Nucleotide sequence and genome organization of tomato leaf curl geminivirus

Ian B. Dry, Justin E. Rigden, Leslie R. Krake, Philip M. Mullineaux† and M. Ali Rezaian*

CSIRO Division of Horticulture, GPO Box 350, Adelaide, SA 5001, Australia

The genome of tomato leaf curl virus (TLCV) from Australia was cloned and its complete nucleotide sequence determined. It is a single circular ssDNA of 2766 nucleotides containing the consensus nona-nucleotide sequence present in all geminiviruses. It has six open reading frames with an organization resembling that of certain other dicotyledonous plant-infecting monopartite geminiviruses, i.e. tomato yellow leaf curl and beet curly top viruses. The regulatory sequences present indicate a bidirectional mode of transcription. A dimeric TLCV DNA clone was constructed in a binary vector and used to agroinoculate three different host species. Typical virus infections were produced, confirming that the single DNA component is sufficient for infectivity.

Geminiviruses are characterized by their small (20 × 30 nm) twinned isometric particles and circular ssDNA genomes of approximately 2.7 kb, and have been classified into two major subgroups based on host range, insect vector and number of DNA components (Davies & Stanley, 1989). One subgroup contains viruses possessing a single DNA component which infect monocotyledonous plants and are leafhopper-transmitted. The second subgroup consists of viruses with a bipartite genome (designated DNA A and B) which infect dicotyledonous plants and are transmitted by the whitefly Bemisia tabaci. Beet curly top virus (BCTV; Stanley et al., 1986) and tobacco yellow dwarf virus (TobYDV; Thomas & Bowyer, 1980) demonstrate characteristics of both groups (i.e. monopartite, leafhopper-transmitted and dicotyledonous plant-infecting), and may be potential members of a third subgroup of geminiviruses. However, recent evidence indicates that the genome organization of TobYDV is unlike that of BCTV and more similar to that of the monocotyledonous plant-infecting geminiviruses (Morris et al., 1992).

Tomato leaf curl virus (TLCV) causes severe crop damage in the northern parts of Australia (B. Conde, personal communication). Plants infected with TLCV contain the characteristic geminate virus particles (Thomas et al., 1986) and show disease symptoms similar to those caused by tomato yellow leaf curl virus (TYLCV) (Makkouk & Laterrot, 1983; Czosnek et al., 1988). An isolate of whitefly-transmitted TYLCV from Thailand has been reported to contain a bipartite genome (Rochester et al., 1990), but two other isolates from Israel (TYLCV-I; Navot et al., 1991) and Sardinia, Italy (TYLCV-S; Khyer-Pour et al., 1991) have been described which contain only a single DNA component. This DNA structurally resembles DNA A of the bipartite geminiviruses, but is sufficient to initiate a complete cycle of infection, including systemic spread and whitefly transmission.

Here we describe the genome structure of TLCV. It consists of a single DNA species which is infectious as a dimeric construct upon agroinoculation, giving rise to disease symptoms indistinguishable from those of the field isolate.

Infected tomato stems from the Northern Territory, Australia were used to inoculate tomato plants (Lycopersicon esculentum cv. Grosse Lisse) by side-grafting, and plants were maintained in an insect-proof glasshouse under quarantine conditions. Total nucleic acids were prepared from tomato leaves. The tissues were pulverized in liquid nitrogen, mixed with two volumes of 50 mM-Tris-HCl pH 8, 100 mM-NaCl, 10 mM-EDTA, 1% (v/v) SDS and 1% 2-mercaptoethanol and extracted four or five times with phenol:chloroform (4:1), and the nucleic acids were precipitated with ethanol.

