Virus Genes (2006) 32:289–298
DOI 10.1007/s11262-005-6913-1

ORIGINAL PAPER

The complete nucleotide sequence and genomic organization of Citrus Leprosis associated Virus, Cytoplasmatic type (CiLV-C)

Renata C. Pascon · João Paulo Kitajima · Michèle C. Breton · Laura Assumpção · Christian Greggio · Almir S. Zanca · Vagner Katsumi Okura · Marcos C. Alegria · Maria E. Camargo · Giovana G.C. Silva · Jussara C. Cardozo · Marcelo A. Vallim · Sulamita F. Franco · Vitor H. Silva · Hamilton Jordão Junior · Fernanda Oliveira · Poliana F. Giachetto · Fernanda Ferrari · Carlos I. Aguilar-Vildoso · Fabrício J.B. Franchiscini · José M.F. Silva · Paulo Arruda · Jesus A. Ferro · Fernando Reinach · Ana Cláudia Rasera da Silva

Received: 26 June 2005 / Accepted: 29 August 2005

Abstract The Citrus leprosis disease (CiL) is associated to a virus (CiLV) transmitted by *Brevipalpus* spp. mites (Acari: Tenuipalpidae). CiL is endemic in Brazil and its recently spreading to Central America represents a threat to citrus industry in the USA. Electron microscopy images show two forms of CiLV: a rare nuclear form, characterized by rod-shaped naked particle (CiLV-N) and a common cytoplasmic form (CiLV-C) associated with bacilliform-enveloped particle and cytoplasmic viroplasm. Due to this morphological feature, CiLV-C has been treated as *Rhabdovirus*-like. In this paper we present the complete nucleotide sequence and genomic organization of CiLV-C. It is a bipartite virus with sequence similarity to ssRNA positive plant virus. RNA1 encodes a putative replicase polyprotein and an ORF with no known function. RNA2 encodes 4 ORFs. pl5, p24 and p61 have no significant similarity to any known proteins and p32 encodes a protein with similarity to a viral movement protein. The CiLV-C sequences are associated with typical symptoms of CiL by RT-PCR. Phylogenetic analysis suggests that CiLV-C is probably a member of a new family of plant virus evolutionarily related to Tobamovirus.

R. C. Pascon · J. P. Kitajima · M. C. Breton · L. Assumpção · C. Greggio · A. S. Zanca · V. K. Okura · M. C. Alegria · M. E. Camargo · G. G.C. Silva · J. C. Cardozo · M. A. Vallim · S. F. Franco · V. H. Silva · H. J. Junior · F. Oliveira · P. F. Giachetto · F. Ferrari · C. I. Aguilar-Vildoso · F. J.B. Franchiscini · J. M.F. Silva · P. Arruda · J. A. Ferro · F. Reinach · A. C. R. da Silva (✉)
Alellyx Applied Genomics, Techno Park (Rod. Anhanguera Km 104), Rua James Clerk Maxwell 320, 13069-380, Campinas, SP, Brazil
e-mail: ana.rasera@alellyx.com.br

Keywords Citrus leprosis · Citrus leprosis associated virus · Citrus disease · *Citrus sinensis* · *Tobamovirus* · Genome sequence

Introduction

Citrus leprosis (CiL) is an economically relevant virus disease that affects *Citrus* spp., mainly sweet orange (*C. sinensis*) and mandarins (*C. reticulate*, *C. reshni*, *C. deliciosa*). The disease has been reported in South, Central and North America, particularly in Argentina, Brazil, Paraguay, Uruguay, Bolivia, Venezuela, Panama and Costa Rica [1–7]. In North America it was detected in Mississippi and Florida, but has been absent from these sites since 1920 [8]. CiL is transmitted by the false spider mite *Brevipalpus* spp. (Acari: Tenuipalpidae), specifically by *B. phoenicis* (8–11). The control of CiL is done by vector extermination and the cost of acaricides represents around 100 million dollars per year in Brazil. CiL also represents a threat to Central and North America since the vector is present in these areas.

A virus, CiLV (Citrus Leprosis Virus), has been associated with CiL, and the infection is located in tissues where the mites feed, resulting in non-systemic, chlorotic and necrotic lesions that are observed in leaves, fruits and stems [4, 9, 12, 13]. There are various degrees of symptoms severity, which depend on citrus variety and age of the lesions. For instance, sweet orange is more sensitive than mandarins and the infection can lead to defoliation, fruit dropping and eventually tree loss. Virus particles are routinely found by electron microscopy in lesions of citrus leaves and viruliferous mites [10, 11, 14–16]. CiLV is circulative in the vector, and it accumulates and replicates within mite organs [8]. However, CiLV is not transmitted to the progeny by viruliferous female-mite [8].

CiLV was described as *Rhabdovirus*-like due to its morphological similarity to the members of this virus family [10]. Electron microscopy reveals two distinct morphological types: a cytoplasmic (CiLV-C) and a nuclear (CiLV-N) type, CiLV-C is observed as short bacilliform particles, 120–130 nm long and 50–55 nm wide, surrounded by a membrane. It is the most common form and causes cytoplasmic alterations such as large electron dense viroplasms [10, 17]. The nuclear type (CiLV-N) is characterized by rod-shaped, naked particles about the same size as CiLV-C. These particles are frequently found associated to the endoplasmic reticulum, nuclear membrane and the presence of translucent nuclear viroplasms [8, 17]. CiLV-N was detected only once in Brazil [10] and recently in Panama [6].

Locali et al. [18] developed two-set of primers to amplify CiLV-C sequences by RT-PCR using RNA extracted from lesions of symptomatic leaves. These primers amplify two short DNA fragments with sequence similarity to a viral movement protein (338 nt) and a viral replicase (401 nt) related to *Furovirus, Tobamovirus, Tobravirus* and *Pomovirus*. Their experiments also showed a broad and constant association between CiLV-C sequences and CiL symptoms and the presence of viral particles and/or viroplasm.

Because of the relevance of CiL it is necessary to develop a deeper knowledge of the disease and its causative agent. Here, we present the complete nucleotide sequence of CiLV-C genome. The assembly of the genome yielded two contigs: RNA1 (8,730 nt) and RNA2 (4,975 nt), which are specific to symptomatic citrus leaves. Phylogenetic analysis of the sequence suggests that CiLV-C is a tentative member of a new family of plant virus evolutionarily related to *Tobamovirus*.

