



Molecular characterization reveals Brazilian *Tomato chlorosis virus* to be closely related to a Greek isolate

Leonardo C. Albuquerque^{1,2,3}, Francisco Villanueva², Renato O. Resende¹, Jesús Navas-Castillo³, Júlio César Barbosa⁴ & Alice K. Inoue-Nagata¹

¹Embrapa Hortaliças, Km 09, BR060, Cx. Postal 218, 70359-970, Brasília, DF, Brazil; ²Instituto de Hortofruticultura Subtropical y Mediterránea “La Mayora” (IHSM-UMA-CSIC), Consejo Superior de Investigaciones Científicas, Estación Experimental “La Mayora”, 29750 Algarrobo-Costa, Málaga, Spain; ³Universidade de Brasília, 70910-970, Brasília, DF, Brazil; ⁴Departamento de Fitopatologia e Nematologia, ESALQ, USP, 13418-900, Piracicaba, SP, Brazil

Author for correspondence: Alice K. Inoue-Nagata, email: alice.nagata@embrapa.br

ABSTRACT

Tomato chlorosis virus (ToCV, genus *Crinivirus*, family *Closteroviridae*) is a whitefly-transmitted crinivirus with a bipartite RNA genome. This virus is emerging as a serious threat to tomato crops worldwide. To date, only three complete genomic sequences of ToCV have been described from North America, Spain, and Greece isolates. In this study, we present the fourth complete nucleotide sequence of the RNA 1 (8594 nt) and RNA 2 (8242 nt) components of a Brazilian ToCV isolate (ToCV-BR). The complete genome sequences of RNA 1 and RNA 2 have been deposited in the GenBank database under the accession numbers JQ952600 and JQ952601, respectively. The sequences of RNA 1 and RNA 2 shares the highest nucleotide identity of 99.6% and 99.5%, respectively, with the Greek isolate sequences. Phylogenetic analysis confirmed that both RNA 1 and RNA 2 of the Brazilian isolate are most closely related to the Greek isolate of that virus. These results suggest that ToCV may have been recently introduced to Brazil from Europe.

Key words: *Crinivirus*, *Solanum lycopersicum*, ToCV, tomato.

Tomato chlorosis virus (ToCV) is an emerging virus that currently threatens the tomato production worldwide. It is a typical member of the genus *Crinivirus* (family *Closteroviridae*) with a bipartite genome of single-stranded RNA of positive polarity. Both RNA molecules are separately encapsidated in long, flexuous virions that are transmitted in nature by the whiteflies *Bemisia tabaci* (Gennadius), *Trialeurodes vaporariorum* Westwood and *T. abutilonea* Haldeman in a semipersistent manner (Wisler et al., 1998). ToCV was first described in tomato plants cultivated in Florida, USA, in 1989, and the disease was named “yellow leaf disorder” (Wisler et al., 1998). Since then, ToCV has been reported in many countries from the Americas, Europe, Africa and Asia (Fiallo-Olivé et al., 2011; Navas-Castillo et al., 2011). In Brazil, the presence of ToCV has been confirmed at least in five states (Barbosa et al., 2011), which are important tomato growing areas. Symptoms of ToCV infection in tomato plants include interveinal chlorotic areas especially in the lower leaves, which become brittle and may show necrotic spots. Complete genome sequences of ToCV are available only for isolates from the USA (Wintermantel et al., 2005), Spain (Lozano et al., 2006, 2007) and Greece (Kataya et al., 2008). Genetic information about ToCV in Brazil has been limited to only one partial sequence of the heat shock protein 70 family homolog gene (Hsp70h) (Barbosa et al., 2008). To better understand this important emerging tomato

pathogen and relationships between ToCV isolates, we cloned and sequenced the complete genome of both RNA 1 and RNA 2 of an isolate of ToCV from Brazil.

