk/services/DeepLoc-1.0/index.php

[8]https://www.ebi.ac.uk/Tools/msa/mafft/

TABLE 1 | List of specific and generic primers for the detection of typical cileviruses and cile-like viruses described in this study.

Virus[a]	ORF target	Primer sequence (5′-3′)	Ta[b](°C)	Amplicon size (bp)	References
CiLV-C	*p24*	F: CGCAGTTTCCTAATAACACC	54	322	Chabi-Jesus et al., 2021
		R: GGGAGTTCAGCATAAAGC			
CiLV-C CRD	*p29*	F: CAGAAGGCCGAGGTTGTAAAG	56	330	Ramos-González et al., 2016
		R: GTAGTGATCACTGAACTCGAATACC			
	p24	F: ATGTTGGCAACGGAAAGT	54	522	Chabi-Jesus et al., 2021
		R: AACTTTTTCAACCCTGTTCAC			
CiLV-C SJP	*p29*	F: GTAARCAAAAGGTCGAGGTTGTCC	56	456	Ramos-González et al., 2016
		R: TCTGTTGTCTAGCAGCRAGTAATG			
	p24	F: CTCATGATATCCTTGATGACC	54	393	Chabi-Jesus et al., 2021
		R: CAACCTTCTCAACCTTATTAGTC			
CiLV-C2	*p29*	F: ATGAGTAACATTGTGTCGTTTTCGTTGT	56	795	Roy et al., 2013
		R: TCACTCTTCCTGTTCATCAACCTGTT			
PfGSV	*p29*	F: ACACCAAGAGTACTATCGATC	54	452	Ramos-González et al., 2020
		R: CATCAAGTGGAGCAAGTTC			
SvRSV	*RdRp*	F: TGTCGAACTTTGGTATGAGTCG	54	596	Ferreira et al., 2007
		R: CCGGTTCGTCAAATAACTCC			
	p31	F: CACGTCGTTCAGCAGAA	54	490	This study
		R: ACCTCTTGGTCATCGACT			
	p23	F: GGCTGTTCTGGTATTATTTG	54	474	
		R: GAACTCAAGGTACTAGAAG			
LigCSV	*p31*	F: TGGTTACCGTTACTTTTTCTC	54	123	
		R: TTTTAGACTTCAACGCCTTC			
		F: GGTCACGAATTATAAGGCAG	54	373	
		R: TGGAATGGCTTTGATAGTCT			
	p23	F: TCGGATTGATTGTCTCTGTG	54	420	
		R: AAACCGGATTTGAATTATATG			
		F: CTCTCATAAATTGGGCGAAG	54	304	
		R: GGCAACTCCTTGTAAAGTTT			
	p33	F: GGATATACTCTCGAGCGATT	54	318	
		R: CCCATTTCAGAACCAACATT			
LigLV	*RdRp*	F: AAAACCCACACTTTCTGATG	54	303	
		R: TTGCACTCGAATAACAAGAC			
	p32	F: AAATCAGGCTGTTAATGTCG	54	435	
		R: AGGACACGCAAATTCTTATG			
	p34	F: TTGTCTCTAATGGATCCGAG	54	391	
		R: GCATTTTCATTTACGCTGTC			
	p24	F: CATGTATGTAGCAGTGTTGG	54	316	
		R: GAGAATTCGCGTTATTGGAT			

[a]*CiLV-C, citrus leprosis virus C; CiLV-C CRD, citrus leprosis virus C strain CRD; CiLV-C SJP, citrus leprosis virus C strain SJP; CiLV-C2, citrus leprosis virus C2; PfGSV, passion fruit green spot virus; SvRSV, Solanum violifolium ringspot virus; LigCSV, Ligustrum chlorotic spot virus; LigLV, Ligustrum leprosis virus.* [b]*Ta: PCR annealing temperature. F and R indicate forward and reverse primers, respectively.*

(Criscuolo and Gribaldo, 2010) implemented in NGPhylogeny[9] (Lemoine et al., 2019). BMGE parameters were set as follows: estimated matrix BLOSUM 62, sliding windows size = 7, maximum entropy threshold = 0.5, gap rate cut-off = 0.5, and minimum block size = 0.5. The substitution models with the lower Bayesian information criterion scores for RdRp and P24 MSAs and the Maximum Likelihood trees were obtained using W-IQ-TREE software v. 1.6.12[10] (Trifinopoulos et al., 2016). The reliability of the inferred evolutionary relationships was assessed by 1,000 bootstrap replications. Sequences of the methyltransferase and helicase domains, and RdRp encoded by the RNA1 and RNA2 molecules, respectively, of the blunerviruses blueberry necrotic ring blotch virus of the strains Georgia and RL (BNRBV_Georgia and BNRBV_RL, respectively), TPNRBV, and tomato fruit blotch virus (ToFBV) were concatenated as previously described (Quito-Avila et al., 2013; Ramos-González et al., 2020). Trees were edited and visualized using Interactive Tree Of Life (iTOL) v 5 (Letunic and Bork, 2021).

Identification of Conserved 3′Untranslated Regions and *in silico* Search for Novel Viral Segments

Viral segments of SvRSV, LigCSV, LigLV, and those of typical cileviruses were aligned using MAFTT (Katoh and Standley, 2013) and visualized using the Jalview program

[9]https://ngphylogeny.fr/

[10]http://iqtree.cibiv.univie.ac.at/

(Procter et al., 2021). Sequences comprising 90–120 nts of the highly conserved 3′ untranslated regions (UTR) of SvRSV, LigCSV, and LigLV were used to search for the presence of putative non-detected viral segments in the HTS-contigs libraries using the UTR-backed iterative BLASTN approach (Zhang et al., 2022). Previously, BLASTN databases were built from the HTS-contig libraries of each sample using Geneious software package v 11.1.4 (Kearse et al., 2012).

Experimental Viral Transmission

Brevipalpus mites collected from branches of the privet sample SPa1 were directly transferred to healthy *Arabidopsis thaliana* Col-0 plants. Ten to fifteen mites were distributed onto each plant at a maximum of five mites per leaf. Seven to ten days after the transfer, mites were collected and kept in 90% ethanol solution while waiting for further morphological characterization. Arabidopsis leaves were collected when chlorotic and yellowish blotches were visually detected. Fragments of the symptomatic areas were conserved in Karnowski solution for TEM analyses or processed for RNA extraction and viral detection. Arabidopsis plants were kept at 23 ± 1°C and 14 h of light in a controlled growth chamber (Adaptis AR A1000, Conviron, Winnipeg, MB, Canada) from their germination to the end of the experiments. Two independent experiments were conducted in two seasons in which groups of four and seven arabidopsis plants were used for mite infestations.

Morphological Identification of *Brevipalpus* Mites

Mites prefixed in 90% ethanol solution were mounted in Hoyer's medium before the examination by differential interference contrast in a Zeis Axioimager II microscope (Carl Zeis AG., Jena, Germany). When required, other specimens were processed for observation under a JEM IT 300 scanning electron microscope (JEOL) following protocols previously described (Ramos-González et al., 2017). Distinctive external structures of mites, i.e., number of dorsolateral setae, number of solenidia on the tarsus of leg II, dorsal and venter reticulations patterns, the shape of vesicle of the spermatheca, and morphology of microplates were visually evaluated under microscopy and compared with the holotypes considering the detailed descriptions available (Beard et al., 2015). All images were digitally recorded.

RESULTS

Detection of Solanum Violifolium Ringspot Virus and Cilevirus-Like Particles in Symptomatic Samples

Plant samples collected in this study were selected based on their infection by partially characterized or unknown viruses likely to be members of the family *Kitaviridae* according to previous reports (Vergani, 1942; Lima et al., 1991; Lima Neto et al., 1994; Ferreira et al., 2007; Kitajima et al., 2010). Regardless of the plant species, these samples showed chlorotic or yellowish blotches in leaves (**Figure 1**). Nearly concentric rings of necrotic tissues were frequently observed in *S. violifolium* infected leaves likely in advanced stages of disease progression (**Figure 1A**). Small necrotic points were also observed in the central area of lesions of privet leaves of the samples Crb1 and Crb2 (**Figures 1C,D**). In leaves of *S. violifolium* and *Ligustrum* spp. plants, regular chlorotic concentric rings coexisted with asymmetrical chlorotic spots. Generally, symptomatic leaves were unevenly distributed across the plants suggesting non-systemic diseases.

RNA extracts from all samples were tested by RT-PCR using specific primers for the detection of the cileviruses CiLV-C, CiLV-C2, PfGSV, and the partially characterized SvRSV. Bands of approximately 600 bp were obtained in samples Prb1, Crb1, and Crb2 from the reactions using SvRSV-specific primers (**Figure 1F**). Reactions using primers for the detection of typical cileviruses were negative (results not shown).

In parallel analyses using TEM, presumed virus particles were observed in all symptomatic samples (**Figure 2**). Characteristic cytopathologic features of cilevirus infections, including the presence of viroplasm, were distinguished (**Figures 2A,B,D,E,I,J,L**). Viroplasms are electron-dense vacuolated structures of wide-ranging sizes and shapes that are believed to be formed by aggregates of virus-encoded proteins. Non-vacuolated viroplasms were observed in some samples (**Figures 2G,H**). In a large proportion of cells of those samples, viroplasms could not be detected despite the high number of presumed viral particles. Short, bacilliform or spheroidal, membrane-bounded particles, likely virions, were observed inside cisternae of the endoplasmic reticulum (**Figures 2C,F,K**). Viral particles were also observed in the cytosol of *S. violifolium* plants (pictures not shown). Cytopathic effects in *S. violifolium* cells were similar to those previously observed (Ferreira et al., 2007; Arena et al., 2017).

Plant Virome Analyses Reveal the Complete Genomes of Solanum Violifolium Ringspot Virus and Two Other Related but Distinct Types of Cile-Like Viruses

The composition of the RNA extracts from the samples Crb1, SPa1, Cdb1, and Prb1 was investigated through HTS and independent libraries with more than 12 million reads were obtained (**Table 2**). After *de novo* assembly of reads, generated contigs were examined using BlastN and BlastX for global detection of viral sequences, and in particular, of kita-like viruses. Contigs of approximately 8 kb with a high-quality match (*E-value* ≈ 0) to the previously described genomic fragments of SvRSV (597 nts of the ORF *RdRp,* RNA1, GenBank accession numbers DQ514336 and KT733671) were detected in samples Crb1 and Prb1. In these samples, contigs of ≈ 3 kb in length were also detected. They showed the best matches (*E*-value ≈ 10^{-30}) with the RNA2 segments of cileviruses. The sample Crb1 was particularly relevant since another pair of contigs, also comprising kitavirus-like sequences, were identified in its library. The largest contig comprised cilevirus RNA2-like sequences, whereas the smallest one, with approximately 1.0 Kb, contained sequences with certain nt sequence identity with the ORF

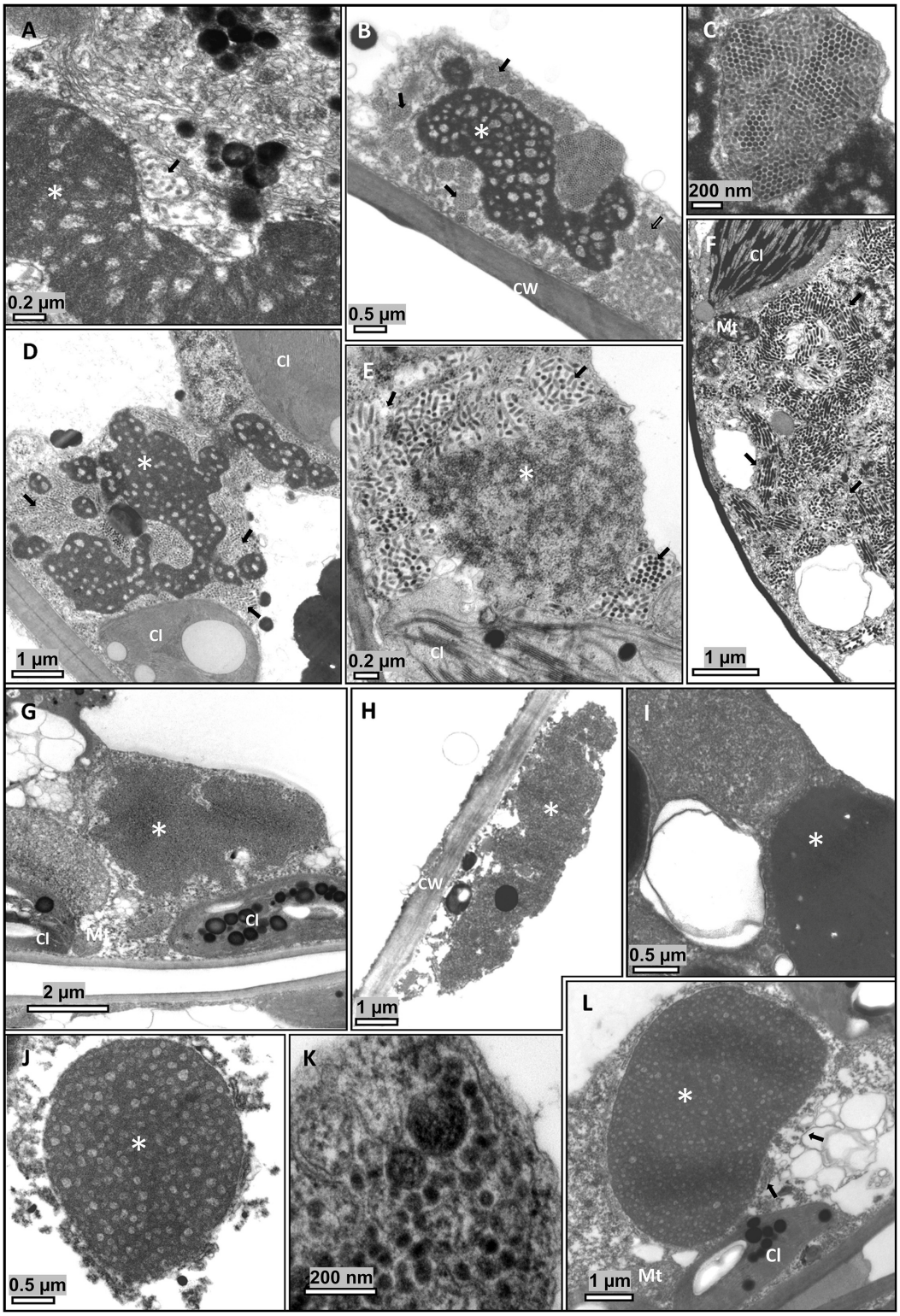


FIGURE 2 | Transmission electron micrographs of ultrathin sections of symptomatic tissues from leaves of *Ligustrum* spp. plants. Short, bacilliform, membrane-bounded particles, presumed virions, are usually observed close to the viroplasm, contained in membrane-bound cisternae of the endoplasmic reticulum. **(A–C)** *Ligustrum lucidum,* sample Crb1. **(D–F)** *L. japonicum,* sample Crb2. **(G–I)** *L. sinense,* sample Cdb1. **(J–L)** *L. japonicum,* sample SPa1. **(A–L)** Electron dense, vacuolated masses of varying shapes and bacilliform particles are observed in the cytoplasm of cells of the epidermis, mesophyll parenchyma, and, occasionally, in the xylem parenchyma. **(G–H)** Non-vacuolated viroplasms. Cl, chloroplast; CW, Cell wall; Mt, mitochondrion; V, virion. *Viroplasm.

TABLE 2 | Description of high-throughput sequencing libraries obtained in this study.

Field sample	**Crb1**				**Prb1**		**SPa1**		**Cdb1**	
Total number of reads recovered by library	13,181,765				16,000,302		15,055,185		12,574,635	
Identified virus_isolate	**SvRSV_Crb1[a]**		**LigCSV_Crb1**		**SvRSV_Prb1**		**LigCSV_SPa1**		**LigLV_Cdb1**	
Viral segment	**RNA1**	**RNA2**	**RNA1**	**RNA2**	**RNA1**	**RNA2**	**RNA1**	**RNA2**	**RNA1**	**RNA2**
Length of viral genomic segments (nts)	8,635	3,641	8,373	3,612	8,658	3,622	8,383	3,681	8,893	3,743
Reads mapped to viral genome	89,906	59,636	1,058	6,260	5,378,852	774,604	42,519	134,316	229,140	214,082
Viral-derived reads in the library (%)	0.682	0.452	0.008	0.047	33.618	4.841	0.282	0.892	1.822	1.702

[a]*GenBank accession numbers of each molecule are the following: SvRSV_Prb1 (OK626439 and OK626440), SvRSV_Crb1 (OK626441 and OK626442), LigCSV_SPa1 (OK626447 and OK626448), LigCSV_Crb1 (OK626449 and OK626450), and LigLV_Cdb1 (OK626451 and OK626452).*

p29 of cileviruses. According to the low nt sequence identity detected, the two pairs of contigs in the sample Crb1 likely belonged to viruses with markedly different genomic sequences. Interestingly, the second pair of contigs identified in the sample Crb1 displayed the best matches with the kitavirus-related contigs of approximately 8 and 3.5 kb that were independently recovered from libraries of the samples SPa1 and Cdb1. In summary, bioinformatics analyses revealed the near-complete genome sequences of two isolates of the partially characterized kitavirus SvRSV and three unknown kitaviruses.

Kitavirus-related contigs were used to generate specific primers (**Supplementary Table 1**) whereby the cDNA extracts from each plant sample were tested. Sanger-obtained sequences of the amplicons were aligned with the sequence of contigs recovered from the HTS method. Successive heuristic approaches based on the testing all the studied plant samples with the specific primer sets and nucleotide sequence comparisons of the generated amplicons revealed the full picture of the kitavirid diversity in the samples as follows: isolates of SvRSV were detected in samples Prb1 and Crb1, whereas isolates of a tentative novel virus called Ligustrum chlorotic spot virus (LigCSV) were identified in samples SPa1 and Crb1, and an isolate of a second tentative novel virus called Ligustrum leprosis virus (LigLV) was identified in the sample Cdb1. 5′- RACE analyses were conducted for the isolates SvRSV_Prb1, LigCSV_SPa1, and LigLV_Cdb1. The near-complete sequence of LigCSV isolate Crb1 was assembled from a mix of HTS- and Sanger-obtained contigs.

Excluding the polyA tails detected in the sequences, the genome of SvRSV_Prb1 comprises 12,232 nts split into two segments identified as RNA1 of 8,622 nts and RNA2 of 3,610 nts (GenBank accession numbers OK626439 and OK626440, respectively) (**Figure 3**). Similarly, the genome of LigCSV_SPa1 has 12,022 nts distributed in the segments RNA1 of 8,369 nts and RNA2 of 3,653 nts (GenBank acc. num. OK626447 and OK626448, respectively), whereas the genome of LigLV_Cdb1, with 12,611 nts in sum, contains 8,879 nts in the RNA1 and 3,732 in its RNA2 (GenBank acc. num. OK626451 and OK626452, respectively). Without RACE analyses, the near-complete genomes of SvRSV of the isolate Crb1 (GenBank # OK626441 and OK626442) and LigCSV_Crb1 (OK626449 and OK626450) were found to be slightly shorter than those of the isolate Prb1 and SPa1, respectively. Nucleotide sequences from different isolates of the same virus showed very high identity values (>93%), i.e., among the two isolates of SvRSV, and the two isolates of LigCSV (**Table 3**). Percent identities of the comparisons of the genomic sequences from different viruses ranged from 49.83 to 69.07%. Higher values corresponded to the comparisons involving the RNA1 molecules of LigCSV and LigLV (68–69%). Overall, the RNA2 molecules were more divergent than the RNA1 molecules.

Since three clusters of viral sequences could be consistently observed upon the analysis of genomic comparisons, one isolate of each virus was chosen for further *in silico* studies. The genomes of SvRSV_Prb1, LigCSV_SPa1, and LigLV_Cdb1 contain five ORFs larger than 500 nts distributed two in the RNA1 segment and three in the RNA2 (**Figure 3**). The RNA1 molecules of SvRSV, LigCSV, and LigLV have a large ORF (>7,200 nts) that according to BlastX encode the RdRp protein and a second ORF, likely an ortholog of the *p29* in typical cileviruses. The RdRp of these viruses had the motifs RdRP_2 (PF00978), Viral_helicase1 (PF01443), V_methyltransferase (PF01660), and AAA_30 (PF13604) in common, whereas other motifs were exclusively found in either the RdRp of SvRSV or LigLV, e.g., Chropara_Vmeth (PF19223, motif present in a family of chroparavirus proteins which are likely methyltransferases involved in mRNA capping), and DUF5488 (PF17590, motif in proteins of unknown function found in orthopoxvirus) (**Supplementary Table 2**).

