Received: 6 December 2020 | Accepted: 27 August 2021

DOI: 10.1111/ppa.13472

Plant Pathology WILEY

ORIGINAL ARTICLE

Biological and molecular characterization of two closely related carlaviruses affecting brassica plants

Agatha M. Oliveira[1] | Leilane K. Rodrigues[1] | Camila Chabi-Jesus[2,3] | Alexandre L. R. Chaves[1] | Elliot W. Kitajima[4] | Ricardo Harakava[2] | Alexander Banguela-Castillo[2] | Pedro L. Ramos-González[2] | Marcelo Eiras[1]

[1]Laboratorio de Fitovirologia e Fisiopatologia, Centro de Pesquisa de Sanidade Vegetal, Instituto Biológico, São Paulo, SP, Brazil

[2]ULR Biologia Molecular Aplicada, Instituto Biológico, São Paulo, SP, Brazil

[3]PPG Microbiologia Agrícola, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, Brazil

[4]Núcleo de Apoio à Pesquisa em Microscopia Eletrônica Aplicada a Agricultura, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, Brazil

Correspondence
Pedro L. Ramos-González, ULR Biologia Molecular Aplicada, Instituto Biológico, São Paulo, SP, 04014-900, Brazil.
Email: plrg1970@gmail.com

Funding information
Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant/Award Number: 2014/22594-2, 2015/50076-9, 2017/18910-4, 2017/50334-3, 2018/17287-4 and 2019/06853-1; Conselho Nacional de Desenvolvimento Científico e Tecnológico, Grant/Award Number: 133331/2017-2; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

Abstract

Carlaviruses are plant-infecting viruses with flexuous filamentous particles of approximately 650 nm in length and a positive single-stranded hexacistronic RNA molecule as the genome. In this study, we analysed 14 samples of brassicas plants, reportedly affected by carlaviruses, that were collected in distant and edaphoclimatic distinct regions in Brazil. The genomes of four viral isolates detected in leaf kale (*Brassica oleracea* var. *acephala*) plants displayed the typical genomic organization of carlaviruses, subdivided into two contrasting identity profiles. Isolates T25, T107, and T110 showed a higher sequence identity among them (c.93%) than with T90 (c.67%), which had the highest nucleotide sequence identity (85.2%) with the only documented genomic fragment, approximately 1 kb of the RNA 3′-end, of cole latent virus (CoLV). Identity values among the other three isolates and CoLV were consistently lower than 78.6%. Nucleotide and amino acid sequence identity values of the replicase cistron of the isolates T25 and T90 are below the threshold for species demarcation in the genus *Carlavirus*. Upon mechanical transmission, these two isolates induced different symptoms in brassicas and solanaceous plants. Overall, data indicated that these viruses belong to two related but distinct species of carlaviruses. While T90 was recognized as an isolate of CoLV, the remaining isolates were considered members of a new tentative carlavirus named Cole mild mosaic virus (CoMMV). Current work provides convincing support for the taxonomic status of the species *Cole latent virus*, enlarges the known diversity of carlaviruses, and makes available new molecular tools to improve surveys of brassica-infecting viruses.

KEYWORDS
brassicas, *Carlavirus*, high-throughput sequencing, leaf kale, mixed infection, *Nicotiana megalosiphon*

1 | INTRODUCTION

The genus *Brassica*, family Brassicaceae, comprises herbaceous crop plants distributed worldwide, which together provide the greatest diversity of food products used by humans from a single plant genus (Melo et al., 2019). Brazil ranks among the largest vegetable producers in the world (Carvalho et al., 2016). In 2016, the national production of brassica crops reached about

200,000 tonnes and generated over $59 million in gross revenue (Anonymous, 2016).

Globally, crucifers are infected by a wide diversity of RNA and DNA viruses (Rodrigues et al., 2019, 2021; Sastry et al., 2019). Viral diseases usually decrease the marketability of the affected plants, mainly due to the reduced quality of, for instance, the leaves from leaf kale (*Brassica oleracea* var. *acephala*), mustard (*B. juncea*), and rocket (*Eruca sativa*); the inflorescences from cauliflower (*B. oleracea* var. *botrytis*) and broccoli (*B. oleracea* var. *italica*); and roots from turnip (*B. rapa*), radish (*Raphanus sativus*), and horseradish (*Armoracia rusticana*) (Kitajima, 2020; Sastry et al., 2019).

In Brazil, a partially characterized RNA virus known as cole latent virus (CoLV) has been intermittently detected affecting some varieties of leaf kale, cauliflower, and horseradish plants (Belintani et al., 2002; Eiras et al., 2008; Kitajima et al., 1970). Infected plants may show vein clearing and mild mosaic in their leaves or, although less frequently, the infection remains latent. CoLV has elongated, nonenveloped flexible particles of 650 nm in length, which are transmitted plant to plant by different species of aphids (order Hemiptera) in a noncirculative manner (Belintani et al., 2002; Bragard et al., 2013; Costa et al., 1972; Eiras et al., 2008; Mello et al., 1987). The unique combination of its biological traits and the relatively low nucleotide (nt) sequence identity with other carlaviruses of a fragment with 1040 nt from the 3′-terminal region of its genome (RefSeq accession NC_038322) led to the early assignment of CoLV to the species *Cole latent virus*, genus *Carlavirus*, in 2004 (Mayo, 2005).

The genus *Carlavirus*, family *Betaflexiviridae*, includes viruses assigned to different species that have a specific natural host range, do not cross-protect in infected common host plant species, are usually readily differentiated by serological procedures, and at the genetic level, have less than 72% nt sequence identity (or 80% deduced amino acid sequence identity) in the open reading frames (ORFs) encoding the replicase or coat protein (CP) (Adams et al., 2012). The genome of carlaviruses consists of a single-stranded RNA (ssRNA) of approximately 8 kb, 3′-end polyadenylated, and with six ORFs (Adams et al., 2004; Foster, 1992). ORF1 encodes the replicase protein (223 kDa), which comprises four conserved domains: methyltransferase (Mt), papain-like protease (P-Pro), helicase (Hel), and RNA-dependent RNA polymerase (RdRp). ORFs 2, 3, and 4, known as the triple gene block (TGB), encode proteins of 25 kDa, 12 kDa, and 7 kDa, respectively, involved in viral movement, whereas ORF5 encodes the coat protein (CP, 34 kDa). Nearest to the 3′-end, the ORF6 encodes a cysteine-rich protein with nucleic acid-binding activity (RBP, 11 kDa) that counteracts the plant gene silencing (Fujita et al., 2018). ORFs 2, 3, 4, 5, and 6 are expressed from two subgenomic RNAs of 2.6 and 1.3 kb (Foster, 1992). Viral replication occurs exclusively in the cytoplasm of the infected cells (Mello et al., 1987).

In this study, we report a survey of brassica-infecting carlaviruses throughout Brazil and the biological and molecular characterizations of some of the isolates found. We provide the first complete genome sequence of a virus of the species *Cole latent virus*, and the genome of a tentative new species we named Cole mild mosaic virus (CoMMV). Moreover, we describe the suitable use of *Nicotiana megalosiphon* plants as an indicator host for CoLV and CoMMV infections, and as an alternative way to separate mixed infections between caulimovirus cauliflower mosaic virus (CaMV) and the studied carlaviruses.

2 | MATERIALS AND METHODS

2.1 | Viral samples

Fourteen leaf samples of cultivated brassicas were collected during a survey in vegetable production areas in different regions of Brazil (Table 1, Figure 1). Leaves were cut into small fragments, transferred to Petri dishes containing $CaCl_2$ (1.6 g per gram of fresh leaf), and kept at −20°C until their use in biological tests, transmission electron microscopy observations, and molecular analyses.

2.2 | Viral transmission

Leaf extracts of leaf kale (*B. oleracea* var. *acephala*) plants of the samples T25 and T90, collected in Divinolândia, 21°30′41″S, 46°44′09″W, state of São Paulo, Brazil, in 2015, and Arapiraca, 09°45′09″S, 36°39′40″W, state of Alagoas, Brazil, in 2017, respectively, were used to inoculate a set of four plants each belonging to 27 species of the families Aizoaceae, Amaranthaceae, Brassicaceae, Cucurbitaceae, Fabaceae, Solanaceae, and Poaceae (Table 2). Leaf extracts were prepared in 0.01 M phosphate buffer, pH 6.0, containing 5% sodium sulphite, and they were mechanically inoculated into new plants using carborundum powder as abrasive. For each treatment, a mock-inoculated plant was included as a negative control. Plants were kept under greenhouse conditions (anti-aphid net, the ambient temperature in the range 20–26°C, and controlled irrigation) and any observed morphological alteration was recorded after a daily inspection. Thirty days postinoculation (dpi), leaf samples from both plants with and without symptoms were collected for further analyses. In particular, leaves from *N. megalosiphon* plants inoculated with the extract from sample T25 were considered the source of the carlavirus isolate T25 in the following characterization experiments. Free of caulimovirus infection, these plants were identified as the purified T25 (pT25) sample. Finally, the host range of isolate T25 was reassayed using pT25 leaf extracts and the same set of plant species described for the inoculation of T25 and T90 samples.

