MOLECULAR ANALYSIS OF THE DNA-A OF THE BEGOMOVIRUS TOMATO MOTTLE LEAF CURL VIRUS (ToMoLCV)

Carolina S. Rocha, César A.D. Xavier, Alison T.M. Lima, Fábio N. Silva, F. Murilo Zerbini*

Departamento de Fitopatologia/BIOAGRO (Phytopathology Department), Universidade Federal de Viçosa, Viçosa, MG, Brasil.

ABSTRACT

Begomoviruses cause economic losses in many crops, mainly in tropical and subtropical regions. Their genome is formed by one or two components and are transmitted by the whitefly Bemisia tabaci to dicotyledonous plants. In Brazil, a viral complex comprised of at least eight species is responsible for severe losses in tomato crops. Tomato and weed samples were collected in tomato-growing regions of the state of Minas Gerais, in southeastern Brazil in 2008 and 2010. Previously described viruses were prevalent in the samples. Two isolates of the partially sequenced Tomato mottle leaf curl virus (ToMoLCV) were associated with tomato plants collected in the city of Jaíba in northern Minas Gerais. Here, we describe the virus's complete DNA-A sequence and its molecular characterization. Genome analysis indicates that the ToMoLCV is a typical New World bipartite begomovirus with greater sequence identity with begomoviruses from Brazil and Central America. Phylogenetic analysis confirms that ToMoLCV clusters with New World begomoviruses from Brazil and Central America. Like all Brazilian begomoviruses, ToMoLCV has a recombinant nature. Together, these results support the classification of ToMoLCV as a distinct species in the genus Begomovirus.

Keywords: diversity; geminivirus; recombination; tomato.

INTRODUCTION

Begomovirus diseases are a major limiting factor to crop yields in tropical and subtropical regions (Moriones et al. 2011). The genus Begomovirus belongs to the Geminiviridae family, and includes viruses with one or two genomic components which infect dicotyledonous plants and are transmitted by the whitefly Bemisia tabaci (Homoptera: Aleyrodidae) (Brown et al. 2011). The B biotype of B. tabaci was first reported in Brazil in the early 1990s, and due to its characteristics of greater adaptability it has spread rapidly throughout the hot and dry regions of the country (Lourenção & Nagai 1994). Following the dissemination of the B biotype of B. tabaci, begomovirus epidemics have greatly increased in Brazil. It is believed that the insect vector transferred indigenous viruses infecting wild and weed hosts to tomato.

The initial characterization of begomoviruses associated with epidemics in tomato crops in Brazil indicated a high genetic diversity, with the description of several new species (Ribeiro et al. 2003). Surveys conducted over the past five years (Castillo-Urquiza et al. 2007; Cotrim et al. 2007; Fernandes et al. 2008) indicate that certain species have become prevalent in different regions of the country. However, Castillo-Urquiza et al. (2008) recently reported six new begomoviruses species in tomato and weeds, indicating that new species continue to emerge.

The emergence of new viruses is dependent on mutation, recombination and, in viruses with divided genomes such as most begomoviruses, pseudorecombination events. Mixed infections are common in the field, favoring recombination and pseudorecombination and facilitating the emergence of new strains or species better adapted to new hosts (Chakraborty et al. 2008; Davino et al. 2009; Patil & Fauquet 2009; Pita et al. 2001).

MATERIAL AND METHODS

As part of an ongoing study of the genetic diversity of tomato-infecting begomoviruses in Brazil, 117 tomato and 23 weed samples were collected in July 2008 in tomato fields located around Jaíba, (15o11’01”S; 43o49’07”W). Total DNA was extracted as described by Doyle & Doyle (1987) and full-length viral genomes were amplified by rolling-circle amplification (Inoue-Nagata
et al. 2004). After monomerization with the restriction enzymes Apa I, BamH I, Cla I, Hind III, Kpn I, Pst I, Ssp I or Sac I, samples that displayed a restriction pattern including a 2,600 bp band were selected for cloning of full-length viral genome components in pBLUESCRIPT KS+ (Stratagene). Recombinant plasmids with inserts corresponding to full-length begomovirus components were identified by restriction analysis, and the viral inserts were completely sequenced at Macrogen, Inc. (Seoul, South Korea). Pairwise p-distance comparisons of the nucleotides sequences of the complete DNA-A and of the five genes in the DNA-A (Replication-associated protein, Rep; Transactivating protein, Trap; Replication enhancer protein, Ren; AC4; and Coat protein, CP) were performed using MEGA 5 with exclusion of alignment gaps (Tamura et al. 2011). The deduced amino acid sequences of the five proteins were compared using EMBOSS (http://www.ebi.ac.uk/Tools/psa/) with default settings. A maximum likelihood phylogenetic tree was inferred using PAUP 4.0 (Wilgenbusch & Swofford 2003), using a full-length genome dataset. The program ModelTest 3.7 (Posada & Crandall 1998) was used to predict the best-fit model GTR+I+G. The tree was viewed using FigTree version 1.3.1 and edited using Corel Draw X3.