RNA2 molecules in SvRSV_Prb1, LigCSV_SPa1, and LigLV_Cdb1 comprise three major ORFs encoding proteins possessing features similar to those in cileviruses (**Figure 3**). The larger protein encoded in RNA2 is a putative glycoprotein. *In silico* analyses revealed the presence of a signal peptide that in the case of P62 from SvRSV_Prb1 has a predicted cleavage site between the aa positions 17 and 18: VL**K**|**F**E. In P61 and P62 of LigCSV_SPa1 and P62 of LigLV_Cdb1, the signal peptides are larger with cleavage sites between positions 24 and 25: VN**A**|**T**I, and 20 and 21: T**L**|**Y**V, respectively. These three proteins have theoretical N-glycosylation sites predicted in position $\mathbf{N}_{142}$ASR in SvRSV_Prb1, $\mathbf{N}_{152}$HSE, $\mathbf{N}_{175}$LTR, and $\mathbf{N}_{235}$FTN in LigCSV_SPa1, and N_{91}VST, $\mathbf{N}_{152}$HSE, $\mathbf{N}_{175}$LTR, and $\mathbf{N}_{235}$FTN in LigLV_Cdb1. Invariably, these putative glycoproteins from the three viruses also show transmembrane domains near their COOH ends. Immediately downstream of the ORFs *p61-62*,

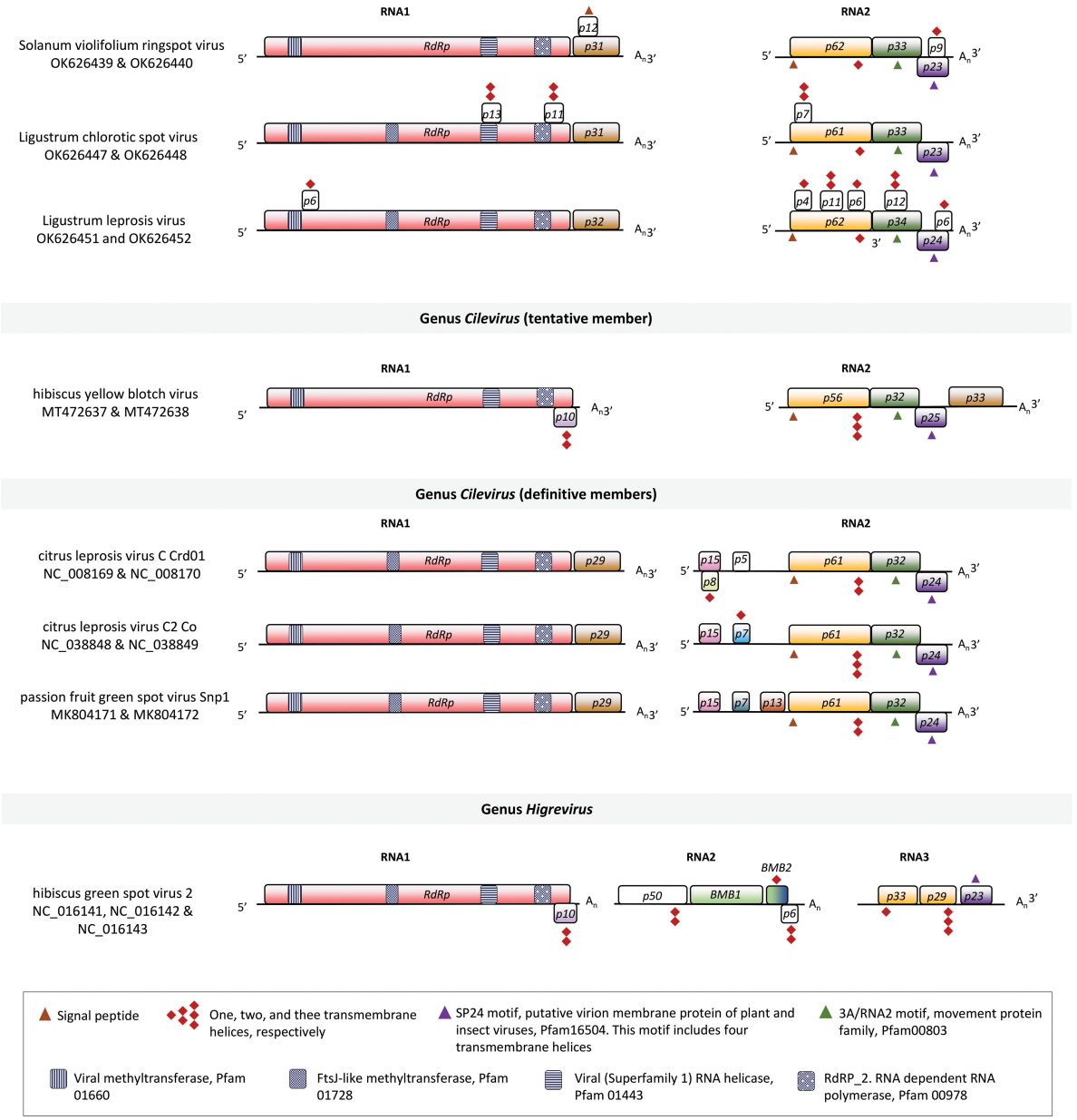


FIGURE 3 | Linear genomic maps of Solanum violifolium ringspot virus, Ligustrum chlorotic spot virus, Ligustrum leprosis virus, and other tentative and definitive members of the family *Kitaviridae*. Open reading frames (ORFs) are represented by boxes. Colors indicate a putative conserved functional or structural relationship between ORFs from different viruses. In citrus leprosis virus C, genus *Cilevirus, RdRp*: RNA dependent-RNA polymerase, *p29*: putative coat protein, *p15*: putative RNA silencing suppressor, *p61*: putative glycoprotein, likely a structural protein of the virion, *p32*: movement protein, *p24*: putative structural protein of the virion. White boxes indicate unknown features. Symbols of black pattern-filled boxes, triangles, and diamonds depict relevant amino acid motifs as described in the legend box.

the ORFs *p33-34* encode proteins with the characteristic 3A motif (PF00803) of viral movement proteins. The 3′-end of ORFs *p33-34* partially overlaps the 5′-end of the ORFs *p23-24*. The proteins encoded by ORF *p23-24* harbor the motifs SP24 (PF16504) which is centrally placed in the protein and at the structural level possesses four transmembrane helixes.

Other fully or partially overlapping ORFs smaller than 500 nts were detected over the genome of the studied viruses (**Figure 3**). In the three viruses, *in silico* analyses of the putative structure of the predicted encoded polypeptides indicated the presence of up to two transmembrane helices per molecule (**Supplementary Table 3**). Predicted small transmembrane (TM) proteins were more frequently detected in the RNA2 of LigLV, i.e., three with one TM helices, and two showing two TM helices. A putative basic protein (pI = 12.9) of approximately 12 kDa encoded in the RNA1 of SvRSV_Prb1 (ORF of 315 nts,

TABLE 3 | Nucleotide sequence identity amongst the RNA segments of viruses described in this study.

RNA1/RNA2	LigLV_Cdb1	LigCSV_Crb1	LigCSV_SPa1	SvRSV_Crb1	SvRSV_Prb1
SvRSV_Prb1[a]	52.91	52.52	52.60	**98.12**	
SvRSV_Crb1	52.63	52.59	52.65		**98.65**
LigCSV_SPa1	69.07	**98.87**		50.66	50.55
LigCSV_Crb1	68.97		**93.96**	50.66	50.61
LigLV_Cdb1		61.63	62.25	49.83	49.90

Figures above and below the diagonal depict the RNA1 and RNA2 identity values, respectively. The identity values higher than 90% are highlighted in boldface. [a]GenBank accession numbers of each molecule are the following: SvRSV_Prb1 (OK626439 and OK626440), SvRSV_Crb1 (OK626441 and OK626442), LigCSV_SPa1 (OK626447 and OK626448), LigCSV_Crb1 (OK626449 and OK626450), and LigLV_Cdb1 (OK626451 and OK626452).

positions 7,605–7,919) shows a signal peptide, which according to *in silico* analysis using DeepLoc, might suggest its interaction with the secretory pathway machinery. None of these predicted proteins had a significant identity with any protein described in public databases.

Paired comparisons among ORFs and their deduced proteins of SvRSV, LigCSV, and LigLV denoted the highest identity values between genes from LigCSV and LigLV (**Table 4**). Maximum identity percentages for each genomic segment corresponded with ORFs *RdRp* (nts: 70.5; aa: 74.7) and *p23-24* (nts: 75.7; aa: 83.1), in the RNA1 and 2, respectively. The highest values of identity in the evaluations including SvRSV were concentrated on the same ORFs, but they were consistently lower. For instance, amino acid sequence identities in the comparisons of RdRp and P23 from SvRSV with LigCSV and LigLV reached no more than 45 and 57%, respectively. When these analyses were expanded to definitive and tentative kitaviruses of the genera *Cilevirus* and *Higrevirus*, the higher identity values were invariably observed in the comparisons involving genes from the typical cileviruses, i.e., CiLV-C, CiLV-C2, and PfGSV (**Table 4**). In comparisons where orthologous genes from the higrevirus hibiscus green spot virus 2 (HGSV2) were included, the obtained values were almost as low as those obtained in the comparison concerning ORFs from the cile-like virus HYBV.

The alignment of RNA1 and RNA2 of SvRSV, LigCSV, and LigLV revealed the presence of conserved nucleotide stretches of ≈90–120 nts between the 3′-termini of each segment (**Figure 4**). In SvRSV isolates, these regions were more similar (98% nt sequence identity) than in LigLV and LigCSV (83–88% nt sequence identity). When the 3′-termini of the cileviruses CiLV-C, CiLV-C2, and PfGSV were included in the alignment, the presence of highly conserved nucleotide sequences in all these viruses became evident, likely denoting the putative functional role of these sequences as described in other viruses (Dreher, 1999; Liu et al., 2020; Rasekhian et al., 2021). The implementation of the UTR-backed iterative BLASTN approach (Zhang et al., 2022) using the 3′-terminus conserved nucleotide sequences of SvRSV, LigCSV, and LigLV, resulted in the recovery of contigs corresponding to the RNA1 and RNA2 segments already known, and new putative genomic segments were not identified.

A maximum-likelihood tree based on the deduced aa sequences of RpRp clustered SvRSV, LigCSV, and LigLV into two branches (**Figure 5**). They were placed in an intermediary position between branches comprising typical cileviruses, the cile-like virus HYBV, and the partially characterized cile-like virus Pistachio virus Y (GenBank accession numbers MT362606 and MT362605). Isolates of LigCSV and LigLV were grouped closer to typical cileviruses, whereas isolates of SvRSV were grouped in a more basal position. Overall, kitaviruses from the three genera were monophyletically clustered in a highly bootstrap-supported branch (>80%). Members of the family *Kitaviridae* shared a branch with three viruses found associated with the thrips *Frankliniella occidentalis* in Italy (Chiapello et al., 2021). *Frankliniella occidentalis* associated negev-like virus 1 (Foanegev1), Foanegev2, and Foanegev3 may belong to a new clade of arthropod-infecting viruses that, although related to negeviruses, have a distinct genomic organization. Subsequently, kitaviruses and *Frankliniella occidentalis* associated viruses clustered in a larger branch which also includes nelorpiviruses and sandewaviruses of the proposed taxon Negevirus, the centiviruses Wuhan insect virus 8, and Wuhan house centipede virus 1, and several other kita-like viruses, e.g., Saiwaicho virus and Beihai barnacle virus 2 (Kondo et al., 2020). This larger branch was not supported by bootstrap values > 60%.

In the tree using the P24 protein, SvRSV, LigCSV, and LigLV had almost the same phylogenetic pattern observed in the RdRp tree (**Supplementary Figure 1**). They were subclustered in a branch comprising the typical cileviruses and, in more basal positions, the cile-like virus HYBV, and the higrevirus HGSV2. In contrast to the RdRp-based tree, blunerviruses grouped in a sister branch that also included nelorpiviruses, centiviruses, aphiglyviruses, and other nege-kita/related arthropod-infecting viruses, e.g., Tetranychus urticae kitavirus.

Arabidopsis thaliana Plants Are Susceptible to Infection by Ligustrum Chlorotic Spot Virus

Brevipalpus mites from the privet sample SPa1 were transferred to healthy arabidopsis plants. Seven to ten days after infestation, leaves showing yellow spots, likely symptoms of the infection, were collected (**Figure 6A**). RT-PCR tests using specific primers detected the presence of LigCSV in 22 leaf samples (**Figure 6B**). Mites found after leaf collection were kept in ethanol solution for further identification.

Mites in the Infected Plants Belong to Three Species of the Genus *Brevipalpus*

Flat mites collected from field samples and arabidopsis plants used in the transmission experiments were morphoanatomically

TABLE 4 | Nucleotide (nt) and deduced amino acid (aa) identities between SvRSV_Prb1, LigCSV_SPa1 and LigLV_Cdb1, and other tentative and definitive members of the family *Kitaviridae*.

Virus	Genome segment	ORF (nts)[a]	Theoretical pI/MW (kDa)[b]	Cile-like viruses				Typical cileviruses			Higrevirus
				LigLV_Cdb1	LigCSV_SPa1	HYBV	PisVY	CiLV-C_Crd1	CiLV-C2_Co	PfGSV_Snp1	HGSV2
				Nucleotide/deduced amino acid sequence identity (%)							
SvRSV_Prb1[c]	RNA1	*RdRp* (7,419)	6.25/283.14	52.1/**44.5**	51.8/**44.2**	47.6/37.5	50.3/40.2	51.7/43.5	51.8/**44.7**	51.3/**44.3**	45.8/32.3
		p31 (849)	8.64/30.65	41.6/27.8	43.4/28.6	38.1/21.2	37.5/21.0	42.7/27.5	39.6/27.9	41.7/**31.0**	*[d]
	RNA2	*p62* (1,665)	7.13/62.15	41.8/25.4	43.1/**26.2**	38.5/17.8	35.4/18.5	42.1/24.0	42.0/24.1	41.1/24.4	*
		p33 (900)	7.81/32.97	51.7/44.4	51.8/43.0	50.7/43.4	54.0/44.7	57.5/**48.6**	57.4/**50.2**	54.7/**49.8**	*
		p23 (633)	9.56/23.43	58.8/**55.1**	58.2/**56.6**	47.7/37.1	55.8/49.7	58.7/52.1	58.3/52.4	59.2/**53.4**	42.7/25.1
LigCSV_SPa1	RNA1	*RdRp* (7,230)	7.41/276.11	70.5/**74.7**		49.7/37.0	50.3/39.4	58.0/54.7	57.0/53.2	58.4/54.4	44.4/32.2
		p31 (849)	9.02/30.63	57.4/**47.5**		38.4/20.8	38.6/22.7	46.8/33.2	44.3/34.1	45.6/31.1	*
	RNA2	*p61* (1,596)	5.81/60.56	58.4/**50.1**		41.3/19.6	39.3/19.6	43.9/28.5	43.8/28.7	42.6/29.9	*
		p33 (915)	7.67/33.34	63.3/**59.9**		51.4/42.6	54.8/40.8	56.2/53.9	57.2/50.5	56.2/53.0	*
		p23 (624)	9.36/23.23	75.7/**83.1**		47.9/39.6	56.5/51.3	64.2/68.3	63.6/65.4	64.8/66.0	46.2/31.4
LigLV_Cdb1	RNA1	*RdRp* (7,578)	6.80/288.80		70.5/**74.7**	48.9/36.2	51.0/40.0	57.8/53.7	57.4/52.0	57.7/52.3	45.0/31.5
		p32 (885)	9.05/31.98		57.4/**47.5**	38.3/21.9	36.4/20.5	45.0/29.8	43.7/31.4	41.0/29.0	*
	RNA2	*p62* (1,605)	6.33/61.55		58.4/**50.1**	38.8/17.6	40.3/17.9	44.5/31.0	45.7/33.3	45.8/31.7	*
		p34 (924)	7.64/33.83		63.3/**59.9**	50.3/42.4	52.5/41.3	56.6/54.3	57.3/53.1	55.1/52.1	*
		p24 (630)	9.47/23.57		75.7/**83.1**	47.7/37.19	58.1/51.2	62.7/66.5	63.7/62.9	65.3/63.6	45.0/27.5

The highest values of nt and aa identity in each row are underlined and highlighted in bold. [a]*Length (nts) of each ORF including the stop codon.* [b]*The isoelectric point (pI) and molecular weight (MW) in kDa of the deduced polypeptides were assessed using the Compute pI/Mw tool available in https://web.expasy.org/compute_pi/.* [c]*GenBank or RefSeq accession numbers of each molecule are the following: SvRSV_Prb1 (OK626439 and OK626440), SvRSV_Crb1 (OK626441 and OK626442), LigCSV_SPa1 (OK626447 and OK626448), LigCSV_Crb1 (OK626449 and OK626450), and LigLV_Cdb1 (OK626451 and OK626452), CiLV-C Crd1 (NC008169 and NC008170), CiLV-C2_Co (NC038848 and NC038849), PfGSV_Snp1 (MK804171 and MK804172), HYBV (MT472637 and MT472638), PisVY (MT362606 and MT362605), and HGSV2 (NC_016141, NC_016142, and NC_016143).* [d]*Absence of orthologues genes.*

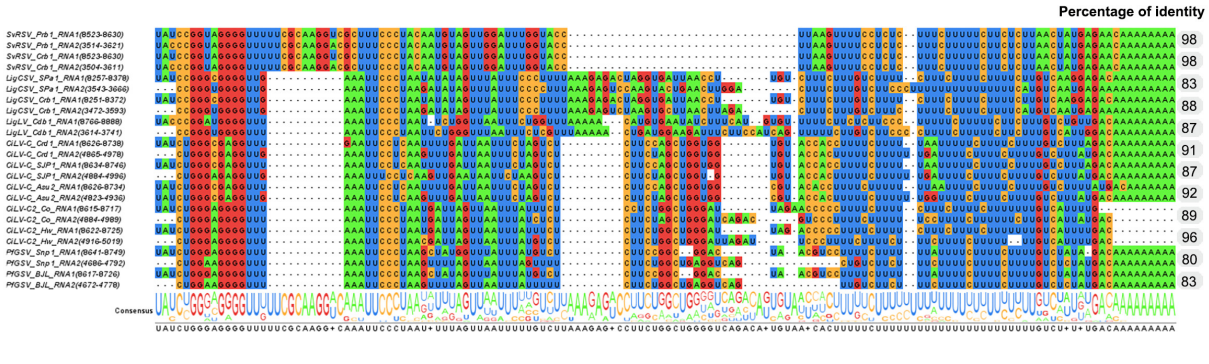


FIGURE 4 | Identification of conserved nucleotide sequences in the 3′-end genomic segments of isolates of Solanum violifolium ringspot virus (SvRSV), Ligustrum chlorotic spot virus (LigCSV), and Ligustrum leprosis virus (LigLV). Values in the column on the right indicate the nucleotide sequence identity between the 3′-termini of RNA1 and RNA2 viral segments. Genomic sequences of several isolates of the cileviruses CiLV-C, CiLV-C2, and PfGSV were also included in the study highlighting conserved nucleotide stretches across definitive and putative members of the genus *Cilevirus*.

examined. Analyzed mites were distributed as follows: sample Cdb1, field-collected mites: 79; sample Prb1, field-collected mites: 407, after transmission experiments: 50; sample SPa1, field-collected mites: 25, after transmission experiments: 15; and sample Crb1, field-collected mites: 54.

Mites from samples Crb1 and SPa1 were assigned to *B. papayensis* (**Figures 7A,D,H**), whereas those from the sample Cdb1 to *B. tucuman* (**Figures 7B,E**), and from sample Prb1 to *B. obovatus* (**Figures 7C,F,G**). The following traits were observed: (i) dorsal ophistosoma with six setae, *f2* absent, and palps four segmented, with one solenidion and two eupathidia on distal segment on mites of the three species; (ii) presence of two solenidia on tarsus II in *B. papayensis* and *B. tucuman*, and one solenidium in *B. obovatus*; (iii) prodosum with central cuticle with broad areolae on *B. papayensis* and *B. tucuman* and smooth on *B. obovatus*; (iv) in the dorsal opisthosoma, *d1-d1* to *e1-e1* weakly reticulate anteriorly, and short transverse fold posterior to *e1* on *B. obovatus*, weakly wrinkled on the anterior part, transversal folds becoming longitudinal folds toward *h1* on *B. papayensis*, and with some folds and wrinkles anteriorly, few

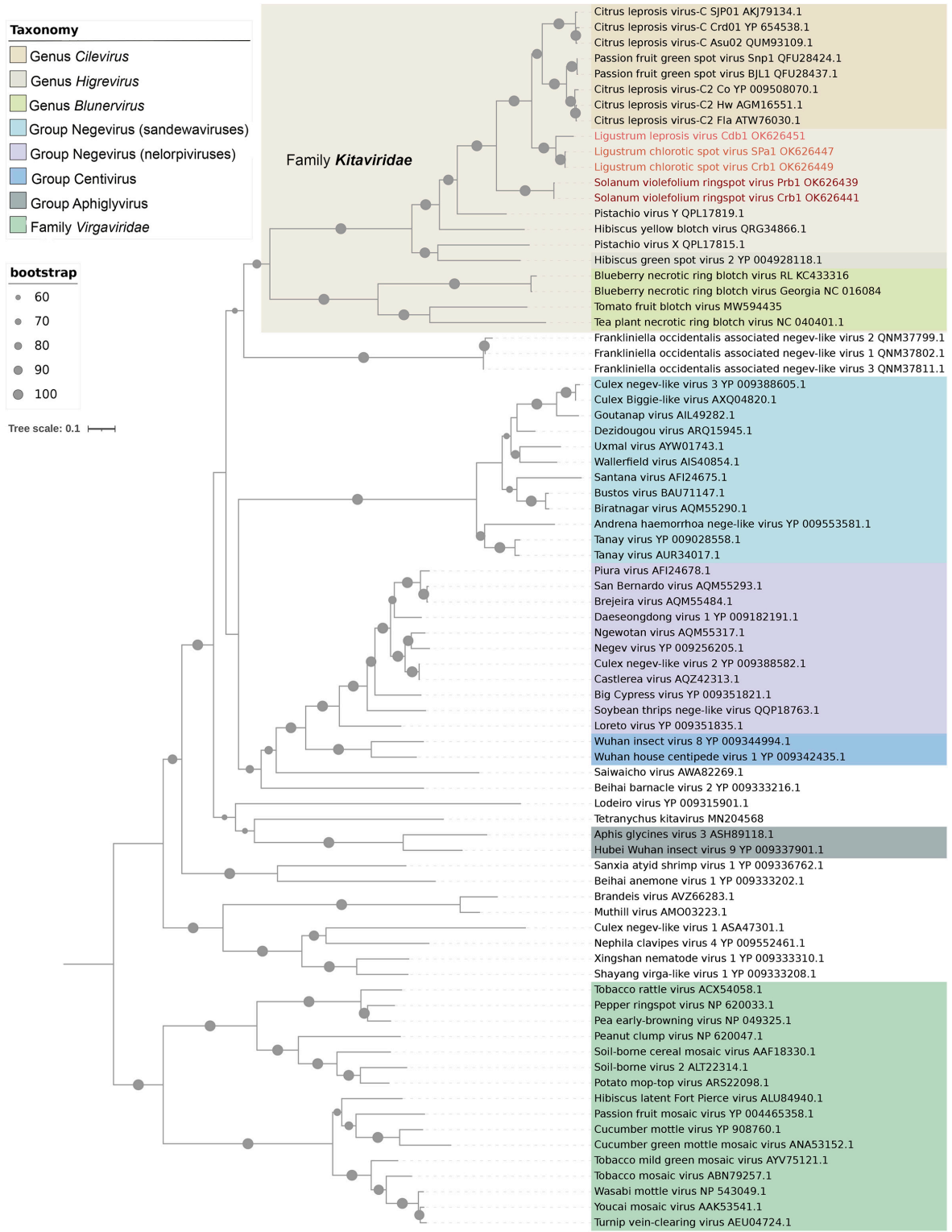


FIGURE 5 | Phylogenetic reconstruction for viruses of the family *Kitaviridae*. Isolates of Solanum violifolium ringspot virus, Ligustrum chlorotic spot virus, and Ligustrum leprosis virus are highlighted in different red tones. The maximum-likelihood phylogenetic tree is based on the deduced amino acid sequences of the RNA-dependent RNA polymerase. The tree was rooted using viruses of the family *Virgaviridae* as an external group. Phylogenetic informative regions of the multiple sequence alignment included 546 residues that were selected using BMGE software (Criscuolo and Gribaldo, 2010) and its evolutionary history was inferred based on the model LG+F+I+G4 (Le and Gascuel, 2008). The bootstrap support values (1,000 replications) of branches greater than 60% are indicated with solid black circles next to the corresponding nodes. The scale bar specifies the average number of amino acid substitutions per site.