2.3 | Transmission electron microscopy

Leaves of Chinese cabbage cv. Natsume (*B. rapa* subsp. *pekinensis*) plants with symptoms, previously inoculated with sap extracts of the samples pT25 and T90, were analysed by transmission electron microscopy (TEM) after a negative contrast staining (leaf dip) (Kitajima & Nome, 1999). For the examination of the cytoplasmic inclusions, leaf tissues were fixed in paraformaldehyde (2%) plus glutaraldehyde (2.5%) in 0.05 M cacodylate

TABLE 1 List of brassica samples collected in this work

Sample (code)	Host of collection	Collection		Symptoms
		Date	Place	
T4 (1)	*Brassica oleracea* var. *acephala* (leaf kale)	02/2012	Socorro, SP	NS
T6 (2)	*Armoracia rusticana* (horseradish)	06/2012	Divinolândia, SP	NS
T9 (3)	*B. oleracea* var. *acephala*	07/2012	Monte Alegre do Sul, SP	NS
T15 (4)	*B. oleracea* var. *acephala*	—	Distrito Federal, DF	Leaf mosaic
T25 (5)	*B. oleracea* var. *acephala*	05/2015	Divinolândia, SP	Vein clearing/leaf mosaic
T36 (6)	*Brassica rapa pekinensis* (Chinese cabbage)	09/2015	Santo Antônio do Pinhal, SP	Vein clearing/leaf mosaic
T62 (7)	*B. oleracea* var. *acephala*	09/2015	Pindamonhangaba, SP	Vein clearing, mosaic, and deformation
T63 (8)	*B. oleracea* var. *acephala*	09/2015	Pindamonhangaba, SP	Mosaic, deformation, and chlorotic spots
T64 (9)	*B. oleracea* var. *italica* (broccoli)	09/2015	Pindamonhangaba, SP	Mosaic, deformation, and chlorotic spots
T89 (10)	*B. oleracea* var. *acephala*	—	Ubatuba, SP	Mosaic, whitening, and rib necrosis
T90 (11)	*B. oleracea* var. *acephala*	07/2017	Arapiraca, AL	Vein clearing/leaf mosaic
T107 (12)	*B. oleracea* var. *acephala*	08/2018	Pinhais, PR	NS
T109 (13)	*Raphanus raphanistrum* (radish)	08/2018	Pirenópolis, GO	NS
T110 (14)	*B. oleracea* var. *acephala*	08/2018	Pirenópolis, GO	NS

Note: Samples selected for further analyses are highlighted in bold.

Abbreviations: AL, Alagoas; DF, Distrito Federal; GO, Goiás; NS, no symptoms were observed; PR, Paraná; SP, São Paulo.

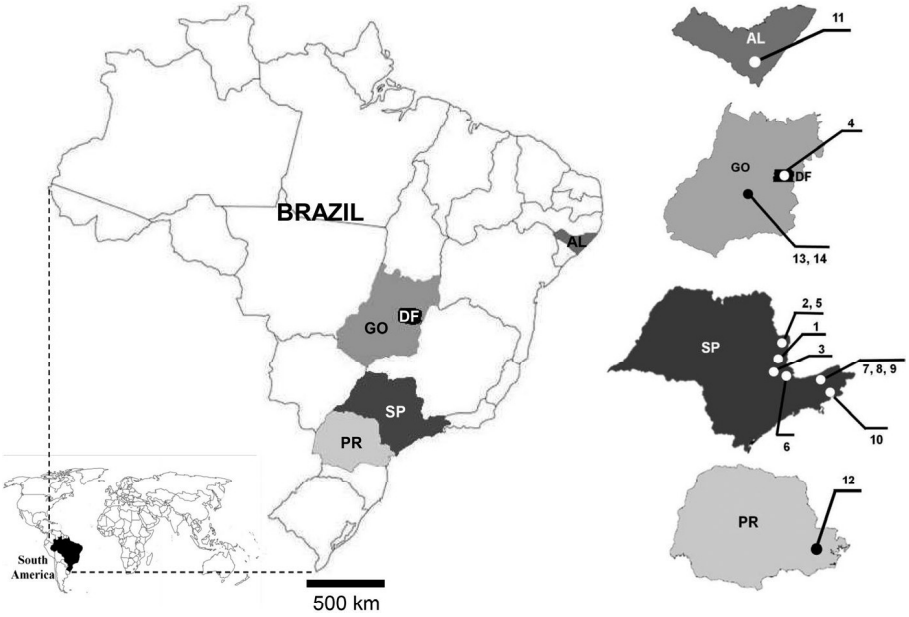

FIGURE 1 Distribution of fields where the brassica samples analysed in this work were collected during the period 2012–2017. Numbers indicate the codes of each sample as shown in Table 1. Brazilian states are identified as follows: AL, Alagoas; GO, Goiás; SP, São Paulo; PR, Paraná; DF, Distrito Federal

buffer. The fixed tissue was postfixed in 1% OsO_4, dehydrated in ethanol, infiltrated, and embedded in epoxy Spurr medium (Kitajima & Nome, 1999). Ultrathin sections were cut with a diamond knife (DiATOME) in a UTC ultramicrotome (Leica), stained with uranyl acetate and Reynold's lead citrate, and examined under a transmission electron microscope (JEM 1011; JEOL). The images were recorded digitally.

TABLE 2 Biological tests on plants of different botanical families inoculated with leaf extracts from samples pT25 and T90

Family	Species	Symptoms		RT-PCR	
		pT25	T90	pT25	T90
Aizoaceae	*Tetragonia tetragonioides* 'Nova Zelândia' (spinach)	NS	NS	–	–
Amaranthaceae	*Chenopodium giganteum*	LL	LL	+	+
	Gomphrena globosa (globe amaranth)	NS	NS	–	–
Brassicaceae	*Brassica oleracea* var. *italica* 'Precoce Piracicaba' (broccoli)	NS	NS	+	+
	B. rapa subsp. *pekinensis* 'Natsume' (Chinese cabbage)	NS	NS	+	+
	B. oleracea var. *botrytis* 'Juliana' (cauliflower)	M	NS	+	+
	B. oleracea var. *gongylodes* (kohlrabi)	M	M	+	+
	Eruca sativa 'Astro' (rocket)	NS	NS	+	+
	Raphanus sativus 'No. 19 Sakata' (radish)	NS	NS	–	–
Cucurbitaceae	*Citrullus lanatus* 'Charleston Gray' (watermelon)	NS	NS	–	–
	Cucurbita pepo 'Caserta' (zucchini)	NS	NS	–	–
	C. lanatus 'Crimson Sweet' (watermelon)	NS	NS	–	–
	C. lanatus 'Fairfax' (watermelon)	NS	NS	–	–
	Cucumis sativus 'Caipira' (cucumber)	NS	NS	–	–
Fabaceae	*Pisum sativum* 'Triofin' (pea)	NS	NS	–	–
Solanaceae	*Capsicum annuum* 'Cascadura Ikeda' (pepper)	NS	NS	–	–
	Datura stramonium	NS	NS	–	–
	Nicotiana debneyi	NS	NS	–	–
	Nicotiana glutinosa	NS	NS	–	–
	Nicotiana megalosiphon	M/DD/LD	M	+	+
	Nicotiana tabacum 'Samsun'	NS	NS	–	–
	N. tabacum 'Samsun NN'	NS	NS	–	–
	Nicotiana sylvestris	NS	NS	–	–
	N. tabacum 'White Burley'	NS	NS	–	–
	N. tabacum 'Xanthi'	NS	NS	–	–
	Solanum melongena 'Embu' (eggplant)	NS	NS	–	–
Poaceae	*Zea mays* 'Itapua 700' (maize)	NS	NS	–	–

Note: Carlavirus presence was verified by RT-PCR.

Abbreviations: DD, developmental delay, LL, local lesion; LD, leaf deformation (systemic); M, systemic mosaic; NS, no symptoms were observed.

2.4 | Extraction of total RNA and DNA

Plant total RNA extracts were obtained from dried leaves (0.03–0.05 g) conserved in $CaCl_2$ and fresh leaves (0.1 g) of plants inoculated under controlled conditions. Leaf samples were ground in a mortar in the presence of liquid N_2 and processed using TRIzol reagent following the manufacturer's recommendation (Life Technologies). DNA extraction was performed as previously described (Dellaporta et al., 1983). Both DNA and RNA extracts were resuspended in Milli-Q water, treated with DEPC for RNA solutions, and stored at −80°C.

2.5 | Reverse transcription-PCR

Approximately 500 ng of the RNA templates were used for cDNA synthesis using 50 pM of either oligo$(dT)_{21}$ or CoLV-II-specific primers (Table 3) and RNase-free water up to a final volume of 10 μl. Mixes were incubated at 70°C for 3 min and immediately placed on ice for 2 min. cDNA synthesis was completed after the addition of the reverse transcriptase enzyme buffer (1×), deoxynucleotide solution mix (dNTPs, 0.1 mM), and Moloney murine leukemia virus (MMLV) reverse transcriptase (8 U/μl). The reaction mix was further incubated at 37°C for 60 min according to the recommendation in the MMLV Reverse Transcriptase kit (Promega). Viruses were detected by PCR using cDNA solutions as templates (2 μl), specific primer pairs (0.1 μM), and GoTaq G2 Green Master Mix (1×; Promega) in a PTC100 thermal cycler (MJ Research Inc.). Thermocycling conditions were as follows: 5 min at 94°C; 35 cycles of 1 min at 94°C, 30 s at the specific annealing temperature (Table 3), and a period of extension (1 min per 1 kb of amplicon length) at 72°C; with a final extension of 2 min at 72°C. Amplicons were visualized in 1.2% agarose gel in the presence of