Analysis of DNA extracts on 1.2% agarose gels revealed the presence of a DNA species in TLCV-infected tissue which was absent from healthy tissue. This DNA species was isolated by two rounds of agarose gel electrophoresis/electroelution, digested with EcoRI or XbaI, and cloned into an appropriately cut pBluescript-SK+ vector (Stratagene). The complete sequence of both strands of a full-length XbaI clone, pTLC4, was obtained.
Fig. 1. Nucleotide sequence of the virion-sense strand of an infectious clone of TLCV. Numbering begins with the first base of the conserved nonanucleotide of the stem–loop sequence. The virion-sense strand was identified by comparison with published geminivirus sequences.

using the dideoxynucleotide chain termination method (Sanger et al., 1977) by subcloning and the use of TLCV-specific oligonucleotide primers.

The sequence of pTLC4 (an infectious clone; see below) is shown in Fig. 1. The first nine nucleotides of the 2766 base sequence (TAATATTAC), which form the
**Table 1. TCLV ORFs**

<table>
<thead>
<tr>
<th>ORF*</th>
<th>Nucleotide</th>
<th>Protein Mr,†</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>148-492</td>
<td>13 250</td>
</tr>
<tr>
<td>V2</td>
<td>308-1075</td>
<td>29 713</td>
</tr>
<tr>
<td>C1</td>
<td>2615-1530</td>
<td>41 197</td>
</tr>
<tr>
<td>C2</td>
<td>1627-1223</td>
<td>15 304</td>
</tr>
<tr>
<td>C3</td>
<td>1479-1078</td>
<td>16 114</td>
</tr>
<tr>
<td>C4</td>
<td>2464-2159</td>
<td>11 410</td>
</tr>
</tbody>
</table>

* V and C denote virion-sense and complementary-sense ORFs, respectively, as shown in Fig. 2.
† Size of putative translation products.

**Table 2. Amino acid sequence similarity between putative translation products of TCLV and selected geminiviruses**

<table>
<thead>
<tr>
<th>Virus*</th>
<th>V1</th>
<th>V2</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGMVA</td>
<td>–</td>
<td>71</td>
<td>64</td>
<td>45</td>
<td>51</td>
<td>52</td>
</tr>
<tr>
<td>PYMVA</td>
<td>–</td>
<td>71</td>
<td>68</td>
<td>52</td>
<td>52</td>
<td>NA†</td>
</tr>
<tr>
<td>ABMVA</td>
<td>–</td>
<td>73</td>
<td>68</td>
<td>47</td>
<td>52</td>
<td>60</td>
</tr>
<tr>
<td>ACMVA</td>
<td>–</td>
<td>74</td>
<td>69</td>
<td>58</td>
<td>55</td>
<td>51</td>
</tr>
<tr>
<td>ACMV</td>
<td>63</td>
<td>75</td>
<td>78</td>
<td>62</td>
<td>63</td>
<td>48</td>
</tr>
<tr>
<td>TYLCV-I</td>
<td>62</td>
<td>75</td>
<td>83</td>
<td>61</td>
<td>68</td>
<td>50</td>
</tr>
<tr>
<td>TYLCV-S</td>
<td>58</td>
<td>73</td>
<td>79</td>
<td>56</td>
<td>59</td>
<td>63</td>
</tr>
<tr>
<td>BCTV</td>
<td>15</td>
<td>17</td>
<td>61</td>
<td>20</td>
<td>59</td>
<td>55</td>
</tr>
<tr>
<td>TobYDV</td>
<td>10</td>
<td>11</td>
<td>25</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>MSV</td>
<td>14</td>
<td>10</td>
<td>25</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>WDV</td>
<td>17</td>
<td>6</td>
<td>24</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

* ABMVA, Abutilon mosaic virus DNA A (Frischmuth et al., 1990); BGMVA, bean golden mosaic virus DNA A (Howarth et al., 1985); TGMVA, TGMV DNA A (Hamilton et al., 1984); PYMVA, potato yellow mosaic virus DNA A (Coutts et al., 1991); ACMVA, ACMV DNA A (Stanley & Gay, 1983) – TYLCV-I (Navot et al., 1991); TYLCV-S (Khyer-Pour et al., 1991); BCTV (Stanley et al., 1986); TobYDV (Morris et al., 1992); MSV, maize streak virus (Mullineaux et al., 1984); WDV, wheat dwarf virus (MacDowell et al., 1985).
† Amino acid identity was calculated using the PALIGN program of the PC/GENE sequence analysis package using the structure–genetic comparison matrix (open gap cost value, 10; unit gap cost value, 10).