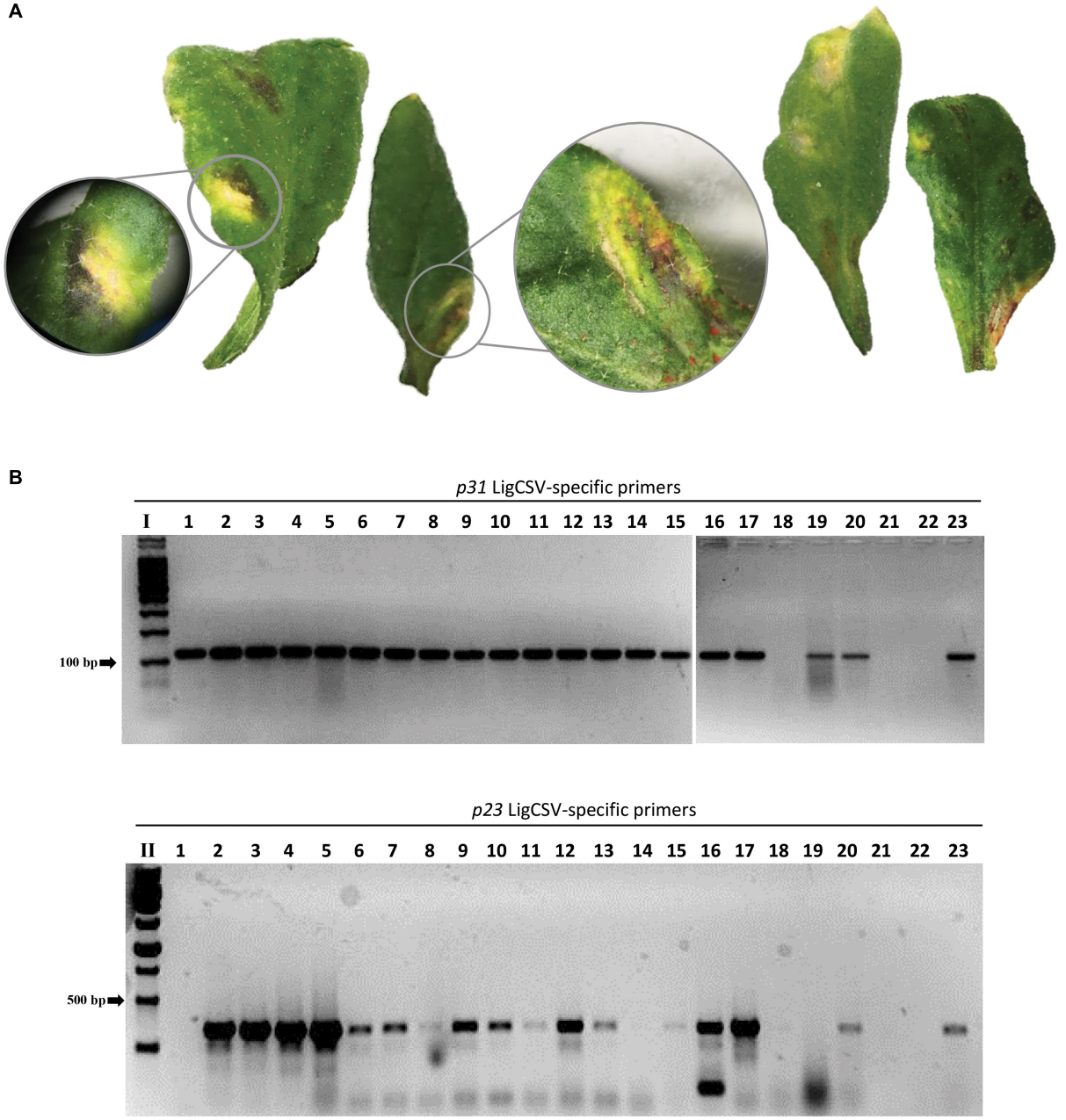


FIGURE 6 | Experimental transmission of Ligustrum chlorotic spot virus (LigCSV) to healthy *Arabidopsis thaliana* plants using viruliferous *Brevipalpus papayensis* mites. **(A)** Arabidopsis leaves showing chlorotic and necrotic lesions after 7 days of infestation with *Brevipalpus papayensis* found in symptomatic *Ligustrum japonicum* plants infected with LigCSV, collected in São Paulo, SP, Brazil. **(B)** 1% agarose gel electrophoresis of RT-PCR products for the specific detection of LigCSV RNA1 (primer pair *p31*, expected amplicon size: 123 bp) and LigCSV RNA2 (primer pair *p23*, expected amplicon size: 420 bp) of LigCSV. I: Molecular weight marker, 100 bp DNA Ladder (Promega, Madison, WI, United States); II: Molecular weight marker, 1 kb DNA Ladder (Promega). Lanes 1–20: Selected symptomatic individual lesions from mite-infested arabidopsis plants; lane 21: Reverse-transcription reaction negative control; lane 22: PCR negative control; and lane 23: LigCSV-infected *Ligustrum japonicum* plant, positive control of the RT-PCR tests.

transverse folds between *d1-d1* to *e1-e1*, with transverse folds posteriorly on *B. tucuman*; (v) ventral plate with small rounded cells becoming fused centrally on *B. obovatus*, bands of mixed orientations on *B. papayensis*, and irregularly shaped cells on *B. tucuman* (data not shown); (vi) spermathecal apparatus ending in small tick vesicle covered on its entire surface with small projection on *B. obovatus*, big spherical vesicle with projection all around the surface on *B. papayensis*, and rounded vesicle similar in size with *B. papayensis* covered with tiny projections on *B. tucuman* (data not shown).

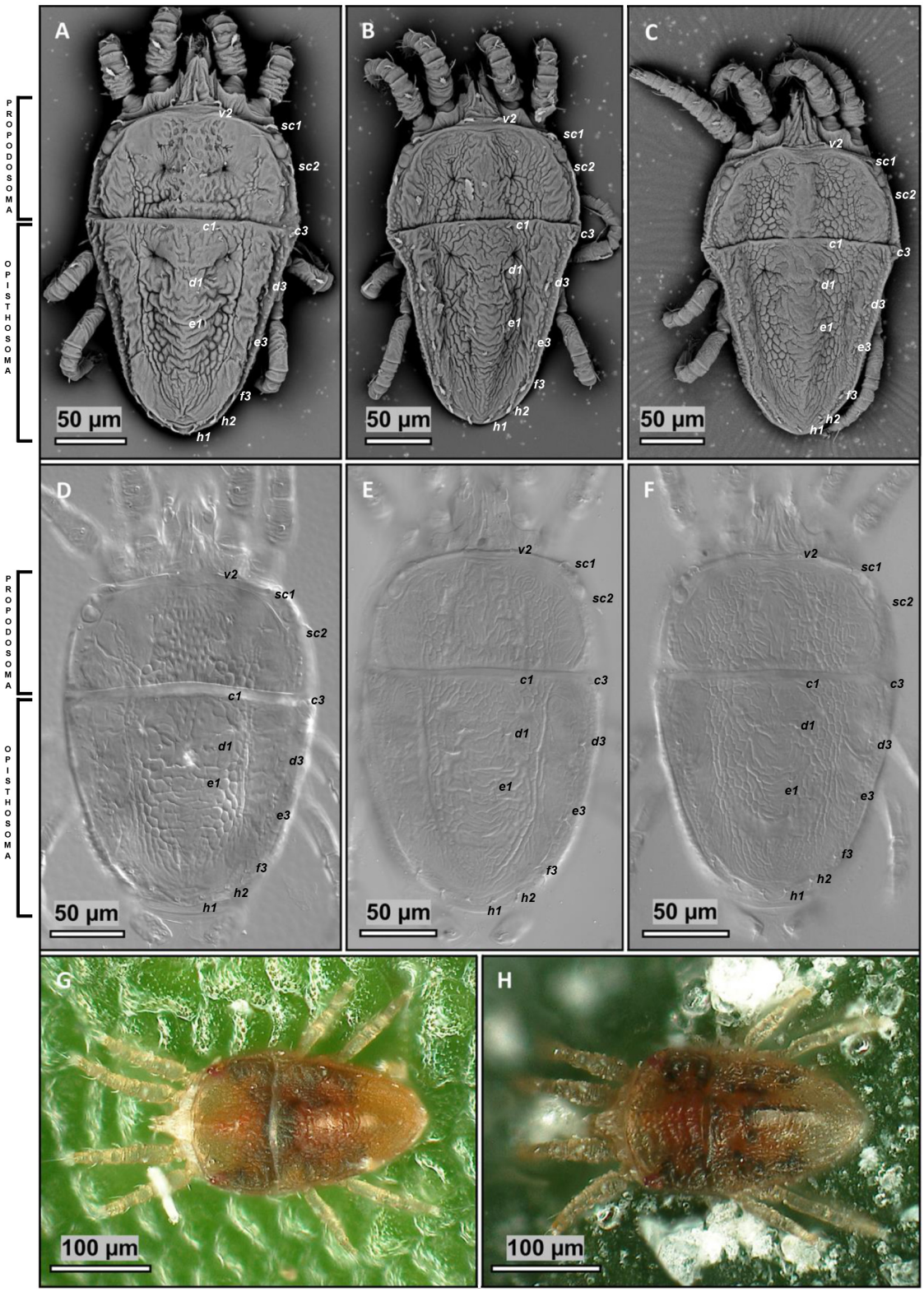
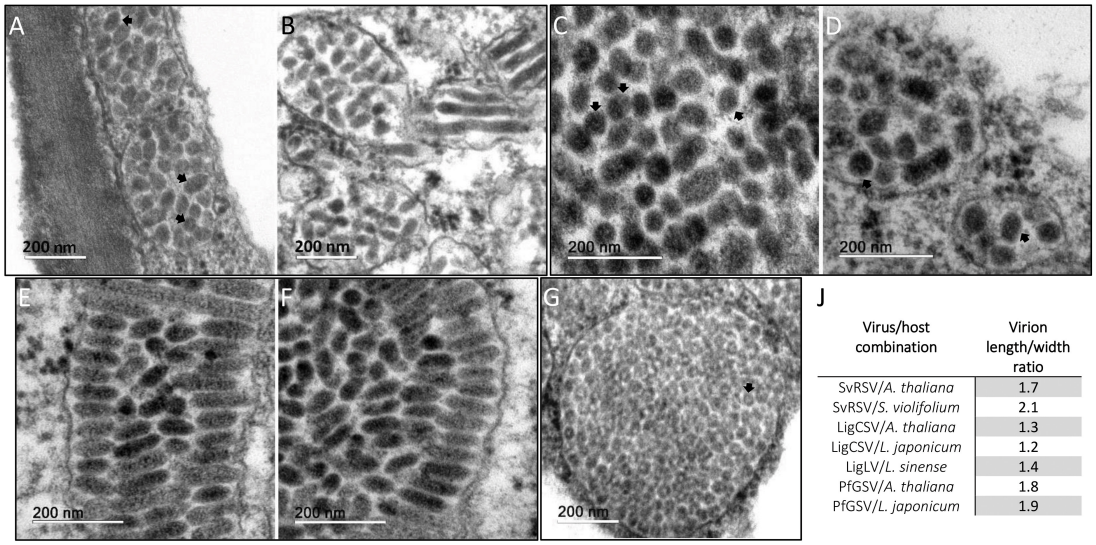
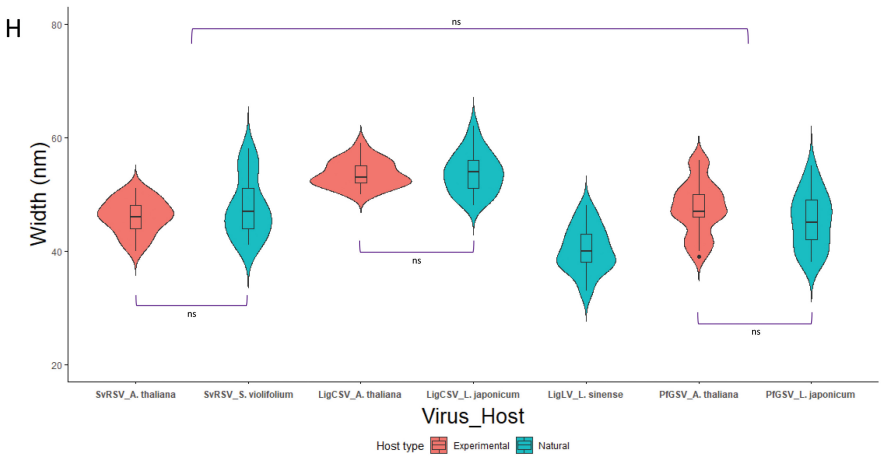


FIGURE 7 | Morphoanatomical identification of *Brevipalpus* mites. **(A,B,H)** *Brevipalpus papayensis* mite collected from *Ligustrum* spp. plants, samples Crb1 and SPa1, **(B,E)** *Brevipalpus tucuman* collected from *Ligustrum sinense* plant, sample Cdb1. **(C**,**F**,**G)** *Brevipalpus obovatus* gathered from *Solanum violifolium* plant, sample Prb1. **(A–C)** Scanning electron microscopy (SEM) micrographs of the cuticle of the mite dorsum. **(D–F)** Differential interference contrast (DIC) micrographs of dorsal view of the reticulation pattern. **(G,H)** Stereoscopic light photographs.



J

Virus/host combination	Virion length/width ratio
SvRSV/*A. thaliana*	1.7
SvRSV/*S. violifolium*	2.1
LigCSV/*A. thaliana*	1.3
LigCSV/*L. japonicum*	1.2
LigLV/*L. sinense*	1.4
PfGSV/*A. thaliana*	1.8
PfGSV/*L. japonicum*	1.9



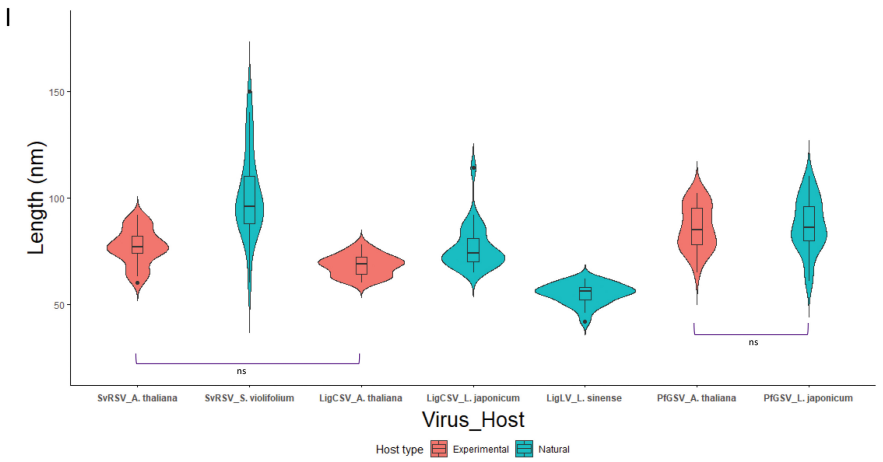


FIGURE 8 | Size and morphology of virions from SvRSV, LigCSV, and LigLV. Microphotographs of ultrathin sections of plants infected by SvRSV_Prb1, LigCSV_SPa1, and LigLV_Cdb1. **(A,B)** SvRSV in *Arabidopsis thaliana* and *Solanum violifolium* plants, respectively. **(C,D)** LigCSV in *A. thaliana* and *Ligustrum japonicum* plants. **(E,F)** PfGSV in *A. thaliana* and *L. japonicum* plants. **(G)** LigLV in *L. sinense* plant. **(H,I)** Violin plots represent the width and length of 25 viral particles measured in at least three micrographs of each virus/host combination. Values of each group were compared using the Wilcoxon signed-rank test. Horizontal brackets depict the results of the statistical test. ns: non-significant difference. In all other comparisons, significant differences were observed ($p < 0.05$). A detailed description of statistical analysis results is described in **Supplementary Tables 4, 5**. **(J)** Length/width ratio of virions from SvRSV, LigCSV, LigLV, and PfGSV in natural and experimental plant hosts.

Solanum Violifolium Ringspot Virus, Ligustrum Chlorotic Spot Virus, and Ligustrum Leprosis Virus Show Bacilliform, Enveloped Virion Particles of Different Sizes

Preliminary examination by TEM of ultrathin sections of plant tissues revealed differences in width and length among particles from SvRSV, LigCSV, and LigLV (**Figure 2**). To get insight into virion dimensions, 25 particles of the three viruses were measured from TEM microphotographs taken at the same magnification (200,000×) (**Figures 8A–G**). The diameter and length of virions of SvRSV and LigCSV were assessed in particles detected in both their natural hosts and arabidopsis plants. For the sake of comparisons, the same evaluations were carried out with virions of the typical cilevirus PfGSV in ultrathin sections of arabidopsis and *L. japonicum* plants. In the case of LigLV_Cdb1, analyses were circumscribed to its natural host, *L. sinense*.

Regardless of the analyzed host, virions from SvRSV and PfGSV had a similar width, approximately 45–46 nm ($p < 0.05$), but they were wider than those from LigLV (40.3 ± 4.0 nm) and narrower than particles from LigCSV in arabidopsis (53.6 ± 2.3 nm) and *L. japonicum* (53.8 ± 3.6 nm) ($p < 0.05$) (**Figure 8H** and **Supplementary Table 4**). In mean values, virions of SvRSV and LigCSV observed in arabidopsis plants were shorter than those observed in their corresponding natural hosts (**Figure 8I**). Although uncommon, virus particles of up to 120–150 nm long were observed in *S. violifolium* and *L. japonicum* plants infected by SvRSV and LigCSV, respectively (**Figures 8B,D**), and their mean length values were 101.6 ± 20.7 nm and 65.9 ± 10.6 nm, respectively. With a mean length of 55.0 ± 4.5 nm, virions of LigLV were the shortest particles evaluated in this study (**Supplementary Table 5**). Virions of PfGSV infecting both *L. japonicum* and arabidopsis had the same length, approximately 86–87 nm (**Figures 8E,F**).

Detailed analyses of ultrathin sections of the infected plants also revealed the presence of an outer membranous layer surrounding the particles of the three studied viruses (**Figures 8A,C,D,G**). These structures were best distinguished during the observation directly in the TEM.

Detection of Mixed Infection of Solanum Violifolium Ringspot Virus and Ligustrum Chlorotic Spot Virus

Both HTS and RT-PCR tests revealed the presence of mixed infection by SvRSV and LigCSV in the sample Crb1. Detailed analysis of the relative abundance of reads corresponding to these two viruses in the HTS library suggested that SvRSV viral load was greater than that of LigCSV (**Table 2**). In addition, in the case of LigCSV, RNA1-derived reads were almost sixfold more abundant than those from its RNA2. To best characterize the distribution of viral genomic molecules in the sample Crb1, the presence of RNA1 and RNA2 molecules of these two viruses were tested in individual lesions by RT-PCR.