TABLE 3 List of primers used in RT-PCR, PCR, and 5′ RACE analyses in this work

Virus/genomic portion	Primer	Sequence (5′–3′)	Ta (°C)	Expected amplicon size (bp)	Reference
Carlaviruses/ORF6	Carla-Uni –F	GGAGTAACCGAGGTGATACC	50	120	Badge et al. (1996)
	OligodT-21 –R	TTTTTTTTTTTTTTTTTTTT			
Carlaviruses/coat protein	Carla-CP –F	GGBYTNGGBBTNCCNCANGA	50	940	Gaspar et al. (2008)
	OligodT-21 –R	TTTTTTTTTTTTTTTTTTTT			
CoLV	CoLV-I –F	ATGCGCCCAGATCCC	50	363	This work
	CoLV-II -R	TACATGTTRACGCACAACGC			
CoMMV/replicase	CoMMV F- Rep C	GAGCGCTGCAACCATAGAAA	54	532	
	CoMMV R- Rep C	CCGCATGAACGTGAATAGCA			
CoMMV/coat protein	CoMMV F- CP	TGCACGCRCAAAGCAAGAAGC	54	669	
	CoMMV F- CP	GTGTTCTCCACGTAATCAAAGCA			
CoLV/replicase	CoLV F – Rp N	ACTGAGAAGGATGTGGTGCA	54	845	
	CoLV R – Rp N	TCATCCACCACTTTCGACCAG			
	CoLV F – Rep C	GAAGTACTTACGCCTCCCCA	54	486	
	CoLV R – Rep C	ACCTCCTCCTCACTCATCCT			
CoLV/coat protein	CoLV F – CP	AACGAAAAGAGCGAGCAGGA	54	810	
	CoLV R – CP	CTGATTCCGATTCGCACTGC			
CoMMV RACE 5′-end	RACE-CoMMV	GATTACGCCAAGCTTCCAGGCCACGGCCATGCTTAACC	68	–	
	Nested-CoMMV	TCCAGACATACACTGCAATC	52	–	
CoLV RACE 5′-end	RACE-CoLV	GATTACGCCAAGCTTCGCTTCTAGACCTCTGCCATGACGCAC	68	–	
	Nested-CoLV	ACTATATTCGGACTACGTTGC	52	–	
–	Short primer-F (used in RACE analyses)	CTAATACGACTCACTATAGGGC	52	–	
TuMV/CI	CI-F	GGIVVIGTIGGIWSIGGIAARTCIAC	54	700	Ha et al. (2008)
	CI-R	ACICCRTTYTCDATDATRTTIAGTIGC			
CaMV/ORF1	CaMV-F	AGCAAGCAATCGCATAATCT	55	361	López-Moya et al. (1992)
	CaMV-R	GCTCAAAGATGAACCATGGA			
CaMV/ORF6	$CaMV_{5738}$-F	GCGGACTTCCTTTCAAGAGAATTC	55	1517	This work
	$CaMV_{7235}$-R	GGCAATGGAATCCRAGRAGG			

ethidium bromide (0.5 μg/ml), under ultraviolet light (Eagle Eye II; Stratagene Inc.). Appropriate primer pairs for the detection of turnip mosaic virus (TuMV), carlaviruses, and cauliflower mosaic virus (CaMV) (Ha et al., 2008; Jenner et al., 2002; López-Moya et al., 1992) are described in Table 3. The primer pairs CoMMV/Replicase, CoMMV/CP, CoLV/Replicase, and CoLV/CP were designed in this study. For this, viral sequences were aligned using MUSCLE, implemented in MEGA v. X (Kumar et al., 2018), and primers were designed using Primer3 (Untergasser et al., 2012) implemented in Geneious software platform v. 11.1.4 (Kearse et al., 2012; Table 3).

2.6 | High-throughput sequencing, assembly, and analysis of viral sequences

Leaves from the Chinese cabbage cv. Natsume plants inoculated with extracts of samples pT25 and T90 were collected at 20 dpi. Leaves of Chinese cabbage plants infected with isolates T25 and T90 and kale leaf plants infected with isolates T107 and T110 were individually ground into a fine powder in liquid N_2 with a pestle and mortar. RNA extracts were obtained using the TRizol reagent and further purified using the RNeasy Mini Kit (Qiagen). Quantification and estimation of the A_{260}/A_{280} ratio were carried out using a NanoDrop ND-8000 microspectrophotometer (Thermo Scientific). Five hundred nanograms of RNA were sent to the Animal Biotech Laboratory at Escola Superior de Agricultura Luiz de Queiroz, University of São Paulo (Piracicaba, SP, Brazil). Poly(A) enrichment of the RNA extracts and cDNA libraries were prepared with TruSeq Stranded mRNA Library Prep Kit (Illumina). Sequencing was performed in an Illumina HiSeq 2500 system and paired-end reads of 2 × 125 bp were generated using HiSeq SBS v. 4 High Output Kit (Illumina). Read analysis, assembly, and processing were carried out using several pieces of bioinformatics software and following previously described workflows (Chabi-Jesus et al., 2018, 2021; Ramos-González et al., 2018). Briefly, the read quality was verified with FastQC v. 0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Bases with low quality and the sequences of adapters were filtered using the package BBDuk implemented in Geneious package v. R11.2 (Kearse et al., 2012). Then, reads were assembled de novo using Trinity (Haas et al., 2013) implemented on the Galaxy platform (https://usegalaxy.org; Afgan et al., 2016). Contigs were annotated using BlastX, implemented in the Geneious software, using a plant viral genome database (https://www.ncbi.nlm.nih.gov/genbank/). The largest contigs from each library producing the best E-value score (≈ 0) with the fragment of 1040 nt from the CoLV genome, RefSeq accession number NC_038322.1 (Belintani et al., 2002), and other carlaviruses, were selected. Viral ORFs were identified in silico by the ORF finder (https://www.ncbi.nlm.nih.gov/orffinder/). The total of virus-specific reads in each high-throughput sequencing (HTS) library was carried out using BBMap (Bushnell, 2014).

2.7 | Amplification of the 5′-end of viral genomes

The 5′-end sequences of the carlavirus genomes of the isolates T25 and T90 were obtained using the Rapid Amplification of cDNA Ends (RACE) technique and specific primers (Table 3), as recommended by the manufacturer (SMARTer Kit; Clontech Laboratories). Sequentially, the amplified DNA fragments were visualized in 1.2% agarose gel in the presence of ethidium bromide (0.5 μg/ml) under ultraviolet light. Amplification products of the expected size were collected, purified using the Wizard SV gel and PCR Clean-Up System (Promega), and ligated to pGEM-T Easy vector (Promega). Recombinant plasmids from at least five colonies per viral isolate were purified and sent for Sanger sequencing at URL Applied Molecular Biology, Instituto Biológico, SP, Brazil.

2.8 | Recombination analysis

Putative recombination events between viruses of the genus *Carlavirus* were assessed using the methods RDP, GENECONV, Bootscan, Maxchi, Chimaera, SiScan, and Topal, implemented in RDP v. 5 (Martin et al., 2021). Sequences were aligned using the MUSCLE, MAFFT, and Clustal software, implemented in Geneious. Signals were only considered when detected by at least three algorithms at $p \leq 0.05$.

2.9 | Sequence comparisons and phylogenetic analyses

The complete genome and individual ORF nt sequences, and deduced amino acid (aa) sequences of viral proteins, were aligned with those of reference sequences belonging to other species in the genus *Carlavirus*, and members of the family *Betaflexiviridae* using MAFFT multiple sequence alignment software v. 7 (Katoh & Standley, 2013). The genetic distances (*D*) were calculated using MEGA X (Kumar et al., 2018). Phylogenetically informative regions of the multiple sequence alignments were selected using Noisy software (Dress et al., 2008) implemented in http://ngphylogeny.fr (Lemoine et al., 2019). Best evolutionary models for CP (LG+F+I+G4) and RdRp (LG+F+I+G4) proteins and maximum-likelihood phylogenetic trees were calculated using W-IQ-TREE software v. 1.5.5 (http://iqtree.cibiv.univie.ac.at; Trifinopoulos et al., 2016). The reliability of the inferred evolutionary relationships was assessed by 1500 bootstrap replications. Trees were viewed and edited using iTOL v. 4 (Letunic & Bork, 2019).

3 | RESULTS

3.1 | Identification of carlaviruses in field-collected samples

Fourteen brassica samples were collected from vegetable production areas throughout Brazil (Figure 1, Table 1). Molecular tests performed on these samples indicated the presence of carlaviruses

mostly in mixed infections with TuMV and CaMV, that is, two samples tested positive for co-infections with TuMV and nine with CaMV (Figure 2). Some amplicons from TuMV and CaMV detection assays were purified, sequenced, and their identities confirmed after comparisons with the reference sequences for each virus (i.e., TuMV amplicons of the isolate T4, GenBank acc. no. MK249372 and MK357827; isolate T110, MK307995 and MK370485; CaMV amplicons of the isolates T9, T25, T36, T62, T63, and T89, MK861855, MK962761, MK962764, MK962766, MK962767, and MK962768, respectively). Single infection by a carlavirus was only detected in samples T15, T64, and T90 (Figure 2, lanes 4, 9, and 11), whereas the radish sample T109 was the only one in which carlaviruses were not detected (Figure 2, lane 13). Besides T90, collected in the State of Alagoas, in the north-east of Brazil, samples T25 (State of São Paulo, south-east region), T107 (State of Paraná, south region), and T110 (State of Goiás, mid-west region) (Figure 2, lanes 5, 12, and 14) were selected for further characterization, as between them they cover contrasting edaphoclimatic and distant regions in Brazil.

