Complete nucleotide sequence, genomic organization and phylogenetic analysis of Citrus leprosis virus cytoplasmic type

E. C. Locali-Fabris,1,2 J. Freitas-Asstúa,1,3 A. A. Souza,1 M. A. Takita,1,4 G. Astúa-Monge,1 R. Antonioli-Luizon,1 V. Rodrigues,1 M. L. P. N. Targon1 and M. A. Machado1

Correspondence
E. C. Locali-Fabris eliane@centrodecitricultura.br

1Centro APTA Citros Sylvio Moreira, IAC, Rod. Anhanguera, Km 158, CP 4, 13490-970 Cordeirópolis/SP, Brazil
2UNESP, Universidade Estadual Paulista, Instituto de Biociências, Rod. D. Sartori 21-500, Distrito Rubião Jr 18618-000, Botucatu/SP, Brazil
3EMBRAPA Milho e Sorgo, Rod. MG 424, Km 65, CP 151, 35701-970 Sete Lagoas/MG, Brazil
4Centro de Recursos Genéticos Vegetais/IAC, CP 28, 13001-970 Campinas/SP, Brazil

The complete nucleotide sequence of the genomic RNA 1 (8745 nt) and RNA 2 (4986 nt) of Citrus leprosis virus cytoplasmic type (CiLV-C) was determined using cloned cDNA. RNA 1 contains two open reading frames (ORFs), which correspond to 286 and 29 kDa proteins. The 286 kDa protein is a polyprotein putatively involved in virus replication, which contains four conserved domains: methyltransferase, protease, helicase and polymerase. RNA 2 contains four ORFs corresponding to 15, 61, 32 and 24 kDa proteins, respectively. The 32 kDa protein is apparently involved in cell-to-cell movement of the virus, but none of the other putative proteins exhibit any conserved domain. The 5' regions of the two genomic RNAs contain a 'cap' structure and poly(A) tails were identified in the 3'-terminals. Sequence analyses and searches for structural and non-structural protein similarities revealed conserved domains with members of the genera Furovirus, Bromovirus, Tobravirus and Tobamovirus, although phylogenetic analyses strongly suggest that CiLV-C is a member of a distinct, novel virus genus and family, and definitely demonstrate that it does not belong to the family Rhabdoviridae, as previously proposed. Based on these results it was proposed that Citrus leprosis virus be considered as the type member of a new genus of viruses, Cilevirus.

INTRODUCTION

Citrus leprosis virus (CiLV) is the causal agent of leprosis, one of the most destructive diseases affecting citrus plants in the Americas. In the Brazilian citrus industry, this virus is responsible for yearly costs of $75 to 100 (US) millions for the chemical control of its vector, the tenuipalpid mite Brevipalpus phoenicis Geijskes (Rodrigues, 2000).

This disease affects various species within the genus Citrus and is characterized by the appearance of localized chlorotic lesions, commonly with necrotic ringspots, in leaves and stems, and depressed or salient chlorotic or brownish lesions in the fruits. Affected leaves and fruits may drop prematurely. Dieback can be observed in stems, and high viruliferous mite infestation can lead to the death of a tree (Bitancourt, 1955; Fawcett, 1936).

For years, the aetiology of leprosis was controversial. However, the constant association of the disease with infestation by mites of Brevipalpus species (Knorr, 1950), its transmission to healthy plants by grafting (Chagas et al., 1983; Knorr, 1968) and sap inoculations (Colariccio et al., 1995), the presence of virions and viroplasms in cells of diseased plants (Colariccio et al., 1995; Kitajima et al., 1972) and mites (Rodrigues et al., 1997), and detection of the virus only in infected plants and viruliferous mites by RT-PCR using specific primer pairs (Locali et al., 2003) have demonstrated that CiLV is indeed the causal agent of leprosis.