A total of 54 randomly selected lesions including spots of different sizes and morphologies, e.g., absence or presence of necrotic tissues, yellowish blotches, or markedly chlorotic spots, were independently analyzed (**Figure 9A**). Viral RNAs were detected in 32 (59%) of the lesions. Single infection by SvRSV was most common, with SvRSV RNA1 + RNA2 detected in 18 of the lesions. Conversely, LigCSV RNA1 + RNA2 alone were only detected in 3 lesions (**Figure 9B** and **Supplementary Figure 2**). In some lesions, only one RNA molecule, e.g., RNA1 or RNA2 of SvRSV, or RNA2 of LigCSV, could be detected, but RNA1 of LigCSV was never detected alone. The heterologous combination SvRSV RNA1 + LigCSV RNA2 and the mixed infection SvRSV with the RNA2 of LigCSV were detected in one and two lesions, respectively. Any association between RNA composition and morphology of the lesions could not be observed.

Non-target Viruses Detected in the High-Throughput Sequencing Libraries

The near-complete sequences of two new narna-like viruses were recovered from samples Prb1 (contigs of 1,670 nts) and Crb1 (contig of 2,462 nts). In addition, the sequence of a novel mitovirus was identified in the sample SPa1 (contig of 2,969 nts). These viruses might have originated from fungi present in the samples. Since non-kitaviruses were outside the scope of this study, further examination was not performed as part of this work.

DISCUSSION

In this study, we revisited some plant viral diseases in which the molecular identity of their causal agents remained partially characterized or completely unknown (Vergani, 1942; Lima et al., 1991; Ferreira et al., 2007; Kitajima et al., 2010). Leaf samples from four plant species, collected in the southern and southeastern regions of Brazil and Pampas in Argentina, showed chlorotic and/or necrotic spots comparable to those generally found in plants affected by *Brevipalpus*-transmitted viruses (Freitas-Astúa et al., 2018). In addition, flat mites of the genus *Brevipalpus* were recovered from the collected plant samples and in three of these plants, we preliminarily detected fragments of the tentative cilevirus SvRSV by RT-PCR tests using specific primers (Ferreira et al., 2007).

Based on a combined strategy of HTS and conventional Sanger sequencing, analyses of the plant viromes enabled us to recover the genomic sequences of several isolates of three viruses. In agreement with the preliminary RT-PCR tests, two isolates corresponded with SvRSV, while the other ones were assigned to two novel kita-like viruses that we identified as LigCSV and LigLV. Genomic data indicated that the highest nt sequence identity of these viruses always corresponded to genes also found in either one of the other two viruses described in this study or members of the genus *Cilevirus*. Still, the values of nucleotide sequence identities obtained in the comparisons of every viral ORF were below 76%.

Besides the relatively low identity values with other kitaviruses, SvRSV, LigCSV, and LigLV display a distinguishing genomic organization. In comparison with typical cileviruses, the RNA2 segments of SvRSV, LigCSV, and LigLV are shorter, lacking a

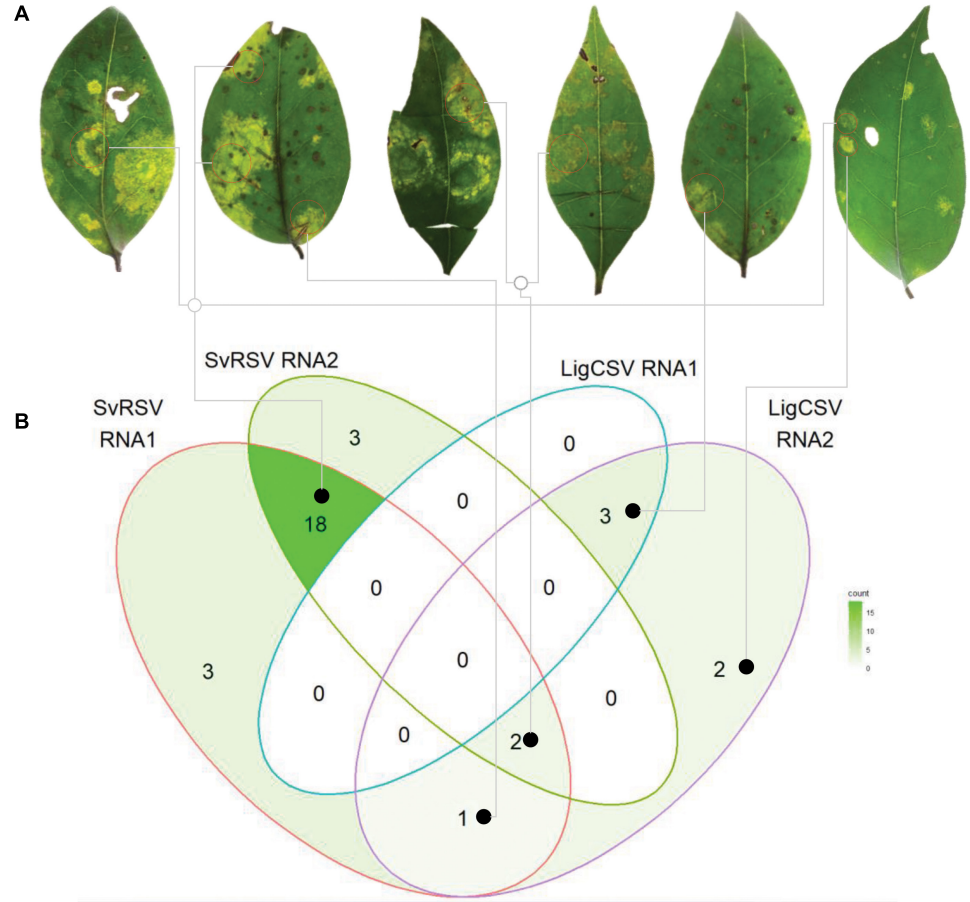


FIGURE 9 | Distribution of SvRSV and LigCSV in isolated lesions found in leaves of *Ligustrum lucidum* plants of the sample Crb1. **(A)** Leaf discs comprising viral symptoms were collected and independently processed for the specific detection of RNA1 and RNA2 of each virus. **(B)** The Venn diagram represents the general distribution of the detected viral genome RNA molecules across the 54 analyzed lesions. Gray lines link lesions to the viral RNA detection indicated in the diagram.

region of approximately 1.5 kb at the 5′-end of the molecule. In cileviruses, the genomic region upstream of the ORF *p61* includes the ORF *p15* and other small ORFs encoding putative proteins with predicted TM helixes (Ramos-González et al., 2021). In PfGSV, the region also harbors orphan ORFs encoding predicted proteins > 11 kDa of unknown functions (Ramos-González et al., 2020). An RNA2 genomic molecule resembling a 5′-end truncated version of the RNA2 of typical cileviruses was also described in the cile-like virus HYBV (Olmedo-Velarde et al., 2021). However, the information in the RNA2 of HYBV is not completely equivalent to that in the RNA2 of SvRSV, LigCSV, and LigLV. While the RNA2 of these viruses comprise genes that are likely orthologs of the ORFs *p61*, *p32,* and *p24* found in typical cileviruses, HYBV RNA2 has an additional ORF at its 3′-end, an ortholog of *p29.* In typical cileviruses, the ORF *p29* is located at the 3′-end of the RNA1.

HYBV has been suggested as an intermediary link between members of the genera *Cilevirus* and *Higrevirus* (Olmedo-Velarde et al., 2021). While the information encoded by RNA2 of HYBV is more similar to that encoded by RNA2 of typical cileviruses, its RNA1 looks like that of the higrevirus HGSV2. If the same logic is extended to our current analysis, the genomes of SvRSV, LigCSV, and LigLV could be regarded as intermediary arrays between those displayed by HYBV and typical cileviruses. Evolutionarily, SvRSV, LigCSV, and LigLV could be considered descendants of a common ancestor of HYBV and typical cileviruses. This transitional position of SvRSV, LigCSV, and LigLV is also supported by the phylogenetic reconstruction based on the RdRp protein presented in this study. Branches encompassing these viruses are flanked by those of HYBV and Pistachio virus Y (PisVY), both placed in a relatively more basal position, and those of the typical cileviruses. PisVY is a partially characterized virus found in pistachio plants (*Pistacia vera*) of the cultivar Ohadi collected in Rafasnjan, Iran. Since the genome organization of PisVY is similar to that in SvRSV, LigCSV, and LigLV, these four viruses may represent new sublineages of kitaviruses. The relatively large genetic distances among viruses of these sublineages and other members of the family *Kitaviridae* suggest the existence of phylogenetic gaps and reinforce the hypothesis that a great diversity of kitaviruses might remain unidentified. However, while further sampling is likely needed, what is relevant after

the inclusion of these viruses in the phylogenetic analysis based on the RdRp protein is their contribution to conciliate the evolutionary history among members of the three accepted genera of kitaviruses. In the RdRp phylogenetic tree presented in this study, kitaviruses are distributed by genera from a single node describing a monophyletic clade supported by more than 80% of the trees sampled in the bootstrap analysis. It should be also noted, however, that in the phylogenetic reconstruction using the P24 protein, blunerviruses clustered in a branch also including nelorpiviruses, centiviruses, aphiglyviruses, and other arthropod-infecting viruses, separately from that with the definitive and tentative members of the genera *Cilevirus* and *Higrevirus*. Subdivision of kitaviruses in two sister branches is likely a consequence of the chimeric origin of genomic segments of kitaviruses, as previously described for blunerviruses (Quito-Avila et al., 2013; Morozov et al., 2020).

It has been theorized that the origin of the 5′-end of the RNA2 in typical cileviruses could have involved its acquisition by horizontal gene transfer from a heterologous source (Ramos-González et al., 2021). The finding of SvRSV, LigCSV, and LigLV could add novel hints for the reconstruction of a plausible evolutionary pathway of cileviruses. While RNA1 molecules of SvRSV, LigCSV, LigLV, and typical cileviruses have similar compositions and likely share an ancestral root, the "truncated" RNA2 of an ancestor of SvRSV, LigCSV, and LigLV could have actively participated as a receptor of the 5′-end region of the RNA2 of current cileviruses.

Virion-like particles identified in plants infected with the blunervirus TPNRBV and the cile-like HYBV have spherical shapes with approximately 85 nm and 50–60 nm in diameter, respectively (Hao et al., 2018; Olmedo-Velarde et al., 2021). Typical cileviruses show enveloped bacilliform virions having 40–70 nm wide and 110–120 nm long (Kitajima et al., 2003; Roy et al., 2013). Detailed analysis of ultrathin sections from infected tissues carried out in this study demonstrated that virions of SvRSV, LigCSV, and LigLV are enveloped bacilliform particles with significant differences in their sizes. Particles of SvRSV and LigLV are narrower than those of LigCSV, and particles of LigLV are shorter than those from SvRSV and LigCSV. Hence, with a lesser length/width ratio (**Figure 8J**), LigLV and LigCSV particles seem more spherical than those of SvRSV. Particles of SvRSV are shorter than those of PfGSV but they show the same width. Taken together, kitavirus virions seem to represent a spectrum of morphologies ranging from quasi-spherical to typical short bacilliform particles. It should be mentioned, however, that even though a large number of virions were analyzed in more than one host, our work is not free of technical constraints which may produce biased results. Measurement of sizes of enveloped virions in fixed tissues may add novel challenges since these particles can be expanded or retracted depending on cellular physiological conditions in the sample at the time of collection, and/or during its conservation and processing for transmission electron microscopy.

The absence of systemic movement is a central topic in the biology of kitaviruses, therefore, successful viral spread even from two distant points in the same plant leaf depends on their arthropod vectors. Evidence indicates that the transmission of the higrevirus HGSV2 and the cile-like virus HYBV is mediated by mites of the genus *Brevipalpus* (Melzer et al., 2012; Olmedo-Velarde et al., 2021), but the direct involvement of these mites in the transmission of kitaviruses has only been verified for typical cileviruses and SvRSV (Ferreira et al., 2007). In this study, we have demonstrated that mites of the species *B. papayensis* collected from infected *L. japonicum* plants can transmit LigCSV to arabidopsis plants. Besides, morphoanatomical characterization of mites collected in *S. violifolium* plants infected with SvRSV indicated the presence of *B. obovatus*, which is in agreement with previous reports of transmission of this virus by this mite species (Ferreira et al., 2007). Finally, the identification of *B. tucuman* in the *L. sinense* plant infected by LigLV would represent the first evidence of the viral vector activity of this mite species, which, however, needs further confirmatory experiments. More than 300 species of the genus *Brevipalpus* have been described but the number of species with virus vector activity is restricted to less than a dozen (Kitajima and Alberti, 2014; de Lillo et al., 2021). Among them, *B. yothersi* has been identified as the main vector of the typical cileviruses CiLV-C, CiLV-C2, and PfGSV (Roy et al., 2015; Ramos-González et al., 2016; de Lillo et al., 2021). CiLV-C and PfGSV can be also vectored by *B. papayensis* (Nunes et al., 2018; Tassi, 2018), which according to the results obtained in this study can also vector LigCSV. Furthermore, since only *B. papayensis* mites were identified in the sample Crb1, infected by LigCSV and SvRSV, our results suggest the transmission of SvRSV by these mites. However, this finding may also emphasize that a wider number of mites from the sample Crb1 need to be analyzed to detect any *B. obovatus* specimen present, which may be in lower frequency. Anyway, the possibility that SvRSV can be also transmitted by *B. papayensis* cannot be ruled out. Both *B. yothersi* and *B. papayensis* have been demonstrated as vectors of more than one species of cilevirus and one species of the genus *Dichorhavirus*, family *Rhabdoviridae* (Chabi-Jesus et al., 2018; de Lillo et al., 2021).

In addition to the mites involved in viral transmission, the examination of *L. lucidum* sample Crb1, where SvRSV and LigCSV were detected, also suggested some clues about kitavirus ecology that will require further study before conclusions can be drawn. Particularly, the use of more sensitive methods would be necessary to detect the presence, if any, of some genomic segments, perhaps at very low concentrations in some lesions, and undetected using conventional RT-PCR tests. Nonetheless, the analysis of the isolated lesions showed (*i*) the predominance of infection loci where only SvRSV could be detected, (*ii*) the lack of lesions with the four genomic RNAs, and (*iii*) the existence, although in low frequency, of heterodox mixes of genomic RNAs, e.g., RNA1 and RNA2 of SvRSV + RNA2 of LigCSV. These findings could be only a reflection of the best transmission efficiency of SvRSV, the highest loads of SvRSV in *L. lucidum*, or the widest distribution and prevalence of SvRSV in that sample. Conversely, although under different experimental conditions, the isolate SPa1 of LigCSV seemed to be efficiently transmitted by *B. papayensis* to arabidopsis. In the experiment with the isolate SPa1, the virus was acquired from a different host, i.e., *L. japonicum*, and particularly, it was in a single infection. Altogether, the snapshot obtained from the sample Crb1

may suggest the existence of undescribed interspecific interaction between kitaviruses in both plants and mites during mixed infections including, perhaps, genomic rearrangements. Mixed infections between a cilevirus and a dichorhavirus infecting citrus have been reported (Roy et al., 2014) but among kitaviruses, only mixed infections of two strains of the same viral species have been described (Chabi-Jesus et al., 2021).

The molecular and biological characterization of SvRSV, LigCSV, and LigLV has also exposed some data and experimental results that might be used as the basis for future analyses. First, the existence of extra-large virions detected in the ultrathin sections of plant tissues infected by SvRSV and LigCSV. In SvRSV, the existence of these particles seems to be an intrinsic feature, since they were also described in previous studies (Ferreira et al., 2007). Indeed, morphological differences between virions of SvRSV and other cileviruses were suggested as evidence supporting its classification in a separated taxon (Ferreira et al., 2007). Whether the atypical virions correspond to a viral species not detected by our virome analysis or they result from defective maturation processes, and whether they have any role in the replication process, are aspects worth studying. Second, a thorough examination of orthologous genes among HYBV, PisVY, typical cileviruses, and the isolates of the three viruses described in this study reveals a common trend to the presence of larger *p29*, *p32*, and *p24* orthologous ORFs in PisVY or HYBV, mid-sized versions of these ORFs in SvRSV, LigCSV or LigLV, and short versions in typical cileviruses (**Supplementary Table 6**). Finally, the high frequency of small ORFs found in the genome of these viruses and particularly, the large proportion of them that potentially encode peptides with predicted TM helices. Since genome size reduction and the presence of overlapping genes and TM small proteins might be linked and represent central elements to understand processes such as gene expression regulation and virus-host interaction (Belshaw et al., 2007; Wu et al., 2020), we speculate they could be valuable attributes to study the evolutionary process of viruses of the family *Kitaviridae*.

Currently, viruses of the family *Kitaviridae* are assigned to one of three genera according to the number of genomic segments and their genomic organization. Within the genus *Cilevirus*, demarcation of new species ponders the extent of the serological relationship, proteome with less than 85% amino acid sequence identity with other members, natural and experimental host range reactions, and vector species and transmission (Quito-Avila et al., 2021). Based on their genome organization, it seems evident that SvRSV, LigCSV, and LigLV best fit in the genus *Cilevirus*, but considering the relatively low identity between their proteins and the distinct virion morphology of LigCSV and LigLV, it is tentative to speculate that these viruses may require a new taxonomic accommodation. While expecting for the analyses of new kitaviruses that will likely allow getting a better insight into the phylogenetic relationship of kitaviruses, we propose that the viruses described in this study be considered tentative members of the genus *Cilevirus*.

DATA AVAILABILITY STATEMENT

The virus sequences derived from this study can be found in GenBank under the accession numbers: SvRSV_Prb1 (OK626439 and OK626440), SvRSV_Crb1 (OK626441 and OK626442), LigCSV_SPa1 (OK626447 and OK626448), LigCSV_Crb1 (OK626449 and OK626450), and LigLV_Cdb1 (OK626451 and OK626452).

AUTHOR CONTRIBUTIONS

PR-G, CC-J, and AT: conceptualization. PR-G, CC-J, AT, RFC, and EK: formal analysis. RH, EK, and JF-A: funding acquisition. PR-G, CC-J, AT, RFC, CN, RH, EK, and JF-A: investigation. PR-G, CC-J, AT, RFC, RH, and EK: methodology. PR-G, EK, and JF-A: supervision. PR-G: writing—original draft. PR-G, CC-J, AT, RFC, EK, and JF-A: writing—review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

The authors were grateful to CAPES, CNPq, and FAPESP for scholarships, research fellowships, and research grants associated with this work (CAPES PNPD20132154-33141010001P4-PNPD-IBSP, 88882.157041/2017-01; CNPq 312528/2020-5; FAPESP 2017/50222-0, 2017/50334-3, 2018/12252-8, 2019/25078-9, and 2020/15413-2).

ACKNOWLEDGMENTS

We thank the reviewers for their careful reading of the manuscript and useful and constructive comments on our work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.836076/full#supplementary-material

REFERENCES

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46, W537–W544. doi: 10.1093/nar/gky379

Alberti, G., and Kitajima, E. W. (2014). Anatomy and fine structure of brevipalpus mites (*Tenuipalpidae*) – economically important plant-virus vectors. *Zoologica* 160, 1–192.

Arena, G. D., Ramos-González, P. L., Falk, B. W., Casteel, C. L., Freitas-Astúa, J., and Machado, M. A. (2020). Plant immune system activation upon citrus leprosis virus c infection is mimicked by the ectopic expression of the p61 viral protein. *Front. Plant Sci.* 11:1188. doi: 10.3389/fpls.2020.01188

Arena, G. D., Ramos-González, P. L., Nunes, M. A., Alves, M. R., Camargo, L. E. A., Kitajima, E. W., et al. (2016). Citrus leprosis virus C infection results in hypersensitive-like response, suppression of the JA/ET plant defense pathway and promotion of the colonization of its mite vector. *Front. Plant Sci.* 7:1757. doi: 10.3389/FPLS.2016.01757

Arena, G. D., Ramos-González, P. L., Nunes, M. A., Jesus, C. C., Calegario, R. F., Kitajima, E. W., et al. (2017). *Arabidopsis thaliana* as a model host for *Brevipalpus* mite-transmitted viruses. *Sci. Agric.* 74, 85–89. doi: 10.1590/1678-992x-2015-0380

Armenteros, J. J. A., Sønderby, C. K., Sønderby, S. K., Nielsen, H., and Winther, O. (2017). DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics* 33, 3387–3395. doi: 10.1093/bioinformatics/btx431

Armenteros, J. J. A., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. doi: 10.1038/s41587-019-0036-z

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Bastianel, M., Novelli, V. M., Kitajima, E. W., Kubo, K. S., Bassanezi, R. B., Machado, M. A., et al. (2010). Citrus leprosis: centennial of an unusual mite–virus pathosystem. *Plant Dis.* 94, 284–292. doi: 10.1094/PDIS-94-3-0284

Beard, J. J., Ochoa, R., Braswell, W. E., and Bauchan, G. R. (2015). *Brevipalpus phoenicis* (Geijskes) species complex (Acari: Tenuipalpidae)-a closer look. *Zootaxa* 3944, 1–67. doi: 10.11646/zootaxa.3944.1.1

Belshaw, R., Pybus, O. G., and Rambaut, A. (2007). The evolution of genome compression and genomic novelty in RNA viruses. *Genome Res.* 17, 1496–1504. doi: 10.1101/gr.6305707

Bitancourt, A. (1940). Citrus leprosis. (a leprose dos *Citrus*.). *Biologico* 6, 39–45.

Bitancourt, A. (1955). Estudos sobre a leprose dos citros. *Arq. Inst. Biol.* 22, 161–231.

Bushnell, B. (2014). *BBMap?: A Fast, Accurate, Splice-Aware Aligner: Technical Report*. Berkeley, CA: Lawrence Berkeley National Lab. (LBNL), 3–5.