The Brazilian ‘JB’ isolate, referred to here as ToCV-BR, partially characterized by Barbosa et al. (2008), was selected for this study. This isolate was obtained in 2010 from a naturally infected tomato plant in Sumaré, São Paulo, Brazil and was maintained in a tomato plant at the University of São Paulo. Total RNA was extracted from leaf tissue by using Trizol reagent (Invitrogen, Carlsberg, CA, USA) as described by the manufacturer. ToCV-specific primers (Table 1) were designed based on available sequences of ToCV (Wintermantel et al., 2005; Lozano et al., 2006, 2007; Kataya et al., 2008) and used in RT-PCR to obtain the complete viral genome. Reverse transcription was performed using Superscript III (Invitrogen), and PCR amplification was conducted using Expand High-Fidelity polymerase (Roche, Mannheim, Germany). cDNA corresponding to the 5’ ends of both RNA 1 and RNA 2 were obtained using the 5’ RACE System for Rapid Amplification of cDNA Ends Version 2.0 (Invitrogen) and specific primers MA464 and MA407 in the PRO domain of ORF1a for RNA 1 and MA341 and MA316 in the 5’ UTR and Hsp70h gene, respectively, for RNA 2. To clone the 3’ end, total RNA was polyadenylated and used as template for cDNA synthesis. The cDNA was subsequently amplified using an oligo (dT) and the degenerate primer

TABLE 1 - Primers used for amplification of RNA 1 and RNA 2 cDNA

Primer	Sequence (5' to 3')	Position ^a	Region
RNA 1			
MA406 (+)	GAAATAGTATTCGTGTGATTAC	1-22	5' UTR
MA464 (-)	CCGTCTAAGTACGTTTCATGC	519-538	ORF1a (Pro)
MA407 (-)	TCTTACTAAAAACTCTGGAGTA	1043-1023	ORF1a (Pro)
MA408 (+)	GGCTTGATTTTTCTGTGGTG	888-909	ORF1a (Pro-MTR)
MA412 (+)	CGGTGGTTTTGTAAGTGGTTTC	3713-3734	ORF1a
MA458 (-)	GGTCGTCACATCCACAACCTG	4788-4769	ORF1a
MA329 (+)	TGYGATGARRTYTTYATGRTKCAT	5456-5479	ORF1a (HEL)
MA350 (-)	ATATCATCAAGGTCGTAAACAC	7201-7180	ORF1b (RdRp)
MA324 (+)	GRACDGGDTDCDCRAAYACKTGG	7120-7142	ORF1b (RdRp)
MA325 (-)	ATRWARCACAAHGCGYGCRTABGA	7585-7566	ORF1b (RdRp)
MA279 (-)	CCGGATCCTCTAGAGCGGCCGC(T) ₁₇ V	Oligo (dT)	
RNA 2			
MA341 (-)	AAATGCTCAACCACTCAACC	533-511	5' UTR
MA260 (+)	GCTTCCGAAACTCCGTCTTG	907-926	Hsp70h
MA316 (-)	GGAACAGACACGTTGAGATCGG	1172-1151	Hsp70h
MA318 (+)	GATCGACTTATCTCGAACCTGC	5005-5023	CP
MA311 (-)	TTAGCAACCAGTTATCGATGC	5089-5109	CP
MA378 (+)	TCCTGTTGAGGTAAGAGCGAAT	7613-7634	p27
MA379 (-)	GTACCAGAGTATCTGAACAC	7786-7767	p27
MA279 (+)	CCGGATCCTCTAGAGCGGCCGC(T) ₁₇ V	Oligo (dT)	

^aNucleotide position of the primers in the sequence of RNA 1 (DQ983480) and RNA 2 (DQ136146) of ToCV.