**RESULTS AND DISCUSSION**

Out of 140 samples, only six tomato samples and one sample of the common weed *Sida* sp. tested positive for the presence of a begomovirus (data not shown). This was an unexpected result, since all sampled fields had high levels of whitefly infestation and only symptomatic samples were collected. It is possible that the begomovirus-negative samples were infected with other whitefly-transmitted viruses, such as criniviruses (Navas-Castillo et al. 2011).

Out of the seven begomovirus-positive samples, a total of seven DNA-A and three DNA-B clones were obtained. Sequence comparisons indicated that the DNA-A components in four tomato samples corresponded to the previously described begomoviruses *Sida micrantha mosaic virus* (SiMMV; one sample), *Tomato common mosaic virus* (ToCMoV; one sample) and *Tomato severe rugose virus* (ToSRV; two samples) (data not shown). The DNA-A component in the *Sida* sp. sample corresponded to SiMMV (data not shown). These sequences will be described in detail elsewhere, as part of a country-wide population study of begomoviruses.

Two tomato samples had a mixed infection with *Tomato yellow spot virus* (ToYSV) and the previously described *tomato mottle leaf curl virus* (ToMoLCV). The ToMoLCV sequences shared 97% identity with the partial sequence (GenBank access number AY049226) reported by Ribeiro et al. (2003), and 97-99% with six complete sequences which have recently become available (JF803246-JF803251). The two isolates (BR:Jai13:08 and BR:Jai56:08) share a 99% nucleotide sequence identity. The highest nucleotide sequence identity with other begomoviruses is 80% with *Passionfruit severe leaf distortion virus* (PSLDV, FJ972767) and *Tomato chlorotic mottle virus* (ToCMoV, AF490004) (Figure 1A). Although the DNA-B has not been cloned, the two ToMoLCV isolates have a DNA-A organization typical of New Worlds bipartite begomoviruses, with five ORFs: coat protein (CP), replication-associated protein (Rep), trans-activating protein (Trap), replication-enhancer protein (Ren) and AC4. The CP gene is the most conserved in relation to other begomoviruses, with 92% amino acid sequence identity with *Potato yellow mosaic Panama virus* (PYMPV, Y15034) (Figure 1B). The Rep gene is the least conserved in terms of nucleotide sequence, showing the greatest identity with PSLDV (Figure 1C).

Phylogenetic analysis based on the DNA-A component placed ToMoLCV (the two isolates described here plus the six sequences from GenBank) in a monophyletic cluster, supported by a 93% bootstrap value, which includes the passionfruit-infecting PSLDV and the tomato-infecting ToCMoV (Figure 2).

Recombination analysis using RDP3 (Martin & Rybicki 2000) indicated that isolates BR:Jai13:08 and BR:Jai56:08, are recombinants. The single recombination event was detected by all methods of the RDP3 package using a data set including all American begomoviruses. PSLDV and ToYSV (DQ336350) were identified as the recombinant parents (Table 1). The same event was also detected for the six newly available ToMoLCV sequences (Table 1). RDP3 identifies putative parents based on the detected recombination signal, and both the strength and the source of the signal may change depending on the sequences which are included in the analysis. Larger data sets are preferable (as long as pairwise identities among the sequences are >70%, otherwise the analysis will be unreliable as <70% identity generates a poor alignment), because exploratory searches for recombination require repetitive statistical testing, with the number of tests performed increasing exponentially with the number of sequences. Therefore, we decided to use a data set including all begomovirus sequences from the Americas. Even so, it must be pointed out that parent identification is not completely reliable, as the program looks for the sequences which are most similar to the input sequence. The actual parents might be unknown (undetected) viruses, and therefore absent from the data set.

The tomato samples from which BR:Jai13:08 and BR:Jai56:08 were cloned were also infected with ToYSV (data not shown). Detailed analysis of the common regions (CR) of isolates BR:Jai13:08 and BR:Jai56:08 indicated that they share similar iterons with PSLDV, ToCMoV and ToYSV. Interestingly, one of the direct repeats (GGGG) is identical to the one from PSLDV and ToCMoV, whereas the other (GGTG) is identical to the one from ToYSV (Figure 3). The inverted repeat (CCAC) is the same for all four viruses (Figure 3). These features
indicate that the formation of viable pseudorecombinants among these viruses may occur. In fact, PSLDV (DNA-A) and ToCMoV (DNA-B), which have identical iterons but share only 70% nucleotide sequence identity in their CRs, form viable pseudorecombinants in *Nicotiana benthamiana* (Ferreira et al. 2010). Also, a PSLDV-A and ToYSV-B pseudorecombinant was viable in 20% of the inoculated *N. benthamiana* plants (Ferreira et al. 2010). These results confirm the close relationship among these viruses, and are an additional line of evidence pointing at a common origin. Evidently, in the case of ToMoLCV this must be confirmed with the production of infectious clones and plant inoculations with the mixtures of its genomic components with those from PSLDV, ToCMoV and ToYSV.