Materials and methods

Plant material

Leaves from *C. sinensis* (Hamlin and Valencia) and *C. reshni* (Cleopatra) containing typical CiL chlorotic lesions were collected from Analândia, Piracicaba, Matão, Rio Claro, and Capela do Alto in the State of São Paulo, Brazil. Leaves were washed; dried and foliar lesions were carefully excised using sharp blades. The material was frozen at −80°C until RNA preparations or kept at 4°C for virus enrichment procedures up to 3 days.

RT-PCR

For Total RNA extraction, 3 g of CiL foliar lesions were macerated in the presence of liquid Nitrogen and extracted with Trizol (Invitrogen) according to the protocol described previously [19]. RT-PCR was carried out using Superscript II Reverse transcriptase (Invitrogen) according to the instructions of the manufacture. First strand synthesis primers were either 570 ηg of Oligo-dT or 300 ηg of GDP rnix (Genome Directed Primers) or 250 ηg of Random Primers (Invitrogen). GDP mix was designed based on alignment of *Rhabdovirus* and *Furovirus* sequences (GDP available upon request). PCR was performed with gene specific primers (GSP) (Table 1). The PCR program varied in the annealing temperature and extension time according to primer combination and length of the DNA fragment, respectively. In general, the amplification was achieved with: 1 cycle at 94°C for 5 min, 25 cycles at 94°C for 30 s, 50°C for 30 s, 72°C for 1 min and one final extension at 72°C for 7 min. An aliquot of 5 µl of PCR reaction was loaded in 1% agarose gel.

CiLV-C detection in symptomatic leaves

Viral sequences were detected by RT-PCR using the same conditions described above. Primers for detection covered the vast majority of the viral genome in both RNA1 and RNA2. RNA from healthy leaves was used as negative control. Primer combinations for viral detection were: REP13 (5′-ACAGACTACGTGAAATATACC-3′) and REP14 (5′-CGTGAAACTCCGAATCCATT-3′), REP15 (5′-GGAAATTGCTTCCTCACTTG-3′) and REP16 (5′-GGTGTTGTGGTACCACCACC-3′), REP19 (5′-GTACC GCACTTGAGCCTATC-3′) and REP18 (5′-AACCTCGC CCAGCTGACAAC-3′), MP21 (5′-GGTTGCTTTACGTT CGAGTGTGA-3′) and MP13 (5′-CGTTGTGGAGACCC AGAGCA-3′) and MP15 (5′-GGTAGTGATATCATGTC ATGTG-3′) and MP16 (5′-GGTACACCACCAGCTAGA AG-3′) (Fig. 4a). Citrus chalcone synthase primers were used as internal controls (CHS-F 5′-CCAGCCCAAGTCT AAGATCACC-3′) and CHS-R (5′-CGCAAGTGTCCGT-CAATTGC-3′). These primers amplify a 423 bp fragment from citrus genome. Diagnostic primers (Table 1) described earlier [18] were also used to detect CiLV-C sequences in the same RNA samples.

Cloning and sequencing of DNA fragments

DNA fragments were isolated from 1% agarose gel, purified with GFX (Amersham Biosciences, UK), cloned in pGEM-T easy Vector (Promega) and introduced into competent *E. coli* DH10B by electroporation. Plasmids were sequenced with Big Dye 3.1 in an ABI 3730 XL from Applied Biosystem.

Table 1 Primer list

Primer	GSP – sequence (5′–3′)	Source
MPPD-F	GCGTATTGGCGTTGGATTTCTGAC	[18]
MPExt-R	GTCAGAAATCCAACGCCAATACGC	This work
MPPD-R	TGTATACCAAGCCGCCTGTGAACT	[18]
MPExt-F	AGTTCACAGGCGGCTTGGTATACA	This work
RepPD-F	GATACGGGACGCATAACA	[18]
RepExt-R	TGTTATGCGTCCCGTATC	This work
RepPD-R	TTCTGGCTCAACATCTGG	[18]
RepExt-F	CCAGATGTTGAGCCAGAA	This work
REP1	TTGTGTAAACGAGCAGAATGG	
REP2	CGCAGCCTTAGCTTCAGTGG	
REP3	CGCGACTCTGGCGAAGATGAG	
REP4	CATTCTCATGCACAAGAAGG	
REP5	TTGACCGATGTTCATTGC	
REP6	TTGCAACTTACGTACAGG	
REP7	CTAACTTCAACTTGCGCAGATC	
REP8	TGATAGTCAGAATTTCTCTG	
REP9	AAGGTGCAGGTGCAACGCCG	
REP10	CTCATAGCGTTGTGTGCACG	
REP11	TAACGTCCAGAGTTTCAACGG	
REP12	CAATGTGATGAAGCCTCACGATGGC	
REP13	ACAGACTACGTGAAATATACC	
REP14	CGTGAAACTCCGAATCCATT	
REP15	GGAAATTGCTTCCTCACTTG	
REP16	GGTGTTGTGGTACCACCACC	
REP17	ACGCTGAATATGTTGTTGTC	
REP18	AACCTCGCCCAGCTGACAAC	
REP19	GTACCGCACTTGAGCCTATC	
MP1	TGACTTGACGGTTAAGAAGAG	
MP2	CATCTCAACATAAAATTCACC	
MP3	AATACTTTGAAGACTGTACTC	
MP4	TCTTGGTCATTAACAGCC	
MP5	TGTTTTAGAGCTTATCGC	
MP6	GATTGTCACTCACCACTCCTC	
MP7	TTGGCGGAAGCAACATCCACC	
MP8	ATAGGCAAGCACTAATAGCGG	
MP9	GTCTACTGAAGGCACCGTCAACTTTGG	
MP10	ACCGGTGTTATGCGGAGGGTT	
MP11	CACATGACATGATATCACTACC	
MP12	GGTACCAATATGTAAACTATAC	
MP13	CGTTGTGGAGACCCAGAGCA	
MP14	TCGACCTTACGGCGATAATT	
MP15	GGTAGTGATATCATGTCATGTG	
MP16	GGTACACCACCAGCTAGAAG	
MP17	GTTTCAGATGCATCTGGCCCGG	
MP18	TGAGAAAGGCAACTTGTCTTCG	
MP19	GAATGATAAATCGCGTAATCATGC	
MP20	TCACACTCGAACGTAAAGCAACC	
MP21I	GGTTGCTTTACGTTCGAGTGTGA	

Extraction of double strand RNA

dsRNA was extracted from 6 g of chlorotic lesions from *C. sinensis* (Hamlin) using the protocol described previously [20] with one round of CF11 chromatograph.