MA324 in the RdRp gene for RNA 1 amplification, and the specific primer MA378 in the p27 gene for RNA 2. All DNA fragments amplified by PCR (Figure 1) were cloned by ligation into pGEM-T Easy (Promega, Fitchburg, USA) and transformed into *Escherichia coli* DH5 α by electroporation. Two recombinant clones for each virus cDNA fragment were selected, and inserts were sequenced using primer walking at Macrogen, Inc. (Seoul, South Korea). SeqMan (MEGALIGN, DNASTar Inc., Madison, WI, USA) was used for analysis and assembly of the sequences. Identification of open reading frames (ORFs) and protein translation were conducted with ORF finder (<http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi>), and MEGA 5 (Tamura et al., 2011) was used for phylogenetic inference by the neighbor-joining method. The complete genome sequences of RNA 1 and RNA 2 have been deposited in the GenBank database under the accession numbers JQ952600 and JQ952601, respectively.

The bipartite genome of the Brazilian ToCV isolate (ToCV-BR) consists of 8594 and 8242 nucleotides (nt) for RNA 1 and 2, respectively, and has the typical characteristics of other described ToCV isolates (Wintermantel et al., 2005; Lozano et al., 2006, 2007; Kataya et al., 2008) (Figure 1). The RNA 1 sequence contains four ORFs (Figure 1a) and shares the highest nucleotide identity with the Greek isolate (Kataya et al., 2008) (99.6%), followed by the North American (Wintermantel et al., 2005) (99.2%) and Spanish (Lozano et al., 2007) (97.4%) isolates. Based on homology to other viral replication proteins, the first two ORFs of RNA 1 encode proteins involved in replication of viral RNA. ORF 1a extends from 304 to 6141 nt and encodes the conserved closterovirus protein with the protease,

methyltransferase, and helicase (Pro-MTR-HEL) domains. ORF 1b (nt 6212-7657) encodes the RNA-dependent RNA polymerase (RdRp). The region between the ToCV MTR and HEL domains contained no significant similarities with any other known proteins, and the ORF 1b is presumed to be expressed by a +1 translational frameshift, as suggested for other closteroviruses. ORF 2 (nt 7664-8245) encodes a 22 kDa protein, which is a suppressor of gene silencing (Cañizares et al., 2008). A small ORF 3 (nt 8265-8420) is located at the 3' end of RNA 1 and encodes a putative 6-kDa protein with a transmembrane domain, similar to proteins at the 3' end of RNA 1 of other criniviruses, also of unknown function. Additionally, RNA 1 contains a 303-nt 5' UTR and a 174-nt 3' UTR. The RNA 2 sequence also shares the highest nucleotide identity with the Greek isolate (99.5%), followed by the American (99.0%) and the Spanish (97.6%) isolates. The genomic structure of the RNA 2 is schematically represented in Figure 1b. RNA 2 contains nine ORFs and encodes the proteins p4 (ORF 1, nt 239-340), Hsp70h (ORF2, nt 733-2397), p8 (ORF 3, nt 2407-2610), p59 (ORF 4, nt 2562-4115), p9 (ORF 5, nt 4097-4333), CP (the coat protein, ORF 6, nt 4333-5106), CPm (the minor coat protein, ORF 7, nt 5112-7121), p27 (ORF 8, nt 7125-7823) and p7 (ORF 9, nt 7832-8029), which are likely to be involved in viral movement and encapsidation. The putative 5' and 3' UTRs of RNA 2 are 238 and 213 nt in length, respectively. As observed with other criniviruses, including sequenced isolates of ToCV, the 5' and 3' ends of both RNAs are remarkably similar. For the ToCV-BR isolate, the first six nucleotides of the 5' UTR of RNA 1 and 2 were identical to those of other ToCV isolates, except