Leaf extract of sample T25 from a leaf kale plant was mechanically inoculated into plants of different species to investigate a potential biological separation of the mixed infection carlavirus–CaMV. Thirty days postinoculation, PCR and RT-PCR tests were performed and the results confirmed the absence of CaMV and the infection by a carlavirus in *N. megalosiphon* plants (Figure 3). For further analyses, T25-infected *N. megalosiphon* plants free of the caulimovirus were

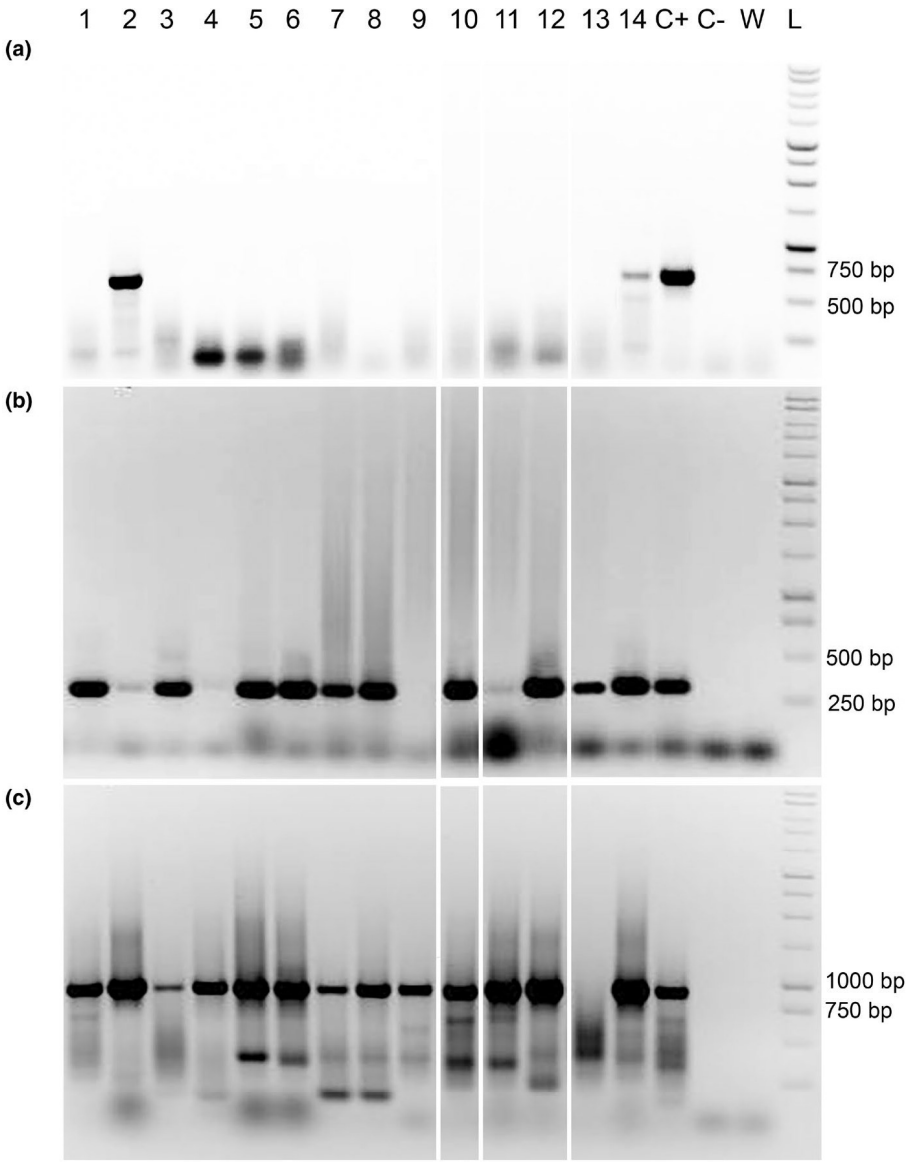

FIGURE 2 Virus detection in field-collected brassica samples by reverse transcription (RT)-PCR or PCR. Electrophoresis in 1.2% agarose gel of amplicons of (a) turnip mosaic virus (primers: CI-F/CI-R, expected amplicon: size: 700 bp), (b) cauliflower mosaic virus (primers: CaMV-F/CaMV-R, expected amplicon size: 361 bp), and (c) carlaviruses (primers: Carla-CP/OligodT-21, expected amplicon size: 940 bp). Lanes 1–14: samples identified according to the codes shown in Table 1. C+, infected leaf kale (*Brassica oleracea* var. *acephala*); C−, healthy leaf kale plant; W, blank; L, Molecular weight marker (DNA 1 kb ladder; Promega)

considered as the source of the carlavirus in sample T25, hereinafter referred to as sample pT25 (purified T25).

Leaf kale plants infected with the carlavirus in samples pT25 and T90 showed vein clearing and mosaic symptoms 30 days after mechanical inoculation (Figure 4a,b). Flexuous filamentous particles of approximately 650 nm in length were observed in leaf extract preparations (leaf dip) from these plants (Figure 4c). In ultrathin sections of leaf tissues, viral particles were observed as spiral aggregates in the cytoplasm of the infected cells (Figure 4d).

3.2 | Host range and symptoms of virus isolates T25 and T90

Leaf extracts from the leaf kale plants infected with isolate T25, derived from the sample pT25, and T90 were mechanically inoculated into a set of potential indicator host plants of the families Aizoaceae, Amaranthaceae, Brassicaceae, Cucurbitaceae, Fabaceae, Solanaceae, and Poaceae (Table 2). Both isolates induced local lesions in the inoculated leaves of *Chenopodium giganteum* plants (Amaranthaceae) (Figure 5). Inoculation with the extract of sample pT25 induced mosaic in three out of the four inoculated plants each of *B. oleracea* var. *botrytis* (cauliflower cv. Juliana) and *B. oleracea* var. *gongylodes* (kohlrabi), while plants inoculated with the extract of sample T90 expressed practically imperceptible symptoms or remained symptomless. In this case, the infection by T90 was only detected by RT-PCR (Table 2). It is worth noting that inoculation with extracts of both pT25 and T90 resulted in symptomless infections of *B. rapa* and *E. sativa*. The remaining inoculated plants of the families Solanaceae (except for *N. megalosiphon* plants), Cucurbitaceae, Fabaceae, and Poaceae neither showed symptoms nor were latently infected, as confirmed by molecular tests (Table 2).

N. megalosiphon plants inoculated with the extract of sample pT25 showed mosaic symptoms at 15 dpi, while those infected with isolate T90 remained symptomless for at least 20 dpi. After that, local chlorotic lesions appeared in the T90-inoculated leaves. Consistently, these lesions gradually turned into necrotic areas and mosaic appeared in the systemically infected leaves in the four challenged plants. However, despite the presence of these lesions in T90-infected *N. megalosiphon* plants, the symptoms caused by T25 were more severe (Figure 5). Besides the typical mosaic, T25-infected plants displayed leaf deformations, stunting, and in general showed a delayed development when compared with those infected by T90. All infections were confirmed by molecular tests.

3.3 | Molecular characterization of carlaviruses

Bioinformatics analysis of the HTS libraries from Chinese cabbage plants inoculated with extracts of samples T25, T90, T107, and T110 indicated the presence of carlavirus-specific sequences in all of them, CaMV in samples T107 and T110, and TuMV in T110. According to the analysis of nt sequence identity, related but

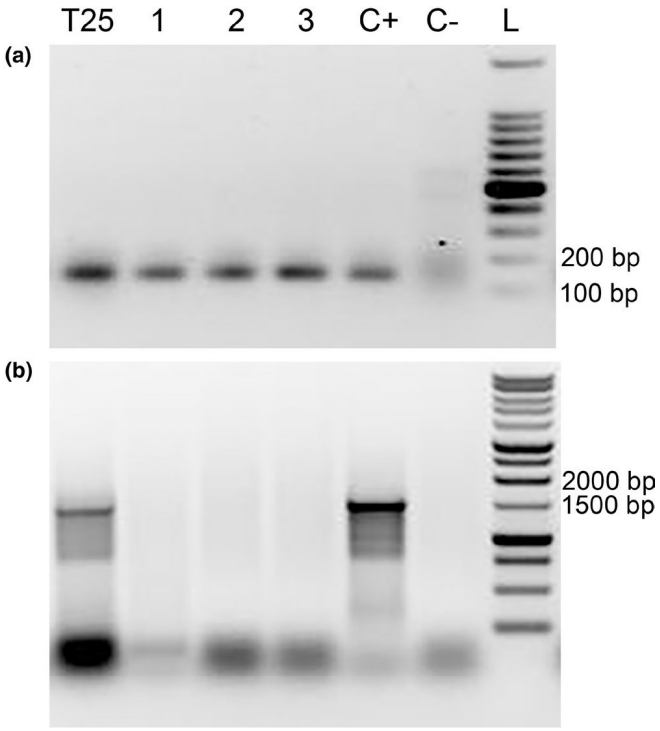

FIGURE 3 Separation of mixed infection of cole mild mosaic virus (CoMMV) and cauliflower mosaic virus (CaMV) in *Nicotiana megalosiphon* plants. Leaf extract from the leaf kale (*Brassica oleracea* var. *acephala*) plant infected with CoMMV and CaMV, sample T25, was mechanically inoculated in *N. megalosiphon* plants. Electrophoresis in 1.2% agarose gel of reverse transcription-PCR or PCR products for the detection of (a) CoMMV, and (b) CaMV, respectively. Lane T25, leaf kale T25 leaf extract. Lanes 1–3, three *N. megalosiphon* plants inoculated with extract of sample T25 containing a mixed infection CoMMV–CaMV. Lane C+, CoMMV-infected brassica plant (a) and CaMV-infected brassica plant (b). Lane C−, healthy *N. megalosiphon* plant. L, Molecular weight marker (DNA 1 kb ladder; Promega)

distinct carlaviruses were detected. The virus in sample T90 markedly diverged from the carlavirus sequences recovered from the remaining samples, that is, pT25, T107, and T110, which showed higher nt sequence identities among themselves. The carlavirus isolates T25 and T90 were further analysed by RACE to obtain the complete viral genomes. Considering the genomic sequences of carlaviruses in samples pT25 and T90 as references, we determined that the HTS libraries of samples pT25, T90, T107, and T110 contained 3.7% (921,344 out of 25.1 million reads), 27.4% (7,914,013 out of 28.8 million reads), 1.4% (201,734 out of 15.2 million of reads), and 2.7% (337,316 out of 12.5 million reads) of carlavirus-derived reads, respectively.