Chabi-Jesus, C., Ramos-González, P. L., Postclam-Barro, M., Fontenele, R. S., Harakava, R., Bassanezi, R. B., et al. (2021). Molecular epidemiology of citrus leprosis virus c: a new viral lineage and phylodynamic of the main viral subpopulations in the americas. *Front. Microbiol.* 12:986. doi: 10.3389/fmicb.2021.641252

Chabi-Jesus, C., Ramos-González, P. L., Tassi, A. D., Guerra-Peraza, O., Kitajima, E. W., Harakava, R., et al. (2018). Identification and characterization of citrus chlorotic spot virus, a new dichorhavirus associated with citrus leprosis-like symptoms. *Plant Dis.* 102, 1588–1598. doi: 10.1094/PDIS-09-17-1425-RE

Chiapello, M., Bosco, L., Ciuffo, M., Ottati, S., Salem, N., Rosa, C., et al. (2021). Complexity and local specificity of the virome associated with tospovirus-transmitting thrips species. *J. Virol.* 95:e0059721. doi: 10.1128/JVI.00597-21

Collins, T. J. (2007). ImageJ for microscopy. *Biotechniques* 43, S25–S30. doi: 10.2144/000112517

Criscuolo, A., and Gribaldo, S. (2010). BMGE (block mapping and gathering with entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* 10:210. doi: 10.1186/1471-2148-10-210

de Lillo, E., Freitas-Astúa, J., Kitajima, E. W., Ramos-González, P. L., Simoni, S., Tassi, A. D., et al. (2021). Phytophagous mites transmitting plant viruses: update and perspectives. *Entomol. Gen.* 41, 439–462. doi: 10.1127/ENTOMOLOGIA/2021/1283

Dreher, T. W. (1999). Functions of the 3'-untranslated regions of positive strand RNA viral genomes. *Annu. Rev. Phytopathol.* 37, 151–174. doi: 10.1146/ANNUREV.PHYTO.37.1.151

Ferreira, P., de, T. O., Locali-Fabris, E. C., Freitas-Astúa, J., Antonioli-Luizon, R., Gomes, R. T., et al. (2007). Caracterização de um vírus baciliforme isolado de Solanum violaefolium transmitido pelos ácaros *Brevipalpus phoenicis* e *Brevipalpus obovatus* (Acari: Tenuipalpidae). *Summa Phytopathol.* 33, 264–269. doi: 10.1590/S0100-54052007000300009

Freitas-Astúa, J., Ramos-González, P. L., Arena, G. D., Tassi, A. D., and Kitajima, E. W. (2018). Brevipalpus-transmitted viruses: parallelism beyond a common vector or convergent evolution of distantly related pathogens? *Curr. Opin. Virol.* 33, 66–73. doi: 10.1016/J.COVIRO.2018.07.010

Frezzi, M. S. (1940). "La lepra explosiva del naranjo-Investigaciones realizadas por el laboratorio de patología de Bella Vista (Corrientes)," in *Boletin Frutas y Hortalizas*, ed. C. Laboratorio de Fitopatología (Buenos Aires: Ministerio de la Agricultura, Dirección de Sanidad Vegetal), 1–16.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi: 10.1038/nprot.2013.084

Hao, X., Zhang, W., Zhao, F., Liu, Y., Qian, W., Wang, Y., et al. (2018). Discovery of plant viruses from tea plant (*Camellia sinensis* (L.) O. Kuntze) by metagenomic sequencing. *Front. Microbiol.* 9:2175. doi: 10.3389/fmicb.2018.02175

Hatcher, E., Zhdanov, S., Bao, Y., Blinkova, O., Nawrocki, E., Ostapchuck, Y., et al. (2017). Virus variation resource – improved response to emergent viral outbreaks. *Nucleic Acids Res.* 45, D482–D490. doi: 10.1093/NAR/GKW1065

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kitajima, E. W., and Alberti, G. (2014). Anatomy and fine structure of *Bevipalpus* mites (Tenuipalpidae) – economically important plant virus vectors. Part 7. Ultrastructural detection of cytoplasmic and nuclear types of *Brevipalpus* transmitted viruses. *Zoologica* 160, 174–192.

Kitajima, E. W., and Nome, C. F. (1999). "Microscopia eletrônica em virologia vegetal," in *Métodos Para Detectar Patógenos Sistémicos*, eds D. M. Docampo and S. L. Lenardón (Córdoba: IFFIVE/INTA-JICA), 59–87.

Kitajima, E. W., Chagas, C. M., and Rodrigues, J. C. V. (2003). Brevipalpus-transmitted plant virus and virus-like diseases: cytopathology and some recent cases. *Exp. Appl. Acarol* 30, 135–160. doi: 10.1023/b:appa.0000006546.55305.e3

Kitajima, E. W., Rodrigues, J. C. V., and Freitas-Astua, J. (2010). An annotated list of ornamentals naturally found infected by *Brevipalpus* mite-transmitted viruses. *Sci. Agric.* 67, 348–371. doi: 10.1590/S0103-90162010000300014

Kondo, H., Fujita, M., Hisano, H., Hyodo, K., Andika, I. B., and Suzuki, N. (2020). Virome analysis of aphid populations that infest the barley field: the discovery of two novel groups of nege/kita-like viruses and other novel RNA viruses. *Front. Microbiol.* 11:509. doi: 10.3389/fmicb.2020.00509

Le, S. Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix. *Mol. Biol. Evol.* 25, 1307–1320. doi: 10.1093/molbev/msn067

Lemoine, F., Correia, D., Lefort, V., Doppelt-Azeroual, O., Mareuil, F., Cohen-Boulakia, S., et al. (2019). NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Res.* 47, W260–W265. doi: 10.1093/nar/gkz303

Letunic, I., and Bork, P. (2021). Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–W296. doi: 10.1093/NAR/GKAB301

Lima Neto, V. C., Lima, M. L. R. Z. C., and Sousa, V. B. V. (1994). O agente causador do anel do *Ligustrum*. *Rev. Set. Ciênc. Agrárias* 13, 221–224.

Lima, M. L. R. Z. C., Lima Neto, V. C., and Souza, V. B. V. (1991). The causal agent of the *Ligustrum* ringspot disease. *Phytopathology* 81:1216.

Liu, Y., Zhang, Y., Wang, M., Cheng, A., Yang, Q., Wu, Y., et al. (2020). Structures and functions of the 3' untranslated regions of positive-sense single-stranded RNA viruses infecting humans and animals. *Front. Cell. Infect. Microbiol.* 10:453. doi: 10.3389/FCIMB.2020.00453

Locali-Fabris, E. C., Freitas-Astúa, J., Souza, A. A., Takita, M. A., Astúa-Monge, G., Antonioli-Luizon, R., et al. (2006). Complete nucleotide sequence, genomic organization and phylogenetic analysis of citrus leprosis virus cytoplasmic type. *J. Gen. Virol.* 87, 2721–2729. doi: 10.1099/vir.0.82038-0

Marques, J. P., Kitajima, E., Freitas-Astúa, J., and Apezzato-da-Glória, B. (2010). Comparative morpho-anatomical studies of the lesions causedby citrus leprosis virus on sweet orange. *Ann. Brazilian Acad. Sci.* 82, 501–511. doi: 10.1590/s0001-37652010000200025

Melzer, M. J., Sether, D. M., Borth, W. B., and Hu, J. S. (2012). Characterization of a virus infecting *Citrus volkameriana* with citrus leprosis-like symptoms. *Phytopathology* 102, 122–127. doi: 10.1094/PHYTO-01-11-0013

Morozov, S. Y., Lazareva, E. A., and Solovyev, A. G. (2020). Sequence relationships of RNA helicases and other proteins encoded by *Blunervirus* RNAs highlight recombinant evolutionary origin of kitaviral genomes. *Front. Microbiol.* 11:561092. doi: 10.3389/fmicb.2020.561092

Nunes, M. A., de Carvalho Mineiro, J. L., Rogero, L. A., Ferreira, L. M., Tassi, A. D., Novelli, V. M., et al. (2018). First report of *Brevipalpus papayensis* Baker (Acari:

Tenuipalpidae) as vector of *Coffee ringspot* virus and *Citrus leprosis* virus C. *Plant Dis.* 10, 1046. doi: 10.1094/PDIS-07-17-1000-PDN

Olmedo-Velarde, A., Hu, J., and Melzer, M. J. (2021). A virus infecting hibiscus rosa-sinensis represents an evolutionary link between cileviruses and higreviruses. *Front. Microbiol.* 12:988. doi: 10.3389/fmicb.2021.660237

Procter, J. B., Carstairs, G. M., Soares, B., Mourão, K., Ofoegbu, T. C., Barton, D., et al. (2021). Alignment of biological sequences with jalview. *Methods Mol. Biol.* 2231, 203–224. doi: 10.1007/978-1-0716-1036-7_13

Quito-Avila, D. F., Brannen, P. M., Cline, W. O., Harmon, P. F., and Martin, R. R. (2013). Genetic characterization of *Blueberry necrotic* ring blotch virus, a novel RNA virus with unique genetic features. *J. Gen. Virol.* 94, 1426–1434. doi: 10.1099/vir.0.050393-0

Quito-Avila, D. F., Freitas-Astúa, J., and Melzer, M. J. (2021). " Bluner-, cile-, and higreviruses (Kitaviridae)," in *Encyclopedia of Virology*, eds D. Bamford and M. Zuckerman (Amsterdam: Elsevier), 247–251. doi: 10.1016/B978-0-12-809633-8.21248-X

Ramos-González, P. L., Chabi-Jesus, C., Arena, G. D., Tassi, A. D., Kitajima, E. W., and Freitas-Astúa, J. (2018). Citrus leprosis: a unique multietiologic disease. *Citrus Am.* 1, 4–19. doi: 10.6084/m9.figshare.13513857.v2

Ramos-González, P. L., Chabi-Jesus, C., Guerra-Peraza, O., Breton, M. C., Arena, G. D., Nunes, M. A., et al. (2016). Phylogenetic and molecular variability studies reveal a new genetic clade of *Citrus leprosis* virus C. *Viruses* 8:153. doi: 10.3390/v8060153

Ramos-González, P. L., Chabi-Jesus, C., Guerra-Peraza, O., Tassi, A. D., Kitajima, E. W., Harakava, R., et al. (2017). Citrus leprosis virus N: a new dichorhavirus causing *Citrus Leprosis* disease. *Phytopathology* 107, 963–976. doi: 10.1094/PHYTO-02-17-0042-R

Ramos-González, P. L., dos Santos, G. F., Chabi-Jesus, C., Harakava, R., Kitajima, E. W., and Freitas-Astúa, J. (2020). Passion fruit green spot virus genome harbors a new orphan ORF and highlights the flexibility of the 5'-end of the RNA2 segment across cileviruses. *Front. Microbiol.* 11:206. doi: 10.3389/fmicb.2020.00206

Ramos-González, P. L., Pons, T., Chabi-Jesus, C., Arena, G. D., and Freitas-Astua, J. (2021). Poorly conserved P15 proteins of cileviruses retain elements of common ancestry and putative functionality: a theoretical assessment on the evolution of cilevirus genomes. *Front. Plant Sci.* 12:771983. doi: 10.3389/FPLS.2021.771983

Rasekhian, M., Roohvand, F., Habtemariam, S., Marzbany, M., and Kazemimanesh, M. (2021). The role of 3'UTR of RNA viruses on mRNA stability and translation enhancement. *Mini Rev. Med. Chem.* 21, 2389–2398. doi: 10.2174/1389557521666210217092305

Roy, A., Choudhary, N., Guillermo, L. M., Shao, J., Govindarajulu, A., Achor, D., et al. (2013). A novel virus of the genus *Cilevirus* causing symptoms similar to citrus leprosis. *Phytopathology* 103, 488–500. doi: 10.1094/PHYTO-07-12-0177-R

Roy, A., Hartung, J. S., Schneider, W. L., Shao, J., Leon, M. G., Melzer, M. J., et al. (2015). Role bending: complex relationships between viruses, hosts, and vectors related to citrus leprosis, an emerging disease. *Phytopathology* 105, 872–884. doi: 10.1094/PHYTO-12-14-0375-FI

Roy, A., Leon, M. G., Stone, A. L., Schneider, W. L., Hartung, J., and Brlansky, R. H. (2014). First report of citrus leprosis virus nuclear type in sweet orange in Colombia. *Plant Dis.* 98:1162. doi: 10.1094/PDIS-02-14-0117-PDN

Rstudio_Team (2021). *RStudio | Open Source & Professional Software For Data Science Teams – Rstudio*. Available online at: https://www.rstudio.com/ [Accessed June 9, 2021].

Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 6, 175–182.

Tassi, A. D. (2018). *Diversidade Morfológica E Genética De Diferentes Espécies De Brevipalpus (Acari: Tenuipalpidae) E Suas Competências Como Vetores De Vírus*. Available online at: http://www.teses.usp.br/teses/disponiveis/11/11135/tde-17072018-160552/pt-br.php [Accessed September 10, 2018].

Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., and Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44, W232–W235. doi: 10.1093/nar/gkw256

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., et al. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Res.* 40:e115. doi: 10.1093/nar/gks596

Vasilakis, N., Forrester, N. L., Palacios, G., Nasar, F., Savji, N., Rossi, S. L., et al. (2013). Negevirus: a proposed new taxon of insect-specific viruses with wide geographic distribution. *J. Virol.* 87, 2475–2488. doi: 10.1128/JVI.00776-12

Vergani, A. R. (1942). La transmisión de la lepra explosiva de la ligustrina por ácaros. *Rev. Agri. Agron.* 9, 292–294.

Welbourn, W. C., Ochoa, R., Kane, E. C., and Erbe, E. F. (2003). Morphological observations on Brevipalpus phoenicis (Acari: Tenuipalpidae) including comparisons with B. californicus and B. obovatus. *Exp. Appl. Acarol.* 30, 107–133. doi: 10.1023/b:appa.0000006545.40017.a0

Wu, Q., Wright, M., Gogol, M. M., Bradford, W. D., Zhang, N., and Bazzini, A. A. (2020). Translation of small downstream ORFs enhances translation of canonical main open reading frames. *EMBO J.* 39:e104763. doi: 10.15252/EMBJ.2020104763

Zhang, S., Yang, C., Wu, J., Qiu, Y., Xuan, Z., Yang, L., et al. (2022). Conserved untranslated regions of multipartite viruses: natural markers of novel viral genomic components and tags of viral evolution. *bioRxiv* [Preprint] bioRxiv: 2022.01.16.476546, doi: 10.1101/2022.01.16.476546

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.

frontiers | Frontiers in Microbiology

HYPOTHESIS AND THEORY
published: 06 April 2022
doi: 10.3389/fmicb.2022.836743

Circulative Transmission of Cileviruses in *Brevipalpus* Mites May Involve the Paracellular Movement of Virions

Aline Daniele Tassi[1,2], Pedro Luis Ramos-González[1], Thais Elise Sinico[1,3], Elliot Watanabe Kitajima[2] and Juliana Freitas-Astúa[1,4]*

[1] Laboratório de Biologia Molecular Aplicada, Instituto Biológico, São Paulo, Brazil, [2] Escola Superior de Agricultura Luiz de Queiroz (ESALQ), Universidade de São Paulo, Piracicaba, Brazil, [3] Centro de Citricultura Sylvio Moreira, Cordeirópolis, Brazil, [4] Embrapa Mandioca e Fruticultura, Cruz das Almas, Brazil

Plant viruses transmitted by mites of the genus *Brevipalpus* are members of the genera *Cilevirus,* family *Kitaviridae,* or *Dichorhavirus,* family *Rhabdoviridae*. They produce non-systemic infections that typically display necrotic and/or chlorotic lesions around the inoculation loci. The cilevirus citrus leprosis virus C (CiLV-C) causes citrus leprosis, rated as one of the most destructive diseases affecting this crop in the Americas. CiLV-C is vectored in a persistent manner by the flat mite *Brevipalpus yothersi*. Upon the ingestion of viral particles with the content of the infected plant cell, virions must pass through the midgut epithelium and the anterior podocephalic gland of the mites. Following the duct from this gland, virions reach the salivary canal before their inoculation into a new plant cell through the stylet canal. It is still unclear whether CiLV-C multiplies in mite cells and what mechanisms contribute to its movement through mite tissues. In this study, based on direct observation of histological sections from viruliferous mites using the transmission electron microscope, we posit the hypothesis of the paracellular movement of CiLV-C in mites which may involve the manipulation of septate junctions. We detail the presence of viral particles aligned in the intercellular spaces between cells and the gastrovascular system of *Brevipalpus* mites. Accordingly, we propose putative genes that could control either active or passive paracellular circulation of viral particles inside the mites.

Keywords: virus vector relationship, *Kitaviridae*, virus movement, septate junctions, flat mite, citrus leprosis virus C

OPEN ACCESS

Edited by:
Elvira Fiallo-Olivé, La Mayora Experimental Station (CSIC), Spain

Reviewed by:
Miguel A. Aranda, Spanish National Research Council (CSIC), Spain
Jaqueline Franciosi Della Vechia, São Paulo State University, Brazil

****Correspondence:***
Aline Daniele Tassi
alinetassi@gmail.com

Specialty section:
This article was submitted to Microbe and Virus Interactions with Plants, a section of the journal Frontiers in Microbiology

Received: *15 December 2021*
Accepted: *15 February 2022*
Published: *06 April 2022*

Citation:
Tassi AD, Ramos-González PL, Sinico TE, Kitajima EW and Freitas-Astúa J (2022) Circulative Transmission of Cileviruses in Brevipalpus Mites May Involve the Paracellular Movement of Virions. Front. Microbiol. 13:836743. doi: 10.3389/fmicb.2022.836743

INTRODUCTION

Plant diseases caused by *Brevipalpus*-transmitted viruses (BTV) result in non-systemic infections that produce local necrotic and chlorotic lesions on leaves, stems, and fruits (Kitajima et al., 2003, 2010). Early studies based on ultrastructural analyses of BTV-infected tissues revealed two types of viruses which were further recognized as BTV-Cytoplasmic and -Nuclear types (Kitajima et al., 2003). BTV-C and -N have contrasting molecular biology but they still display some common features suggesting a possible convergent evolution (Freitas-Astúa et al., 2018).

The infection by BTV-N induces the formation of electron-lucent viroplasms in the nucleus. Virions are naked, short rod-like particles (40 nm × 100-110 nm) that accumulate both in the nucleus and the cytoplasm of plant and mite cells (Freitas-Astúa et al., 2018; **Figure 1A**). The

genomes of BTV-N comprise two ssRNA molecules (∼ 6 kb each) of negative sense with six open reading frames (Dietzgen et al., 2018). Five BTV-N have been molecularly characterized and are definitive members of the genus *Dichorhavirus,* family *Rhabdoviridae* (Dietzgen et al., 2018; Amarasinghe et al., 2019; Kuhn et al., 2020).

Virions of the BTV-C type are short, enveloped bacilliform particles of 70-80 nm wide and 100-120 nm long. Virus particles are commonly found inside cisternae of the endoplasmic reticulum, and form electron-dense, vacuolated inclusions (viroplasms) in the infected plant cell cytoplasm (**Figure 1B**). BTV-C genomes consist of two (+) sense single-stranded (ss) RNA molecules of ∼5 and 9 kb. They are assigned to the genus *Cilevirus*, family *Kitaviridae* (Freitas-Astúa et al., 2018; Quito-Avila et al., 2021). Citrus leprosis virus C (CiLV-C) is the best-characterized cilevirus at both molecular and epidemiological levels (Chabi-Jesus et al., 2021).

Besides the genus *Cilevirus*, the family *Kitaviridae* comprises other two plant-infecting virus genera: *Higrevirus* and *Blunervirus* (Quito-Avila et al., 2021). Overall, kitaviruses show a heterogeneous genome organization and share a phylogenetic relationship with arthropod-infecting viruses of the groups negevirus (including nelorpiviruses and sandewaviruses), centivirus, and aphiglyvirus (Kondo et al., 2020; Ramos-González et al., 2020).

Few species of *Brevipalpus* are known as vectors of cileviruses and dichorhaviruses (de Lillo et al., 2021). Dichorhaviruses are transmitted in a circulative propagative manner, whereas cileviruses are transmitted in a circulative manner, but whether cileviruses replicate in the mites is still unclear (Roy et al., 2015; Tassi et al., 2017). Little is known about the mode of transmission of negeviruses and kitaviruses other than cileviruses. The isolation of Okushiri virus from mosquito larvae suggests the vertical transmission of negeviruses (Vasilakis et al., 2013; Carapeta et al., 2015; O'Brien et al., 2017). The expansion of perinuclear spaces filled with vesicles or microtubules, sometimes in paracrystalline arrays, and the accumulation of cytoplasmic vacuoles similar to those detected during the alphavirus infection, are indicators of the multiplication of the sandewavirus Tanay virus in C6/36 mosquito cells (Vasilakis et al., 2013; Zhao et al., 2019). *Brevipalpus* mites have been associated with the transmission of hibiscus green spot virus 2, genus *Higrevirus*, and the cilevirus-like hibiscus yellow blotch virus (Melzer et al., 2013; Olmedo-Velarde et al., 2021).

Brevipalpus Mites as Viral Vectors

The genus *Brevipalpus* groups almost 300 valid species (Castro et al., 2020) of obligatory phytophagous, mostly polyphagous red-brownish mites, which are distributed across the subtropical and equatorial regions of the world. *Brevipalpus* mites are flattened, of approximately 0.3 mm long, move slowly, and display five developmental phases, *i.e.*, egg, larvae, protonymph, deutonymph, and adult (Welbourn et al., 2003; Alberti and Kitajima, 2014; Tassi et al., 2017; Dietzgen et al., 2018).