Subtraction library constructions

Clontech PCR-Select cDNA subtraction kit was used to construct a subtraction library for CiLV-C genome sequencing. mRNA was isolated from symptomatic and asymptomatic leaves using PolyATract mRNA isolation system (Promega). Two micrograms of mRNA from each sample were used to construct the subtraction library using oligo-dT as the first strand synthesis primer. The library construction was made according to instructions of the manufacturer. The subtracted fragments obtained were cloned in pGEM-T easy Vector (Promega) and sequenced.

cDNA library construction

cDNA libraries were constructed from 2 μg of either dsRNA or from total RNA extracted directly from virus enriched extracts. The dsRNA was denatured at 100°C for 5 min and RNA from virus-enriched extracts was denatured at 65°C for 5 min. The First strand synthesis was performed with 1 μg of Random Primer (Invitrogen). Libraries were made according to the Superscript III cDNA kit manufacturer's instructions (Invitrogen). After size-fractioning, larger fragments of cDNA (1–2 kb) were ligated in pCR4Blunt-TOPO (Invitrogen, California) and introduced into *E. coli* DH10B (Invitrogen) by electroporation.

Cloning of 5′and 3′ends

5′ End amplifications were done with 5′ RACE system for rapid amplification of cDNA ends version 2.0 (Invitrogen) according to instructions of the manufacturer using 5 μg of total RNA from lesion fragments. 3′ End sequence was obtained by RT-PCR using Oligo-dT as the First Strand Primer. PCR was conducted with Oligo-dT and Gene Specific Primer (Table 1). PCR conditions were as described above.

Virus enrichment procedure and RNA extraction

Around 15 g of fresh chlorotic CiL lesions excised from leaves of *C. reshni* were completely disrupted by maceration in the presence of 25 ml TACM buffer (50 mM Tris–HCl pH 8.0, 0.1% ascorbic acid, 0.1% L-cysteine, 0.5%, β-mercaptoethanol) and around 5 g of clean autoclaved 2 mm glass beads. All solutions, tubes and instruments were kept on ice. The lysate was filtered through cheesecloth and centrifuged 2,000×*g* at 4°C for 10 min. The supernatant was collected and divided into 2 separate batches. About 0.8% NaCl and 5% PEG 8000 were added to one batch and 0.8% NaCI and 8% PEG 8000 to the other. The extracts were incubated for 1 h on ice and centrifuged for 12,000×*g* during 15 min at 4°C. The resulting pellets were resuspended in 1 ml of ice cold TACM buffer. Three hundred microliter of these extracts were treated with 10 ηg/μl of RNAse (Invitrogen, California) for 45 min at room temperature in order to digest plant RNA and viral free RNA. The RNA extraction using Trizol was conducted as described before [19] and the presence of CiLV-C sequences in the samples was ensured by RT-PCR. The viral enriched RNA was used in RT-PCR, Northern blot and cDNA library construction.

Northern blot

Ten micrograms of total RNA was resolved in a 1% agarose-MOPS gel, transferred to Hybond-N+ nylon membranes (Amersham) and hybridized against various probes labeled with α ^{32}P-dCTP. The RNA separation, transfer, hybridization and washes were conducted according to standard protocol [21].

Transmission electron microscopy (TEM)

Samples of chlorotic foliar lesions from Piracicaba and Capela do Alto were subjected to TEM according to the methodology described before [10].

Bioinformatics

Sequence analysis and alignment were performed using programs based on BLAST [22] and Cross_match [23]. CiLV-C genome sequence was assembled with Phred [24, 25]/Phrap /consed [23, 26] programs. (www.phrap.org) CiLV-C Open reading frames (ORFs) annotation resulted from BLAST alignment with the NCBI non-redundant (NR) protein database. Protein domain searches were done using Prosite scan and PFAM analysis (http://myhits.isb-sib.ch/cgi-bin/motif_scan). Additional alignments were generated using ClustalX [27] with PAM protein weight matrix to calculate amino acid similarity scores. Phylogenetic analysis was performed by PAUP 4.0 b 10 [28] with Maximum Parsimony method (bootstrap with 1,000 resamplings, one random sequence addition per replicate, MAXTREES=100 and 50% majority rule consensus tree).

Nucleotide sequence accession number and utility patent application

The complete CiLV-C nucleotide sequence was submitted to GenBank and has been assigned Accession No. DQ 157465 and DQ 157466. Utility applications of these sequences are under protection of international patent laws [29].

Results and discussion

Cloning and sequencing the CiLV-C genome

Several strategies were designed to clone and sequence the CiLV-C genome. Initially, RT-PCR was performed using the 2 sets of diagnostic primers (Table 1) described previously [18], generating two short DNA fragments with sequence similarity to a viral replicase (401 nt, Accession No. AY289191) and a viral movement protein (338 nt, Accession No. AY289190). Forward and reverse primers based on the sequence of these fragments were designed to extend the sequence at the 5′ and 3′ ends by primer walking. In parallel, three cDNA libraries were made and

sequenced: (i) a subtractive library from mRNA extracted from lesions and from healthy leaves; (ii) dsRNA cDNA library and (iii) viral enriched cDNA library. A large number of reads (1,797) were assembled into the 2 initial contigs generated by RT-PCR and the diagnostic primers described above. A combination of cDNA library screening, RT-PCR, primer walking and 5′ RACE allowed the extension and completion of the CiLV-C genome. In RNA1 and RNA2 a total of 1,075 and 3,032 bases were obtained by 5′ RACE, respectively. At least two 5′ RACE reactions were performed for each RNA1 and 2 in order to provide the best prediction of the sequence at the 5′ end. The only significative gap (4,885 bp) closed by non-library method (RT-PCR) was located in RNA1. All other sequences were represented by library reads. A total of 9,209 reads were sequenced to complete the genome with coverage of 15- and l8-fold for RNA1 and RNA2, respectively. From this number, 5,570 reads were from libraries. The libraries reads classification is presented in Table 2. It is noteworthy to mention that 32.3% of these reads were CiLV-C sequences, indicating that the libraries were enriched for these sequences, specially the subtractive library. Among other viruses (1.1% of total sequences) that were encountered in these libraries, most of the reads were from CTV (*closteroviridae*) which is endemic in Brazil and the remaining have similarity to several other viruses: *Retroviridae, Caulimoviridae, Poxviridae, Baculoviridae, Herpesviridae* and *Myoviridae*. However, none of the analyzed sequences had similarity to *Rhabdoviridae* family.