Excluding the poly-A tails at their 3′-ends, the complete genome sequences of the isolates T25 and T90 have a length of 8547 and 8512 nt, respectively. The genomes of both isolates contain six ORFs, arranged in a typical organization of carlaviruses (Figure 6). The genomes of T25 and T90 show 68.0% nt sequence identity between them, and they show the highest nt sequence identities, 78.6% and

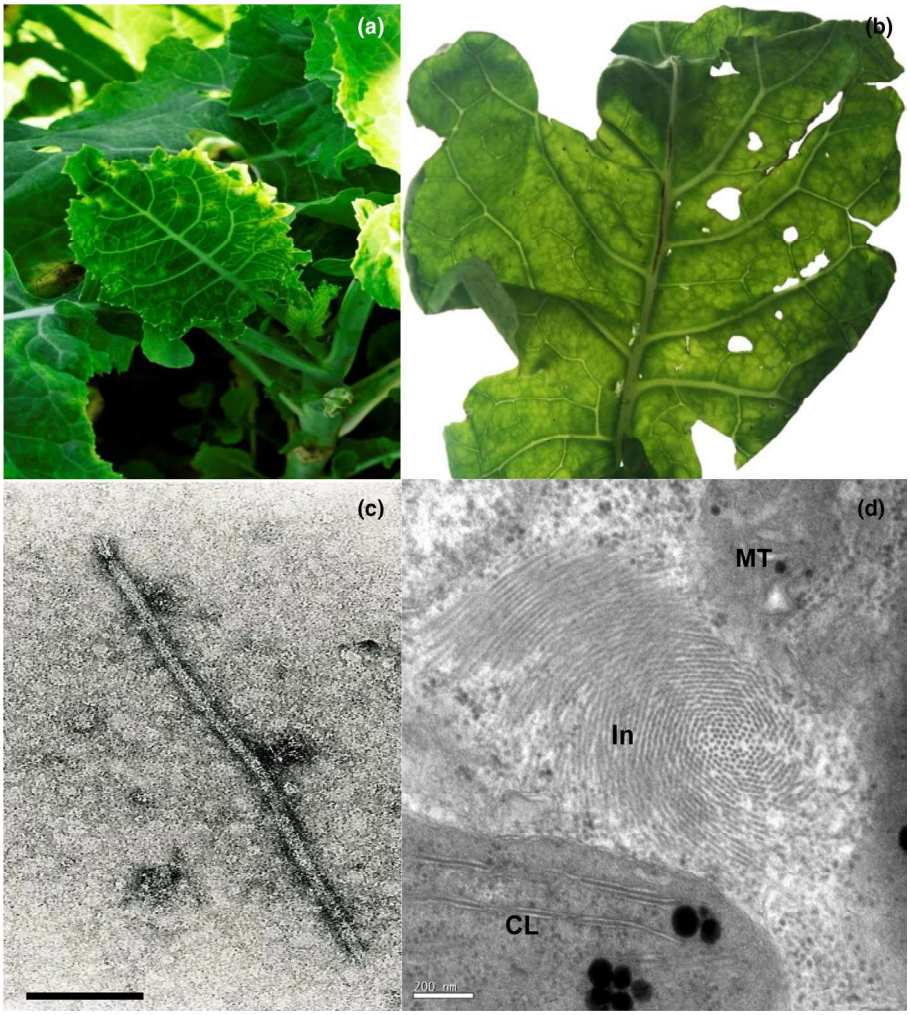

FIGURE 4 Kale (*Brassica oleracea* var. *acephala*) leaves showing symptoms of vein clearing and mosaic after infection with the carlavirus isolates (a) T25 from Divinolândia, State of São Paulo, and (b) T90, from Arapiraca, State of Alagoas. (c) Transmission electron microscopy from a plant inoculated with isolate T25, showing an elongated and flexuous particle, approximately 650 nm long, presumed virion of CoMMV; bar 150 nm. (d) Ultrathin sections of a leaf of a kale plant inoculated with an extract of sample T90 depicting typical cytoplasmic inclusions formed by aggregates of elongated particles, putative CoLV virions organized in a spiral shape, in a leaf parenchyma cell; bar 200 nm. MT, mitochondrion; CL, chloroplast; In, cytoplasmic inclusion [Colour figure can be viewed at wileyonlinelibrary.com]

85.2%, respectively, with the only known genomic fragment, approximately 1 kb of the RNA 3′-end, of CoLV (RefSeq NC_038322) (Belintani et al., 2002). Considering the portion of the ORF5 inside the CoLV sequence available (642 nt, approximately two-thirds of this gene in carlaviruses), the corresponding fragments in the genomes of T25 and T90 show 78.0% and 81.5% nt sequence identities, and 97.6% and 99.2% aa sequence identities, respectively. The ORF6 (342 nt) of CoLV, the only known complete ORF of that isolate, shows higher identity to the ortholog of T90 (nt: 90.8% and aa: 92.1%) than to that of T25 (nt: 77.6% and aa: 77.2%).

Paired comparisons of the nt and deduced aa sequences from each ORF of isolates T25 and T90 revealed that ORF5 (nt: 73.8% and aa: 83.8%) and ORF6 (nt: 76.3% and aa: 77.8%) are the most conserved between them, whereas ORF1 (nt: 66.7% and aa: 71.0%), and the three genes of the TGB (nt sequences of TGB1: 64.8%, TGB2: 69.1%, and TGB3: 61.5%; and aa sequences of TGB1: 67.2%, TGB2: 74.1%, and TGB3: 61.5%) showed the maximum divergences (Table 4, Table S1). Given the higher identity of T90 with the known fragment of CoLV than with the genomic sequence of T25, T90 was considered an isolate of CoLV, while the isolate T25 was assigned to a tentative new species of carlavirus named Cole mild mosaic virus (CoMMV). The complete genome sequences of CoMMV_T25 and CoLV_T90 were deposited in the GenBank as accessions MK684348 and MK770418, respectively.

The near-completed sequences of the isolates T107 and T110 comprise 8543 nt (GenBank accession MZ189739) and 8546 nt (MZ189738), respectively. They show a high nt sequence identity (>92%) with the genome of CoMMV_T25, whereas with the genome of CoLV_T90, the values are lower than 68.0%. The identity profiles of the isolates T107 and T110 are almost identical and show the same trend of the profiles shown by the comparison between CoMMV_T25 and CoLV_T90. The identity profiles of the comparison among the genomes of T107 and T110 with CoMMV_T25, except for three regions in the first half of the ORF1, are always over a threshold of 90% identity (Figure 6b,d). According to these results, we considered T107 and T110 as isolates of CoMMV.

In the genome of CoMMV_T25, in the 5′–3′ direction, ORF1 (replicase, 1964 aa, nt positions 87–5981) is followed by a sequence of three genes, ORF2 (229 aa, nt positions 6030–6719), ORF3 (108 aa, nt positions 6697–7023), and ORF4 (62 aa, nt positions 7023–7211), recognized as the triple gene block. ORF5 (coat protein, 305 aa, nt positions 7236–8153) and ORF6 (RBP, 113 aa, nt positions 8153–8494) are located at the 3′-end of the genome, before the poly-A tail.

The genomes of CoMMV_T25 and CoLV_T90 were further compared with the carlavirus sequences publicly available. When ORFs were independently analysed, the values of the nt and deduced aa sequence identities were always lower than 71.3% and 75.7%,

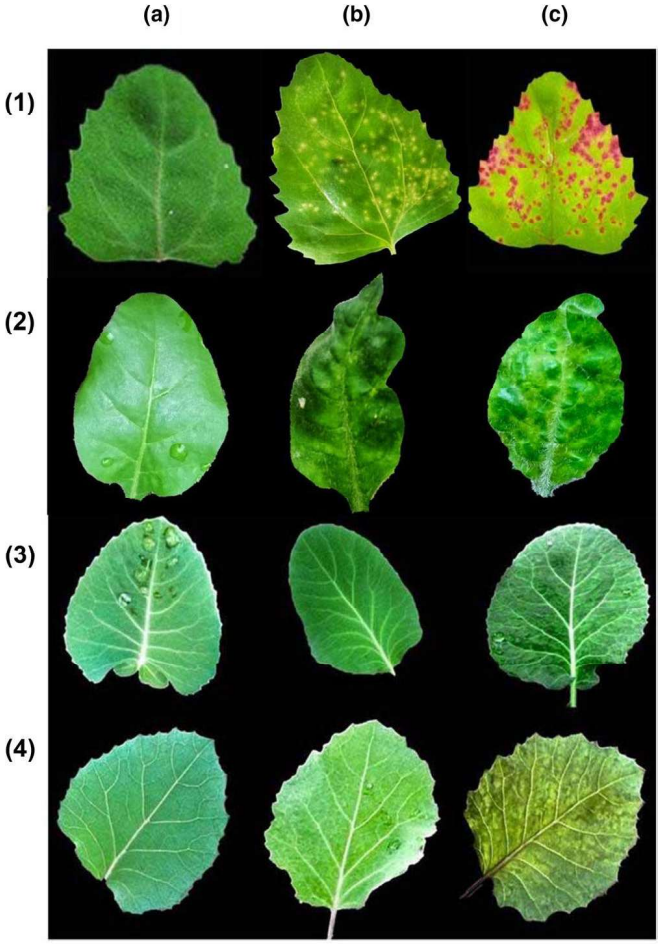

FIGURE 5 Symptoms after inoculation with the carlaviruses (b) cole latent virus_T90, and (c) cole mild mosaic virus_T25. (1) *Chenopodium giganteum* local lesions, (2) *Nicotiana megalosiphon* systemic infection, (3) *Brassica oleracea* var. *botrytis* systemic infection, and (4) *B. oleracea* var. *gongylodes* systemic infection. Column (a) depicts healthy leaves for each plant species [Colour figure can be viewed at wileyonlinelibrary.com]