The coat protein ORFs of AMBVA, BGMVA, TGMV, PYMVA, ACMV and TYLCV-I have been defined as V2 for this comparison. No values are given where comparable ORFs do not exist.

The similarity between the sequences of TCLV ORFs and the equivalent ORFs from a range of other geminiviruses is shown in Table 2. The high degree of similarity between ORF V2 and the viral coat proteins of the other whitefly-transmitted geminiviruses [tomato golden mosaic virus (TGMV), ACMV, TYLCV] strongly implies a similar mode of insect transmission for

Fig. 2. Proposed genome organization of TCLV. ORFs on the virion-sense (clockwise) strand and the complementary-sense (anticlockwise) strand (as described in Table 1) are displayed by arrows. The positions of the conserved stem–loop structure (→), intergenic region (IR), potential TATA boxes (▲) and polyadenylation signals (▶) are also marked.

loop of a stem–loop structure, are strictly conserved in all geminiviruses examined. Alignment of this sequence with published geminivirus DNA sequences (using NALIGN from the PC/GENE sequence analysis package, Intelligenetics Inc.) shows it to be most similar to the two isolates of TYLCV, i.e. TYLCV-I (75.9% similarity) and TYLCV-S (73.2%) and the DNA A component of African cassava mosaic virus (ACMV; Stanley & Gay, 1983) (71.9%).

Organization of open reading frames (ORFs; Table 1) and potential regulatory sequences on the circular TCLV genome are shown in Fig. 2. The location of two ORFs in the virion sense and four in the complementary sense is very similar to that observed for certain other dicotyledonous plant-infecting monopartite geminiviruses, i.e. TYLCV and BCTV. Of the sequences conforming to the consensus for TATA box promoter elements [TATA(A/T)A(A/T); Breathnach & Chambon, 1981] shown in Fig. 2, those located within the intergenic region appear most likely to account for the production of both virion- and complementary-sense transcripts. It should also be noted that a closely related sequence element, TATAAAG, is located 52 nucleotides upstream of the initial AUG of the C2 ORF, and that this may potentially be involved in the production of a complementary-sense transcript covering the C2/C3 region. Polyadenylation signals (AATAAA) map to a region at the convergence of the V2 and C3 ORFs (Fig. 2) suggesting that both virion- and complementary-sense transcripts may terminate at this point.

The similarity between the sequences of TCLV ORFs and the equivalent ORFs from a range of other geminiviruses is shown in Table 2. The high degree of similarity between ORF V2 and the viral coat proteins of the other whitefly-transmitted geminiviruses [tomato golden mosaic virus (TGMV), ACMV, TYLCV] strongly implies a similar mode of insect transmission for
TLCV. This is also supported by serological studies showing good cross-reactivity between TLCV and a monoclonal antibody which specifically detects whitefly-transmitted geminiviruses, but which is unreactive with geminiviruses transmitted by leafhoppers (Thomas et al., 1986). However, direct experimental evidence of whitefly transmission of TLCV is still lacking.

The comparison of various potential translation products of the geminiviruses listed in Table 2 and the similar degree of overall nucleotide sequence identity between TLCV, TYLCV and the DNA A component of ACMV indicates that sequence comparison alone may not be a sufficient criterion for establishing the relationship between these viruses. For example, based on the amino acid identity between a number of selected putative ORF products (Table 2), TLCV DNA appears to be more closely related to the DNA A component of the bipartite geminivirus ACMV than to TYLCV-S DNA. Thus, the nature of the relationship between TLCV and TYLCV will need to be further established by comparative biological characterization such as host range studies (Makkouk & Laterrot, 1983).