CiLV-C genome has 2 replicons (Fig. 1): RNA1 (8,730 nt) and RNA2 (4,975 nt). Analysis of the coding properties of RNA1 revealed 2 ORFs. One is 7,539 nt long and starts with an AUG at position 108 and terminates with an UAA at position 7,646, encoding a 2,512 amino acids polypeptide (286.4 kDa polyprotein) (Fig. 1). Similarity searches at amino acid level showed that this sequence has relevant similarity to functional domains of the replicase polyprotein of several ssRNA positive plant viruses deposited in the GenBank. Among them, the highest similarities found were to the *Furovirus, Tobamovirus, Tobravirus* and *Pomovirus*, which confirms the previous data obtained [18]. Similarity searches allowed the identification of conserved domains to viral methyltransferase (from residues 128 to 325), viral helicase (from residues 1,521 to 1,697) and RNA-dependent RNA polymerase (from residues 2,221 to 2,458) (Fig. 1). Also, the polyprotein has a conserved putative OTU protease (ovarian tumor protease) (from residues 683 to 838) found by Prosite and PFAM analysis. This domain belongs to the cysteine protease super family present in ovarian tumors and some human virus, such as Crimean-Congo hemorrhagic fever (CCHF) virus [30–32]. The OTU domain is characterized by 4 conserved motifs (I–IV). A conserved cysteine and aspartate at motif I, and a histidine at motif IV are thought to be the catalytic residues. All three amino acids are conserved in the putative OTU protease domain of CiLV-C RNA1 (D726, C729 and H831). This is the first report of this kind of protease in a plant virus. Whether this is functionally relevant or not remains to be proven.

A second ORF (792 nt, position 7,709 AUG to 8,500 UAG) was identified in RNA1 and encodes a 263 amino acids putative protein (p29) (Fig. 1). Searches at GenBank revealed no sequence similarity to any known viral gene.

RNA2 is 4,975 nt long and encodes 4 ORFs (Fig. 1). p15 (start codon at position 67 and stop at 459) encodes 130 aminoacids. p61 (start at 1,590 and stop codon at 3,203) encodes 537 aminoacids. None of them showed significant similarity to any known gene. The third ORF (p32) has 894 nt with the start codon at position 3,228 (AUG) and stop at 4,121 (UAA) and encodes a 297 amino acids protein with a predicted molecular weight of 32.5 kDa. Searches at the GenBank showed considerable similarity (22% identity in 186 amino acids) to movement proteins of several ssRNA positive plant viruses, especially those from *Furovirus* and *Bromovirus.* The last ORF in RNA2 (p24, 645 nt) is transcribed in a different frame than p61 and p32, overlapping with the terminal part of p32 by 29 nt (Fig. 1). Its start codon is located at position 4,093 (AUG) and the stop codon at position 4,737 (UGA). Similarity search at GenBank showed no significant similarity to any known genes.

Since CiLV-C replicates within the mite, it will likely need specific proteins to fit in this environment. ORFs (p29, pl5, p61 and p24) that have no sequence similarity to plant viral genes, and therefore have no obvious function for a plant virus, could be necessary for virus replication and survival within the vector.

Regardless if the identified ORFs are functional in the plant or in mite, we observed a peculiar genome organization with a high number of unknown proteins. The functionality of these putative proteins remains to be further tested, as well as their roles in the plant or in the mite.

Table 2 Classification of valid reads* from various cDNA libraries sequenced

Category	Number of valid reads	Percentage
Ribosomal	1,810	32.5
Citrus EST	1,904	34.2
Other virus	59	1.1
CiLV-C	1,797	32.3
Total	5,570	–

*Valid reads are longer than 400 bp with base quality greater than 20 (Phred quality)

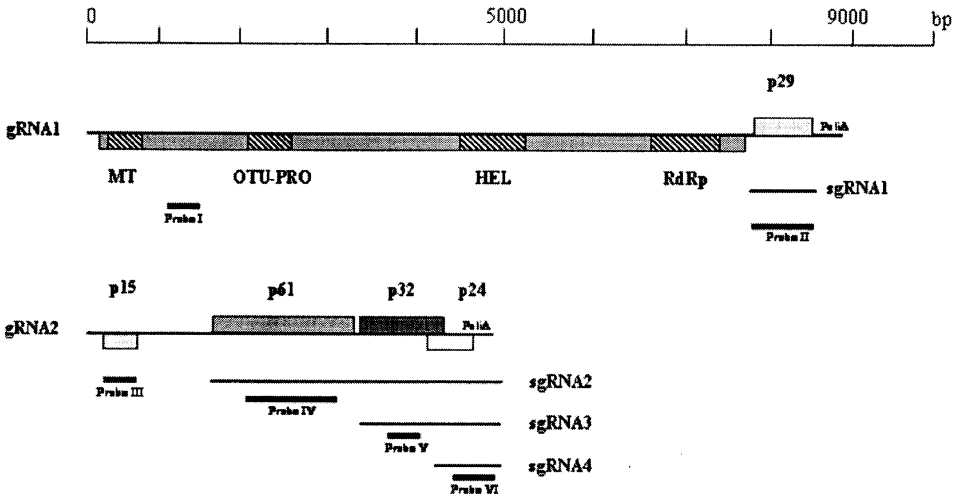

Fig. 1 Genome organization of the CiLV-C. RNA1 (8,730 nt excluding polyA) encodes the replicase of polyprotein and p29. RNA2 (4,975 nt excluding polyA) encodes 4 ORFs (p15, p61, p32 and p24). MT, methyl transferase domain; OUT-PRO, ovarian tumor protease domine; HEL, helicase domain; RdRp, RNA dependent RNA polymerase domain; MP, movement protein. Boxes indicate ORFs while dashed boxes indicate protein domains. The thin bars indicate sgRNA and thicker bars, the localization of the probes used in the Northern blots (Fig. 5)