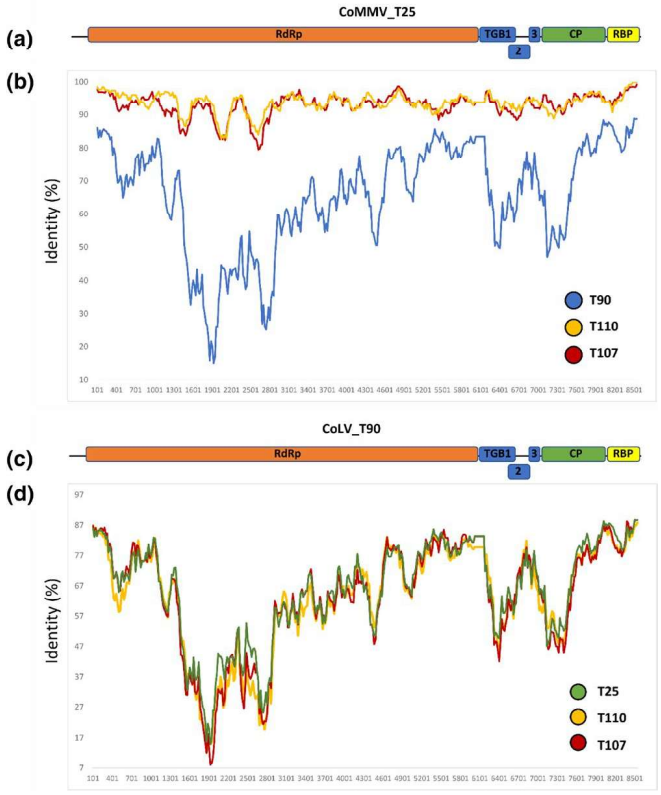

FIGURE 6 Genomic organization of (a) cole mild mosaic virus (CoMMV) isolate T25, and (c) cole latent virus (CoLV) isolate T90 (GenBank accession numbers MK684348 and MK770418, respectively). Nucleotide sequence identity among the genomes of (b) CoMMV_T25, and (d) CoLV_ T90, and isolates T107 (MZ189739) and T110 (MZ189738) of CoMMV. Identity profiles were generated by using Simplot (Lole et al., 1999). Window size: 300 nucleotides (nt). Step size: 20 nt. In (a) and (c) coloured boxes indicate the viral open reading frames. RdRp, RNA-dependent RNA polymerase; TGB1, 2, and 3, triple gene block; CP, coat protein; RBP, RNA-binding protein [Colour figure can be viewed at wileyonlinelibrary.com]

respectively (Table 4). The ORF1 and ORF5 from CoMMV_T25 and CoLV_T90 showed the highest percentage of identity with the homologous sequences from potato virus M (PVM) and narcissus common latent virus (NCLV). Values corresponding to ORF5 were always the highest. Phylogenetic analyses using the deduced aa sequences of the replicase and coat proteins of betaflexiviridids revealed two large branches separating members of the subfamilies *Quinvirinae* and *Trivirinae*. CoMMV_T25, T107, and T110, and CoLV_T90 were grouped in a subclade where PVM was also included (Figure 7). In the tree using the coat proteins, the sequence corresponding to the isolate of CoLV (NC_038322), described in 2002, was also included in the same subclade defined by the sequences described in this study.

3.4 | In silico analysis of CoMMV in an interspecific recombination event

The alignment of the complete genomes of all the isolates of CoLV and CoMMV was processed using the software RDP5. A second alignment was prepared also including the genomes of PVM, NCLV, and Aconitum latent virus, which, according to the phylogenetic analyses, are the carlaviruses most closely related to CoLV and CoMMV.

No recombination signals were detected in the alignment comprising the sequences of CoLV and CoMMV. In the analysis of the second alignment, three out the seven methods in the package RDP5 (MaxChi, $p = 1.8 \times 10^{-3}$; Chimaera, $p = 0.02$; and SiScan, 1.8×10^{-89}) revealed a putative recombination event involving CoMMV_107 and NCLV as major and minor parents of PVM, respectively. The presumed recombinant segment spans a fragment of approximately 330 nt covering the ORF6 of PVM.

3.5 | Detection of CoMMV and CoLV using specific primers

Specific primers for the detection of CoMMV (two primer pairs) and CoLV (three primer pairs), designed from the sequences generated in

TABLE 4 Nucleotide (nt) and deduced amino acid (aa) sequence identity between cole mild mosaic virus (CoMMV) isolate T25, cole latent virus (CoLV) isolate T90, and selected viruses of the subfamilies *Quinvirinae* and *Trivirinae*

			Identity values (%)							
			Replicase (ORF1)		Coat protein (ORF5)		Cysteine-rich protein (ORF6)		Genome	
			CoMMV[a]	CoLV[b]	CoMMV	CoLV	CoMMV	CoLV	CoMMV	CoLV
Genus, subfamily	Virus	GenBank accession number	nt/aa	nt/aa	nt/aa	nt/aa	nt/aa	nt/aa	nt	nt
Carlavirus, *Quinvirinae*	CoMMV_T25	MK684348	–	66.7/71.1	–	73.8/83.8	–	76.3/77.8	–	68.0
	CoLV_T90	MK770418	66.7/71.1	–	73.8/83.8	–	76.3/77.8	–	68.0	–
	CoLV	AY340584	–	–	78.0/97.6[c]	**81.5/99.2**[c]	77.6/77.2	**90.8/92.1**	78.6[d]	**85.2**[d]
	Aconitum latent virus	AB051848	60.1/55.4	59.3/55.7	65.7/72.7	66.8/71.3	44.7	43.0	60.0	59.9
	Narcissus common latent virus	AM158439	61.9/60.0	61.2/60.0	68.1/74.6	66.2/73.9	52.6	48.2	62.4	61.4
	Potato virus M	D14449	61.6/60.9	62.4/60.1	69.4/75.7	71.3/75.0	51.8	49.1	62.0	63.3
Foveavirus, *Quinvirinae*	Apple stem pitting virus	KY242757	54.5/44.0	53.7/42.5	48.5/31.4	52.6/32.9	–	–	53.9	52.0
Robigovirus, *Quinvirinae*	Cherry necrotic rusty mottle virus	EU188438	51.5/38.7	51.9/37.4	48.4/30.9	51.1/29.6	–	–	51.7	51.7
Capillovirus, *Trivirinae*	Apple stem grooving virus	KX686100	48.8/33.5	50.7/33.3	62.4/21.7	62.2/23.1	–	–	51.5	50.9
Trichovirus, *Trivirinae*	Apple chlorotic leaf spot virus	KU960942	50.1/34.4	51.6/36.0	53.5/29.5	55.1/27.5	–	–	50.0	50.2
Vitivirus, *Trivirinae*	Grapevine virus A	AF007415	50.6/32.7	51.2/33.4	54.8/25.9	52.1/27.0	–	–	50.0	50.0

Note: The highest values between CoMMV and CoLV, and viruses of other species, are highlighted in bold and underlined, respectively.

[a]Isolate T25.

[b]Isolate T90.

[c]Comparison involving only 642 nt, approximately two-thirds of the ORF5 in carlaviruses.

[d]Fragment of 1040 nt available in GenBank from an isolate of CoLV identified in Brazil in 2002.

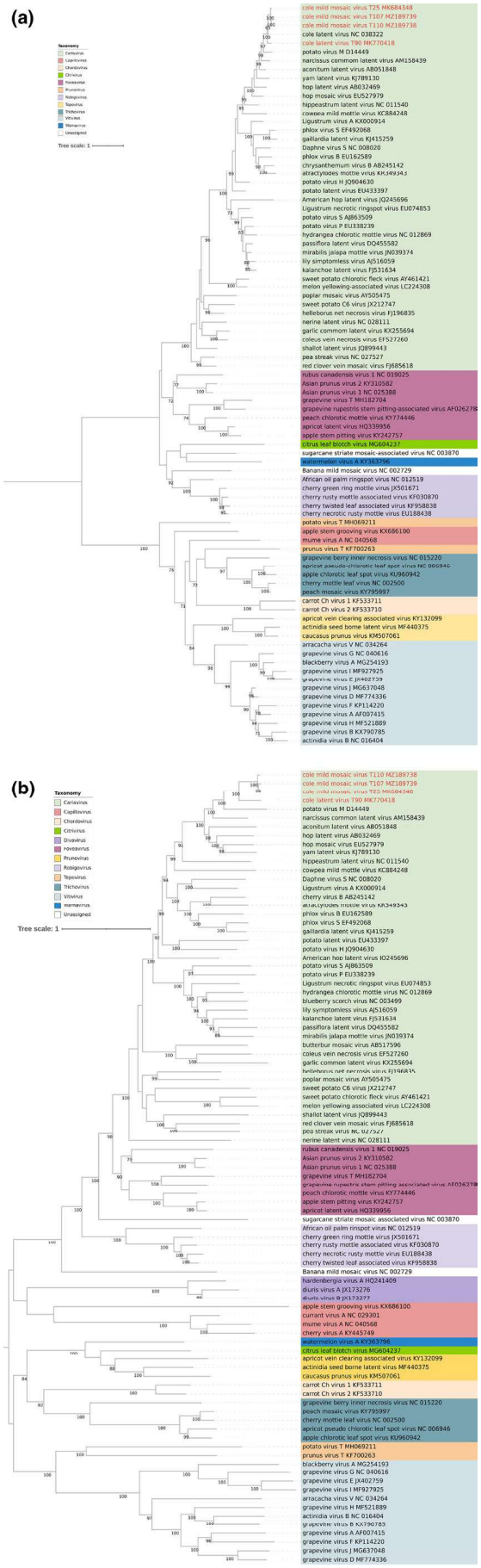

FIGURE 7 Phylogenetic reconstruction for viruses of the family *Betaflexiviridae*. Midpoint-rooted maximum-likelihood phylogenetic trees are based on the deduced amino acid sequences of the (a) coat protein, and (b) replicase. Bootstrap support values (1500 replications) of branches greater than 70% are indicated near the corresponding nodes. The isolates of cole latent virus and cole mild mosaic virus characterized in this study are highlighted in red. The scale bar in each tree indicates the average number of amino acid substitutions per site [Colour figure can be viewed at wileyonlinelibrary.com

this study (Table 3), were used to identify the carlaviruses present in the collected samples not evaluated by HTS (Table 1). Samples pT25, T90, T107, and T110 were used as control.