Data on *Brevipalpus*-dichorhavirus relationships are almost limited to studies derived from the pathosystem orchid fleck virus (OFV)-*B. californicus* (Kondo et al., 2003). Upon acquisition, OFV transmission has a latent period of three weeks, the inoculation access period is approximately 30 min, and viral retention in mites occurs for at least three weeks (Kondo et al., 2003). Nymphs and adults, but not the larvae, have vector activity, suggesting a persistent propagative type of transmission (Kondo et al., 2003).

Different species of *Brevipalpus* colonizing dichorhavirus-infected plants exhibit electron-lucent viroplasms in the nucleus, and short rod-like particles in both the nucleus and the cytoplasm of midgut and anterior podocephalic gland cells (Alberti and Kitajima, 2014; Ramos-González et al., 2018). In the nucleus, viral particles may appear dispersed within nucleoplasm or viroplasm and arranged perpendicularly onto the inner membrane of the nuclear envelope (**Figure 1**). In the cytoplasm, they are commonly seen associated with endoplasmic reticulum membranes, occasionally radially arranged, forming the so-called "spoke wheel" configuration (**Figure 1**). The accumulation patterns of viral particles in viruliferous mites are essentially similar to those observed in dichorhavirus-infected plant cells, suggesting that dichorhaviruses replicate in the mite (Kitajima et al., 2003). No accumulation of particles is observed between adjacent cells of dichorhavirus-infected mites.

All the active life stages of *B. yothersi* can transmit the cilevirus CiLV-C, but no transovarial passage occurs (Chiavegato, 1996; Tassi et al., 2017). Using common bean (*Phaseolus vulgaris*) as indicator plant (Garita et al., 2013), CiLV-C acquisition and inoculation access periods by *B. yothersi* are 4- and 2 h, respectively, with a latent period of 7 h (Tassi et al., 2017). Once the virus is acquired, *B. yothersi* remains viruliferous for at least 20 days (Tassi et al., unpublished). Viral particles are consistently found between adjacent cells at the basal part of the caeca and in the anterior podocephalic gland of mites (**Figure 1C**; Alberti and Kitajima, 2014). The load of viral particles in the intercellular spaces seems to increase proportionally with the number of acquisitions (Tassi et al., 2017). Frequently, lines of viral particles are interrupted by septate junctions (**Figures 2B,C**). Differently from what is easily seen in plant cells infected by CiLV-C, viroplasms are not observed in CiLV-C viruliferous mites. To improve the viral identification, *in situ* immunogold labeling using polyclonal antibodies against the P29 protein of CiLV-C (Calegario et al., 2013) was carried out as previously described for plant tissues (Calegario et al., 2013) and *Brevipalpus* mites (Alberti and Kitajima, 2014). Sets of virion particles aligned up to 10 μm in length were detected in paracellular spaces of *Brevipalpus* mites that fed on CiLV-C-infected oranges (**Figure 2**).

The detection of anti-genomic (complementary strand) RNA of the cileviruses CiLV-C and CiLV-C2 in viruliferous *B. yothersi* was considered evidence of cilevirus replication within the vector (Roy et al., 2015). However, anti-genomic RNA molecules of CiLV-C may have arisen in the mite body upon feeding on infected plants or may have been generated as a result of self-primed genomic molecules during the RT-PCR detection. Therefore, further assays, including new controls, a time-course experiment, and the search for putative replication sites in

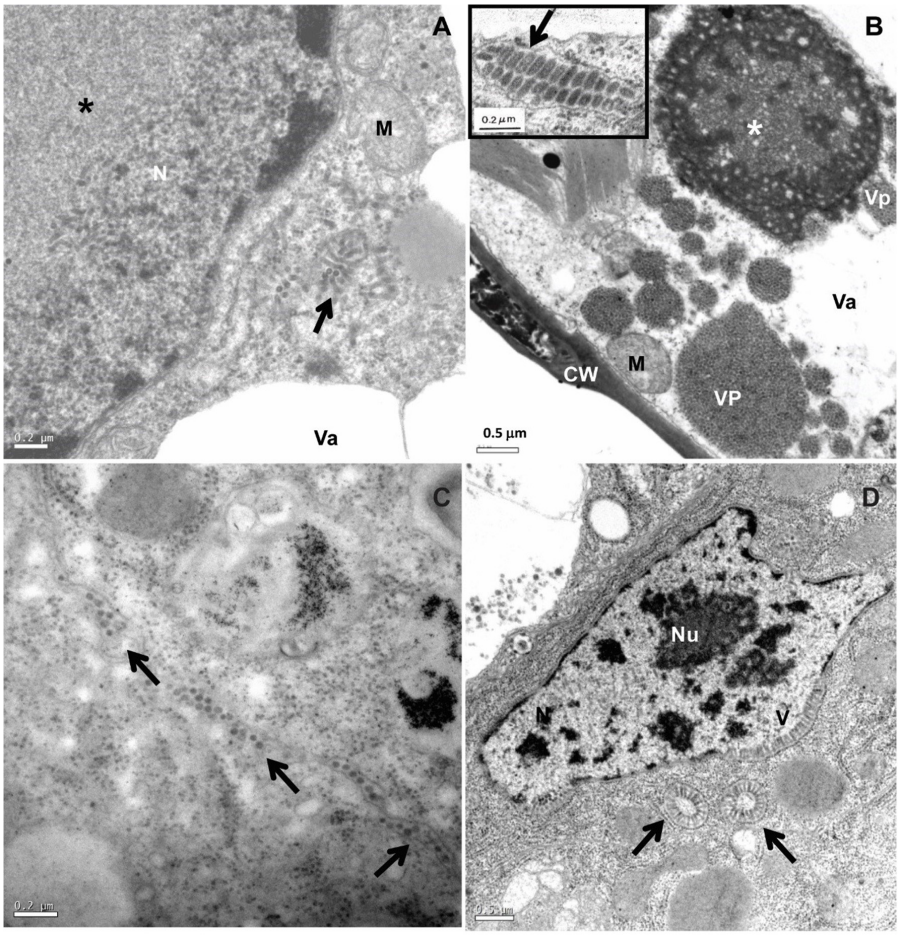


FIGURE 1 | Transmission electron micrographs of BTV-infected leaf lesion cell and mite vector *Brevipalpus yothersi* tissues. **(A)** *Clerodendrum thomsonae* infected by the dichorhavirus clerodendrum chlorotic spot virus (ClCSV), with electron lucent viroplasm (*) and short, rod-like particles (arrow) in the cytoplasm. **(B)** *Ligustrum sinense* infected with the cilevirus passion fruit green spot virus (PfGSV) that induces large electron dense and vacuolated viroplasm (*) in the cytoplasm; short bacilliform virions (VP) are present within endoplasmic reticulum cisternae (insert). **(C)** Sections from adults *Brevipalpus yothersi* collected from *L. sinense* infected with PfGSV showing virus particles (arrows) in the extracellular space. **(D)** Sections from adults *Brevipalpus yothersi* collected from *C. thomsonae*, infected by ClCSV showing evidences of viral replication in their tissues, large nuclei of the anterior podocephalic gland cell, with virus-like particles (V) at the nuclear periphery. These particles are also present in the cytoplasm, forming the spoke wheel configuration (arrows). C, chloroplast; CW, cell wall; M, mitochondrion; N, nucleus; Va, vacuole; Nu, nucleolus.

specific mite tissues not yet visualized by transmission electron microscopy are ongoing (Tassi et al., unpublished).

CILEVIRUSES MOVEMENT WITHIN THEIR MITE VECTORS: A CRITICAL EVALUATION OF THE ALTERNATIVES

Transcytosis in Circulative Non-propagative Viruses

In addition to whether cileviruses replicate in their mite vectors or not, the mechanisms that promote virion internalization, movement, and their release into the stylet canal also remain uncertain. Transcytosis is a cellular mechanism in which extracellular materials, enclosed in vesicles generated by endocytosis, move across the cell and eject the content in the distal section of the cells by exocytosis (Whitfield et al., 2015). Transcytosis has been also described as a form of circulation of plant viruses in their vectors, but the underlying mechanisms are not fully elucidated (Brault et al., 2007; Hogenhout et al., 2008; Blanc et al., 2014; Di Mattia et al., 2020). The internalization in the vector body of several viruses of the families *Luteoviridae, Geminiviridae,* and *Nanoviridae* occurs without replication of the viral genome. Virus particles are transported across cells into membrane vesicles, preventing any contact between viruses and the insect cell cytoplasm in the epithelia of the gut and salivary gland (Brault et al., 2007; Hogenhout et al., 2008; Blanc et al., 2014). The vesicles formed during transcytosis seemingly follow the early endosomal pathway before the appearance of non-coated tubular vesicles. Inside these vesicles, virions likely reach the

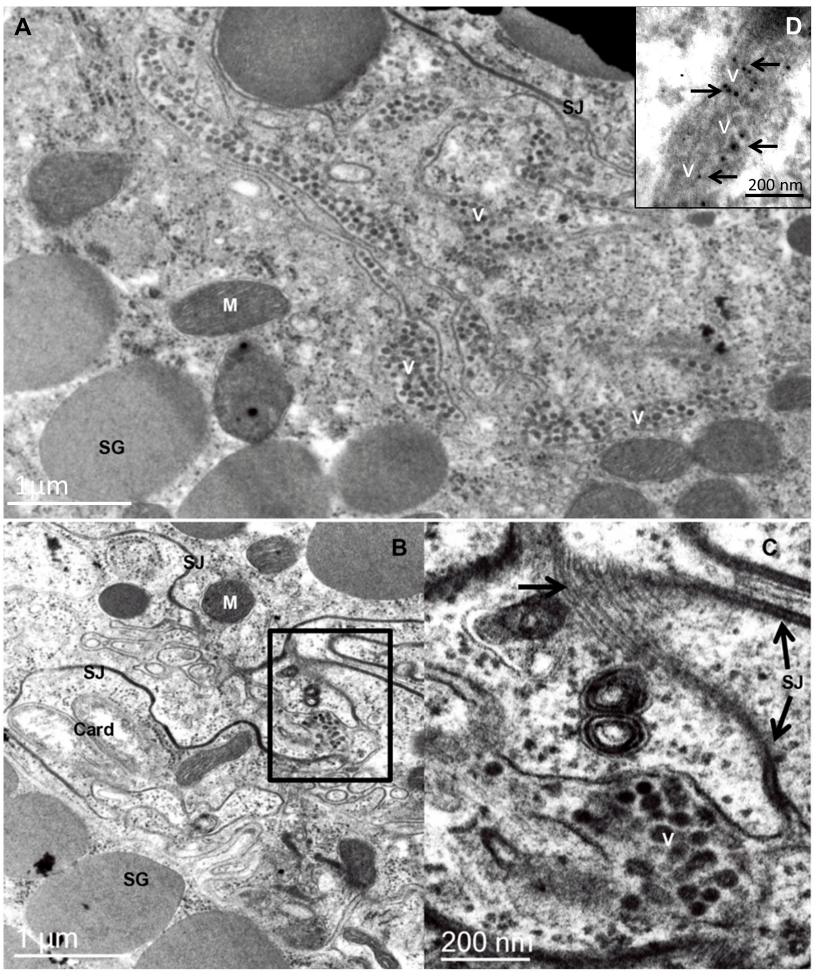


FIGURE 2 | Transmission electron micrographs of sections of the prosomal region of an adult female *Brevipalpus yothersi*, viruliferous for citrus leprosis virus C (CiLV-C). **(A)** Basal part of midgut caeca, showing several rows of virions (V), aligned in the extracellular space formed by four layers of cells. It is presumed that once internalized, crossing the midgut epithelial cell barrier, these particles move passively in the direction of the anterior podocephalic gland (= salivary gland) following the celomic flux, where they will reach the stylet channel, after overtaking the gland cell barrier. **(B)** An area of the branched ceaca, revealing a labyrinth of membranes running between adjacent cells. A small group of virions (V) is present in one of these intercellular spaces. **(C)** An enlarged region of figure B in which septate junctions (SJ) are well depicted. The large arrow points to a tangential section through a septate junction, revealing the rows of intermembrane proteins. **(D)** *in situ* immunolabeling using anti-p29 polyclonal antibody in aldehyde-fixed and LRWhite embedded *B. yothersi* viruliferous for CiLV-C. Card, bacterial endosymbiont Cardinium; M, mitochondrion; SG, secretion granules.

basal membrane and exit the gut cells into the hemolymph (Ali et al., 2018). A transcytosis process is also observed when luteovirids cross the cellular barrier of the accessory salivary gland (Brault et al., 2007). For cileviruses, however, although transcytosis should not be completely disregarded, it seems an unlikely route of circulation in *Brevipalpus* mites. Viral particles have been observed neither inside cells of the anterior midgut epithelium nor in cells of anterior and dorsal podocephalic glands of mites feeding on cilevirus-infected plants (Alberti and Kitajima, 2014).

It is important to notice that anatomic differences between *Brevipalpus* mites and insects may account for the nonexistence of transcytosis in these mites. During feeding, *Brevipalpus* mites use stylets to perforate the epidermal layer of plant organs and reach the underlying parenchymal cell content after punctuating its wall and membrane. Saliva produced by the anterior podocephalic gland is injected to pre-digest the cellular content (Alberti and Kitajima, 2014). The ingested material then follows to the esophagus that crosses the synganglion and ends into the anterior midgut through a small valve, the ventriculus, which consists of a small lumen and the highly branched caeca (Alberti and Kitajima, 2014). The caeca, formed by large epithelial cells, extend both to anterior and posterior parts of the mite body, occupying every space among the organs, producing a complex labyrinth of cell membranes and intercellular spaces, many of which are joined by septate junctions (**Figures 2A-C**), leaving the hemolymph confined to small and restricted cavities. This complex of cells comprises the so-called gastrovascular system which may directly irrigate several organs with digested nutrients and, probably, virus particles in viruliferous mites

(Alberti and Kitajima, 2014). *Brevipalpus* mites lack a pulsating organ, so the circulation of the hemolymph depends on the movement generated by their muscles and internal organs, diverging from insects that have a circulatory system and, therefore, a more active circulation of nutrients and other fluids (Alberti and Coons, 1999).

Paracellular Route of CiLV-C in *Brevipalpus yothersi*: A Hypothesis

The persistent circulatory transmission of cileviruses by *Brevipalpus* mites poses a challenge to explain how the cilevirus movement occurs. A raising question is how virions get access from the midgut lumen to the hemolymph space, and later to the stylet channel, since two epithelial barriers hamper it: the midgut (caeca) and the anterior podocephalic gland. Ultrastructural observations of viruliferous *B. yothersi* mites feeding on CiLV-C-infected plants reveal the presence of viral particles between cells (Alberti and Kitajima, 2014), which led us to ponder the existence of a paracellular pathway of virion movement within its vector.

In the epithelium of invertebrates, occluding structures named septate junctions (SJ) act as a barrier that separates distinct compartments, limiting the paracellular passage of fluids (Izumi and Furuse, 2014). Unlike tight junction (TJ), the structure sealing the apical part of epithelial cells in the vertebrates, the SJ forms circumferential belts around the apicolateral regions of the cells. Visually, they appear as a ladder-like septum, with 15-20 nm of spacing (Izumi and Furuse, 2014). Studies of the morphophysiology of the SJ in the fly *Drosophila melanogaster* exposed two types of SJ, i.e., the pleated septate junctions (pSJ), present in ectodermal-derived epithelium, i.e., epidermis, fore- and hindgut, salivary gland, etc., and the smooth septate junctions (sSJ), which occurs in the endodermal-derived epithelium, i.e., midgut, gastric caeca (Izumi and Furuse, 2014; Hall and Ward, 2016). Mutations of the SJ proteins may be lethal or produce functionally deficient junctions, but the exact mechanisms underlying the transient opening of SJ have not been described. To our knowledge, there are no reports of arthropod viruses interacting with tight or septate junctions.

In vertebrates, TJs could act as physical barriers from the innate immune defense system, especially on the respiratory tract. The coordination of TJ opening is mediated by chemical signals and membrane receptors, as happens during the paracellular passage of lymphocytes through the walls of capillary vessels (Yonekawa and Harlan, 2005; Yumine et al., 2019). In humans, the permeability of TJ between cells of the wall of the digestive tract is mediated by zonulin, a pre-haptoglin protein, and by gliadin, a component of gluten, in patients suffering from celiac disease (Fasano, 2012). Similarly, to replicate or transit through epithelia, viruses take advantage of the structural proteins that form the TJ and adherent junctions (AJ) as their receptors (Mateo et al., 2015). Viruses evolved selecting strategies that could counter the antiviral function of TJs, by degradation processes, e.g., it is suggested that mosquito-borne viruses like west nile virus and Japanese encephalitis virus establish infections in vertebrate hosts by degrading TJ molecules to disrupt the epithelial barriers (Medigeshi et al., 2009; Yumine et al., 2019).

The gastrovascular system of *Brevipalpus* mites, composed of the highly branched caeca that fill all spaces left by other tissues on the hemolymph, may serve as a pathway to the circulation of viruses inside the mites.

Based on our transmission electron microscopy studies following the methodology proposed by Alberti and Kitajima (2014), observations of the accumulation of virus particles between cells and the possible role of SJ limiting or coordinating the circulation of viruses to new tissues, we pose alternative hypotheses to explain virus-gastrovascular system interaction including either passive or active mechanisms. The passive interaction involves the SJs transient openings which allow the transport of nutrients between cells. Following the flow, the virus could reach the spaces between cells, leading to a circulation of virus particles within the hemolymph. Alternatively, virus-encoded proteins could recognize the SJs components and induce transient openings, allowing the passage of viruses in an active process, as seen in some vertebrates-infecting viruses, or even in other physiological or pathological processes, i.e., zonulin and gliadin-like interaction (**Figure 3**). The viral protein P61, a putative membrane glycoprotein, or P24, a putative viral structural protein, might trigger a transient local opening of the SJ, ensuing the paracellular traffic of cileviruses in *Brevipalpus* mites.

Orthologue Proteins of SJ and sSJ Factors in Mites

The involvement of TJs and AJs in cell-to-cell viral movement and their role as receptors have been reported for human viruses (Cifuentes-Munoz et al., 2020), but there is no information on viruses using this route in arthropods. In humans, for instance, after primary infection of the respiratory airway, measles virus, family *Paramyxoviridae,* spreads laterally into the epithelium *via* AJs (Mühlebach et al., 2011; Nakatsu et al., 2013). Besides, human TJ components act as viral receptors, e.g., the glycoprotein of hepatitis C virus (HCV) uses occludin and claudin as co-receptors to enter hepatocytes (Reynolds et al., 2008; Timpe et al., 2008; Brimacombe et al., 2011; Yumine et al., 2019), the coxsackieviruses (positive-strand RNA viruses) and adenoviruses (double-stranded DNA viruses), although exploring different strategies, target the integral chimeric antigen receptor (CAR) protein associated with TJ (Bergelson et al., 1997; Bergelson, 2009; Mateo et al., 2015), reoviruses use the TJ protein junctional adhesion molecule A (JAM-A) as receptor (Barton et al., 2001), claudin 1 is involved on dengue virus entry by the interaction of the viral protein prM and TJ component (Che et al., 2013).

Proteins of the family LY6_uPAR, also called three-finger proteins (TFP), are found associated with the membrane by a glycosylphosphatidylinositol anchor and play essential roles in cell adhesion, signaling, and lipid metabolism (Vasilyeva et al., 2017). TFPs display one or several domains consisting of 60-90 amino acids which have an β-structural core stabilized by a system of four invariant disulfide bonds. Proteins with similar structural characteristics have been found in the mite *Tetranychus urticae* and insects (Grbić et al., 2011; Zhang et al., 2019). In flies, TFP-like proteins are involved in the formation of SJ adhesion

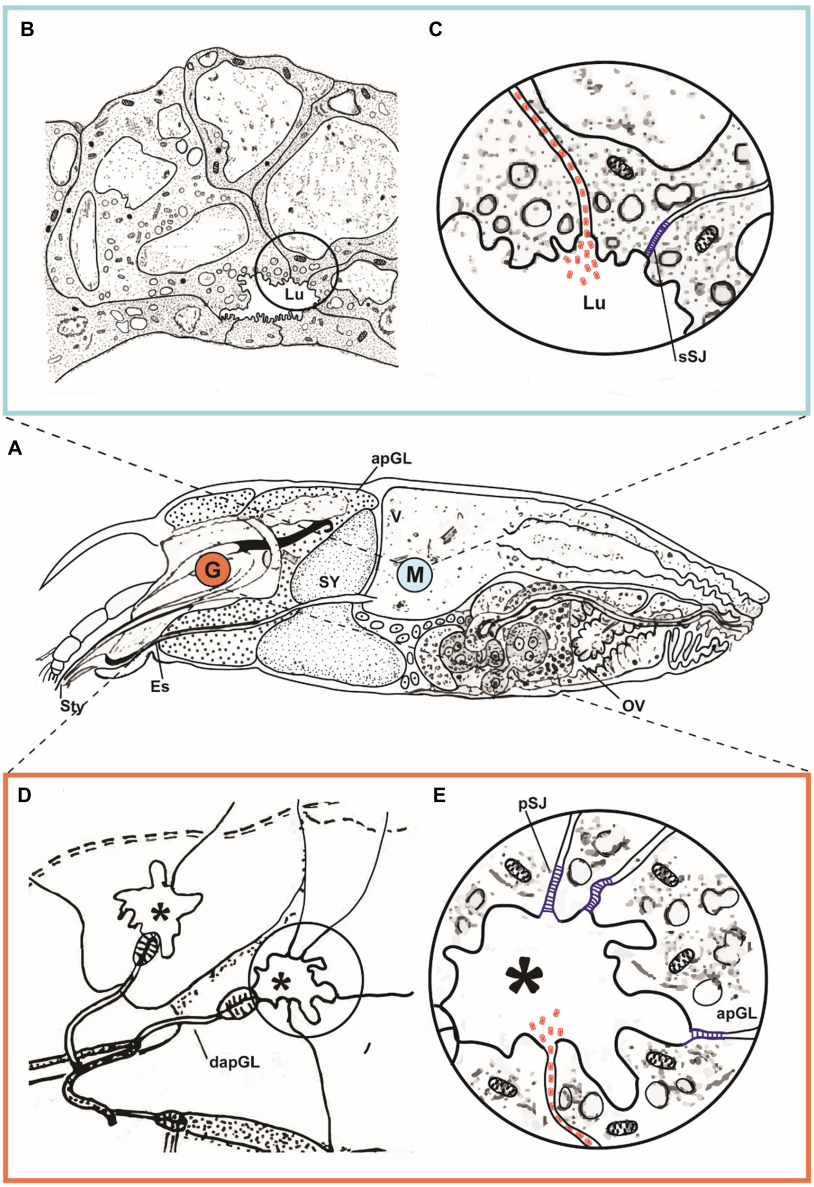


FIGURE 3 | Hypothetical pathway of cileviruses circulating within the mite vector. **(A)** Schematic representation of sagittal section of an adult female of *Brevipalpus yothersi*, revealing details of its internal anatomy [stylet complex, prosomal glands, synganglion, digestive tract (ventriculus, caeca), ovary]. Details of ventriculus (V) and anterior podocephalic gland (apGL) are shown indicated in the marked area M and detailed in **(B,C)**. Portion of the anterior podocephalic gland (equivalent to salivary gland) marked by G, is detailed in **(D,E)**. **(B)** Part of ventriculum (V) exhibiting epithelial cells of caeca with small lumen (Lu). In the detail **(C)**, hypothetical transitional opening of the smooth septate junction (sSJ) at the apical part of the epithelial cell, induced by the presence of viral particles represented in red. Virions gain access to the intercellular space (Lu), being carried out by the celomic flux, to the apical part of the anterior podocephalic gland. **(D)** Schematic drawing of collecting reservoir (*) of the apGL secretions. **(E)** Detail depicted in **(D)**. Glandular cells are held together by pleated septate junctions (pSJ) represented in blue, which hypothetically open due to the presence of viral particles (in red) arriving by the intercellular space, releasing them into the reservoir (*), and subsequently to the stylet channel. dapGL, duct of the anterior podocephalic gland (adapted from Alberti and Kitajima, 2014).

structures, suggesting a common ancient role for these proteins in arthropods (Hijazi et al., 2009).