We carried out the molecular characterization of the begomovirus Tomato mottle leaf curl virus (ToMoLCV) detected in two tomato samples collected in northern Minas Gerais state, Brazil. This virus had not yet been recognized as a distinct species, because its DNA-A had not been completely sequenced (Ribeiro et al. 2003). Our results support the classification of ToMoLCV as a species in the genus *Begomovirus*.

![Figure 1](image-url). Percent nucleotide sequence identities between the full-length DNA-A (A), and percent nucleotide (below the diagonal) and deduced amino acid (above the diagonal) sequence identities of the (B) CP, (C) Rep, (D) Trap, (E) Ren and (F) AC4 genes of the BR:Jai13:08 and BR:Jai56:08 isolates and the most closely related begomoviruses. For simplicity, the six additional ToMoLCV sequences are not shown, as they have 97-99% nucleotide sequence identity with the two isolates described in this study.
Table 1. Recombination events detected between ToMoLCV isolates (including isolates BR:Jai13:08 and BR:Jai56:08 described in this study) and begomoviruses infecting tomato and weeds in Brazil and in the Americas. Results based on data set comprising all begomoviruses from the Americas.

<table>
<thead>
<tr>
<th>Parents</th>
<th>Breakpoints</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parents</td>
<td>Major</td>
</tr>
<tr>
<td>BR:Jai13:08</td>
<td>ToYSV</td>
<td>1964</td>
</tr>
<tr>
<td>BR:Jai56:08</td>
<td>ToYSV</td>
<td>1964</td>
</tr>
<tr>
<td>BR:PADFM:04</td>
<td>ToYSV</td>
<td>2028</td>
</tr>
<tr>
<td>BR:PA2143:04</td>
<td>ToYSV</td>
<td>2045</td>
</tr>
<tr>
<td>BR:Turv2911:04</td>
<td>ToYSV</td>
<td>2028</td>
</tr>
<tr>
<td>BR:Turv2904:04</td>
<td>ToYSV</td>
<td>2028</td>
</tr>
<tr>
<td>BR:Jua2586:04</td>
<td>ToYSV</td>
<td>2002</td>
</tr>
<tr>
<td>BR:Bez2665:04</td>
<td>ToYSV</td>
<td>2027</td>
</tr>
</tbody>
</table>

1Numbering starts at the first nucleotide after the cleavage site at the origin of replication and increases clockwise.

2R, RDP; G, GeneConv; B, Bootscan; M, MaxChi; C, CHIMAERA; S, SisScan; 3, 3SEQ.

3ToYSV, DQ336350; PSLDV, FJ972767; BR:PADFM:04, JF803246; BR:PA2143:04, JF803247; BR:Turv2911:04, JF803248; BR:Turv2904:04, JF803249; BR:Jua2586:04, JF803250; BR:Bez2665:04, JF803251.

Figure 2. Maximum likelihood tree obtained from the alignment of the full-length DNA-A sequences of begomoviruses from the Americas, including the ToMoLCV isolates BR:Jai13:08 and BR:Jai56:08 (highlighted in red, bold font) as well as additional ToMoLCV sequences from GenBank (highlighted in red, normal font). Numbers on branches indicate bootstrap values (2,000 replications).
Molecular Analysis of the DNA-A of the Begomovirus Tomato Mottle Leaf Curl Virus (ToMoLCV)

Figure 3. (A) Alignment of the common regions of the ToMoLCV isolates BR:Jai13:08, BR:Jai56:08, BR:PADFM:04 (JF803246) and BR:Juaz2586:04 (JF803250) with those from Tomato yellow spot virus (ToYSV, DQ336350), Tomato chlorotic mottle virus (ToCMoV, AF490004) and Passionfruit severe leaf distortion virus (PSLDV, FJ972767). The TATA box and the conserved nonanucleotide are highlighted in grey. Iterated direct and inverted repeats (iterons) are boxed. The arrows indicate the direction of the repeats. Asterisks indicate nucleotide positions which are conserved among all four aligned sequences. Nucleotide differences in the ToMoLCV iterons are highlighted in yellow. The sequences of ToMoLCV isolates BR:PA2143:04 (JF803247), BR:Turv2911:04 (JF803248) and BR:Turv2904:04 (JF803249) are identical to BR:PADFM:04, and the sequence of isolate BR:Bez2665:04 (JF803251) is identical to BR:Juaz2586:04. (B) Partial alignment of the amino acid sequence of the Rep proteins of ToMoLCV, ToYSV, ToCMoV and PSLDV. The domain associated with sequence-specific recognition of iterons (iteron-related domain, IRD) is boxed. Red asterisks indicate nucleotide positions which are conserved among all six aligned sequences. Motif 1, motif 2 and the specificity determinants (SPDs), which according to Arguello-Astorga & Ruiz-Medrano (2001) and Londono et al. (2010) are conserved in rolling-circle replication initiator proteins, including geminivirus Rep proteins, are highlighted in grey.

AKNOWLEDGMENTS

The authors wish to thank Drs. Eduardo S.G. Mizubuti, Sérgio Oliveira de Paula and Simone G. Ribeiro for critical review of the manuscript. This work was funded by Fapemig grants CAG-666-08 and CAG-949-09 to FMZ.

REFERENCES