Amplicons of the expected lengths were obtained using the five primer pairs in six out of the 14 evaluated samples, indicating the simultaneous infection by the two carlaviruses. Notably, all samples detected with the mixed infection CoMMV–CoLV (i.e., T4, T6, T9, T36, T63, and T89) were collected in the State of São Paulo. The remaining samples found to be infected by one of the two carlaviruses, for example, T15 and T62, were infected by CoMMV.

Amplicons obtained from each test and sample were purified and sequenced. Values of the nt sequence identity were always higher than 90% with their respective sequences in CoLV_T90 or CoMMV_T25. These analyses, besides confirming the identity of the virus present in the samples, allowed us to calculate the genetic distances among the isolates of each species, CoLV and CoMMV, and among viruses belonging to different species. By and large, the genetic distances intraspecies were one-third to one-fifth of those observed in the comparison of viruses from the two species (Table 5). For instance, the genetic distance (*D*) based on the CP amplicon among the isolates of CoLV was 0.059 ± 0.032, and among the isolates of CoMMV was 0.053 ± 0.017, whereas, among isolates from the two viruses, the *D* value was five-fold higher (0.252 ± 0.029). These data enlarge our knowledge about the diversity of the two studied carlaviruses, supporting their classification into two different species.

4 | DISCUSSION

The infection of brassica plants by the partially characterized carlavirus cole latent virus (CoLV) has been periodically reported in Brazil (Belintani et al., 2002; Eiras et al., 2008; Kitajima et al., 1970). Based on a field survey conducted from 2012 to 2017, the current study has obtained for the first time the complete genome sequence of CoLV, providing support for its taxonomic status. In addition, we have identified and characterized a new putative brassica-infecting carlavirus named Cole mild mosaic virus. Moreover, we have also documented the susceptibility of *N. megalosiphon* to CoLV and CoMMV and its suitability as an indicator host, in which an infection by these viruses results in the expression of distinguishing systemic symptoms including mosaic, leaf deformation, and developmental delay.

TABLE 5 Genetic distance among viruses belonging to one species (CoLV or CoMMV) and both species

	Genetic distance (*D*)			
	Intraspecies		Interspecies	
Virus	CP[a]	RdRp[b]	CP	RdRp
CoLV	0.059 ± 0.032	0.061 ± 0.023	0.252 ± 0.029	0.218 ± 0.011
CoMMV	0.053 ± 0.017	0.061 ± 0.031		

Note: Calculations were based on the nucleotide sequences of the amplicons corresponding to the ORF1 (RNA-dependent RNA polymerase, RdRp) and ORF5 (coat protein, CP). Because amplicons from the two species comprised different sequence lengths, the largest fragments were trimmed to adjust them to the same length as the smallest ones.

[a]Fragments of 600 nucleotides (nt) of the 3′-end of the open reading frame.

[b]Fragments of 440 nt.

Both the shape and size of the virus particles detected in the leaf kale plants inoculated with the sap extracts of samples pT25 and T90, as well as the cytoplasmic inclusions observed in the infected cells, matched with those previously described in *B. oleracea* var. *acephala* and *Chenopodium quinoa* plants infected by CoLV (Belintani & Gaspar, 2003; Kitajima et al., 1970). Morphometric analyses verified the existence of elongated flexuous particles of approximately 650 nm in length with rounded tips, as typically observed in virions of the genus *Carlavirus* (Adams et al., 2012; Belintani et al., 2002). The genomes of the isolates T25 and T90 are 3′-end poly-A single-stranded RNA molecules that display six ORFs, as detected in known carlaviruses (Adams et al., 2012). However, given the differences of the nucleotide sequence, the isolates T25 and T90 were definitively considered two distinct carlaviruses, identified as CoMMV and CoLV, respectively. In the case of CoMMV, we propose its assignment to a new putative species in the genus *Carlavirus*: Cole mild mosaic virus.

ORF5 (CP) from T25 and T90 show nt and aa identity values higher than 78.0% and 97.6%, respectively, with the only documented fragment of one isolate of CoLV, described in 2002 (NC_038322) (Belintani et al., 2002). Therefore, based on these values, the three isolates could be considered viruses of the same species, that is, *Cole latent virus*. However, it is of note that according to current demarcation criteria followed by the *Flexiviridae* study group of the International Committee on Taxonomy of Viruses, distinct species in the genus *Carlavirus* have less than about 72% nt identity (or 80% aa identity) in the CP (ORF5) or replicase (ORF1) genes (Adams et al., 2012). Consequently, the main rationale behind our proposal for the creation of a new species in the genus *Carlavirus* is the low identity values of the ORF1. When the ORF1 is analysed, the identity values between T25 and T90 (nt: 66.7%, aa: 71.1%) are below the standard threshold, which leads to recognizing T90 and T25 as members of two different species, *Cole latent virus* and the tentative new species Cole mild mosaic virus, respectively.

Other lines of evidence provide additional support to the designations of T90 as an isolate of CoLV and T25 as a member of a new species. (a) While demarcation criteria of species in the genus *Carlavirus* take into account the identity values of ORF1 or ORF5, the ORF1 of carlavirus is six-fold larger than ORF5, representing almost 70% of the viral genome. Accordingly, ORF1 rather than ORF5 might describe more extensively the evolutionary history of carlaviruses. This is particularly relevant in the case of T25 and T90, whose identity values of ORF1, but not those of ORF5, fulfil the requirement established by the species demarcation criteria. (b) ORF6, whose identity is not regarded by the species demarcation rules, of T90 is more closely related to its ortholog in the isolate of CoLV described in 2002 than to the cognate sequence in T25. The value of the identity between ORF6 of CoLV identified in 2002 and T90 surpasses 90% (at both nt and aa levels), whereas between T25 and T90, the figures remain below 80%. (c) The genetic distance of amplicons corresponding to two genomic segments obtained from more than five isolates of viruses of each species indicates a lower range of sequence divergence among viruses of the same species than among viruses from the two studied species. However, despite considering T90 as an isolate of the species *Cole latent virus*, the assignment could be modified should new isolates, more closely related to the isolate of CoLV described in 2002, be identified. This would be the case, for instance, if a putative isolate showed a higher nucleotide sequence identity of its ORF5 with that in CoLV described in 2002, but with an ORF1 whose identity values in comparison with T90 are lower than the threshold for species demarcation in the genus *Carlavirus*.

The complete genomes of CoMMV and CoLV show the highest nt sequence identities with each other (68.0%). The second highest scores in these comparisons corresponded with different carlaviruses: narcissus common latent virus (NCLV, AM158439) in the case of CoMMV (62.4%), and potato virus M (PVM, D14449) for CoLV (63.3%). Accordingly, the phylogenetic reconstruction based on aa sequences of replicase and coat proteins showed the closest relationship between CoMMV and CoLV in a subclade also including PVM and NCLV. Notably, a virus closely related to CoMMV_T107 is suggested as the major parent of PVM in a putative recombination event that includes NCLV as a minor parent.

Previous studies showed that CoLV systemically infects several plant species of the family Brassicaceae, as well as watermelon (*Citrullus lanatus*, Cucurbitaceae), eggplant (*Solanum melongena*, Solanaceae), soybean (*Glycine max*, Fabaceae), and pea (*Pisum sativum*, Fabaceae), but these plants displayed no obvious symptoms (Brunt et al., 1996; Kitajima, 2020). By using a distinct isolate of CoLV, we have not only enlarged the identified spectrum of plant species susceptible to this virus, but also accurately described the development of symptoms in some indicator plants. Our study using commercial varieties of brassicas and a group of experimental host plants of the families Brassicaceae and Solanaceae has revealed the

development of mosaic symptoms in cauliflower and kohlrabi plants (Brassicaceae), and mosaic, leaf deformation, and delayed growth in *N. megalosiphon* plants (Solanaceae). Similarly, our experiments also corroborated the local infection of *C. giganteum* plants and the systemic infection in cauliflower (*B. oleracea* var. *botrytis*) and Chinese cabbage (*B. rapa* subsp. *pekinensis*). Plants of the species *S. melongena*, *C. lanatus*, and *Gomphrena globosa* were not susceptible to isolate T90, results that differ from those observed previously (Belintani et al., 2002; Kitajima et al., 1970; Mello et al., 1987), but that could be biased by either differences between the virus isolates or the genotypes of the plant species assayed. In general, the set of plants susceptible to CoMMV matches with the host range of CoLV_T90. However, symptoms observed in *B. oleracea* var. *botrytis* and *B. oleracea* var. *gongylodes* were consistently different, and in the case of *N. megalosiphon*, the visual signs of infection with CoLV_T90 were delayed as compared to symptom appearance in *N. megalosiphon* plants infected with CoMMV.

In addition to the contribution to a robust taxonomic classification of the species *Cole latent virus* by way of characterization of the full genome of isolate T90, through the identification of CoMMV, we have also enlarged the known virome of brassica plants and provided evidence of the distribution of brassica-infecting carlaviruses in crop areas in Brazil. Moreover, *N. megalosiphon* plants have been proposed as an indicator host for the detection of CoLV and CoMMV, which can also be advantageous for separating the mixed infection CoMMV–CaMV. The new primers and the genomic information generated during the characterization of carlaviruses might be useful to tackle relevant features of the viral epidemiology and for further experiments to evaluate the interactions in mixed infections between the carlaviruses, potyviruses, and caulimovirus commonly affecting brassica crops in Brazil.