In this study, based on literature review and Blast N search[1], we were able to identify putative homologs of proteins that regulate SJ and sSJ in *Drosophila* encoded by the genomes of *T. urticae* (the most studied mite in the superfamily Tetranychoidea) and *B. yothersi* (**Table 1**). One of the genes identified is the polychaetoid (*pyd*) gene, which is a recognized homolog of zonulin (Seppa et al., 2008; Choi et al., 2011; Djiane et al., 2011). The study of the transcriptomic profile of these genes in viruliferous mites will probably add hints on SJ and sSJ opening processes in mites and their possible relationship with cileviruses.

[1]https://blast.ncbi.nlm.nih.gov/Blast.cgi

TABLE 1 | Genes of *Drosophila melanogaster* involved in SJ and sSJ regulation and their respective orthologs in *Tetranychus urticae* and *Brevipalpus yothersi*.

Gene name			Function in *Drosophila*			References
Drosophila	***E*-value**	***T. urticae***	***E*-value**	***B. yothersi***		
Anakonda	0.0	tetur04g07130.1	0.00	bryot81g00190	Putative transmembrane scavenger receptor-like protein that is essential for the maturation SJ	Byri et al., 2015; Hildebrandt et al., 2015; Daniel et al., 2018
Gliotactin	2e-147	tetur30g01560.1	0.00	bryot209g00240	Transmembrane protein localized at tricellular junctions that is necessary for septate junction and permeability barrier formation	Auld et al., 1995; Genova and Fehon, 2003; Schulte et al., 2003; Venema et al., 2004
Cora	2e-151	tetur17g00600.1	0.00	bryot168g00070	Required for SJ integrity with a role in cell-cell interactions, vital for embryonic proper development.	Fehon et al., 1994; Lamb et al., 1998; Ward et al., 2001
Mesh	1e-157	tetur08g07660.1	0.00	bryot35g00820	Transmembrane protein component of smooth SJ organization	Izumi et al., 2012; Chen et al., 2018, 2020
Tetraspanin 2A	2e-08	tetur17g03500.1	1.16e-112	bryot101g00020	Component necessary for the assembly of SJ, on the midgut.	Izumi et al., 2016; Xu et al., 2019
Neuroglian (Nrg)	0.0	tetur19g00920.1	0.00	bryot101g00400	Contributes to the formation of SJ in epithelial cells.	Genova and Fehon, 2003; Godenschwege et al., 2006; Williams, 2009
Nervana 2 (nvr2)	3e-80	tetur35g00730.1	9.65e-154	bryot98g00300	Plays an ion-pump-independent role in junction formation and transport on the plasma membrane	Sun and Salvaterra, 1995; Paul et al., 2003
Lethal (2) giant larvae l(2)gl	1e-156	tetur18g00160.1	0.00	bryot15g00230	Regulates cell polarity, asymmetric cell division. Localized in smooth SJ.	Woods and Bryant, 1991
G protein α i subunit (Gαi)	0.0	tetur05g01580.1	0.00	bryot23g00160	Involved in regulating asymmetric cell division. Localized in SJ.	Schwabe et al., 2005
	0.0	tetur04g03270.1	0.00	bryot32g00500		
	0.0	tetur15g03060.1	1.6e-124	bryot13g00350		
Patj	1E-107	tetur27g00480.1	5.3e-94	bryot140g00140	Paly supporting roles in apico-basal cell polarity and stability of adherens junction	Tanentzapf et al., 2000; Nam and Choi, 2006; Sen et al., 2012
P21-activated kinase	1e-12	tetur13g00020.1	3.9e-172	bryot77g00180	Involved in regulation of cytoskeleton, apical junction assembly.	Conder et al., 2004; Koch et al., 2008; Bahri et al., 2010
Ankyrim 2 (Ank2)	0.0	tetur15g02730.1	2.82e-176	bryot07g00160	Cytoskeletal binding protein, plasma membrane-bounded cell projection organization.	Hortsch et al., 2002; Koch et al., 2008; Bulat et al., 2014
Polychaetoid (pyd)	0.0	tetur33g01420.1	0.00	bryot71g00250	Broadly acting protein that is associated with multiple proteins at the surface and within the cytoskeleton	Seppa et al., 2008; Choi et al., 2011; Djiane et al., 2011

DISCUSSION

Virion particles of CiLV-C are routinely observed between the membrane of adjacent cells of *B. yothersi* (Alberti and Kitajima, 2014). In contrast, neither virions nor structures such as viroplasms, commonly associated with viral multiplication, have been observed inside mite cells. In this context, this study presents elements that guided us to pose the hypothesis of the paracellular movement of CiLV-C inside *Brevipalpus* mites. This unconventional viral movement has been described in the circulation of several viruses in localized organs by inducing disruptions of TJ in vertebrates. An example is the coronavirus SARS-COV-2, which is favored by the disruption of the airway epithelium. This process facilitates the virus paracellular spread into other tissues besides the translocation of endothelial cells (Tugizov, 2021). A similar mechanism is triggered by the rotavirus VP4 capsid protein. VP4 interacts with zonulin 1, occludin, and claudin, stimulating their redistribution and granting access to the junctional areas, which promotes viral spread in a paracellular way (Nava et al., 2004; Tugizov, 2021).

In phytopathosystems comprising (+)ssRNA viruses, a common virus-vector mode of transmission is circulative non-replicative (Hogenhout et al., 2008; Whitfield et al., 2015; Kondo et al., 2019). In many of these systems, the ingested virions must pass through vector cell barriers to reach the salivary glands, including the gut and hemocoel, involving specific interactions between the virus and vector membrane (Bragard et al., 2013; Blanc et al., 2014).

Negeviruses and other insect-borne viruses are likely vertical transmitted to host offprints (Vasilakis and Tesh, 2015;

Kondo et al., 2019). Vertical transmission efficiency of plant-infecting (+)ssRNA viruses is generally low due to biological barriers *via* RNA silencing mechanisms that protect plant germ cells against viral infection (Foster et al., 2002; Martín-Hernández and Baulcombe, 2008). As consequence, most of these viruses depend on arthropod vectors for horizontal transmission in a non-replicative manner (Hull, 2013; Whitfield et al., 2015; Andika et al., 2016).

It is speculated that during virus evolution, some viruses, whose ancestors were arthropods-infecting viruses, adapted to plant hosts, but maintained an intimate relationship with those species of arthropods that eventually became their vectors. In these cases, virus circulation and replication are observed within the arthropod and plant hosts, as are the cases of the plant-infecting reoviruses rice dwarf virus and Southern rice black-streaked dwarf virus and their hemipteran vectors *Nephotettix cincticeps* and *Sogatella furcifera*, respectively, or the bunyavirus tomato spotted wilt virus and its thrips vector *Frankliniella dentalis* (Whitfield et al., 2015, 2018; Dolja and Koonin, 2018; Chen et al., 2019; Lefeuvre et al., 2019).

Biological, cytopathic, and molecular data on the cilevirus-mite relationship suggest that these viruses circulate in the vectors, but whether they replicate still needs to be addressed. It has been proposed that the cileviruses have evolved from an arthropod virus ancestor that somehow was able to infect plants after acquiring a movement protein from a plant virus (Ramos-González et al., 2020). The identification of some nege-like viruses infecting plants gives further support to the arthropod-plant host transitional process particularly involving kitaviruses and kita-like viruses infecting arthropods (Morozov et al., 2020). On this basis, it is tempting to speculate that during adaptation to plants, the presumed ancestor of kitaviruses lost arthropod fitness as it gradually adapted to plant hosts, but still, some viral factors required for its interaction with the arthropod were retained, for instance, those minimal components allowing for the circulative route using paracellular spaces. The study of Tetranychus urticae kitavirus (Niu et al., 2019), the closest kita-like virus infecting mites known, would likely add new elements to the mechanism underlying the movement of nege-kita-like viruses in their hosts.

If the paracellular route of cilevirus circulation in mites may be controlled in an active form, virion-membrane receptor(s) interaction has to be assumed. Accordingly, the transmissibility or not and efficiency of this process for a given pair of virus- *Brevipalpus* species must be dictated by the receptor(s) required for the virus entry into different groups of tissues by interactions with SJ. Studies on the molecular interactions between mite vectors and plant viruses are scarce (de Lillo et al., 2021). Global transcriptomic response of wheat streak mosaic virus (WSMV; genus *Tritimovirus*; family *Potyviridae*) and *Aceria tosichella* (Eriophyidae) showed the upregulation of two gene families that participate in the SJ formation (Gupta et al., 2019). Recent sequencing of the *B. yothersi* genome (Navia et al., 2019) and investigations on the function of CiLV-C-coded proteins (Leastro et al., 2018, 2020; Arena et al., 2020) may provide new insights into the putative participation of orthologues of these genes in *Brevipalpus*-cilevirus interaction.

In case the paracellular route of CiLV-C movement in *Brevipalpus* mites is confirmed, it will be the first example of a plant virus using this unconventional route of cell-to-cell movement in its arthropod vector. The unusual type of interaction with its vector, as also happens during the interaction of CiLV-C with their plant hosts, suggest that this virus, and probably other members of the family *Kitaviridae*, represent unique chimeric genetic systems that are likely reshaping features inherited from their ancestors to adapt to new ecological challenges.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

AT, EK, PR-G, and JF-A: conceptualization and methodology. PR-G, AT, EK, and TS: formal analysis. JF-A and EK: funding acquisition. AT, TS, PR-G, EK, and JF-A: investigation. PR-G, EK, and JF-A: supervision. AT: writing—original draft. AT, PR-G, EK, TS, and JF-A: writing—review and editing. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors are grateful to CAPES, CNPq, and FAPESP for scholarships, research fellowships, and research grants associated with this work (CAPES PNPD20132154 – 33141010001P4 – PNPD – IBSP, 88882.157041/2017-01; CNPq 312528/2020-5; FAPESP 2017/50222-0, 2017/50334-3, 2018/12252-8, and 2019/25078-9).

REFERENCES

Alberti, G., and Coons, L. B. (1999). "Microscopic anatomy of invertebrates. Volume 8c. Chelicerate arthropoda," in *Microscopic Anatomy of Invertebrates*, eds F. W. Harrison and R. Foelix (Chichester: Wiley & Sons), 267–514.

Alberti, G., and Kitajima, E. W. (2014). Anatomy and fine structure of brevipalpus mites (*Tenuipalpidae*) – economically important plant-virus vectors. *Zoologica* 160, 1–192.

Ali, M., Anwar, S., Shuja, M. N., Tripathi, R. K., and Singh, J. (2018). The genus *Luteovirus* from infection to disease. *Eur. J. Plant Pathol.* 151, 841–860. doi: 10.1007/s10658-018-1425-8

Amarasinghe, G. K., Ayllón, M. A., Bào, Y., Basler, C. F., Bavari, S., Blasdell, K. R., et al. (2019). Taxonomy of the order *Mononegavirales*: update 2019. *Arch. Virol.* 164, 1967–1980. doi: 10.1007/s00705-019-04247-4

Andika, I. B., Kondo, H., and Sun, L. (2016). Interplays between soil-borne plant viruses and RNA silencing-mediated antiviral defense in roots. *Front. Microbiol.* 7:1458. doi: 10.3389/fmicb.2016.01458

Arena, G. D., Ramos-González, P. L., Falk, B. W., Casteel, C. L., Freitas-Astúa, J., and Machado Marcos, A. (2020). Plant Immune system activation upon citrus leprosis virus c infection is mimicked by the ectopic expression of the p61 viral protein. *Front. Plant Sci.* 7:1757. doi: 10.3389/fpls.2020.01188

Auld, V. J., Fetter, R. D., Broadie, K., and Goodman, C. S. (1995). Gliotactin, a novel transmembrane protein on peripheral glia, is required to form the blood-nerve barrier in *Drosophila*. *Cell* 81, 757–767. doi: 10.1016/0092-8674(95)90537-5

Bahri, S., Wang, S., Conder, R., Choy, J., Vlachos, S., Dong, K., et al. (2010). The leading edge during dorsal closure as a model for epithelial plasticity: pak is required for recruitment of the Scribble complex and septate junction formation. *Development* 137, 2023–2032. doi: 10.1242/dev.045088

Barton, E. S., Forrest, J. C., Connolly, J. L., Chappell, J. D., Liu, Y., Schnell, F. J., et al. (2001). Junction adhesion molecule is a receptor for reovirus. *Cell* 104, 441–451. doi: 10.1016/S0092-8674(01)00231-8

Bergelson, J. M. (2009). Intercellular junctional proteins as receptors and barriers to virus infection and spread. *Cell Host Microbe* 5, 517–521. doi: 10.1016/j.chom.2009.05.009

Bergelson, J. M., Cunningham, J. A., Droguett, G., Kurt-Jones, E. A., Krithivas, A., Hong, J. S., et al. (1997). Isolation of a common receptor for coxsackie B viruses and adenoviruses 2 and 5. *Science* 275, 1320–1323. doi: 10.1126/science.275.5304.1320

Blanc, S., Drucker, M., and Uzest, M. (2014). Localizing viruses in their insect vectors. *Annu. Rev. Phytopathol.* 52, 403–425. doi: 10.1146/annurev-phyto-102313-045920

Bragard, C., Caciagli, P., Lemaire, O., Lopez-Moya, J. J., MacFarlane, S., Peters, D., et al. (2013). Status and prospects of plant virus control through interference with vector transmission. *Annu. Rev. Phytopathol.* 51, 177–201. doi: 10.1146/annurev-phyto-082712-102346

Brault, V., Herrbach, É, and Reinbold, C. (2007). Electron microscopy studies on luteovirid transmission by aphids. *Micron* 38, 302–312. doi: 10.1016/j.micron.2006.04.005

Brimacombe, C. L., Grove, J., Meredith, L. W., Hu, K., Syder, A. J., Flores, M. V., et al. (2011). Neutralizing antibody-resistant hepatitis c virus cell-to-cell transmission. *J. Virol.* 85, 596–605. doi: 10.1128/jvi.01592-10

Bulat, V., Rast, M., and Pielage, J. (2014). Presynaptic CK2 promotes synapse organization and stability by targeting Ankyrin2. *J. Cell Biol.* 204, 77–94. doi: 10.1083/jcb.201305134

Byri, S., Misra, T., Syed, Z. A., Bätz, T., Shah, J., Boril, L., et al. (2015). The triple-repeat protein anakonda controls epithelial tricellular junction formation in *Drosophila*. *Dev. Cell* 33, 535–548. doi: 10.1016/j.devcel.2015.03.023

Calegario, R. F., Locali, E. C., Stach-Machado, D. R., Peroni, L. A., Caserta, R., Salaroli, R. B., et al. (2013). Polyclonal antibodies to the putative coat protein of CiLV-C (*Citrus leprosis* virus C) expressed in *Escherichia coli*: production and use in immunodiagnosis. *Trop. Plant Pathol.* 38, 188–197. doi: 10.1590/S1982-56762013005000005

Carapeta, S., do Bem, B., McGuinness, J., Esteves, A., Abecasis, A., Lopes, Â, et al. (2015). Negeviruses found in multiple species of mosquitoes from southern portugal: isolation, genetic diversity, and replication in insect cell culture. *Virology* 483, 318–328. doi: 10.1016/j.virol.2015.04.021

Castro, E. B., Mesa, N. C., Feres, R. J. F., de Moraes, G. J., Ocha, R., Beard, J. J., et al. (2020). A newly available database of an important family of phytophagous mites: tenuipalpidae database. *Zootaxa* 4868, 577–583. doi: 10.11646/zootaxa.4868.4.7

Chabi-Jesus, C., Ramos-González, P. L., Postclam-Barro, M., Fontenele, R. S., Harakava, R., Bassanezi, R. B., et al. (2021). Molecular epidemiology of citrus leprosis virus c: a new viral lineage and phylodynamic of the main viral subpopulations in the americas. *Front. Microbiol.* 12:641252. doi: 10.3389/fmicb.2021.641252

Che, P., Tang, H., and Li, Q. (2013). The interaction between claudin-1 and dengue viral prM/M protein for its entry. *Virology* 446, 303–313. doi: 10.1016/j.virol.2013.08.009

Chen, H. J., Li, Q., Nirala, N. K., and Ip, Y. T. (2020). The snakeskin-mesh complex of smooth septate junction restricts yorkie to regulate intestinal homeostasis in *Drosophila*. *Stem Cell Rep.* 14, 828–844. doi: 10.1016/j.stemcr.2020.03.021

Chen, J., Sayadian, A. C., Lowe, N., Lovegrove, H. E., and St Johnston, D. (2018). An alternative mode of epithelial polarity in the *Drosophila* midgut. *PLoS Biol.* 16:e3000041. doi: 10.1371/journal.pbio.3000041

Chen, Y., Dessau, M., Rotenberg, D., Rasmussen, D. A., and Whitfield, A. E. (2019). Entry of bunyaviruses into plants and vectors. *Adv. Virus Res.* 104, 65–96. doi: 10.1016/BS.AIVIR.2019.07.001

Chiavegato, L. (1996). Aspectos biológicos e transmissão de leprose pelo ácaro *Brevipalpus phoenicis* (Geijskes, 1939) (Acari: Tenuipalpidae) em citros. *Laranja* 17, 229–235. doi: 10.1590/S0100-83582012000100011

Choi, W., Jung, K. C., Nelson, K. S., Bhat, M. A., Beitel, G. J., Peifer, M., et al. (2011). The single *Drosophila* ZO-1 protein polychaetoid regulates embryonic morphogenesis in coordination with canoe/afadin and Enabled. *Mol. Biol. Cell* 22, 2010–2030. doi: 10.1091/mbc.E10-12-1014

Cifuentes-Munoz, N., El Najjar, F., and Dutch, R. E. (2020). Viral cell-to-cell spread: conventional and non-conventional ways. *Adv. Virus Res.* 108, 85–125. doi: 10.1016/bs.aivir.2020.09.002

Conder, R., Yu, H., Ricos, M., Hing, H., Chia, W., Lim, L., et al. (2004). DPak is required for integrity of the leading edge cytoskeleton during *Drosophila* dorsal closure but does not signal through the JNK cascade. *Dev. Biol.* 276, 378–390. doi: 10.1016/j.ydbio.2004.08.044

Daniel, E., Daudé, M., Kolotuev, I., Charish, K., Auld, V., and Le Borgne, R. (2018). Coordination of septate junctions assembly and completion of cytokinesis in proliferative epithelial tissues. *Curr. Biol.* 28, 1380–1391. doi: 10.1016/j.cub.2018.03.034

de Lillo, E., Freitas-Astúa, J., Kitajima, E. W., Ramos-González, P. L., Simoni, S., Tassi, A. D., et al. (2021). Phytophagous mites transmitting plant viruses: update and perspectives. *Entomol. Gen.* 41, 439–462. doi: 10.1127/ENTOMOLOGIA/2021/1283

Di Mattia, J., Vernerey, M.-S., Yvon, M., Pirolles, E., Villegas, M., Gaafar, Y., et al. (2020). Route of a multipartite nanovirus across the body of its aphid vector. *J. Virol.* 94:e01998–19. doi: 10.1128/jvi.01998-19