CiLV is one of the many plant viruses transmitted by Brevipalpus species mites. Virions accumulate in low titres only in the lesions and exhibit lability that has generally hindered attempts to purify CiLV and related viruses with few exceptions (Boari et al., 2004; Kondo et al., 2003).
Classified as a tentative member of the family Rhabdoviridae due to its morphology (enveloped, bacilliform virions) and cytopathic effects (Colariccio et al., 1995; Kitajima et al., 1972), two distinct types of CiLV particles are involved in leprosis symptoms: the nuclear ‘CiLV-N’ (100–110 nm in length and 40–50 nm in width) and the cytoplasmic ‘CiLV-C’ types (120–130 nm in length and 50–55 nm in width), based on where they accumulate in infected cells. The cytoplasmic type is prevalent, while the nuclear type is extremely rare (Colariccio et al., 1995; Dominguez et al., 2001; Kitajima et al., 1972, 2004). Recently, it was shown that the CiLV-C and CiLV-N are indeed different viruses and do not seem to share nucleotide sequences (Freitas-Astúa et al., 2005).

Double-stranded (ds) RNA patterns obtained from citrus plants infected with CiLV-C indicated the presence of two molecules of RNA between 8 and 10 kb and 4 and 5 kb (Colariccio et al., 2000; Locali et al., 2003; Rodrigues, 2000), besides other less intense RNA molecules of smaller sizes, possibly related to subgenomic RNAs (Locali et al., 2003). The first CiLV-C deduced protein sequences obtained showed high similarity with those of other viruses, particularly with the conserved helicase (Hel) and RNA polymerase (RdRp) domains, but also with the movement proteins (MP) of furu-, bromo-, tobra- and tobamoviruses and served as the basis for the development of the first molecular-based diagnosis method of infected plant and viruliferous mite tissues (Locali et al., 2003).

Despite the partial sequence similarity amongst CiLV-C and the above-mentioned viruses, most of the CiLV-C genome sequence does not present significant similarity with any sequence available in databases, suggesting that it belongs to a novel genus and a different family. Here, we report the complete CiLV-C genome sequencing, its genomic organization and phylogeny, and propose a new taxonomic classification for the causal agent of citrus leprosis.

METHODS

Plant, mite and virus propagation. Seedlings of sweet orange (Citrus sinensis L. Osb.) ‘Pera’ obtained from seeds and, hence, virus- and viroid-free, were kept under greenhouse conditions and infested every 15 days with CiLV-C viruliferous Brevipalpus phoenicis mites.

The mite population used in all experiments originated from a single female obtained from and reared for several generations onto fruits of the same variety in the Entomological Laboratory of the Centro APTA Citrus Sylvio Moreira. The fruits were covered with paraffin in order to avoid dehydration, with the exception of a region where the mites were reared, called arena, which contained a mixture of wheat flour, plaster, fine sand (1:1:2) and water, aiming to mimic citrus scab lesions and favour mite development.

The CiLV-C isolate used throughout the experiment was obtained from an infected sweet orange ‘Pera’ plant from Cordeirópolis, SP, and consistently induces the appearance of chlorotic and necrotic local lesions in the same sweet orange variety 17 to 21 days after mite inoculation. Non-viruliferous mites were transferred to the sources of inoculum (symptomatic fruits), where they were kept for 72 h for virus acquisition. After that period, the mites were evaluated by RT-PCR (Locali et al., 2003), to confirm the presence of virus and placed onto virus- and viroid-free sweet orange ‘Pera’ plants. After 20 days, young chlorotic lesions in leaves and stems were analysed through RT-PCR using specific primers and transmission electron microscopy to confirm the presence of the virus (Locali et al., 2003), and were used for library construction.

dsRNA extraction. Two grams of young chlorotic lesions was collected and powdered in liquid nitrogen. dsRNA was extracted using the CF-11 column method (Valverde et al., 1990). After the extraction, dsRNA was digested with RNaseA (Ribonuclease A, Bovine Pancreas, Calbiochem), Mung bean nuclease (Gibco-BRL) and DNase I-RNase-free (Boehringer Mannheim) to eliminate contaminating single-stranded (ss) RNA, ssDNA and dsDNA, respectively. The digestions were performed following the manufacturer’s instructions.

ssRNA and DNA extracted from citrus plants were used as controls of the digestions. Aliquots of dsRNA solution were visualized in a 1% agarose gel containing 0.5 µg ethidium bromide ml⁻¹. dsRNA bands were extracted from the low-melting-point (LMP) agarose gel and used as templates for the cDNA library.

cDNA synthesis, cloning and sequencing. Around 100 ng purified dsRNA was denatured at 95 °C during 10 min, and the synthesis of the first cDNA strand was performed using SuperScript II (Invitrogen) and either random primers or oligo(dT).