ACKNOWLEDGEMENTS

The authors thank Dr Juliana Freitas-Astua (Embrapa Mandioca e Fruticultura, Cruz das Almas, BA, Brazil) for her contribution with molecular biology reagents and to Dr Edmilson Santos da Silva (Lab. Acarologia, Universidade Federal de Alagoas, campus Arapiraca) for assistance in collecting some brassica samples. The authors are also grateful to CAPES, CNPq, and FAPESP for scholarships and research grants associated with this work (CNPq 133331/2017-2, CAPES PNPD20132154 – 33141010001P4 – PNPD – IBSP, 88882.157041/2017-01; FAPESP 2014/22594-2; 2015/50076-9, 2018/17287-4, 2017/50334-3, 2017/18910-4 and 2019/06853-1). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) Finance Code 001. EWK, RH, and ME are supported by a CNPq research fellowship.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in GenBank at https://www.ncbi.nlm.nih.gov/ (accession numbers MK684348, MK77041, MZ189739, and MZ189738).

ORCID

Leilane K. Rodrigues https://orcid.org/0000-0001-6492-6505
Camila Chabi-Jesus https://orcid.org/0000-0002-8397-9992
Elliot W. Kitajima https://orcid.org/0000-0002-9138-2918
Ricardo Harakava https://orcid.org/0000-0003-1431-2665
Alexander Banguela-Castillo https://orcid.org/0000-0002-4513-4884
Pedro L. Ramos-González https://orcid.org/0000-0002-9640-6587
Marcelo Eiras https://orcid.org/0000-0001-7901-9107

REFERENCES

Adams, M.J., Antoniw, J.F., Bar-Joseph, M., Brunt, A.A., Candresse, T., Foster, G.D. et al (2004) The new plant virus family *Flexiviridae* and assessment of molecular criteria for species demarcation. *Archives of Virology*, *149*, 1045–1060.

Adams, M.J., Candresse, T., Kreuze, J.F., Martelli, G.P., Namba, S., Pearson, M.N. et al (2012) Genus *Carlavirus*. In: King, A.M.Q., Adams, M.J., Carstens, E.B. & Lefkowitz, E.J. (Eds.) *Ninth Report of the International Committee on Taxonomy of Viruses*. Amsterdam: Elsevier Academic Press, pp. 924–927.

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M. et al (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research*, *44*, W3–W10.

Badge, J., Brunt, A., Carson, R., Dagless, E., Karamagioli, M., Phillips, S. et al (1996) A carlavirus-specific PCR primer and partial nucleotide sequence provides further evidence for the recognition of cowpea mild mottle virus as a whitefly-transmitted carlavirus. *European Journal of Plant Pathology*, *102*, 305–310.

Belintani, P. & Gaspar, J.O. (2003) Coat protein and RNAs of Cole latent virus are not present within chloroplasts of *Chenopodium quinoa*-infected cells. *Fitopatologia Brasileira*, *28*, 84–88.

Belintani, P., Gaspar, J.O., Targon, M.L.P.N. & Machado, M. (2002) Evidence supporting the recognition of Cole latent virus as a distinct carlavirus. *Journal of Phytopathology*, *150*, 330–333.

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. *Annual Review of Phytopathology*, *51*, 177–201.

Brunt, A.A., Crabtree, K., Dallwitz, M.J., Gibbs, A.J., Watson, L. & Zurcher, E.J. (1996) Plant Viruses Online – Cole latent carlavirus. *Plant Viruses Online: Descriptions and Lists from the VIDE Database*.

Bushnell, B. (2014) *BBMap: A Fast, Accurate, Splice-Aware Aligner*. Splice-Aware Aligner, 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. *Frontiers in Microbiology*, *12*, 986.

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 Disease*, *102*, 1588–1598.

Anonymous 2016. CONAB PROHORT. Programa brasileiro de modernização do mercado hortigranjeiro. Available at: https://www.conab.gov.br/info-agro/hortigranjeiros-prohort [Accessed 10th July 2021].

de Carvalho, C., Kist, B.B. & Treichel, M. (2016) *Anuário brasileiro das hortaliças*. Santa Cruz do Sul: Gazeta Santa Cruz.

Costa, A.S., Kitajima, E.W., Costa, C.L. & Nagai, H. (1972) A couve comum como fonte de tres virus que afetam cruciferas. *Revista de Olericultura*, *12*, 82.

Dellaporta, S.L., Wood, J. & Hicks, J.B. (1983) A plant DNA minipreparation: version II. *Plant Molecular Biology Reporter*, *1*, 19–21.

Dress, A.W.M., Flamm, C., Fritzsch, G., Grünewald, S., Kruspe, M., Prohaska, S.J. et al (2008) Noisy: Identification of problematic columns in multiple sequence alignments. *Algorithms for Molecular Biology*, *3*, 1–10.

Eiras, M., Chaves, A.L.R., Colariccio, A., Harakava, R. & Chagas, C.M. (2008) Primeiro relato do Cole latent virus em raiz-forte. *Tropical Plant Pathology*, *33*(supp), S290.

Foster, G.D. (1992) The structure and expression of the genome of carlaviruses. *Research in Virology*, *143*, 103–112.

Fujita, N., Komatsu, K., Ayukawa, Y.U., Matsuo, Y., Hashimoto, M., Netsu, O. et al (2018) N-terminal region of cysteine-rich protein (CRP) in carlaviruses is involved in the determination of symptom types. *Molecular Plant Pathology*, *19*, 180–190.

Gaspar, J.O., Belintani, P., Almeida, A.M.R. & Kitajima, E.W. (2008) A degenerate primer allows amplification of part of the 3′-terminus of three distinct carlavirus species. *Journal of Virological Methods*, *148*, 283–285.

Ha, C., Coombs, S., Revill, P.A., Harding, R.M., Vu, M. & Dale, J.L. (2008) Design and application of two novel degenerate primer pairs for the detection and complete genomic characterization of potyviruses. *Archives of Virology*, *153*, 25–36.

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. *Nature Protocols*, *8*, 1494–1512.

Jenner, C.E., Wang, X., Ponz, F. & Walsh, J.A. (2002) A fitness cost for *Turnip mosaic virus* to overcome host resistance. *Virus Research*, *86*, 1–6.

Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780.

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. (2020) An annotated list of plant viruses and viroids described in Brazil (1926–2018). *Biota Neotropica*, *20*, e20190932.

Kitajima, E.W., Camargo, I.J.B. & Costa, A.S. (1970) Morfologia e aspectos intracelulares do vírus latente da couve. *Bragantia*, *29*, 181–190.

Kitajima, E.W. & Nome, C.F. (1999) Microscopia eletrônica em virologia vegetal. In: Docampo, D.M. & Lenardón, S.L. (Eds.) *Métodos para detectar patógenos sistémicos*. Córdoba: IFFIVE/INTA-JICA, pp. 59–87.

Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*, 1547–1549.

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 Research*, *47*, W260–W265.

Letunic, I. & Bork, P. (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, *47*, W256–W259.

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. *Journal of Virology*, *73*, 152–160.

López-Moya, J.J., Cubero, J., López-Abella, D. & Díaz-Ruíz, J.R. (1992) Detection of cauliflower mosaic virus (CaMV) in single aphids by the polymerase chain reaction (PCR). *Journal of Virological Methods*, *37*, 129–137.

Martin, D.P., Varsani, A., Roumagnac, P., Botha, G., Maslamoney, S., Schwab, T. et al (2021) RDP5: a computer program for analysing recombination in, and removing signals of recombination from, nucleotide sequence datasets. *Virus Evolution*, *7*, 87.

Mayo, M.A. (2005) Changes to virus taxonomy 2004. *Archives of Virology*, *150*, 189–198.

Mello, S.C.M., Cupertino, F.P., Kitajima, E.W. & Lin, M.T. (1987) Propriedades biológicas e microscopia eletrônica do vírus latente da couve. *Fitopatologia Brasileira*, *12*, 353–360.

Melo, R.A.C., Vendrame, L.P.C., Madeira, N.R., Blind, A.D. & Vilela, N.J. (2019) Characterization of the Brazilian vegetable brassicas production chain. *Horticultura Brasileira*, *37*, 366–372.

Ramos-González, P.L., Chabi-Jesus, C., Banguela-Castillo, A., Tassi, A.D., Rodrigues, M.D.C., Kitajima, E.W. et al (2018) Unveiling the complete genome sequence of clerodendrum chlorotic spot virus, a putative dichorhavirus infecting ornamental plants. *Archives of Virology*, *163*, 2519–2524.

Rodrigues, L.K., Chaves, A.L.R., Kitajima, E.W., Calegario, R.F., Brunelli, K.R., da Silva, F.N. et al (2021) Characterisation of turnip mosaic virus isolates reveals high genetic variability and occurrence of pathotype 1 in Brazil. *European Journal of Plant Pathology*, *160*, 883–900.

Rodrigues, L.K., de Oliveira, A.M., Chaves, A.L.R., Kitajima, E.W., Harakava, R. & Eiras, M. (2019) Cauliflower mosaic virus naturally infects wild radish (*Raphanus raphanistrum*) in Brazil. *Australasian Plant Disease Notes*, *14*, 1–4.

Sastry, K.S., Mandal, B., Hammond, J., Scott, S. & Briddon, R. (2019) *Encyclopedia of Plant Viruses and Viroids*. India: Springer.

Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A. & Minh, B.Q. (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research*, *44*, W232–W235.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. et al (2012) Primer3-new capabilities and interfaces. *Nucleic Acids Research*, *40*, e115.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Oliveira, A.M., Rodrigues, L.K., Chabi-Jesus, C., Chaves, A.L.R., Kitajima, E.W., Harakava, R., et al (2022) Biological and molecular characterization of two closely related carlaviruses affecting brassica plants. *Plant Pathology*, 71, 479–493. https://doi.org/10.1111/ppa.13472