Conservation of the 5′ and 3′ UTR of the CiLV-C

Regarding 5′ UTR similarity, sequence alignment demonstrated that only the 6 first nucleotides are identical (5′-GATAAA-3′). At the 3′ UTR, nucleotide sequences are more similar. In RNA1 and RNA2 there are 230 and 238 nt at the 3′ UTRs, respectively, excluding polyA tail. Nucleotide sequence alignment showed striking similarity with 79.6% identity (Fig. 2). Considering only the last 134 nt from the 3′ UTR, the identity goes up to 85%. No specific relevant secondary structure was found at either 5′ or 3′ UTRs. Similar reports are common in the literature. Alignment of all 3 segment of Potato Mop Top Virus (PMTV), a *Pomovirus*, shows a high degree of similarity at the 3′ UTR, whereas the 5′ UTR is not as conserved [33]. The sequence conservation at the 3′ end strongly suggests co-evolution of the two RNA segments of the CiLV-C genome and confirms that two RNA segments belong to the same virus.

None of the two segments showed a canonical polyadenylation signal, such as ''AATAAA'' or ''ATTAAA.'' This is not so surprising since, in general, the efficiency of polyadenylation is not a function of a single element but consists of multiple signals which are more diffuse in plants than in animals [34].

Association of CiLV-C genomic sequences with citrus leprosis symptoms

To create a clear association between CiL symptoms and all parts of CiLV-C genome, we performed RT-PCR with primers designed to generate overlapping PCR fragments of RNA1 and RNA2 (Fig. 3a). As shown in Fig. 3b, all four samples of citrus leprosis (Piracicaba, Rio Claro, Matão and Capela do Alto) were RT-PCR positive for all sets of primers. Healthy leaves were RT-PCR negative for all 5 sets of CiLV-C specific primers. Internal control amplification yielded bands for all samples tested. Also, the 2 sets of CiLV-C specific primers (Table 1) described previously [18] generated RT-PCR positive for our symptomatic samples, but not for the healthy sample (data not shown).

Transmission electron microscopy confirmed the presence of cytoplasmic alterations in two symptomatic samples used in RT-PCR (Piracicaba, Fig. 4a and Capela do Alto, Fig. 4b). Figure 4a and b show the image of parenchyma cells with large electron dense viroplasms (arrows), which are typically found in CiLV-C, the most common form of CiL [17].

In the view of the fact that we start our sequence effort from the diagnostic set of primers characterized by Locali et al. [18], and they had performed a large correlation study between PCR amplification, CiL symptoms and cytopathology, it is out of the scope of this paper to perform a large association study. In addition, our results show that all assembled sequences are in fact contiguous and specific to symptomatic leaves. Therefore, we can conclude that we have sequenced the cytoplasmic type of CiLV and the assembled sequences are specific to the CiLV-C genome and associated with CiL.

Subgenomic RNA (sgRNA) production in CiLV-C

In order to detect CiLV-C sgRNA, various fragments of CiLV-C were labeled and used in Northern blots hybridizations. The location of the probes is presented in Fig. 1.

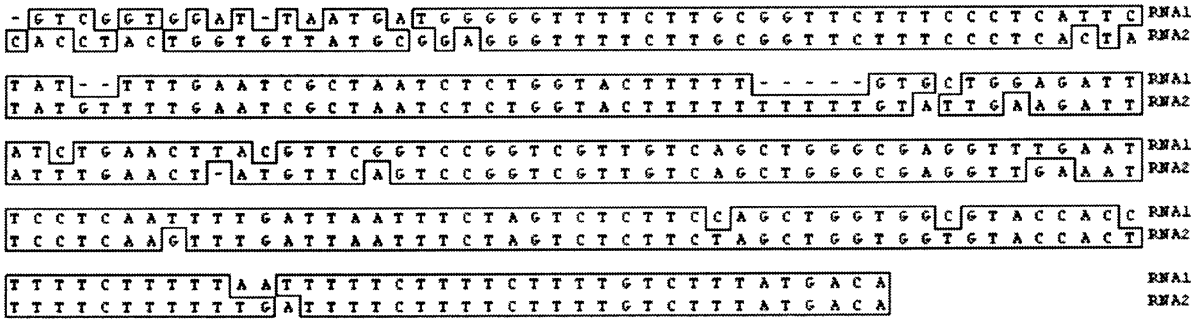

Fig. 2 Nucleotide alignment of the last 230 and 238 nucleotides of the 3′ UTR of RNA1 and RNA2, respectively. In this segment the percentage identity is 79.6%

Both probes made for RNA1, yielded one band of around 9-kb corresponding to gRNA1 (Fig. 5). This size is consistent with the assembled sequence of RNA1 (8,730 nt). However, only Probe II, which covers the 3′ end of RNA1 and p29 nucleotide sequence, detected an abundant sgRNA1 fragment of 792 nt (Fig. 5).

Northern blot analysis of RNA2 revealed 3 sgRNAs. sgRNA2 (Figs. 1, 5) appears as a 3 kb band and corresponds to p61, p32 and p24. sgRNA3 is 1.5 kb in size and transcribes p32 and p24. sgRNA 4 corresponds to p24 only, which is also very abundant, similarly to p29 (Fig. 5). From Northern blot experiments we concluded that RNA1 (8,730 kb) produces one single abundant sgRNA1 of 0.7 kb in size and RNA2 (4,975 kb) generates 3 sgRNA: 3 kb (sgRNA2), 1.5 kb (sgRNA3) and 0.6 kb (sgRNA4).

Since we used RNA extracted from a PEG precipitation, we cannot discard the possibility of identifying double strand sgRNA instead of single strand sgRNA.