The product of the first round of PCR was diluted 10-fold and 5 µl of it was used as template for the second reaction, which was performed as described above. PCR products were purified from a 0-8% LMP gel with the QIAquick Extraction kit (Qiagen) and used for cloning. Fifty nanograms of the purified PCR product, 1 µl pGEM-T (50 ng µl⁻¹; Promega), 1 µl 10X ligation buffer and 5 U T4 DNA ligase were used in the ligation reaction, which was incubated for 16 h at 4 °C. After the ligation was complete, 10 µl of the reaction was used to transform Escherichia coli DH5α competent cells. The inserts carried by the recombinant clones were sequenced in automatic sequencers ABI Prism 3700 and 3730 (Applied Biosystems).

Analysis of the genome sequence. Sequences were analysed using the Sequencing Analysis Program from Perkin Elmer, Seq Man – Lasergene 99 (DNASTAR), Phred-Phrap-Consed using a threshold of 300 bases with Phred quality greater than 20. The assembly was done with the program CAP3. Sequences corresponding to the adapters and vector used for cloning were trimmed off and discarded before the analyses. These bioinformatics tools were used to form a genome scaffold and to verify the orientation and integrity of the contigs formed.

DNA sequences corresponding to the contigs and singletons obtained in the assembly were compared to the non-redundant GenBank database using the BLASTX and BLASTN algorithms. The first sequences with similarity to known viral genes were used to design primers using the Primer select software, part of the Lasergene 99 package (DNASTAR). The Primer3 program (www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) was used to design primers used to fill the sequence gaps by RT-PCR and produce fragments extending towards the 5' and 3' ends.
3’ ends. A combination of primers designed based on the internal sequences was used to amplify, clone and sequence the entire CiLV-C genome. The 3’-terminal sequence was determined using specific primers and oligo(dT), and the 5’-terminal sequence was obtained using a 5’-RACE kit (Invitrogen) following the manufacturer’s instructions.

Validation and confirmation of the CiLV-C sequences. The CiLV-C sequences were confirmed by hybridization with selected, representative cDNA clones through dot- and Northern blot analyses, as well as by PCR amplification using sets of primer pairs that amplify 1 kb fragments throughout the CiLV-C genome.

RNA extraction. Total RNA was extracted from 50 to 100 mg fresh healthy or symptomatic leaf tissues (Gibbs & Mackenzie, 1997). RNA concentration and purity were estimated by spectrophotometry and denaturing agarose gel electrophoresis [1% agarose, 6-7% formaldehyde, 10 × MOPS (200 mM MOPS, 5 mM sodium acetate, 10 mM EDTA) and DEPC-treated H2O].

Dot-blot hybridization. Dot-blot hybridization was carried out using biotin-labelled CiLV-C cDNA cloned sequences from various genomic regions as probes. The following materials were spotted onto membranes at approximately the same concentrations: the cloned sequences of the virus (the same used for labelling the probes), previously denatured at 90 °C for 10 min, ds- and ssRNA of sweet orange symptomatic for leprosis and ssRNA from healthy plants. After UV-cross-linking, the pre-hybridizations and hybridizations were carried out according to the standard protocol (Sambrook et al., 1989) and the specifications of the manual of the kit used for labelling the probes (Dig DNA Labelling kit; Roche). The detection was done by exposing HYPERfilm MP Amersham X-ray films to the membranes.

Northern-blot hybridization. For the Northern blots, 10 µg RNA was separated by formaldehyde gel electrophoresis and transferred to a nylon membrane (Hybond-N; Amersham Biosciences). Further steps were carried out as described above.

RT-PCR. RT-PCR was performed using 200 U M-MLV reverse transcriptase (Invitrogen), 1.5 µl 50 mM MgCl2, 100 ng total RNA and 100 ng random primers (3 µg µl−1). Samples were denatured at 95 °C for 10 min and placed onto ice. Then, 4 µl 5× buffer was added along with 1 µl (10 mM) dNTP mix, 0.5 µl (2 mM) DTT, 15 U RNase inhibitor (Gibco-BRL), 200 U M-MLV-RT (Gibco-BRL) and sterile Milli-Q water to a 20 µl final volume. The reaction was incubated at 37 °C for 2 h. All PCR amplifications were conducted in a PTC 100 thermocycler (MJ Research). The amplification reactions consisted of 2-5 mM MgCl2, 10 mM dNTP mix (Invitrogen), 100 ng of each specific primer, 2 µl cDNA used as template, 1 U of Taq DNA polymerase (Invitrogen), and sterile Milli-Q water for a final volume of 25 µl. An initial denaturing cycle at 94 °C for 2 min was followed by 32 cycles of denaturation at 94 °C for 30 s, annealing at 56-60 °C (according to each specific primer pair) for 30 s and extension at 72 °C for 40 s. A final 5 min extension was added to the cycle.