Dietzgen, R. G., Freitas-Astúa, J., Chabi-Jesus, C., Ramos-González, P. L., Goodin, M. M., Kondo, H., et al. (2018). Dichorhaviruses in their host plants and mite vectors. *Adv. Virus Res.* 102, 119–148. doi: 10.1016/bs.aivir.2018.06.001

Djiane, A., Shimizu, H., Wilkin, M., Mazleyrat, S., Jennings, M. D., Avis, J., et al. (2011). Su(dx) E3 ubiquitin ligase-dependent and -independent functions of polychaetoid, the *Drosophila* ZO-1 homologue. *J. Cell Biol.* 192, 189–200. doi: 10.1083/jcb.201007023

Dolja, V. V., and Koonin, E. V. (2018). Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus Res.* 244, 36–52. doi: 10.1016/j.virusres.2017.10.020

Fasano, A. (2012). Zonulin, regulation of tight junctions, and autoimmune diseases. *Ann. N. Y. Acad. Sci.* 1258, 25–33. doi: 10.1111/j.1749-6632.2012.06538.x

Fehon, R. G., Dawson, L. A., and Artavanis-Tsakonas, S. (1994). A Drosophila homologue of membrane-skeleton protein 4.1 is associated with septate junctions and is encoded by the coracle gene. *Development* 120, 545–557. doi: 10.1242/dev.120.3.545

Foster, T. M., Lough, T. J., Emerson, S. J., Lee, R. H., Bowman, J. L., Forster, R. L. S., et al. (2002). A surveillance system regulates selective entry of rna into the shoot apex. *Plant Cell* 14, 1497–1508. doi: 10.1105/tpc.001685

Freitas-Astúa, J., Ramos-González, P. L., Arena, G. D., Tassi, A. D., and Kitajima, E. W. (2018). Brevipalpus-transmitted viruses: parallelism beyond a common vector or convergent evolution of distantly related pathogens? *Curr. Opin. Virol.* 33, 66–73. doi: 10.1016/J.COVIRO.2018.07.010

Garita, L. C., Tassi, A. D., Calegario, R. F., Kitajima, E. W., Carbonell, S. A. M., and Freitas-Astúa, J. (2013). Common bean: experimental indicator plant for citrus leprosis virus C and some other cytoplasmic-type brevipalpus-transmitted viruses. *Plant Dis.* 97, 1346–1351. doi: 10.1094/PDIS-12-12-1143-RE

Genova, J. L., and Fehon, R. G. (2003). Neuroglian, gliotactin, and the Na+/k+ ATPase are essential for septate junction function in *Drosophila*. *J. Cell Biol.* 161, 979–989. doi: 10.1083/jcb.200212054

Godenschwege, T. A., Kristiansen, L. V., Uthaman, S. B., Hortsch, M., and Murphey, R. K. (2006). A conserved role for *Drosophila* neuroglian and human L1-CAM in central-synapse formation. *Curr. Biol.* 16, 12–23. doi: 10.1016/j.cub.2005.11.062

Grbić, M., Van Leeuwen, T., Clark, R. M., Rombauts, S., Rouzé, P., Grbić, V., et al. (2011). The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. *Nature* 479, 487–492. doi: 10.1038/nature10640

Gupta, A. K., Scully, E. D., Palmer, N. A., Geib, S. M., Sarath, G., Hein, G. L., et al. (2019). Wheat streak mosaic virus alters the transcriptome of its vector, wheat curl mite (*Aceria tosichella* keifer), to enhance mite development and population expansion. *J. Gen. Virol.* 100, 889–910. doi: 10.1099/jgv.0.001256

Hall, S., and Ward, R. E. (2016). Septate junction proteins play essential roles in morphogenesis throughout embryonic development in *Drosophila*. *G3 Genes Genomes Genet.* 6, 2375–2384. doi: 10.1534/g3.116.031427

Hijazi, A., Masson, W., Augé, B., Waltzer, L., Haenlin, M., and Roch, F. (2009). Boudin is required for septate junction organisation in *Drosophila* and codes for a diffusible protein of the Ly6 superfamily. *Development* 136, 2199–2209. doi: 10.1242/dev.033845

Hildebrandt, A., Pflanz, R., Behr, M., Tarp, T., Riedel, D., and Schuh, R. (2015). Bark beetle controls epithelial morphogenesis by septate junction maturation in *Drosophila*. *Dev. Biol.* 400, 237–247. doi: 10.1016/j.ydbio.2015.02.008

Hogenhout, S. A., Ammar, E.-D., Whitfield, A. E., and Redinbaugh, M. G. (2008). Insect vector interactions with persistently transmitted viruses. *Annu. Rev. Phytopathol.* 46, 327–359. doi: 10.1146/annurev.phyto.022508.092135

Hortsch, M., Paisley, K. L., Tian, M. Z., Qian, M., Bouley, M., and Chandler, R. (2002). The axonal localization of large *Drosophila* ankyrin2 protein isoforms is essential for neuronal functionality. *Mol. Cell. Neurosci.* 20, 43–55. doi: 10.1006/mcne.2002.1113

Hull, R. (2013). *Comparative Plant Virology*. Cambridge, MA: Academic Press.

Izumi, Y., and Furuse, M. (2014). Molecular organization and function of invertebrate occluding junctions. *Semin. Cell Dev. Biol.* 36, 186–193. doi: 10.1016/j.semcdb.2014.09.009

Izumi, Y., Motoishi, M., Furuse, K., and Furuse, M. (2016). A tetraspanin regulates septate junction formation in *Drosophila* midgut. *J. Cell Sci.* 129, 1155–1164. doi: 10.1242/jcs.180448

Izumi, Y., Yanagihashi, Y., and Furuse, M. (2012). A novel protein complex, Mesh-Ssk, is required for septate junction formation in the *Drosophila* midgut. *J. Cell Sci.* 125, 4923–4933. doi: 10.1242/jcs.112243

Kitajima, E. W., Chagas, C. M., and Rodrigues, J. C. V. (2003). Brevipalpus-transmitted plant virus and virus-like diseases: cytopathology and some recent cases. *Exp. Appl. Acarol.* 30, 135–160. doi: 10.1023/B:APPA.0000006546.55305.e3

Kitajima, E. W., Rodrigues, J. C. V., and Freitas-Astua, J. (2010). An annotated list of ornamentals naturally found infected by *Brevipalpus* mite-transmitted viruses. *Sci. Agric.* 67, 348–371. doi: 10.1590/S0103-90162010000300014

Koch, I., Schwarz, H., Beuchle, D., Goellner, B., Langegger, M., and Aberle, H. (2008). Drosophila ankyrin 2 Is required for synaptic stability. *Neuron* 58, 210–222. doi: 10.1016/j.neuron.2008.03.019

Kondo, H., Chiba, S., Maruyama, K., Andika, I. B., and Suzuki, N. (2019). A novel insect-infecting virga/nege-like virus group and its pervasive endogenization into insect genomes. *Virus Res.* 262, 37–47. doi: 10.1016/j.virusres.2017.11.020

Kondo, H., Fujita, M., Hisano, H., Hyodo, K., Andika, I. B., and Suzuki, N. (2020). Virome analysis of aphid populations that infest the barley field: the discovery of two novel groups of Nege/Kita-like viruses and other novel RNA Viruses. *Front. Microbiol.* 11:509. doi: 10.3389/fmicb.2020.00509

Kondo, H., Maeda, T., and Tamada, T. (2003). Orchid fleck virus: Brevipalpus californicus mite transmission, biological properties and genome structure. *Exp. Appl. Acarol.* 30, 215–223. doi: 10.1023/b:appa.0000006550.88615.10

Kuhn, J. H., Adkins, S., Alioto, D., Alkhovsky, S. V., Amarasinghe, G. K., Anthony, S. J., et al. (2020). 2020 taxonomic update for phylum *Negarnaviricota* (Riboviria: Orthornavirae), including the large orders *Bunyavirales* and *Mononegavirales*. *Arch. Virol.* 165, 3023–3072. doi: 10.1007/s00705-020-04731-2

Lamb, R. S., Ward, R. E., Schweizer, L., and Fehon, R. G. (1998). Drosophila coracle, a member of the protein 4.1 superfamily, has essential structural functions in the septate junctions and developmental functions in embryonic and adult epithelial cells. *Mol. Biol. Cell* 9, 3505–3519. doi: 10.1091/mbc.9.12.3505

Leastro, M. O., Kitajima, E. W., Silva, M. S., Resende, R. O., and Freitas-Astúa, J. (2018). Dissecting the subcellular localization, intracellular trafficking, interactions, membrane association, and topology of citrus leprosis Virus C proteins. *Front. Plant Sci.* 9:1299. doi: 10.3389/fpls.2018.01299

Leastro, M. O., Ortega-Castro, D. Y., Freitas-Astúa, J., Kitajima, E. W., Pallas, V., Sanchéz-Navarro, J. A. et al. (2020). Citrus leprosis virus C encodes three proteins with gene silencing suppression activity. *Front. Microbiol.* 11:1231. doi: 10.3389/fmicb.2020.01231

Lefeuvre, P., Martin, D. P., Elena, S. F., Shepherd, D. N., Roumagnac, P., and Varsani, A. (2019). Evolution and ecology of plant viruses. *Nat. Rev. Microbiol.* 17, 632–644. doi: 10.1038/s41579-019-0232-3

Martín-Hernández, A. M., and Baulcombe, D. C. (2008). Tobacco rattle virus 16-kilodalton protein encodes a suppressor of rna silencing that allows transient viral entry in meristems. *J. Virol.* 82, 4064–4071. doi: 10.1128/JVI.02438-07

Mateo, M., Generous, A., Sinn, P. L., and Cattaneo, R. (2015). Connections matter - How viruses use cell-cell adhesion components. *J. Cell Sci.* 128, 431–439. doi: 10.1242/jcs.159400

Medigeshi, G. R., Hirsch, A. J., Brien, J. D., Uhrlaub, J. L., Mason, P. W., Wiley, C., et al. (2009). West nile virus capsid degradation of claudin proteins disrupts epithelial barrier function. *J. Virol.* 83, 6125–6134. doi: 10.1128/JVI.02617-08

Melzer, M. J., Simbajon, N., Carillo, J., Borth, W. B., Freitas-Astúa, J., Kitajima, E. W., et al. (2013). A cilevirus infects ornamental hibiscus in Hawaii. *Arch. Virol.* 158, 2421–2424. doi: 10.1007/s00705-013-1745-0

Morozov, S. Y., Lazareva, E. A., and Solovyev, A. G. (2020). Sequence relationships of RNA helicases and other proteins encoded by blunervirus RNAs highlight recombinant evolutionary origin of kitaviral genomes. *Front. Microbiol.* 11:561092. doi: 10.3389/fmicb.2020.561092

Mühlebach, M. D., Mateo, M., Sinn, P. L., Prüfer, S., Uhlig, K. M., Leonard, V. H. J., et al. (2011). Adherens junction protein nectin-4 is the epithelial receptor for measles virus. *Nature* 480, 530–533. doi: 10.1038/nature10639

Nakatsu, Y., Ma, X., Seki, F., Suzuki, T., Iwasaki, M., Yanagi, Y., et al. (2013). Intracellular transport of the measles virus ribonucleoprotein complex is mediated by rab11A-positive recycling endosomes and drives virus release from the apical membrane of polarized epithelial cells. *J. Virol.* 87, 4683–4693. doi: 10.1128/jvi.02189-12

Nam, S. C., and Choi, K. W. (2006). Domain-specific early and late function of Dpatj in *Drosophila* photoreceptor cells. *Dev. Dyn.* 235, 1501–1507. doi: 10.1002/dvdy.20726

Nava, P., López, S., Arias, C. F., Islas, S., and González-Mariscal, L. (2004). The rotavirus surface protein VP8 modulates the gate and fence function of tight junctions in epithelial cells. *J. Cell Sci.* 117, 5509–5519. doi: 10.1242/jcs.01425

Navia, D., Novelli, V. M., Rombauts, S., Freitas-Astúa, J., Santos de Mendonça, R., Nunes, M. A., et al. (2019). Draft genome assembly of the false spider mite *Brevipalpus yothersi*. *Microbiol. Resour. Announc.* 8:e01563–18. doi: 10.1128/mra.01563-18

Niu, J., Zhang, W., Sun, Q.-Z., and Wang, J.-J. (2019). Three novel RNA viruses in the spider mite *Tetranychus urticae* and their possible interactions with the host RNA interference response. *J. Invertebr. Pathol.* 166:107228. doi: 10.1016/j.jip.2019.107228

O'Brien, C. A., McLean, B. J., Colmant, A. M. G., Harrison, J. J., Hall-Mendelin, S., van den Hurk, A. F., et al. (2017). Discovery and characterisation of castlerea virus, a new species of Negevirus isolated in Australia. *Evol. Bioinform. Online* 13:1176934317691269. doi: 10.1177/1176934317691269

Olmedo-Velarde, A., Hu, J., and Melzer, M. J. (2021). A Virus Infecting Hibiscus rosa-sinensis represents an evolutionary link between Cileviruses and Higreviruses. *Front. Microbiol.* 12:660237. doi: 10.3389/fmicb.2021.660237

Paul, S. M., Ternet, M., Salvaterra, P. M., and Beitel, G. J. (2003). The Na+/K+ ATPase is required for septate junction function and epithelial tube-size control in the *Drosophila* tracheal system. *Development* 130, 4963–4974. doi: 10.1242/dev.00691

Quito-Avila, D. F., Freitas-Astúa, J., and Melzer, M. J. (2021). " Bluner-, cile-, and higreviruses (Kitaviridae)," in *Reference Module in Life Sciences*, ed. B. D. Roitberg (Elsevier: Amstersam), 247–251. doi: 10.1016/B978-0-12-809633-8.21248-X

Ramos-González, P. L., Chabi-Jesus, C., Banguela-Castillo, A., Tassi, A. D., Rodrigues, M., da, C., et al. (2018). Unveiling the complete genome sequence of clerodendrum chlorotic spot virus, a putative dichorhavirus infecting ornamental plants. *Arch. Virol.* 163, 2519–2524. doi: 10.1007/s00705-018-3857-z

Ramos-González, P. L., dos Santos, G. F., Chabi-Jesus, C., Harakava, R., Kitajima, E. W., and Freitas-Astúa, J. (2020). Passion fruit green spot virus genome harbors a new orphan ORF and highlights the flexibility of the 5'-end of the RNA2 segment across cileviruses. *Front. Microbiol.* 11:206. doi: 10.3389/fmicb.2020.00206

Reynolds, G. M., Harris, H. J., Jennings, A., Hu, K., Grove, J., Lalor, P. F., et al. (2008). Hepatitis C virus receptor expression in normal and diseased liver tissue. *Hepatology* 47, 418–427. doi: 10.1002/hep.22028

Roy, A., Hartung, J. S., Schneider, W. L., Shao, J., Leon, M. G., Melzer, M. J., et al. (2015). Role bending: Complex relationships between viruses, hosts, and vectors related to citrus leprosis, an emerging disease. *Phytopathology* 105, 872–884. doi: 10.1094/PHYTO-12-14-0375-FI

Schulte, J., Tepass, U., and Auld, V. J. (2003). Gliotactin, a novel marker of tricellular junctions, is necessary for septate junction development in *Drosophila*. *J. Cell Biol.* 161, 991–1000. 161. doi: 10.1083/jcb.200303192

Schwabe, T., Bainton, R. J., Fetter, R. D., Heberlein, U., and Gaul, U. (2005). GPCR signaling is required for blood-brain barrier formation in *Drosophila*. *Cell* 123, 133–144. doi: 10.1016/j.cell.2005.08.037

Sen, A., Nagy-Zsvér-Vadas, Z., and Krahn, M. P. (2012). Drosophila PATJ supports adherens junction stability by modulating Myosin light chain activity. *J. Cell Biol.* 199, 685–698. doi: 10.1083/jcb.201206064

Seppa, M. J., Johnson, R. I., Bao, S., and Cagan, R. L. (2008). Polychaetoid controls patterning by modulating adhesion in the *Drosophila pupal* retina. *Dev. Biol.* 318, 1–16. doi: 10.1016/j.ydbio.2008.02.022

Sun, B., and Salvaterra, P. M. (1995). Characterization of nervana, a drosophila melanogaster neuron-specific glycoprotein antigen recognized by anti-horseradish peroxidase antibodies. *J. Neurochem.* 65, 434–443. doi: 10.1046/j.1471-4159.1995.65010434.x

Tanentzapf, G., Smith, C., McGlade, J., and Tepass, U. (2000). Apical, lateral, and basal polarization cues contribute to the development of the follicular epithelium during *Drosophila oogenesis*. *J. Cell Biol.* 151, 891–904. doi: 10.1083/jcb.151.4.891

Tassi, A. D., Garita-Salazar, L. C., Amorim, L., Novelli, V. M., Freitas-Astúa, J., Childers, C. C., et al. (2017). Virus-vector relationship in the *Citrus leprosis* pathosystem. *Exp. Appl. Acarol.* 71, 227–241. doi: 10.1007/s10493-017-0123-0

Timpe, J. M., Stamataki, Z., Jennings, A., Hu, K., Farquhar, M. J., Harris, H. J., et al. (2008). Hepatitis C virus cell-cell transmission in hepatoma cells in the presence of neutralizing antibodies. *Hepatology* 47, 17–24. doi: 10.1002/hep.21959

Tugizov, S. (2021). Virus-associated disruption of mucosal epithelial tight junctions and its role in viral transmission and spread. *Tissue Barries* 9:194374. doi: 10.1080/21688370.2021.1943274

Vasilakis, N., and Tesh, R. B. (2015). Insect-specific viruses and their potential impact on arbovirus transmission. *Curr. Opin. Virol.* 15, 69–74. doi: 10.1016/J.COVIRO.2015.08.007

Vasilakis, N., Forrester, N. L., Palacios, G., Nasar, F., Savji, N., Rossi, S. L., et al. (2013). Negevirus: a proposed new taxon of insect-specific viruses with wide geographic distribution. *J. Virol.* 87, 2475–2488. doi: 10.1128/JVI.00776-12

Vasilyeva, N. A., Loktyushov, E. V., Bychkov, M. L., Shenkarev, Z. O., and Lyukmanova, E. N. (2017). Three-finger proteins from the Ly6/uPAR family: functional diversity within one structural motif. *Biochemistry* 82, 1702–1715. doi: 10.1134/S0006297917130090

Venema, D. R., Zeev-Ben-Mordehai, T., and Auld, V. J. (2004). Transient apical polarization of Gliotactin and Coracle is required for parallel alignment of wing hairs in *Drosophila*. *Dev. Biol.* 275, 301–314. doi: 10.1016/j.ydbio.2004.07.040

Ward, R. E. IV, Schweizer, L., Lamb, R. S., and Fehon, R. G. (2001). The protein 4.1, ezrin, radixin, moesin (FERM) domain of drosophila coracle, a cytoplasmic component of the septate junction, provides functions essential for embryonic development and imaginal cell proliferation. *Genetics* 159, 219–228. doi: 10.1093/genetics/159.1.219

Welbourn, W. C., Ochoa, R., Kane, E. C., and Erbe, E. F. (2003). Morphological observations on Brevipalpus phoenicis (Acari: Tenuipalpidae) including comparisons with B . californicus and B . obovatus. *Exp. Appl. Acarol.* 30, 107–133. doi: 10.1023/B:APPA.0000006545.40017.a0

Whitfield, A. E., Falk, B. W., and Rotenberg, D. (2015). Insect vector-mediated transmission of plant viruses. *Virology* 479–480, 278–289. doi: 10.1016/j.virol.2015.03.026

Whitfield, A. E., Huot, O. B., Martin, K. M., Kondo, H., and Dietzgen, R. G. (2018). Plant rhabdoviruses—their origins and vector interactions. *Curr. Opin. Virol.* 33, 0198–207. doi: 10.1016/j.coviro.2018.11.002

Williams, M. J. (2009). The Drosophila cell adhesion molecule Neuroglian regulates Lissencephaly-1 localisation in circulating immunosurveillance cells. *BMC Immunol.* 10:17. doi: 10.1186/1471-2172-10-17

Woods, D. F., and Bryant, P. J. (1991). The discs-large tumor suppressor gene of *Drosophila* encodes a guanylate kinase homolog localized at septate junctions. *Cell* 66, 451–464. doi: 10.1016/0092-8674(81)90009-X

Xu, C., Tang, H. W., Hung, R. J., Hu, Y., Ni, X., Housden, B. E., et al. (2019). The septate junction protein Tsp2A restricts intestinal stem cell activity *via* endocytic regulation of aPKC and Hippo signaling. *Cell Rep.* 26, 670–688. doi: 10.1016/j.celrep.2018.12.079

Yonekawa, K., and Harlan, J. M. (2005). Targeting leukocyte integrins in human diseases. *J. Leukoc. Biol.* 77, 129–140. doi: 10.1189/jlb.0804460

Yumine, N., Matsumoto, Y., Ohta, K., Fukasawa, M., and Nishio, M. (2019). Claudin-1 inhibits human parainfluenza virus type 2 dissemination. *Virology* 531, 93–99. doi: 10.1016/j.virol.2019.01.031

Zhang, Y. X., Chen, X., Wang, J. P., Zhang, Z. Q., Wei, H., Yu, H. Y., et al. (2019). Genomic insights into mite phylogeny, fitness, development, and reproduction. *BMC Genomics* 20:954. doi: 10.1186/s12864-019-6281-1

Zhao, L., Mwaliko, C., Atoni, E., Wang, Y., Zhang, Y., Zhan, J., et al. (2019). Characterization of a novel tanay virus isolated from anopheles sinensis mosquitoes in yunnan, China. *Front. Microbiol.* 10:1963. doi: 10.3389/fmicb.2019.01963

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership