Journal of General Virology (1994), 75, 477–485. *Printed in Great Britain*

Complete nucleotide sequence of the geminivirus tomato yellow leaf curl virus, Thailand isolate

Dean E. Rochester,[1] Jacqueline J. DePaulo,[1] Claude M. Fauquet[2] and Roger N. Beachy[3]*

[1] *Washington University, Department of Biology, St Louis, Missouri 63130,*
[2] *ILTAB/ORSTOM-TSRI, Department of Cell Biology and* [3] *Department of Cell Biology, The Scripps Research Institute MRC-7, La Jolla, California 92037, U.S.A.*

The complete genome of a Thailand isolate of the geminivirus tomato yellow leaf curl virus (TYLCV-Th) has been cloned and the nucleotide sequence determined. The genome consists of two DNAs each slightly greater than 2700 nucleotides in length and designated A-DNA and B-DNA. The A-DNA contains six open reading frames (ORFs) capable of encoding proteins with M_rs greater than 10K; two ORFs were located on the B-DNA. The predicted ORFs were found at positions on the genome similar to those identified in other geminiviruses. Based upon computer-assisted sequence comparisons with other geminiviruses, TYLCV-Th maintained the greatest degree of similarity with a tomato pathogen isolated in Australia, tomato leaf curl virus. Furthermore, sequence comparisons with a large number of geminiviruses confirm further divisions of the geminiviruses based upon geographical proximity of the viruses. Agroinoculation of *Nicotiana benthamiana* with a TYLCV-Th A-DNA clone containing a mutation in the capsid protein ORF demonstrated that the coat protein is not essential for TYLCV-Th infection but enhances the progression of the disease.

Introduction

Geminiviruses have circular ssDNA genomes encapsidated in geminate particles. The International Committee on Taxonomy of Viruses (ICTV) divided the geminiviruses into three subgroups based upon genome structure, the type of insect vector and host range (Hull *et al.*, 1991). These subgroups will become three genera in the family *Geminiviridae* in the next report of the ICTV (M. Fauquet, personal communication). Subgroup one contains the geminiviruses that infect monocotyledonous plants with maize streak virus (MSV) as the type member. These viruses are transmitted by leafhoppers and contain a single DNA molecule. Beet curly top virus (BCTV) and tobacco yellow dwarf virus (ToYDV) are the only representatives of the second subgroup and BCTV is the type member. These leafhopper-transmitted viruses infect dicotyledonous hosts and also contain a single molecule DNA genome. Bean golden mosaic virus (BGMV) is the type member of the third subgroup comprising the whitefly-transmitted geminiviruses with bipartite genomes. To date, these geminiviruses have been shown to infect only dicotyledonous plants. Another study identified subsets of viruses within this subgroup according to geographical distribution with BGMV and tomato golden mosaic virus (TGMV) as examples of 'New World' geminiviruses and African cassava mosaic virus (ACMV) and tomato yellow leaf curl virus (TYLCV) as 'Old World' geminiviruses (Howarth & Vandemark, 1989). More recently, several isolates of whitefly-transmitted geminiviruses have been described which infect tomato plants and contain only a single DNA component (Navot *et al.*, 1991; Kheyr-Pour *et al.*, 1991; Dry *et al.*, 1993). These geminiviruses include two isolates of TYLCV, isolated in Israel (TYLCV-I) and Sardinia (TYLCV-S), as well as tomato leaf curl virus (TLCV) from Australia. Geographically, these geminiviruses would be placed in the 'Old World' subset. Here we present the complete genomic sequence of the A- and B-DNAs of TYLCV-Th as well as comparisons between its putative protein products and nucleotide regulatory regions with those of other geminiviruses.

The role of the coat protein (CP) in the infection cycle further differentiates the subgroups of the geminiviruses. It was previously demonstrated that mutations disrupting the CP open reading frame (ORF) of bipartite geminiviruses such as ACMV and TGMV delay the

The nucleotide sequence data reported in this paper will appear in the GenBank Sequence Database under accession numbers M59838 (TYLCV-Th A-DNA) and M59839 (TYLCV-Th B-DNA).

appearance of symptoms and reduce the virus titre in systemic infection but do not prevent infection or spread of the pathogen (Stanley & Townsend, 1986; Gardiner *et al.*, 1988). In contrast, several studies of monopartite geminiviruses with dicotyledonous hosts indicated that the CP is necessary for systemic infection and cell to cell movement (Briddon *et al.*, 1989; Rigden *et al.*, 1993). In the present study we demonstrate that disruption of the CP ORF of TYLCV-Th reduces systemic infection following agroinoculation of *Nicotiana benthamiana* plants. This mutation also reduces the accumulation of the viral DNA in excised leaf tissue following agroinoculation. Similarities and differences between TYLCV-Th and the other geminiviruses are discussed.

Methods

DNA cloning. Cloning of the TYLCV-Th genome has been previously described (Rochester *et al.*, 1990). In brief, a synthetic oligonucleotide was used to prime the synthesis of a second strand on purified single-stranded viral DNA. This double-stranded DNA was subsequently digested with *Eco*RI and ligated into the plasmid pBluescript KS+ previously linearized with *Eco*RI. Two resulting plasmids (pDR9 and pDR7) represent the A- and B-DNAs of TYLCV-Th, respectively. Sequencing was carried out on subclones of pDR7 and pDR9 using the dideoxynucleotide chain-termination method (Sanger *et al.*, 1977). Synthetic oligonucleotides were used when restriction sites were not available for subcloning.

Computer analysis of the TYLCV-Th genome. After the complete sequence of the TYLCV-Th genome had been derived, further analysis was carried out using sequence analysis software from DNASTAR Inc. Optimal alignments (pairwise and multiple) and estimating degrees of sequence identity between analogous DNA and predicted protein sequences of specific geminiviruses were made using the Megalign program. Geminiviruses used for the comparative analyses included: TYLCV-Th, TYLCV-I (Navot *et al.*, 1991), TYLCV-S (Kheyr-Pour *et al.*, 1991), TLCV (Dry *et al.*, 1993), ACMV (Stanley & Gay, 1983), TGMV (Hamilton *et al.*, 1984), BGMV (Howarth *et al.*, 1985), squash leaf curl virus (SqLCV) (Lazarowitz & Lazdins, 1991), Chloris striate mosaic virus (CSMV; Andersen *et al.*, 1988), MSV (Mullineaux *et al.*, 1984), wheat dwarf virus (WDV) (MacDowell *et al.*, 1985), Digitaria streak virus (DSV) (Donson *et al.*, 1987), abutilon mosaic virus (AbMV) (Frischmuth *et al.*, 1990) and BCTV (Stanley *et al.*, 1986). The sequences of the intergenic regions of bean dwarf mosaic virus (BDMV; Gilbertson *et al.*, 1991*a*) and tomato mottle virus (TMoV; Gilbertson *et al.*, 1991*b*) were retrieved from the GenBank Sequence Database (accession numbers M88179 and M90494, respectively). The phylogenetic tree shown in Fig. 3 is derived from using an algorithm (developed by Jotun Hein) from DNASTAR. 'Percentage similarity' used throughout the text is defined as:

$$\text{percentage similarity (i,j)} = (\text{sum of identical matches} \times 100 / [\text{length} - \text{gap residues(i)} - \text{gap residues(j)}]$$

The scale of the dendrogram is expressed as 'distance' calculated as distance (i,j) = sum (residue distances) + (gaps × gaps penalty) + (gap residues × gap length penalty) with a gap penalty of 11 and a gap length penalty of 4. The 'total distance' is the sum of all branch lengths in the dendrogram.

Coat protein mutagenesis. Mutagenesis of the TYLCV-Th CP ORF was accomplished by digesting the unique *Bgl*II restriction site in the CP gene (nucleotide number 798, Fig. 1*a*) followed by end filling with Klenow polymerase and dNTPs and subsequent religation. The mutation was present in both copies of the CP ORF in the final agroinoculation clone, pDR56. The plasmid pDR56 contains 1·6 copies of A-DNA and was constructed in a similar manner to pDR21 (1·6 copies TYLCV-Th A-DNA) as described previously (Rochester *et al.*, 1990). Mutagenesis of the TYLCV-Th CP ORF resulted in an insertion at amino acid 120 of four amino acids followed by a stop codon (TGA). The plasmid pDR37 contains 1·6 copies of the TYLCV-Th B-DNA and has been described previously (Rochester *et al.*, 1990). The plasmids used for agroinoculation were shuttled into *Agrobacterium tumefaciens* GV3111SE (Rogers *et al.*, 1986). The modified strains of *A. tumefaciens* harbouring pDR21, pDR56 and pDR37 are referred to as TYLCV-Th:A, TYLCV-Th:Acp and TYLCV-Th:B, respectively.

Agroinoculation of plant tissues. The agroinoculation protocol on whole plants has been described before (Rochester *et al.*, 1990). In brief, the *A. tumefaciens* cultures were grown overnight at 28 °C. A 21-gauge needle was used to transfer bacteria from liquid cultures to plants by pricking the plants. Plants were transferred to growth chambers (16 h day, 21 °C, 65% humidity) and observed daily for the presence of symptoms. Punches of leaf tissue were taken and processed for analysis of viral DNA as described earlier (Rochester *et al.*, 1990).

Excised leaf tissues were also used in agroinoculation experiments. Surface-sterilized whole leaves were agroinoculated by briefly submersing petioles or cut sites on the leaves in overnight cultures of *A. tumefaciens* for 2 to 3 min. Tissue was then transferred to plant tissue culture media and kept in growth chambers until samples were collected. The antibiotic carbenicillin was used in tissue culture media to restrict the growth of the *A. tumefaciens*.

Results

DNA sequence of TYLCV-Th

Purified TYLCV-Th was obtained (Rochester *et al.*, 1990) and the viral DNA was cloned into pBluescript KS+ as described in Methods. We previously demonstrated that the two virus genome components TYLCV-Th A-DNA and TYLCV-Th B-DNA, were infectious when introduced as 1·6 × multimers into tomato and *N. benthamiana* by agroinoculation (Rochester *et al.*, 1990). The entire sequence of each of the two DNAs of TYLCV-Th was determined and is shown in Fig. 1. More than 90% of the genome was sequenced on both strands. The sense strand of each genome was identified on the basis of sequences found within the putative common regions (Rochester *et al.*, 1990), a region of approximately 185 nucleotides on each genome that share 79% percent similarity (Fig. 1, bold type). As described earlier, these regions contain nucleotide sequences and structural similarities with the intergenic regions of other geminiviruses (Rochester *et al.*, 1990). The degree of sequence identity between the common regions of the two cloned TYLCV-Th DNAs provides sufficient evidence to conclude that they are components of a single virus and therefore were designated A- and B-DNAs. The lengths of TYLCV-Th A- and B-DNA were determined to be 2743 and 2737 nucleotides, respectively.

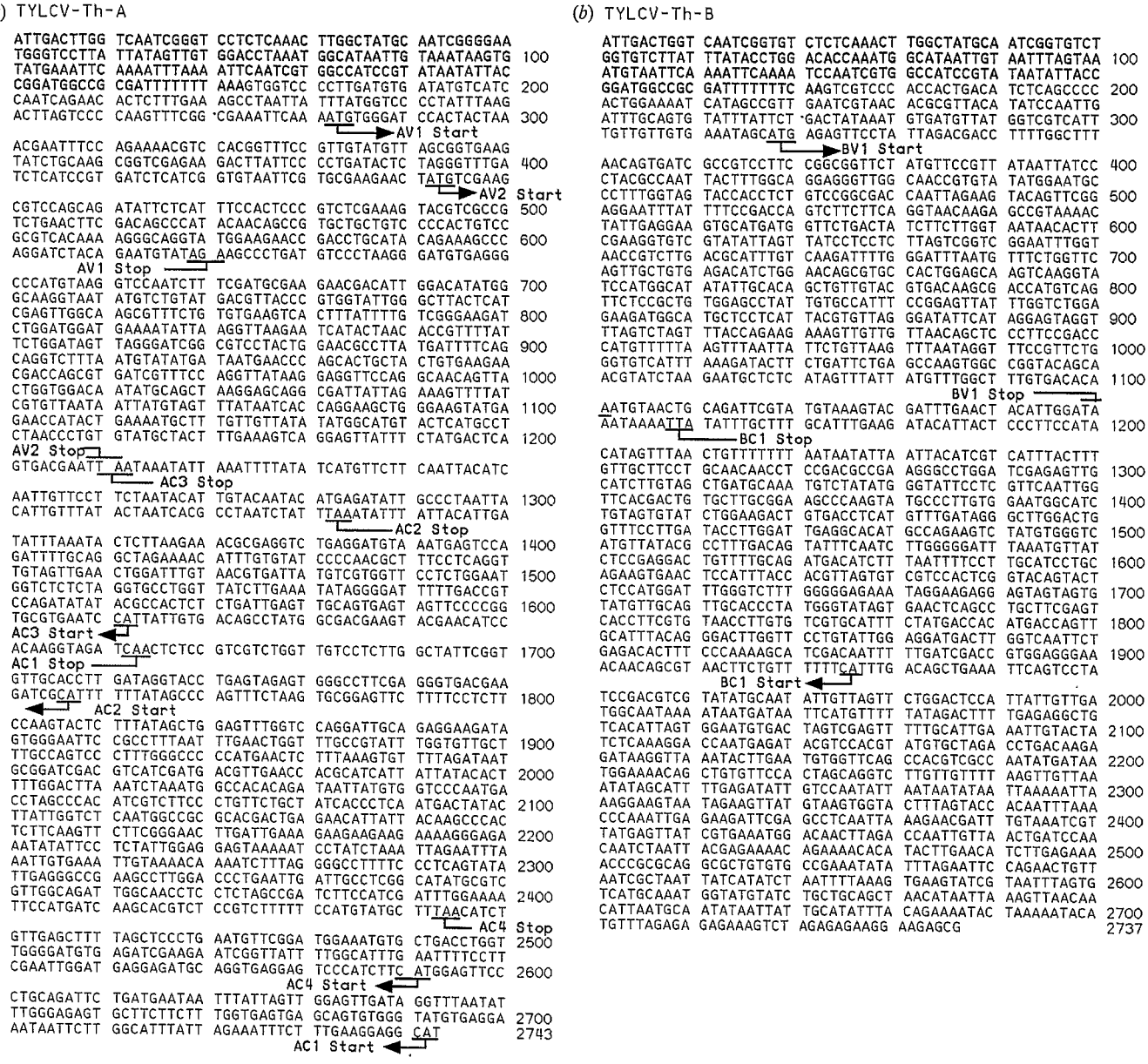

Fig. 1. Nucleotide sequence of the virion strands (5′ to 3′) of TYLCV-Th A-DNA (*a*) and B-DNA (*b*). The first nucleotide begins the IR (Rochester *et al.*, 1990) (bold face). Presumed translational start sites of ORFs are indicated by arrows.

ORFs of TYLCV-Th

Based upon computer analysis and comparisons with other geminiviruses, possible ORFs were located on the two DNAs of TYLCV-Th, as shown in Fig. 2 and Table 1. The A-DNA is capable of encoding six proteins of M_r greater than 10K while the B-DNA contains two such ORFs. The sense strand of A-DNA is capable of encoding two overlapping ORFs with products of 12·9K (AV1) and 29·8K (AV2). Both polypeptides are similar in amino acid sequence and genome position to the precoat and capsid proteins of other geminiviruses with dicotyledonous hosts. The complementary strand is predicted to encode four proteins of M_r greater than 10K. The proteins were designated AC1 (40·9K), AC2 (15·4K), AC3 (16·0K) and AC4 (11·3K) based on size, position and sequence similarities with polypeptides of other geminiviruses (Tables 1 and 2). TYLCV-Th A-DNA ORFs designated AV1, AV2, AC1, AC2, AC3 and AC4 are analogous to the ORFs V1, V2, C1, C2, C3 and C4 described in other geminivirus studies (Dry *et al.*, 1993). The B-DNA contains two ORFs which may encode proteins designated BV1 (31·6K) and BC1 (31·3K).

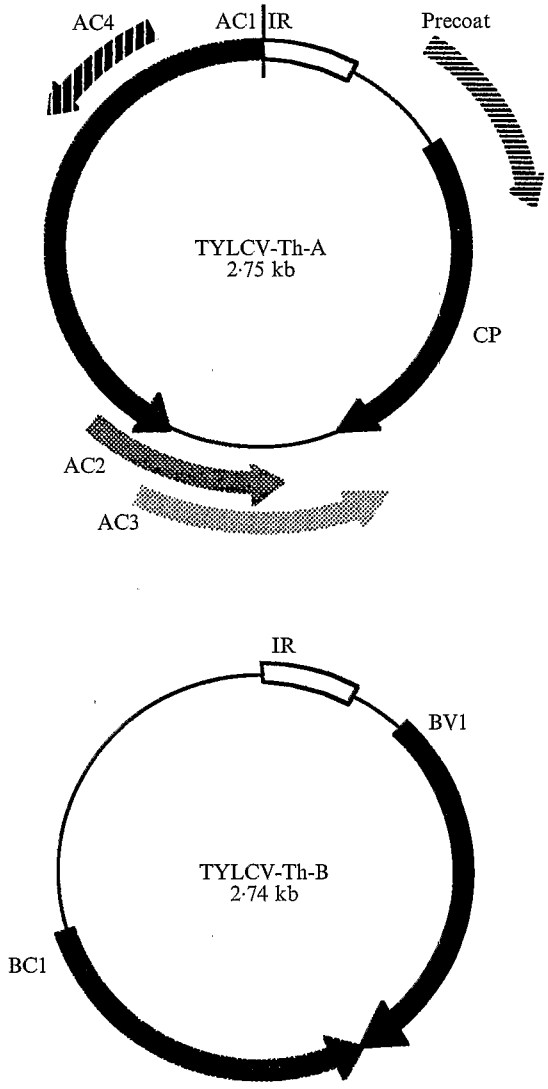

Fig. 2. Genomic organization of the TYLCV-Th A- and B-DNAs. The virion strand orientation is in the clockwise direction. Direction and relative lengths of ORFs are indicated by open boxes.

Table 1. *Predicted ORFs of TYLCV-Th*

ORF*	Nucleotides (start–stop)	M_r† ($\times 10^{-3}$)
AV1	282–620	12·9
AV2	442–1212	29·7
AC1	2743–1660	40·8
AC2	1757–1353	15·3
AC3	1612–1208	15·9
AC4	2591–2443	11·3
BV1	318–1151	31·6
BC1	1990–1158	31·2

* A and B designate whether ORF is found on the A- or B-DNA. V and C denote virion- and complementary-sense ORFs, respectively, as shown in Fig. 2.

† M_r of putative translation products.

Table 2. *Percentage similarity between the products of the different ORFs for the geminiviruses infecting tomatoes in the Old World, and ACMV*

	TLCV	TYLCV-I	TYLCV-S	ACMV
AV1				
TYLCV-Th	6·2	71	66	66
TLCV		60	58	63
TYLCV-I			82	72
TYLCV-S				72
AV2				
TYLCV-Th	74	73	69	71
TLCV		76	73	74
TYLCV-I			86	80
TYLCV-S				76
AC1				
TYLCV-Th	80	78	75	71
TLCV		83	78	76
TYLCV-I			77	72
TYLCV-S				74
AC2				
TYLCV-Th	61	57	55	60
TLCV		61	56	62
TYLCV-I			67	64
TYLCV-S				64
AC3				
TYLCV-Th	68	67	62	64
TLCV-A		47	58	63
TYLCV-I			65	73
TYLCV-S				70
AC4				
TYLCV-Th	46	63	49	23
TLCV		79	63	50
TYLCV-I			76	27
TYLCV-S				37
BV1				
TYLCV-Th				25
BC1				
TYLCV-Th				41

Comparisons of intergenic regions and entire sequences

The intergenic regions (IR) of 15 geminiviruses with members of all described subgroups were compared in a multiple alignment using Megalign as described in Methods. In the case of bipartite geminiviruses, the IR of the A-DNA was used. For the alignments, the IR included the first 200 nucleotides of the virion strand sequence beginning at the first nucleotide 3′ of the ATG start codon of AC1. A phylogenetic tree constructed from the multiple alignment analysis is shown in Fig. 3. One observation resulting from such an analysis is that the geminiviruses used in this comparison are grouped on the basis of the insect vector (i.e. whiteflies or leafhoppers) and whether the natural plant host is dicotyledonous (whitefly-transmitted) or monocotyledonous (leafhopper-transmitted). On the other hand BCTV, a geminivirus of a dicotyledonous host, is transmitted by the leafhopper. The IRs of the viruses

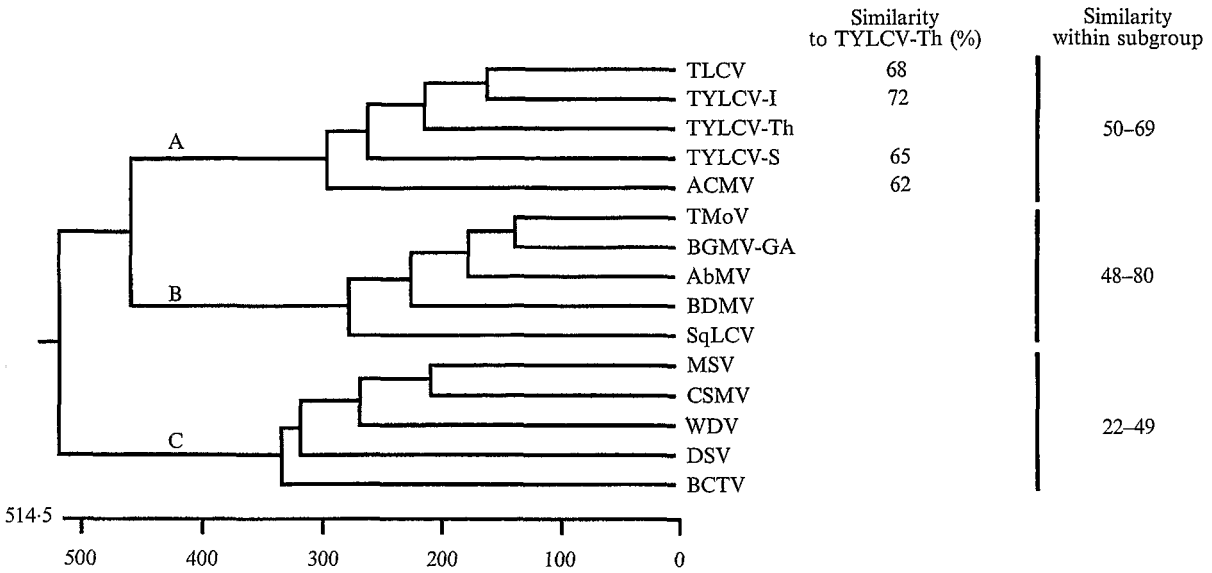

Fig. 3. Phylogenetic tree constructed from pairwise sequence comparison of the intergenic regions of fifteen geminiviruses. Percentage similarity is defined in Methods.

that infect dicotyledonous plants are further divided into two main groups, the 'Old World' and 'New World' viruses based on the geographical location where the virus was found (Howarth & Vandemark, 1989; Abouzid *et al.*, 1992). When the viruses within any of the three major branches (A, B or C, Fig. 3) are compared in a pairwise manner, the greatest overall similarity is found amongst the whitefly-transmitted geminiviruses. The IR of TYLCV-Th maintains the highest degree of percentage similarity with TYLCV-I as shown in Fig. 3. The range of sequence similarities within the IR for any branch of the proposed phylogenetic tree indicates that the greatest degree of diversity occurs in the leafhopper-transmitted geminiviruses. Similar comparisons have been made for sequences of all the ORFs and the three major branches described above can always be found with the exception that one or two viruses are switched from one branch to another (data not shown).

Other sequence comparisons made here were restricted to the Old World geminiviruses, i.e. TYLCV-Th, TYLCV-I, TYLCV-S, TLCV and ACMV. Relationships with other geminiviruses will be discussed where appropriate.

Comparison of ORFs

The predicted products of the ORFs of different Old World geminiviruses that infect tomato plants were compared in a pairwise fashion with the Megalign program to determine the percentage similarity between the viruses as well as relatedness of sequences within the group. The percentage similarity as defined in Methods and shown in Table 2 is derived through multiple alignments of only the Old World geminiviruses.

Comparison of AV1 (precoat) sequences

A small ORF, AV1, is found in a similar position in all Old World geminiviruses and overlaps the amino terminus of the V2 (capsid) ORF. Percentage similarities between the AV1 ORFs is shown in Table 2. TYLCV-Th maintains the highest percentage similarity in the AV1 ORF with TYLCV-I (71%).

Comparison of AV2 (capsid) protein sequences

The CP gene is a highly conserved ORF among geminiviruses with similarity varying between 69 and 83% among the Old World geminiviruses (Table 2). The New World geminiviruses maintain direct amino acid sequence similarities greater than 90% between SqLCV, TGMV and BGMV CP sequences (Lazarowitz & Lazdins, 1991). There are blocks of highly conserved amino acids throughout the ORFs of all CPs compared and C termini of the Old and New World geminiviruses have very high degrees of similarity. In fact, the sequence from amino acids 223 to 248 (TYLCV-Th) is 100% conserved between Old World geminiviruses and the three New World geminiviruses mentioned above (alignment not shown). By pairwise comparisons of the amino acid sequences of the CPs of the Old World geminiviruses, TYLCV-S and TYLCV-I share the greatest similarity (86%). TYLCV-Th maintains the highest similarity with TLCV (74%).

Table 3. *Agroinoculation of* N. benthamiana *with TYLCV-Th CP mutants*

TYLCV-Th DNA agroinoculated	A-DNA*	B-DNA*	Symptoms†
A	−	−	−
A	+	−	A
A	+	−	A
A+B	+	+	A+B
A+B	+	−	A
A+B	+	−	A
Acp	−	−	−
Acp	−	−	−
Acp	−	−	−
Acp+B	+	−	−
Acp+B	+	+	M
Acp+B	+	−	M
B	−	−	−
B	−	−	−
B	−	−	−

* Presence (+) or absence (−) of TYLCV-Th A- or B-DNA in extracts of upper non-inoculated leaves as determined by hybridization with the appropriate insert.

† A and A+B symptoms are described in Rochester *et al.* (1990). A+B, very severe; A, severe; M, mild.

Comparison of AC1, AC2, AC3 and AC4 sequences

Analyses of the leftward-transcribed cistrons of TYLCV-Th and counterparts of the Old World geminiviruses indicate that AC1 maintains the highest sequence conservation amongst these four polypeptides (Table 3). TYLCV-Th shares the highest similarity with TLCV in AC1, AC2 and AC3, although the differences between the viruses are, in general, small. AC4, a small ORF found within the TYLCV-Th AC1 ORF, shows the highest sequence similarity with TYLCV-I (63%). AC4 is the least conserved ORF in terms of sequence similarity amongst the Old World geminiviruses.

Comparison of BC1 and BV1 sequences

To date, TYLCV-Th is the only bipartite geminivirus of the Old World subgroup that readily infects tomato plants. TYLCV-Th BC1 maintains weak sequence similarity with the homologous ORFs of ACMV (43%), TGMV (45%), BGMV (41%) and SqLCV (42%). The most impressive similarity is found in BC1 sequences of the geminiviruses of the New World with several large blocks of identical amino acids (alignments not shown). Overall, the similarities between the BV1 ORFs are less than those found in any of the other ORFs. TYLCV-Th and ACMV BV1 maintain very little similarity with each other (25%), and no blocks of more than four identical amino acids were found. The phylogenetic tree constructed from multiple alignments of each of the polypeptides encoded by the B-DNA consistently separates the Old World and New World geminiviruses (data not shown).

Infectivity of a TYLCV-Th CP mutant

Studies of geminiviruses with a single genomic DNA have shown that the CP is necessary for systemic spread of infection. Therefore, these studies were undertaken to determine whether the CP of TYLCV-Th was required for local and systemic infection. Eleven days after agroinoculation, three of five tomato plants (UC82b) and 13 of 15 *N. benthamiana* plants agroinoculated with wild-type TYLCV-Th:A displayed disease symptoms. All plants with symptoms contained viral DNA (data not shown). On the other hand, at 11 days post-infection, plants agroinoculated with TYLCV-Th:Acp did not display any of the symptoms observed on TYLCV-Th:A infected plants.

None of the tomato plants inoculated with TYLCV-Th:Acp contained detectable amounts of viral DNA. By contrast, three *N. benthamiana* that were similarly agroinoculated accumulated viral DNA at levels five-to-tenfold lower than that found in TYLCV-Th:A infection (data not shown). Because of this result, *N. benthamiana* were then agroinoculated with TYLCV-Th:A, TYLCV-Th:Acp, TYLCV-Th:B and all possible combinations (three plants each treatment). Symptoms were detected on plants infected with TYLCV-Th:A (two of three) and mixed infections of TYLCV-Th:A and TYLCV-Th:B (three of three) at 28 days post-inoculation as expected. Approximately 75 days after inoculation, mild symptoms were observed on two of three plants agroinoculated with a combined culture of TYLCV-Th:Acp and TYLCV-Th:B. Dot blot analysis was carried out on samples from all plants and results are shown in Table 3. In five of six plants agroinoculated with TYLCV-Th:A alone or in combination with TYLCV-Th:B, TYLCV-Th A-DNA was detected in the upper leaves. One of three plants agroinoculated with a mixture of TYLCV-Th:A and TYLCV-Th:B supported B-DNA replication; this plant displayed severe disease symptoms as described before (Rochester *et al.*, 1990). Two of three plants that were agroinoculated with TYLCV-Th:Acp and TYLCV-Th:B contained TYLCV-Th A-DNA sequences. The presence of TYLCV-Th B-DNA was also detected in one of these plants. In these experiments, no significant differences in the amounts of accumulated viral DNA were observed.

These results indicated that the presence of TYLCV-Th CP enhances progression of disease. Excised *N. benthamiana* leaves were subsequently used in other experiments to determine the role of CP in the local movement of TYLCV-Th A-DNA. Leaves were removed from plants such that the petiole was left intact or

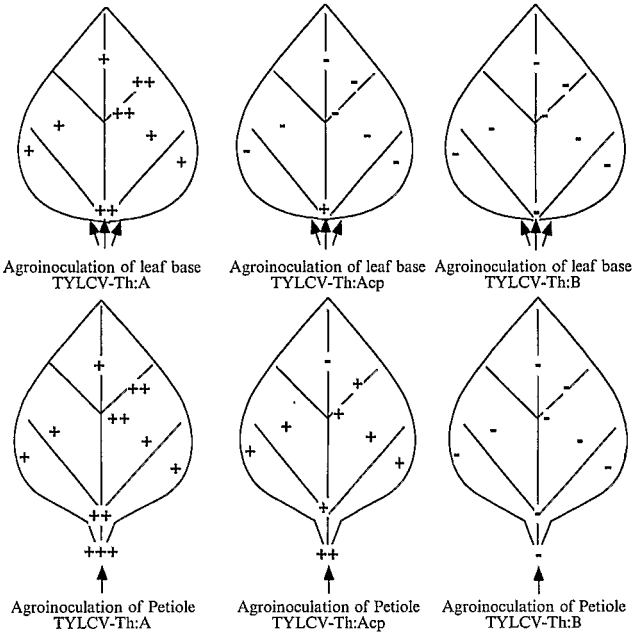

Fig. 4. Agroinoculation site, spread and accumulation of the viral DNA of TYLCV-Th:A and TYLCV-Th:Acp infection on excised *N. benthamiana* leaves. The symbols (+) and (−) designate presence and amount (+ < + + < + + +) or absence of TYLCV-Th A-DNA 4 weeks following initial agroinoculation at the indicated locations (↑). There was no detectable movement of the TYLCV-Th B-DNA following agroinoculation at either location with TYLCV-Th:B (i.e. all locations were negative).

removed, and different sites of the leaf were agroinoculated (as shown in Fig. 4). Treatments included TYLCV-Th:A, TYLCV-Th:Acp, TYLCV-Th:B, and mock inoculation at the tip of the petiole or at the base of the leaf where the petiole and lower portion of the leaf had been removed (see Fig. 4). Samples of leaf tissue were taken 1 month after agroinoculation and the relative levels of accumulation of TYLCV-Th A- or B-DNA in leaf extracts was determined (Fig. 4). Agroinoculation of TYLCV-Th:A at any location resulted in a more extensive spread throughout the leaf and a greater amount in viral DNA than following infection with TYLCV-Th:Acp. Viral DNA from TYLCV-Th:A infection was found at all locations sampled including the vascular system and interveinal areas regardless of the site of inoculation. Agroinoculation at the base of the leaf with TYLCV-Th:Acp led to little spread of the mutant virus, although TYLCV-Th:Acp DNA was detected at the site of agroinoculation, possibly the result of initial infection and replication. In contrast, agroinoculation of TYLCV-Th:Acp at the base end of the petiole allowed subsequent spread of the virus throughout the leaf although virus DNA levels were lower than in tissues agroinoculated with TYLCV-Th:A. The data as presented in Table 3 and Fig. 4 make it unlikely that the accumulation of viral DNA was due to the spread of *A. tumefaciens* since B-DNA was not detected. Furthermore, *A. tumefaciens* did not overgrow the plates.

Discussion

The A- and B-DNAs of TYLCV-Th were cloned and sequenced. By comparing the genomic organization and similarities of the amino acid sequences of the ORFs predicted from the nucleotide sequences we concluded that TYLCV-Th has, not surprisingly, greatest similarity to the geminiviruses that infect dicotyledonous rather than monocotyledonous plants. Construction of a phylogenetic tree based upon a comparison of sequences in the IR divided 15 different geminiviruses into three main groups (Fig. 3). One group contains the leafhopper-transmitted geminiviruses with a single molecule DNA genome. With the exception of BCTV, these geminiviruses have monocotyledonous hosts. Another major branch of this phylogenetic tree contains the New World bipartite geminiviruses with dicotyledonous hosts. The third group, of which TYLCV-Th is a member, comprises the Old World geminiviruses with dicotyledonous hosts. TYLCV-Th exhibited the greatest sequence similarity in the predicted ORFs with TLCV than with the other Old World geminiviruses. In comparisons of the ORFs on the B-DNA, TYLCV-Th and ACMV were found to have weak similarities with all geminiviruses, including each other. In contrast, New World geminiviruses show higher homologies to each other in the ORFs on B-DNA (Lazarowitz & Lazdins, 1991). In phylogenetic trees constructed from multiple alignments containing all the whitefly-transmitted geminiviruses used in the present study, the homologous ORFs of Old World and New World members each clustered together (data not shown).

Three Old World geminiviruses of tomato (TYLCV-I, TYLCV-S and TLCV) appear to contain all the necessary genetic information for natural infection on a single DNA molecule. Agroinoculation of tomato with TYLCV-I and TYLCV-S resulted in symptoms and viral DNA accumulation that was indistinguishable from natural infection. Furthermore, in these studies whiteflies effectively transmitted the disease from agroinoculated plants to healthy plants (Navot *et al.*, 1991; Kheyr-Pour *et al.*, 1991). The genome of TLCV is a single DNA molecule which is infectious on tomato following agroinoculation of a dimer construct (Dry *et al.*, 1993). In these cases, exhaustive searches for a B-like DNA indicate that these geminiviruses are indeed monopartite. TYLCV-Th, on the other hand, contains a bipartite genome similar in structure to those of ACMV and the other whitefly-transmitted New World geminiviruses. However, the A-DNA alone of TYLCV-Th is capable of

causing symptomatic infections in tomato and *N. benthamiana*. Nevertheless, symptoms of TYLCV-Th disease are more severe following agroinoculation of both A- and B-DNAs.

Agroinoculation of *N. benthamiana* plants with TYLCV-Th A-DNA constructs containing a mutation in the CP cistron resulted in delayed and decreased disease symptoms and reduced accumulation of A-DNA. These findings are similar to the results of studies where disruptions in the V2 (CP) cistron of TLCV greatly reduced virus spread following agroinoculation (Rigden *et al.*, 1993). Systemic infection by geminiviruses with bipartite genomes is not severely affected by disruption of their CP genes (Brough *et al.*, 1988; Etessami *et al.*, 1989). Where studied, geminiviruses with a single genomic DNA have no capacity to cause disease in the absence of the CP (Lazarowitz *et al.*, 1989; Boulton *et al.*, 1989; Briddon *et al.*, 1989). Initial infection and movement of the TYLCV-Th A-DNA following agroinoculation is not completely eliminated by a mutation in the CP since systemic infection and mild symptoms occurred in whole plants, although delayed relative to infection by wild-type A-DNA. Interestingly, in one plant which supported TYLCV-Th:Acp and TYLCV B-DNA, accumulation of both viral components was equivalent to the levels found following agroinoculation of both wild-type viral DNAs. In this case, function(s) lost owing to the mutation in the CP may have been provided by the B-DNA. Two research groups demonstrated previously that efficient systemic infection by the bipartite geminiviruses ACMV and TGMV was only possible when both A- and B-DNAs were present (Stanley, 1983; Hamilton *et al.*, 1983). In the case of TYLCV-Th, infection of *N. benthamiana* with both A- and B-DNAs resulted in a more severe infection than infection by A-DNA alone (Rochester *et al.*, 1990). However, infection by A-DNA alone led to significant disease symptoms. In contrast, plants infected with only the A-DNA of ACMV did not develop symptoms and contained 20-fold less viral DNA than those infected with both A- and B-DNAs (Klinkenberg & Stanley, 1990). More recently it was shown that agrobacterium-mediated inoculation of the DNAs of TLCV and two TYLCV isolates on tomato leads to systemic infection indistinguishable from natural whitefly-transmitted disease (Navot *et al.*, 1991; Kheyr-Pour *et al.*, 1991; Dry *et al.*, 1993). In this regard, TYLCV-Th has characteristics somewhat intermediate among the whitefly-transmitted geminiviruses.

Excised intact leaves were used to analyse the role of the CP in the local movement of TYLCV-Th. Inoculation of TYLCV-Th:A at the petiole or the base of the leaf resulted in a more efficient systemic spread than infection by TYLCV-Th:Acp (Fig. 4). Agroinoculation of excised leaves indicated that the location of the initial site of infection dictated and possibly enhanced the degree of infection by TYLCV-Th:Acp. Infection at the petiole resulted in spread of the mutant virus throughout the leaf, although much less viral DNA accumulated than after inoculation with the wild-type A-DNA. However, when TYLCV-Th:Acp was inoculated at the leaf base, there was little or no spread of the mutant virus. A small amount of viral DNA was detected at the site of inoculation possibly due to initial infection and replication. These results are similar to those obtained with TLCV CP mutants (Rigden *et al.*, 1993). The results indicate that the CP of these viruses is not necessary for viral DNA synthesis, yet is needed for cell to cell spread. When excised leaves were inoculated with TYLCV-Th:Acp at the petiole, there was apparently systemic movement through the vascular tissue and limited cell-to-cell movement within the interveinal areas.

TYLCV-Th has genomic and biological characteristics that place it in an intermediate position among the geminiviruses. The bipartite genome structure and arrangement of ORFs of this virus are similar to those of the well characterized whitefly-transmitted geminiviruses. Furthermore, the CP of TYLCV-Th enhances but is not completely necessary for efficient systemic infection yet appears to enhance cell-to-cell movement and accumulation of viral DNAs.

C. M. Fauquet was supported by ORSTOM, 213 Rue La Fayette, 75048 Paris Cedex 10, France. This work was supported by a grant from Monsanto Company.

References

ABOUZID, A. M., POLSTON, J. E. & HIEBERT, E. (1992). The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from tomatoes in Florida. *Journal of General Virology* **73**, 3225–3229.

ANDERSEN, M. T., RICHARDSON, K. A., HARRISON, S. & MORRIS, B. A. M. (1988). Nucleotide sequence of the geminivirus *Chloris* striate mosaic virus. *Virology* **164**, 443–449.

BOULTON, M. I., STEINKELLNER, H., DONSON, J., MARKHAM, P. G., KING, D. I. & DAVIES, J. W. (1989). Mutational analysis of the virion-sense genes of maize streak virus. *Journal of General Virology* **70**, 2309–2323.

BRIDDON, R. W., WATTS, J., MARKHAM, P. G. & STANLEY, J. (1989). The coat protein of beet curly top virus is essential for infectivity. *Virology* **172**, 628–633.

BROUGH, C. L., HAYES, R. J., MORGAN, A. J., COUTTS, R. H. A. & BUCK, K. W. (1988). Effects of mutagenesis *in vitro* on the ability of cloned tomato golden mosaic virus DNA to infect *Nicotiana benthamiana* plants. *Journal of General Virology* **69**, 503–514.

DONSON, J., ACCOTTO, G. P., BOULTON, M. I., MULLINEAUX, P. M. & DAVIES, J. W. (1987). The nucleotide sequence of a geminivirus from *Digitaria sanguinalis*. *Virology* **161**, 160–169.

DRY, I. B., RIGDEN, J. E., KRAKE, L. R., MULLINEAUX, P. M. & REZAIAN, M. A. (1993). Nucleotide sequence and genome organization of tomato leaf curl geminivirus. *Journal of General Virology* **74**, 147–151.