Sequence comparison. Genome structure, identification of open reading frames (ORFs) and conserved domains, as well as translated protein sequences were obtained using the BLASTN, BLASTX, ORF finder and BL2seq, available at http://www.ncbi.nlm.nih.gov/, and Pfam, available at http://pfam.wustl.edu/hmmsearch.shtml. In order to taxonomically classify CiLV-C, phylogenetic analyses were performed on translated sequences of the conserved proteins RdRp, Hel and MP. These CiLV-C sequences were compared with those obtained by Pfam. Analyses were done for the complete and for the conserved sequences (eliminating all the non-aligned ends) using the CLUSTAL W program (Thompson et al., 1994) available at http://clus-talw.genome.jp/. Phylogenetic analysis was carried out using the maximum-parsimony method as implemented by the PAUP* software 4.0b4a (Swofford, 2000). In addition, the data were analysed using the PHYLIP 3.65 program (http://evolution.genetics.washington.edu/phylip.html).

RESULTS

Transmission, symptoms and detection of CiLV

Samples of viruliferous mites reared on symptomatic leaf tissue tested positive for CiLV-C by RT-PCR (Locali et al., 2003). Non-viruliferous mites did not yield any band in a 1% agarose gel, while the viruliferous mites showed bands of sizes expected to CiLV (data not shown). Seventeen to 21 days after viruliferous mite infestation, inoculated citrus seedlings exhibited typical leprosis symptoms in leaves and stems. The presence of bacilliform particles characteristic of CiLV-C was visualized by electron microscopy (Locali et al., 2003) and virus presence was confirmed by RT-PCR using specific primers reported elsewhere (Locali et al., 2003). Asymptomatic, non-inoculated plants did not yield any band in agarose gels (data not shown).

dsRNA isolation

CiLV-C dsRNA migrated as two RNAs with molecular masses of approximately 9 and 5 kb (Fig. 1), consistently detected in preparations from sweet orange-infected tissues but never from healthy samples. The RNAs remained in the gel after the digestions with RNaseA, DNase I and Mung bean nuclease, indicating that they indeed were of dsRNA nature.

cDNA libraries and sequencing

The overall number of valid reads sequenced was 7800, with insert mean size of 500 nt. The sequencing covered approximately 53 times the CiLV-C complete genome.

Fig. 1. Pattern of dsRNA of CiLV-C from sweet orange. Lanes: 1, ‘Pera’ sweet orange infected only with Citrus tristeza virus – CTV (control of the dsRNA extraction); 2, ‘Pera’ sweet orange infected with CiLV-C; 3, ‘Cleopatra’ mandarin infected with CiLV-C; 4, Healthy ‘Pera’ sweet orange (negative control); M, 1 kb ladder (Gibco-BRL).
About 3500 of these reads were obtained from two cDNA genomic libraries constructed from CiLV-C dsRNA and 4300 from the PCR fragments obtained with primer walking or 5’-RACE strategy. These 7800 reads were clustered into two contigs, one of 8745 bp and the other of 4986 bp, corresponding to RNA 1 and RNA 2 of CiLV-C, respectively.

**CiLV-C genome**

The complete nucleotide sequence of the CiLV-C suggests that it is a ~14 kb, likely bipartite, positive-sense, ssRNA virus. RNA 1 contains 8745 nt, excluding the 3’-terminal poly(A) tail, and comprises two ORFs (Fig. 2). RNA 2, of 4986 nt, contains four ORFs (Fig. 2), two of them in-frame +1 and the others in-frame +3.

ORF 1 of RNA 1 (frame +1), begins at AUG and finishes at UAA codons (109–7647 nt positions), and encodes a large polypeptide with 2512 aa (286 kDa) (Fig. 2). Analysis of this translated polyprotein reveals conserved domains for methyltransferase (MTR; aa 126–528), cysteine protease (C-Prot; aa 689–797), Hel (aa 1558–1841) and RdRp (aa 2055–2494). ORF 2 of RNA 1 (frame +3), starts at AUG (7710 nt), finishes at UAG (8501 nt), and encodes a putative protein of 263 aa (29 kDa) (Fig. 2) with no conserved domains. ORF 2 of RNA 1 is followed by a 3’ non-translated region (NTR) of 228 nt and a poly(A) tail of undetermined length.

The first ORF of RNA 2 (frame +1), with start codon AUG at nt 67 and stop codon UAA, at nt 459, encodes a putative protein with 130 aa (15 kDa). ORF 2 in this RNA (frame +3) encodes the largest putative protein of this RNA (537 aa, 61 kDa), which begins at AUG and finishes at UAA (1590–3203 nt positions). ORF 3 of RNA 2 (frame +3) encodes a putative protein with conserved domain to viral MP. This putative protein starts at AUG (3228 nt) and finishes at UAA (4121 nt), and consists of 297 aa (32 kDa). The last ORF (4) in RNA 2 (frame +1), with start codon AUG at nt 4093 and stop codon UGA at nt 4737, encodes a putative protein with 214 aa (24 kDa). None of the putative proteins encoded by ORFs 1, 2 and 4 of RNA 2 present any conserved domains or similarity with other known sequences in databases. ORF 4 of RNA 2 is followed by a 3’ NTR of 232 nt, followed by a poly(A) tail of undetermined length.

The conserved sequence of GAUAAAUCU was found at the 5′ terminus of both RNAs 1 and 2 of CiLV-C.

The complete CiLV-C nucleotide sequence was submitted to GenBank and has been assigned accession nos DQ352194 (CiLV-C RNA 1) and DQ352195 (CiLV-C RNA 2).

**Phylogenetic analysis**

The search for conserved domains of CiLV-C using Pfam resulted in a list of virus species with their corresponding genera or families that share the same domains and were used for the parsimony analyses. Fifty-eight members of several virus groups that shared various levels of similarity with the putative Hel of CiLV-C and their corresponding accession numbers, as well as 38 virus species that presented similarity with the putative CiLV-C’s RdRp and 14 virus sequences with similarity to CiLV-C putative MP, are listed in Supplementary Tables S1, S2 and S3 available in JGV Online.

The data revealed that the CiLV-C putative Hel sequence groups in a clade with representatives of the genera *Benyvirus*, *Hepevirus* [*Hepatitis E virus* (strain Mexico)] and *Vitivirus*. The result of the alignment amongst them can be seen in Fig. 3. CiLV-C RdRp appears in a clade along with members of the genera *Tobamovirus*, *Furovirus*, *Pomovirus*, *Tobravirus* and *Bromovirus*, all of them typically plant virus genera and, with the exception of the last one, of rigid, rod-shaped morphology (Fig. 4). The CiLV-C putative MP appears as the last member of a clade including members of the genera *Furovirus*, *Bromovirus* and *Umbravirus*, being closer to the umbravirus *Carrot mottle mimic virus* (Fig. 5).

---

**Fig. 2.** Schematic drawing of the genome structure and organization of CiLV-C. Each RNA is presented as a line with the corresponding ORFs indicated by rectangles. ORFs located in different frames are represented above (+1) or below (+3) the lines. Functional domains attributed according to the translated sequences are indicated inside the rectangles. For those translated sequences with unidentified function, the approximate molecular mass (in kDa) is indicated inside the rectangles. MTR, methyltransferase domain; C-Prot, cysteine protease domain; Hel, helicase domain; RdRp, RNA-dependent RNA polymerase domain; MP, movement protein domain.
Fig. 3. Most parsimonious tree resulting from the analysis of 58 taxa and 1912 informative characters of the Helicase family by Pfam [analysis 2: length = 27,650; consistent index (CI) = 0.6556; retention index (RI) = 0.5603]. Numbers ahead of each clade are bootstrap values (1000 heuristic searches with 100 pseudoreplicates per search) obtained in analysis 1, which were omitted when lower than 50%. Implemented by PAUP. See Supplementary Table S1 (available in JGV Online) for virus acronyms and further information.
Regardless of the program used or the length of the sequence analysed, the results were basically the same.

**DISCUSSION**

In order to ensure that the citrus plants used in this experiment were originally virus- and viroid-free, they were obtained by seed, since there are no reports of seed transmission of such pathogens. Viruliferous mite CiLV-C inoculation was very efficient and led to a large number of typical leprosis local lesions in all aerial parts of the plants. It was noted that only chlorotic lesions yielded high quality RNA, while necrotic symptoms did not (data not shown). Hence, the use of young, yellowish lesions was very important for achieving the sequence obtained here. Genomic information obtained though BLASTN analysis of RNA 1, revealed 99% identity between the MTR motif sequence (nt 1103–1502) obtained in this study and the partial sequences of CiLV-C available at GenBank [accession nos gi|62184184 (Bolivian isolate) and gi|33621221 (Brazilian isolate)].
Through BLASTX, CiLV-C RNA 1 sequences (nt 6173–7570) exhibited similarity with a gene that encodes a 59 kDa readthrough protein involved in replication of the *Soil-borne wheat mosaic furovirus* (SBWMV) (9e⁻⁴², accession no. gi|7634688), as well as with replication protein genes of other furo- (gi|6018639), tobamo- (gi|19908647), tobra- (gi|21427640) and bromoviruses (gi|2791890), among others. As expected, an RdRp motif was identified in this region of the CiLV-C putative polyprotein.

Four conserved domains were identified in the polyprotein of ORF 1, RNA 1. The first one is a viral MTR, aa 126–528, that is involved in mRNA capping. This domain is found in a large number of ssRNA positive-sense viruses, such as hordei-, tobra-, tobamo-, bromo-, clostero- and caliciviruses, as described at Pfam (accession no. PF01660). The second conserved domain is an OTU (*Ovarian Tumor*)-like C-Prot (aa 689–797, Pfam accession no. PF02338). A third domain corresponds to the viral (superfamily 1) RNA Hel (aa 1558–1841, Pfam accession no. PF01443), which plays several roles in different stages of virus replication and is commonly found in ssRNA positive-sense viruses. Finally, the fourth conserved domain is the RdRp (aa 2055–2494), which correspond to an RdRp containing 2A proteins of bromoviruses, RdRp of tobamoviruses and non-structural polyproteins of togaviruses (Pfam accession no. PF00978).

The RNA 1 ORF 2 sequence did not exhibit significant similarity with any other sequence available in GenBank by BLASTX and its deduced protein did not exhibit any similarity with protein family domains at the Pfam. Conserved domains, despite the high e value, were only identified through ORF finder, at very low similarity, with *Kyzylagach virus* and *Sindbis virus*, both ssRNA positive-sense viruses, genus *Alphavirus*, family *Togaviridae* (accession nos gb|AAO33327|1 and gb|AAC83379|1, respectively). In both cases, the deduced protein corresponding to the nt 33–119 region presented sequence similarity to a structural polyprotein precursor corresponding to the capsid protein (CP) gene. This suggests that the RNA 1 ORF 2 could be involved in virion encapsidation. However, further information is necessary to confirm this hypothesis.

Identities of 99% were found between a small region of the RNA 2 nucleotide sequence of CiLV-C (nt 3313–3762) and those of the same virus available at GenBank (Brazilian, Bolivian and Colombian isolates accession nos gi|33621220, gi|62184186 and gi|82502204, respectively) through BLASTN. This region corresponds to the gene that putatively encodes the viral MP. Through BLASTX, it was possible to determine significant similarities between the CiLV-C nt 3228–3905 and the sequences of the tentative MP of the furovirus *Sorghum chlorotic spot virus* (SgCSV) (accession no. gi|21427640) as well as other ssRNA positive-sense furo- (gi|6018641) and bromoviruses (gi|9626936). The RNA 2 ORF 3 shared a conserved domain with the family 3A MP detected by BLASTX and Pfam. Other members of the 3A family are bromo- and cucumoviruses, which have tripartite RNA genomes, with the third RNA comprising the MP and the CP genes (Pfam accession no. 00803).

No significant similarities were found by BLASTX amongst RNA 2 ORFs 1 (nt 67–459, frame +1), 2 (nt 1590–3203, frame +3) and 4 (nt 4093–4737, frame +1) and sequences available at GenBank. In addition, no conserved domains of protein families were identified by Pfam, suggesting high specificity of the CiLV-C proteins of unknown function.