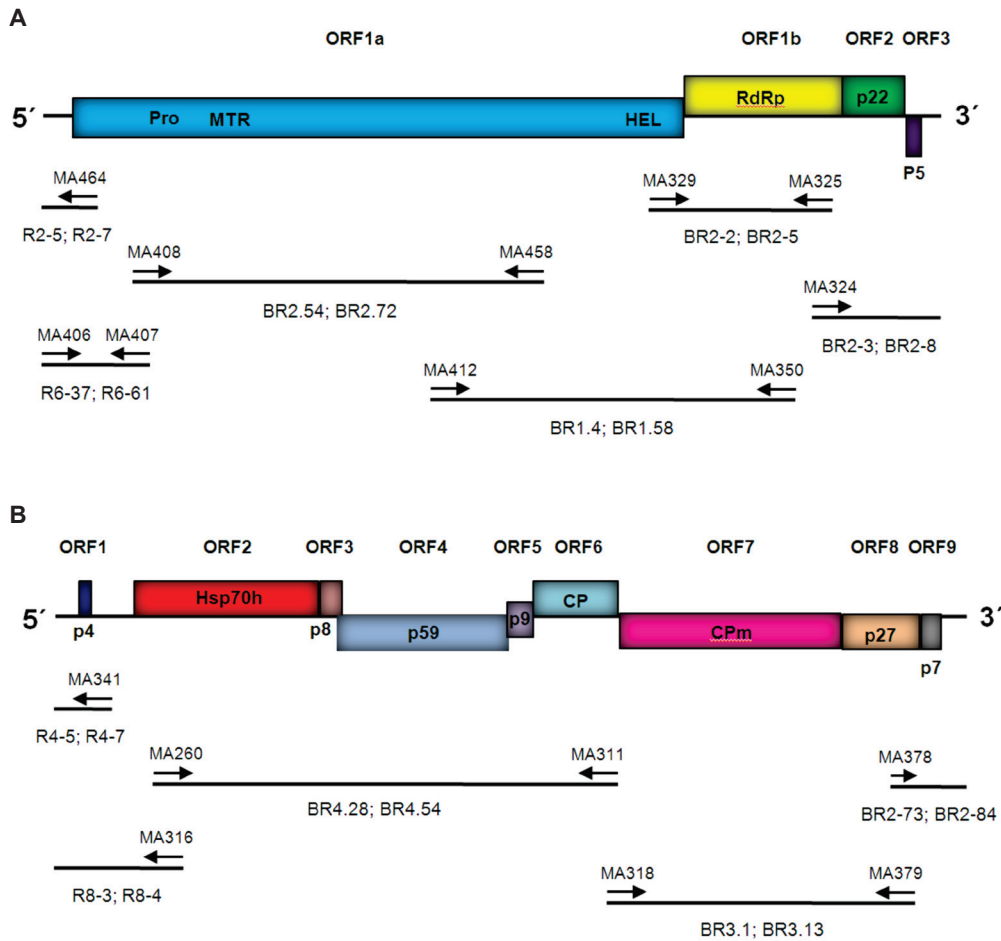


FIGURE 1 – A. Schematic representation of the genomic structure of RNA 1 and **B.** RNA 2 of the Brazilian isolate ToCV-BR of *Tomato chlorosis virus* (ToCV) and the strategy used for cloning and sequencing. Lines represent the genomic RNA; boxes represent ORFs with putative protein products indicated above or below. Arrows represent primers used for PCR amplification (see Table 1), and horizontal bars represent the sequenced clones.

that the RNA-1 of ToCV-BR contained five A's instead of three A's observed in the RNA-2 of the same isolate or in both RNA molecules of the other ones. The function of this conserved region and the reason for the presence of the additional A's is not understood. The first 176 nucleotides located at the 3' UTR of RNA 1 share 89% nt identity with the 3' terminus of RNA 2. This conservation at the 3' end of RNAs 1 and 2 is well known among members of this genus (Wintermantel et al., 2005).

Comparison of the nucleotide and amino acid sequences from the Brazilian ToCV isolate (ToCV-BR) and the fully sequenced ToCV isolates (Table 2) confirmed that both RNA 1 and 2 from the Greek isolate are the most closely related to ToCV-BR. An exception is ORF2 and ORF6, which share the highest amino acid identity of 99.5% and 98.8%, respectively, with those of the North American ToCV isolate (Wintermantel et al., 2005).