Phylogeny of C1LV-C

To attempt a phylogenetic classification of CiLV-C, we constructed an unrooted phylogenetic tree using the replicase polyprotein from CiLV-C and a variety of plant viruses. Some of them are related to CiLV-C, either by sequence similarity (*Furovirus, Tobamovirus, Tobravirus, Bromovirus, Pomovirus, Hordeivirus* and *Cucumovirus*) or morphology (*Rhabdovirus*) and some are unrelated, such as *Tospovirus* and *Nepovirus*. The result in Fig. 6 shows that, regarding the replicase polyprotein, CiLV-C seems to be more closely related to the *Tobamovirus* than any other

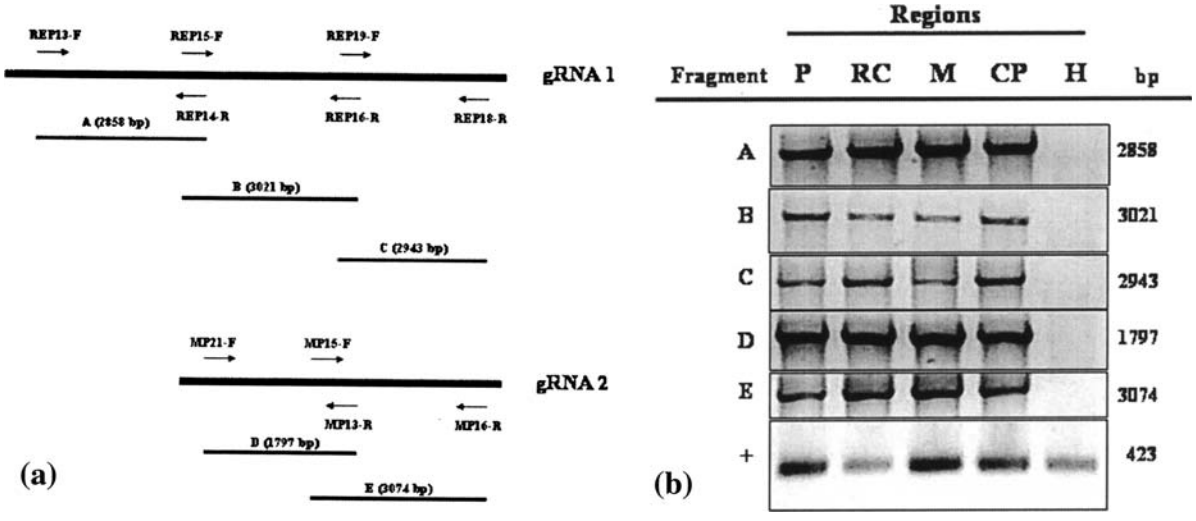

Fig. 3 (**a**) The diagram represents the fragments of the CiLV-C genome amplified by RT-PCR. Thicker bars indicate gRNA1 and 2. Arrows show the position of the primer sets and thinner bars represent the amplified fragments. Letters identify each fragment and numbers in parenthesis indicate the size of them. (**b**) RT-PCR amplifications of the CiLV-C genome. Letters on the left correlate the bands with the fragments schematically described in Fig. 3**a** and numbers on the right indicate band size. RNA samples were extracted from lesions collected in different regions of Sao Paulo State (Brazil). P, Piracicaba; RC, Rio Claro; M, Matão; CP, Capela do Alto; H, healthy leaves; +, amplification with positive control set of primers

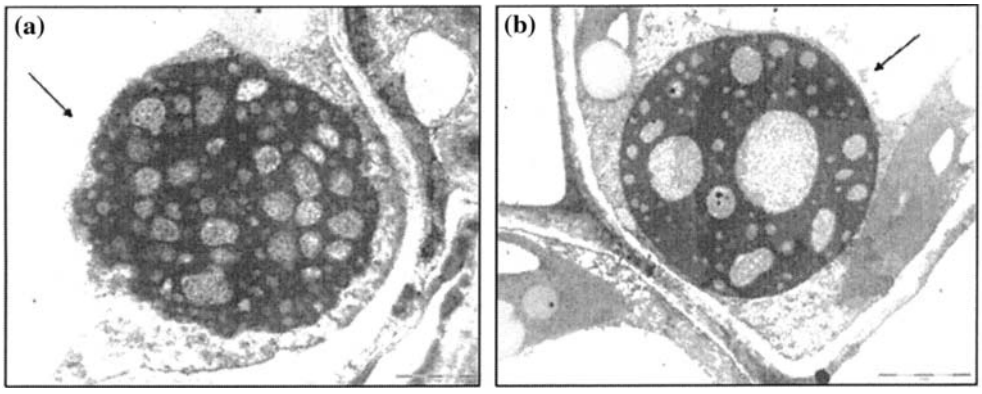

Fig. 4 The transmission electron microscopy of CiLV-C lesions present in leaves from (**a**) Piracicaba and (**b**) Capela do Alto, State of São Paulo, Brazil. The arrows indicate viroplasms typical of CiLV-C. The bar on the right lower corner is 1 μm long

group used in this analysis. Our data suggests that CiLV-C maybe a member of a new family of *Tobamo*-like virus that has not been described before. Initially, CiLV-C was proposed to be a *Rhabdovirus* due to similar particle morphology, however, in this analysis the *Rhabdovirus* grouped together among them, but fall from CiLV-C.

It is important to mention that CiLV-C is observed as bacilliform enveloped virus [10, 14]. No CiLV-C particle free of the envelope has ever been observed, therefore, the shape of the virus given by its coat protein has not been observed. It is possible that the coat protein of CiLV-C assembles into a rod-shape structure, which is not seen, just because the envelope hides it. If CiLV-C is really a rod-shaped virus, that would be more similar to *Tobamovirus*. However, this hypothesis remains to be experimentally proven.

There are other examples of chimerical viruses related to the *Tobamovirus.* For instance, the alga *Chara australis* virus (CAV) resembles Tobacco Mosaic Virus, Soil Born Wheat Mosaic Virus and Potato Mop Top Virus regarding morphology [35]. This virus genome sequence and structure is also a mosaic with some parts related to plant virus and some other parts related to Hepatitis E virus or with no match in GenBank. CiLV-C could come from a *Tobamovirus* evolutionary lineage, like CAV, that co-evolved with their hosts and for this reason acquired diverse sequences.