ETESSAMI, P., WATTS, J. & STANLEY, J. (1989). Size reversion of African cassava mosaic virus coat protein gene deletion mutants during infection of *Nicotiana benthamiana*. *Journal of General Virology* **70**, 277–289.

FRISCHMUTH, T., ZIMMAT, G. & JESKE, H. (1990). The nucleotide sequence of abutilon mosaic virus reveals prokaryotic as well as eukaryotic features. *Virology* **178**, 461–468.

GARDINER, W. E., SUNTER, G., BRAND, L., ELMER, J. S., ROGERS, S. G. & BISARO, D. M. (1988). Genetic analysis of tomato golden mosaic virus: the coat protein is not required for systemic spread of symptom development. *EMBO Journal* **7**, 899–904.

GILBERTSON, R. L., FARIA, J. C., HANSON, S. F., MORALES, F. J., AHLQUIST, P., MAXWELL, D. P. & RUSSEL, D. R. (1991*a*). Cloning of the complete DNA genomes of four bean-infecting geminiviruses and determining their infectivity by electric particle discharge acceleration. *Phytopathology* **81**, 980–985.

GILBERTSON, R. L., HIDAYAT, S. H., ROJAS, M. R. & MAXWELL, D. P. (1991*b*). Infectious DNA clones of a new geminivirus associated with tomatoes in Florida. *Phytopathology* **81**, 1184.

HAMILTON, W. D. O., BISARO, D. M., COUTTS, R. H. A. & BUCK, K. W. (1983). Demonstration of the bipartite nature of the genome of a single-stranded plant virus by infection with cloned DNA components. *Nucleic Acids Research* **11**, 7387–7396.

HAMILTON, W. D. O., STEIN, V., COUTTS, R. H. A. & BUCK, K. W. (1984). Complete nucleotide sequence of the infectious cloned DNA components of tomato golden mosaic virus: potential coding regions and regulatory sequences. *EMBO Journal* **3**, 2197–2205.

HOWARTH, A. J. & VANDEMARK, G. J. (1989). Phylogeny of geminiviruses. *Journal of General Virology* **70**, 2717–2727.

HOWARTH, A. J., CATON, J., BOSSERT, M. & GOODMAN, R. M. (1985). Nucleotide sequence of bean golden mosaic virus and a model for gene regulation in geminiviruses. *Proceedings of the National Academy of Sciences, U.S.A.* **82**, 3572–3576.

HULL, R., STANLEY, J. & BRIDDON, R. W. (1991). The geminivirus group. In *Classification and Nomenclature of Viruses. Archives of Virology* Supplementum 2, pp. 173–177. Edited by R. I. B. Francki, C. M. Fauquet, D. L. Knudson & F. Brown. New York: Springer-Verlag.

KHEYR-POUR, A., BENDAHMANE, M., MATZEIT, V., ACCOTTO, G. P., CRESPI, S. & GRONENBORN, B. (1991). Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartite geminivirus. *Nucleic Acids Research* **19**, 6763–6769.

KLINKENBERG, F. A. & STANLEY, J. (1990). Encapsidation and spread of African cassava mosaic virus DNA A in the absence of DNA B when agroinoculated to *Nicotiana benthamiana*. *Journal of General Virology* **71**, 1409–1412.

LAZAROWITZ, S. G. & LAZDINS, I. N. (1991). Infectivity and complete nucleotide sequence of the cloned genomic components of the bipartite squash leaf curl geminivirus with a broad host range phenotype. *Virology* **180**, 58–69.

LAZAROWITZ, S. G., PINDER, A. J., DAMSTEEGT, V. D. & ROGERS, S. G. (1989). Maize streak virus genes essential for systemic spread and symptom development. *EMBO Journal* **8**, 1023–1032.

MACDOWELL, S. W., MACDONALD, H., HAMILTON, W. D. O., COUTTS, R. H. A. & BUCK, K. W. (1985). The nucleotide sequence of cloned wheat dwarf virus DNA. *EMBO Journal* **4**, 2173–2180.

MULLINEAUX, P. M., DONSON, J., MORRIS-KRSINICH, B. A. M., BOULTON, M. I. & DAVIES, J. W. (1984). The nucleotide sequence of maize streak virus DNA. *EMBO Journal* **3**, 3063–3068.

NAVOT, N., PICHERSKI, E., ZEIDAN, M., ZAMIR, D. & CZOSNEK, H. (1991). Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a single genomic molecule. *Virology* **185**, 151–161.

RIGDEN, J. E., DRY, I. B., MULLINEAUX, P. M. & REZAIAN, M. A. (1993). Mutagenesis of the virion-sense open reading frames of tomato leaf curl geminivirus. *Virology* **193**, 1001–1005.

ROCHESTER, D. E., KOSITRATANA, W. & BEACHY, R. N. (1990). Systemic movement and symptom production following agroinoculation with a single DNA of tomato yellow leaf curl geminivirus (Thailand). *Virology* **178**, 520–526.

ROGERS, S. G., HORSCH, R. B. & FRALEY, R. T. (1986). Gene transfer in plants: production of transformed plants using Ti plasmid vectors. *Methods in Enzymology* **118**, 627–640.

SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, U.S.A.* **74**, 5463–5467.

STANLEY, J. (1983). Infectivity of the cloned geminivirus genome requires sequences from both DNAs. *Nature, London* **305**, 643–645.

STANLEY, J. & GAY, M. R. (1983). Nucleotide sequence of cassava latent virus DNA. *Nature, London* **301**, 260–262.

STANLEY, J. & TOWNSEND, R. (1986). Infectious mutants of cassava latent virus generated *in vivo* from intact recombinant DNA clones containing single copies of the genome. *Nucleic Acids Research* **14**, 5981–5998.

STANLEY, J., MARKHAM, P. G., CALLIS, R. J. & PINNER, M. S. (1986). The nucleotide sequence of an infectious clone of the geminivirus beet curly top virus. *EMBO Journal* **5**, 1761–1767.

(*Received 2 July 1993; Accepted 25 October 1993*)