A phylogenetic tree constructed with the full genome sequence of ToCV isolates and the most related criniviruses clearly indicates that the Brazilian isolate is closely related to the Greek isolate for both RNA 1 (Figure 2A) and RNA 2 (Figure 2B) segments. As only four full sequences are

TABLE 2 – Nucleotide (nt) and amino acid (aa) pairwise identity percentages between the full genome sequence of the Brazilian *Tomato chlorosis virus* (ToCV) isolate and three other ToCV isolates from the USA (ToCV-US), Spain (ToCV-Sp) and Greece (ToCV-Gr)

Virus	ToCV-US		ToCV-Sp		ToCV-Gr	
	aa	nt	aa	nt	aa	nt
ORF/Protein						
RNA 1						
ORF 1a	99.0	99.1	97.7	97.1	99.4	99.6
ORF 1b (RdRp)	100	99.5	99.8	98.2	100	100
ORF 2/ p22	99.5	99.1	95.9	96.9	98.4	99.0
ORF 3/ p5	100	100	100	98.7	100	100
RNA 2						
ORF 1/ p4	90.9	97.1	90.9	97.1	100	99.0
ORF 2/ Hsp70h	98.7	99.2	98.7	98.3	99.5	99.5
ORF 3/ p8	97.0	99.0	95.5	97.5	97.0	99.0
ORF 4/ p59	99.4	99.3	98.3	97.7	99.4	99.6
ORF 5/ p9	100	99.6	100	100	100	100
ORF 6/ CP	98.8	98.7	96.5	97.2	98.1	99.2
ORF 7/ CPm	98.8	99.0	97.5	97.2	99.4	99.5
ORF 8/ p27	99.6	99.3	97.8	97.4	99.6	99.7
ORF 9/ P7	100	100	98.5	99.0	100	100