Since the first micrographs of leprosis-infected citrus cells showed the resemblance between *Citrus* leprosis virus and members of the family *Rhodoviridae* (Kitajima et al., 1972), this and other *Brevipalpus*-transmitted viruses (BTV) such as the Orchid fleck virus (OFV) (Doi et al., 1977) and the Coffee ringspot virus (CoRSV) (Chagas et al., 1983) have been considered tentative members of this family [ICTV 01.062.0.00.359 (GiLV), 01.062.0.00.311 (CoRSV) and Unassigned (OFV) at http://www.ncbi.nlm.nih.gov/ICTVdb/ictv/index.htm].

However, despite the morphology and cytopathic effects they induce in infected cells, there is little additional information supporting this classification (Jackson et al., 2005). The only genomic information on BTV available at

**Fig. 5.** One of the two most parsimonious trees resulting from the analysis of 15 taxa and 368 informative characters of MP by Pfam (analysis 1: length = 1378; CI = 0.9340; RI = 0.9227). Numbers ahead of each clade are bootstrap values (1000 heuristic searches with 100 pseudoreplicates per search) obtained in the analysis 1, which were omitted when lower than 50%. Implemented by PAUP. See Supplementary Table S3 (available in JGV Online) for virus acronyms and further information.
GenBank are accessions of nucleocapsid gene sequences of OFV isolates, and a few Brazilian, Colombian and Bolivian CiLV-C sequences submitted by our group. Even though the OFV sequences present similarity with nucleorhabdoviruses (and OFV indeed accumulates in the nucleus of infected cells), the same was not found for CiLV-C. Instead, no similarity was found between sequences of this virus and other cyto- or nucleorhabdoviruses available at GenBank. On the contrary, similarity was found with ssRNA positive-sense viruses such as furo-, tobamo- and tobraviruses, among others (Locali et al., 2003). In a recent review on plant rhabdoviruses, Jackson et al. (2005) reported that more than 75 viruses (including CiLV, OFV and CoRSV), tentatively classified as rhabdoviruses, do not belong to a specific genus since there is no available information on their molecular properties which, along with their biological properties, could unequivocally classify them as members of this or other virus families. Considering the International Committee on Taxonomy of Viruses (ICTV) key for classification of rhabdoviruses and the genomic information on CiLV-C obtained in this study, it is possible to conclude that CiLV-C should not be considered a member of the family Rhabdoviridae. According to the Committee, all rhabdoviruses share similar biochemical composition of the virions [ssRNA (negative-sense)], replication strategy, nature of the particle structure and genome organization (monopartite). Not only is the CiLV genome likely bipartite and of ssRNA positive-sense, but there is a complete lack of sequence similarity between this virus and actual rhabdoviruses, which clearly excludes it from the family.

The phylogenetic analyses showed, once again, that CiLV-C is not a member of the family Rhabdoviridae, since none of the translated sequences analysed yielded any similarity with those of rhabdoviruses. In addition, they clearly demonstrated that CiLV’s putative proteins exhibit similarity with those of ssRNA positive-sense viruses from several genera and families. However, the results clearly demonstrate that it is not possible to classify the virus taxonomically within any currently accepted genus or family, since it appears to be a unique virus.

When considering all of the available information presented, it becomes clear that it belongs to a novel virus genus. The facts that it is transmitted by a tenuipalp bid mite, Brevipalpus species, and that its morphology resembles that of rhabdoviruses, but is significantly smaller than members of that family, the particular organization of its bipartite genome and the lack of sequence similarity among most of its ORFs and the sequences available at GenBank clearly suggest that this virus cannot be classified within any of the genera currently accepted by the ICTV. Hence, we propose that Citrus leprosis virus be considered as the type member of a new genus, Cilevirus, which would possibly include other members of the bacilliform, cytoplasmic viruses transmitted by Brevipalpus species that share sequence similarity with CiLV-C (data not shown).

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to Luciano Kishi, Eduardo Femino Carlos, Vinicius Lima and Marcelo da Silva Reis for their help with the analyses and discussions on the CiLV-C genome assembly and organization, and to Ana Carla Oliveira da Silva Pinhati for discussions on the phylogenetic analyses. We would also like to thank FAPESP for the financial support (Process nos 2003/05307-5 and 2004/10511-3).

REFERENCES