In summary, we have obtained the sequence and genomic organization, of an important citrus pathogen. We

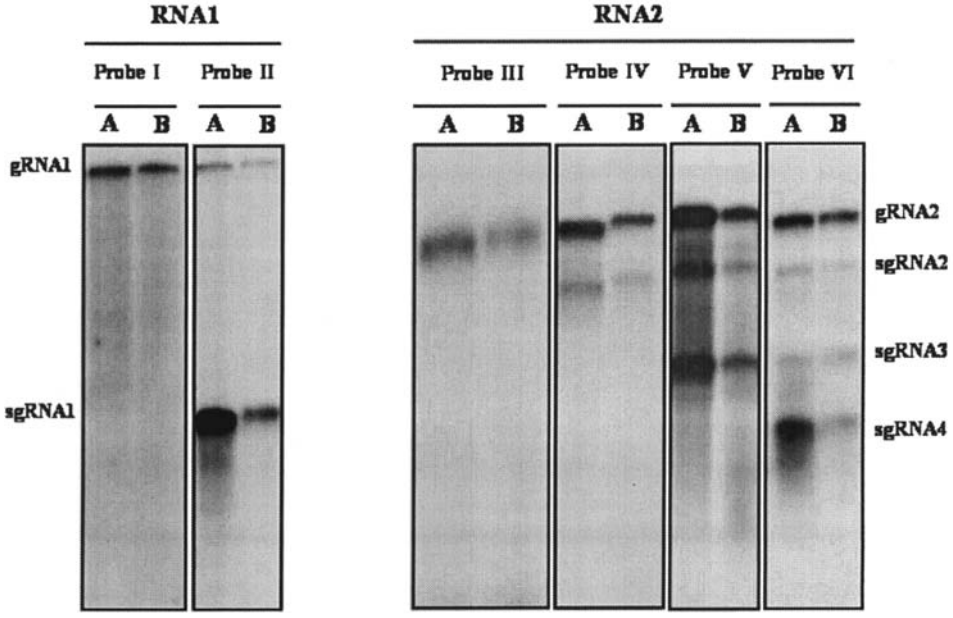

Fig. 5 Northern blots using specific probes designed for RNA1 (probes I, from nucleotides 886 to 2,818 and probe II was made from a library clone that covered the last 1.1 kb of RNA1) and RNA2 (probes III from nucleotides 67 to 459, probe IV from 1,868 to 3,018, probe V from nucleotides 3,375 to 3,689 and probe VI 4,038 to 4,606) were labeled with αP^{32}-dCTP. Each lane contains 10 μg of total RNA isolated from chlorotic lesion extracts precipitated with 8% (**a**) and 5% (**b**) PEG 8000 for viral enrichment. Identification of the sgRNAs is indicated and corresponds to sgRNA represented in Fig. 1

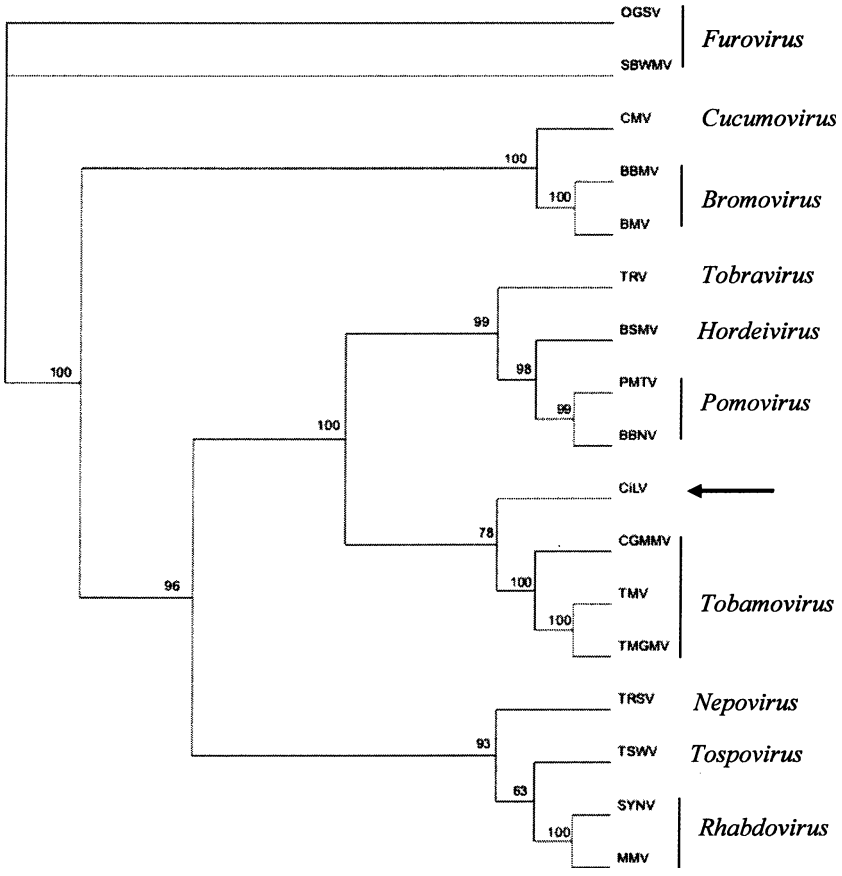

Fig. 6 (**a**) Phylogenetic tree using replicase polyprotein of CiLV-C, 2 *Furovirus*: SBWMV (Soil Born Wheat Mosaic Virus, NP_049336) and OGSV (Oat Golden Stripe Mosaic Virus, CAB57883.1), 2 *Rhabdovirus*: SYNV (Sonchus Yellow Net Virus, NP_042286) and MMV (Maize Mosaic Virus, AAT66757.1), 3 *Tobamovirus*: CGMMV (Cucumber Green Mottle Mosaic Virus, AAM02966), TMV (Tobacco Mosaic Virus, NP_597746) and TMGMV (Tobacco Mild Green Mosaic Virus, NP_062913), 2 *Pomovirus*: PMTV (Potato Mop Top Virus, AAR045180) and BBNV (Broad Bean Necrosis Virus, NP_740760), 2 *Bromovirus*: BBMV (Broad Bean Mottle Virus, NC_004008) and BMV (Brome Mosaic Virus, NP_041196), BSMV (*Hordeivirus*, Barley Stripe Mosaic Virus, AAV54032.1), TRV (*Tobravirus*, Tobacco Rattle Virus, CAE52518), TRSV (*Nepovirus*, Tobacco Ringspot Virus, AAB03785), CMV (*Cucumovirus*, Cucumber Mosaic Virus, CAB77386), TSWV (*Tospovirus*, Tomato Spotted Wilt Virus, BAD86755)

demonstrate that there is an association between CiL symptoms, CiLV-C genome, and the presence of cellular alterations typical of the cytoplasmic type of Leprosis. As a working hypothesis we propose that CiLV-C is the member of a new family of *Tobamo*-like virus.