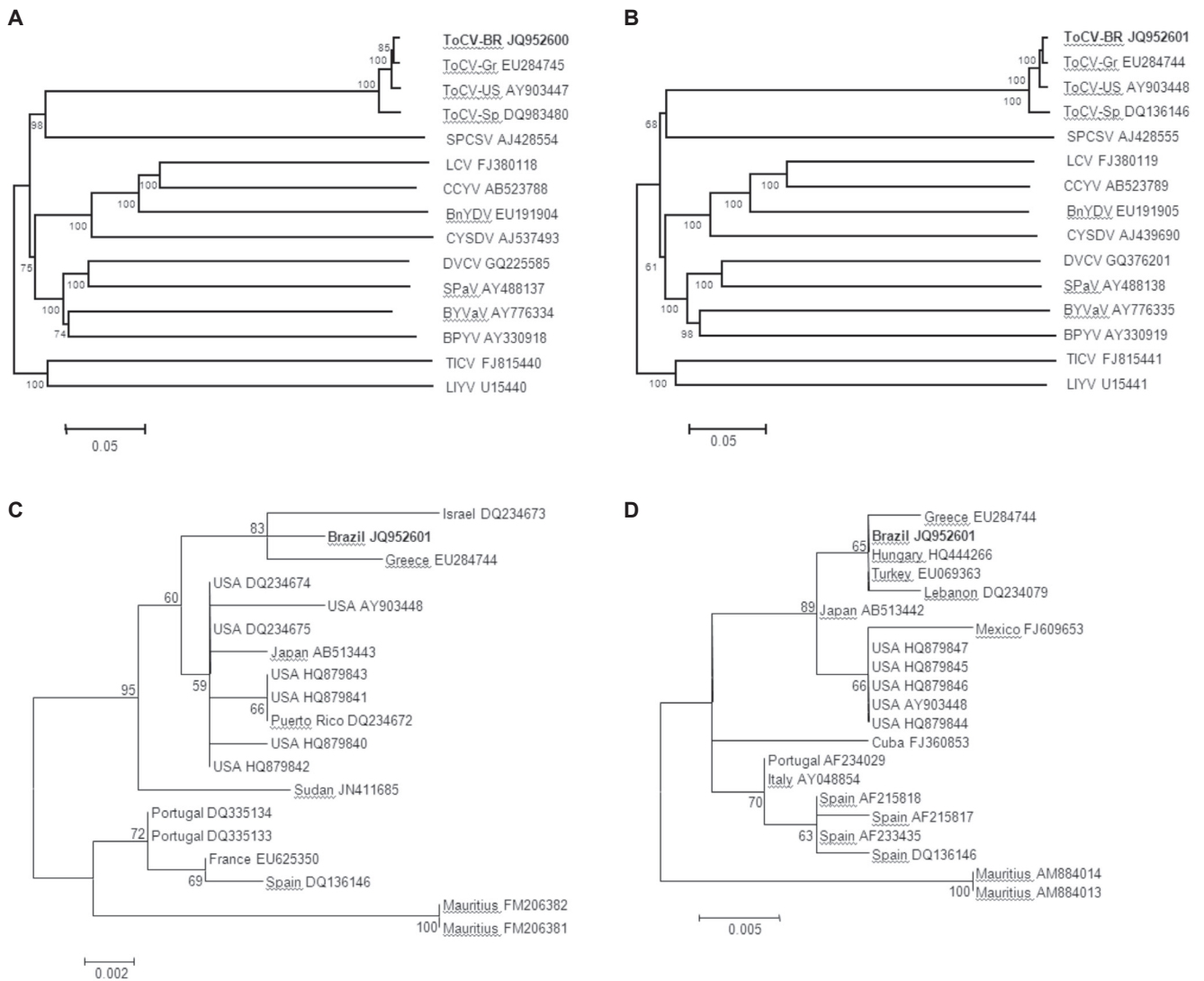


FIGURE 2 - Neighbor-joining phylogenetic trees illustrating the relationships of the Brazilian isolate of *Tomato chlorosis virus* (ToCV-BR) to other members of the genus *Crinivirus*. **A.** Comparisons are shown for complete nucleotide sequences of RNA 1 and **B.** RNA 2. The acronyms and names are as follows: BnYDV, *Bean yellow disorder virus*; BPYV, *Beet pseudo yellows virus*; BYVaV, *Blackberry yellow vein-associated virus*; CCYV, *Cucurbit chlorotic yellows virus*; CYSDV, *Cucurbit yellow stunting disorder virus*; DVCV, *Diodia vein chlorosis virus*; LCV, *Lettuce chlorosis virus*; LIYV, *Lettuce infectious yellows virus*; SPaV, *Strawberry pallidosis-associated virus*; SPCSV, *Sweet potato chlorotic stunt virus*; ToCV-Gr, *Tomato chlorosis virus-Greece*; ToCV-Sp, *Tomato chlorosis virus-Spain*; ToCV-US, *Tomato chlorosis virus-United States* and TICV, *Tomato infectious chlorosis virus*. **C.** Phylogenetic trees based on alignments of the partial sequences of CP and **D.** Hsp70h ORFs of ToCV-BR with corresponding sequences of selected ToCV isolates obtained from GenBank. Accession numbers are shown in the trees. Bootstrap (1000 replicates) values are shown as percentages. Bars below the trees indicate the number of substitutions per site.

presently available, it is not possible to perform a thorough comparison with the isolates found worldwide. Hence, the nucleotide sequences of the two best characterized ToCV ORF's CP and Hsp70h were retrieved from databanks and used for a phylogenetic analysis (Figure 2C and D). Both sequences were trimmed prior to alignment: CP from 4624 to 5064, based on accession JQ952601; Hsp70h from 904

to 1211. Both trees revealed three well supported branches. The first cluster contains all American, Asian and African isolates as well as part of the European isolates (Greece, Hungary, Lebanon and Turkey) and a Sudanese isolate; the second cluster is formed by another group of European isolates (Spain, Portugal, Italy and France); and finally the cluster with the two isolates from Mauritius. Although the

ToCV-BR isolate clustered in the first group, it is clearly separated in a well supported branch from the American isolates (including Mexico and Puerto Rico). This suggests that ToCV-BR may have been introduced from East Europe rather than the geographically closer North America or West Europe.