To test the infectivity of the full-length TLCV clone pTLC4, a dimeric (head-to-tail) clone was constructed in pBin19 to form pTLCBIN1. This construct was transferred into Agrobacterium tumefaciens strain C58 by triparental mating (Ditta et al., 1980), and its correct orientation verified by restriction analysis of binary plasmid DNA (Holmes & Quigley, 1981). Sixteen tomato (L. esculentum cv. Grosse Lisse), 20 tobacco (Nicotiana tabacum cv. Samsun) and 10 jimsonweed (Datura stramonium L.) plants were each agroinoculated by injection with 10 μl of a bacterial suspension (10⁸ to 10⁹ cells/ml) into the apical meristematic regions of tobacco and into the axillary buds of the decapitated main stem of tomato and jimsonweed plants. Severe symptoms (i.e. curling and yellowing on tomato and jimsonweed plants; wrinkling on tobacco plants) were observed on newly developed leaves of all plants inoculated with the dimer within 10 to 14 days. No leaf curl symptoms were observed on plants agroinoculated with pBin19 alone. The presence of viral DNA in plants showing leaf curl symptoms was confirmed by dot blot analysis (data not shown). Attempts to infect nine tomato plants by direct injection with TLCV DNA either as pTLCBIN1 or total nucleic acid extracts from field-infected plants were unsuccessful.

To compare the viral DNA species in plants agroinoculated with pTLCBIN1 to those in field-infected tomato plants, material from each source was grafted onto healthy tomato plants. The rate and severity of symptom development were found to be similar in both groups of grafted plants. Total nucleic acids (derived from 8 mg of infected tissue) were separated on a 1.2% agarose gel for 16 h at 30 V in TBE buffer containing 0.5 μg/ml ethidium bromide. DNA was transferred to a Zetaprobe nylon membrane (Bio-Rad), hybridized overnight with a TLCV probe labelled with ³²P by random hexamer priming, and washed with 0.1 x SSC, 0.1% SDS at 65 °C. Fig. 3 shows that the same viral ds- and ssDNA forms are present in plants infected by graft-transmitted TLCV from both sources. The identity of the virion-sense ssDNA was confirmed by strand-specific probing using labelled RNA transcripts produced by T7 and T3 RNA polymerases from a 754 bp SacI–EcoRV fragment of pTLC4 subcloned into Bluescript SK (P. M. Mullineaux et al., unpublished results).

The above results indicate that TLCV isolated from northern Australia, together with the recently described isolates of TYLCV (Navot et al., 1991; Khyer-Pour et al., 1991) may constitute a new subgroup of geminiviruses that infect dicotyledonous plants and are whitefly-transmitted, but which have only a single genomic component. Clearly transcription of this single component (as must also be the case for BCTV and TobYDV) produces all of the gene products required to
support a complete cycle of infection and spread within dicotyledonous plants. This raises the question as to which TLCV gene product(s) has an equivalent functional role in systemic spread to those produced by the DNA B component of the bipartite geminiviruses (Davies & Stanley, 1989). Of particular interest is the role of the V1 ORF which is present in all monopartite geminiviruses, and which appears to be involved in virus spread in both dicotyledonous (Kyhe-Pour et al., 1991) and monocotyledonous plants (Mullineaux et al., 1988; Boulton et al., 1989).

We thank Mr Barry Conde for a culture of TLCV and for the supply of infected tissues, and Drs Nigel Scott and John Thomas for discussions. This project was funded in part by Gene Shears Pty. Ltd., and the visit by Phil Mullineaux was supported by the Department of Industry, Technology and Commerce.

References


(Received 28 July 1992; Accepted 18 September 1992)