Acknowledgements We thank Dr. Elliot Kitajima for the transmission electron microscopy, Robson F. Souza for the phylogenetic analysis advisement, Dr. Neusa Nogueira for the plant material symptomatic to Citrus Leprosis.

References

1. C. Spegazzini, Ann. Soc. Cient. **90**, 155–188 (1920)
2. H.S. Fawcett, *Citrus Diseases and Their Control*, 2nd edn. (McGraw-Hill, New York, 1936)
3. A.A. Bitancourt, O Biologo **3**, 37–40 (1937)
4. A.R. Vergani, Instituto de Sanidad Vegetal, Ministerio de Agricultura de la Nacion **5**, 3–11 (1945)
5. J.C.V. Rodrigues, Leprose do citros: citopatologia, transmissibilidade e relacao com o vetor Brevipalpus phoenicis Geijskes (Acari: Tenuipalpidae). Dissertation, University of Sao Paulo, Piracicaba, (1995), 79pp
6. F. S. Dominguez, A. Bernal, C.C. Childers, E.W. Kitajima, Plant Dis. **85**, 228 (2001)
7. J. Araya Gonzales, *Informe sobre la prospeccion de la leprosis de los citricos en la zona fronteriza sur* (Ministerio de Agricultura y Ganaderia, Costa Rica, Panama, 2000), pp. 5
8. J.C. Rodrigues, E.W. Kitajima, C.C. Childers, C.M. Chagas, Exp. Appl. Acarol. **30**, 161–179 (2003)
9. L.C. Knorr, Phytopathology **40**, 15 (1950)
10. E.W. Kitajima, G.W. Muller, A.S. Costa, W. Yuki, Virology **50**, 254–258 (1972)
11. J.C.V. Rodrigues, N.L. Nogueira, D.S. Freitas, H. Prates, Anais. Soc. Entomol. Brasil **26**, 391–395 (1997)
12. A.A. Bitancourt, Arquivos Inst. Biol. Sao Paulo **2**, 161–231 (1955)
13. M.R. Musumeci, V. Rossetti, Ciencia e Cultura **15**, 228 (1963)
14. E.W. Kitajima, M.A. Rosillo, M.M. Portillo, G.W. Muller, A.S. Costa, Fitopatologia **9**, 55–56 (1974)
15. A. Colariccio, O. Lovisolo, C.M. Chagas, S.R. Galleti, V. Rossetti, E.W. Kitajima, Fitopatol. Bras. **20**, 208–213 (1995)

16. J.C.V. Rodrigues, A.M. Fioravante, N.L. Nogueira, Fitopatol. Bras. **20**, 309–310 (1995)
17. E.W. Kitajima, C.M. Chagas, J.C. Rodrigues, Exp. Appl. Acarol. **30**, 135–160 (2003)
18. E.C. Locali, J. Freitas-Astua, A.A. Souza, M.A. Takita, G. Astua-Monge, R. Antonioli, E.W. Kitajima, M.A. Machado, Plant Dis. **87**, 1317–1321 (2003)
19. W. Maccheroni, M.C. Alegria, C.C. Greggio, J.P. Piazza, R.F. Kamla, P.R.A. Zacharias, M.B. Joseph, L.C. Assumpção, G. Camarotte, J. Cardozo, E.C. Casagrande, F. Ferrari, S.F. Franco, P.F. Giachetto, A. Girasol, H.J. Jordão, H.A. Silva, C.A. Leonardo Souza, I. Carlos Aguilar-Vildoso, S. Almir Zanca, Paulo Arruda, J.P. Kitajima, F.C. Reinach, J.A. Ferro, A.C.R. Silva, J. Virol. **79**, 3028–3037 (2005)
20. R.A. Valverde, S.T. Nameth, R.L. Jordan, Plant Dis. **74**, 255–258 (1990)
21. J. Sambrook, E.F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd edn. (Cold Spring Harbour Laboratory, Cold Spring Harbour, NY, 1989)
22. S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, J. Mol. Biol. **215**, 403–410 (1990)
23. D. Gordon, C. Abajian, P. Green, Genome Res. **8**, 195–202 (1998)
24. B. Ewing, P. Green, Genome Res. **8**, 186–194 (1998)
25. B. Ewing, L. Hillier, M.C. Wendl, P. Green, Genome Res. **8**, 175–185 (1998)
26. D. Gordon, in *Current Protocols in Bioinformatics*, ed. by A.D. Baxevanis, D.B. Davison (Section 11.2.1–11.2.43, Wliey, NJ,2004)
27. H. Sugawara, T. Koike, R. Lopez, T.J. Gibson, D.G. Higgins, J.D. Thompson, Nucleic Acids Res. **31**(13), 3497–3500 (2003)
28. D.L. Swofford, *PAUP: Phylogenetic Analysis Using Parsimony, version 4.0b 10* (Sinauer Associates, Sunderland, Massachusetts, 2003)
29. R.C. Pascon, A.C.R. Silva, Isolated nucleic acid molecules from the genome of Citrus Leprosis Virus and uses thereof, February 10th. U.S. Patent Serial No. 60/641,335 (2004)
30. J.E. Honig, J.C. Osborne, S.T. Nichol, Virology **321**, 29–35 (2004)
31. E. Kinsella, S.G. Matin, A. Grolla, M. Czub, H. Feldmann, R. Flick, Virology **321**, 23–28 (2004)
32. M.Y. Balakirev, S.O. Techerniuk, M. Jaquinod, J. Chroboczek, EMBO Rep. **4**, 517–522 (2003)
33. E.I. Savenkov, M. Sandgren, J.P. Valkonen, J. Gen. Virol. **80**(10), 2779–2784 (1999)
34. J.H. Graber, C.R. Cantor, S.C. Mohr, T.F. Smith, Proc. Natl. Acad. Sci. USA **96**, 14055–14060 (1999)
35. A. Gibbs, Phyl. Trans. R Soc. Lond. **354**, 593–601 (1999)