ToCV has become an emerging problem in agriculture worldwide, and understanding of the genome and evolution of this species is still limited to a few partial nucleotide sequences and only four complete genome sequences from the USA, Spain, Greece and the Brazilian sequences presented herein. The close relationship of the ToCV-BR isolate to the geographically distant Greek isolate suggests that the Brazilian isolate was recently introduced to the country through the exchange of plant material (or viruliferous whiteflies) from Europe. However, further studies on genetic variability of ToCV isolates from different geographical areas of the country may better clarify the origin and introduction pathway of this virus into Brazil.

ACKNOWLEDGMENTS

LCA is sponsored by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES project Capes/Wageningen (014/07) and Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq project 143257/2008-0. AKIN and ROR are CNPq fellows. This work was partially supported by grants P08-AGR-04045 from Consejería de Economía, Innovación y Ciencia, Junta de Andalucía, Spain (CEIC) and AGL2010-22287-C02-01/AGR from Ministerio de Ciencia en Innovación, Spain (MICINN), both of which are cofinanced by FEDER-FSE. JNC is a member of the Research Group AGR-214, which is partially funded by CEIC. FV was the recipient of a fellowship from CEIC.

REFERENCES

- Barbosa JC, Costa H, Gioria R, Rezende JAM (2011) Occurrence of Tomato chlorosis virus in tomato crops in five Brazilian states. *Tropical Plant Pathology* 36:256-258.
- Barbosa JC, Teixeira APM, Moreira AG, Camargo LEA, Filho AB, Kitajima EW, Rezende JAM (2008) First report of Tomato chlorosis virus infecting tomato crops in Brazil. *Plant Disease* 92:1709-1709.
- Cañizares MC, Navas-Castillo J, Moriones E (2008) Multiple suppressors of RNA silencing encoded by both genomic RNAs of the crinivirus, *Tomato chlorosis virus*. *Virology* 379:168-174.
- Fiallo-Olivé E, Hamed AA, Moriones E, Navas-Castillo J (2011) First report of Tomato chlorosis virus infecting tomato in Sudan. *Plant Disease* 95:1592-1592.
- Kataya ARA, Stavridou E, Farhan K, Livieratos IC (2008) Nucleotide sequence analysis and detection of a Greek isolate of Tomato chlorosis virus. *Plant Pathology* 57:819-824.
- Lozano G, Moriones E, Navas-Castillo J (2006) Complete nucleotide sequence of the RNA2 of the crinivirus tomato chlorosis virus. *Archives of Virology* 151:581-587.
- Lozano G, Moriones E, Navas-Castillo J (2007) Complete sequence of the RNA1 of a European isolate of tomato chlorosis virus. *Archives of Virology* 152:839-841.
- Navas-Castillo J, Fiallo-Olive E, Sanchez-Campos S (2011) Emerging virus diseases transmitted by whiteflies. *Annual Review of Phytopathology* 49:219-248.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28:2731-2739.
- Wintermantel WM, Wisler GC, Anchieta AG, Liu HY, Karasev AV, Tzanetakis IE (2005) The complete nucleotide sequence and genome organization of tomato chlorosis virus. *Archives of Virology* 150:2287-2298.
- Wisler GC, Li RH, Liu HY, Lowry DS, Duffus JE (1998) Tomato chlorosis virus: A new whitefly-transmitted, phloem-limited, bipartite closterovirus of tomato. *Phytopathology* 88:402-409.

TPP 2012-0111 - Received 16 October 2012 - Accepted 26 February 2013
Section Editor: Jorge Alberto Marques Rezende