Nucleotide sequence of a new bipartite geminivirus isolated from the common weed *Sida rhombifolia* in Costa Rica

Petra Höfer, Margit Engel, Holger Jeske and Thomas Frischmuth

Biologisches Institut, Abteilung für Molekularbiologie und Virologie der Pflanzen, Universität Stuttgart, Pfaffenwaldring 57, D-70550 Stuttgart, Germany

The nucleotide sequence of infectious clones of a geminivirus from Costa Rica that infects *Sida rhombifolia* was determined. Sida golden mosaic virus (SiGMV-Co) has a bipartite genome (DNAs A and B). Computer analysis showed that the bipartite genome of SiGMV-Co resembles that of other whitefly-transmitted geminiviruses. The DNA A (2605 nt) and DNA B (2587 nt) components have little sequence homology other than within the common region (CR). Analysis of DNAs A and B showed that SiGMV-Co is closely related to bean dwarf mosaic virus (BDMV). SiGMV-Co was introduced via agroinoculation into seven plant species, including tomato and bean.

Introduction

Geminiviruses are serious pathogens of diverse crops that are widely distributed throughout tropical, subtropical and, to a more limited extent, temperate regions of the world. Members of the geminivirus family have been divided into three subgroups (Murphy *et al*., 1995). The majority of group members that infect dicotyledonous plants are whitefly-transmitted and have bipartite genomes (DNAs A and B). DNA A encodes the coat protein as well as proteins required for replication and gene regulation. DNA B is essential for disease production but plays no role in DNA replication. The two gene products encoded by DNA B are involved in spread of the virus throughout the plant, development of symptoms and in determining the host range of the virus (Lazarowitz, 1992; Timmermans *et al*., 1994; Noueiry *et al*., 1994; Ingham *et al*., 1995).

Diseases caused by whitefly-transmitted geminiviruses impose important constraints on crop production in tropical and subtropical regions. In some cases the incidence of geminiviruses can be so severe that certain economically important crops are not grown, e.g. African cassava mosaic virus (ACMV; Stanley & Gay, 1983) in cassava (Thresh *et al*., 1994). Until recently, tomato and bean plants grown in the western hemisphere were not known to be seriously affected by geminiviruses. However, geminivirus-associated epidemics are currently threatening tomato, bean and pepper production in Central America (e.g. Mexico, Puerto Rico and Costa Rica) and the southern United States (Simone *et al*., 1990; Hidayat *et al*., 1993; Torres-Pacheco *et al*., 1993; Brown *et al*., 1995).

*Sida rhombifolia* is a common weed which grows all over Latin America and the southern United States. Often plants are infected with whitefly-transmitted geminiviruses which might be the source of geminivirus-associated epidemics (Gilbertson *et al*., 1993). In this paper we describe the genome of a new bipartite geminivirus isolated from *S. rhombifolia* collected in Costa Rica.

Methods

- **Virus sources and isolation of double-stranded DNA forms.** *S. rhombifolia* plants infected with SiGMV-Co (kindly provided by Peter Markham, John Innes Centre, Norwich, UK) were collected in Costa Rica. The virus was sap-transmitted to *Nicotiana benthamiana* and viral double-stranded DNA was isolated from infected plants as described previously (Frischmuth & Stanley, 1991).

- **Cloning and sequencing of SiGMV-Co DNAs A and B.** Recombinant DNA techniques were performed as described by Sambrook *et al*., 1989). Restriction endonucleases and DNA-modifying enzymes were used as recommended by the manufacturers. The sequences of DNAs A and B were determined by either the dideoxynucleotide chain termination method (Pharmacia) with [α-35S]dATP or by automatic sequence analysis with the Li-Cor system, according to the manufacturers’ instructions. Sequence analyses were performed with WDNASIS, PROSIS and CLUSTAL W computer programs.

For cloning of full-length DNA A the unique PstI site was used. Viral double-stranded DNA was digested with PstI and linearized DNA was cloned into pBluescript KS(−) (Stratagene). Full-length DNA B was
cloned as an EcoRV fragment. Firstly, an SstI fragment, containing the unique EcoRV site of DNA B, was cloned into pBluescript KS(−) and sequences adjacent to the EcoRV site were determined in order to design primers for PCR amplification. Full-length DNA B was produced with the primer pair SGMV8 (5′ AGC GAATTC GAT ATC 3′) and SGMV9 (5′ ATC GAATTC GAT ATC 3′) introducing an extra EcoRI site (underlined) at the ends of the PCR product. The PCR fragment was cloned into the EcoRI site of pBluescript KS(®) at the unique EcoRI site (underlined) at the ends of the PCR product. The PCR fragment was cloned into the EcoRI site of pBluescript KS(−) and full-length DNA B could be liberated by EcoRV digestion (EcoRV site in italics).

A partial repeat of SiGMV-Co DNA A was produced by cloning a BamHI–SstI fragment into pBluescript KS(−) and subsequent introduction of full-length DNA A into the unique BamHI site. The SiGMV-Co DNA B partial repeat consisted of an EcoRV–SstI fragment and the full-length EcoRV DNA B fragment. The partial DNA A repeat was transferred as an XbaI–KpnI fragment into pBin19. The partial repeat of DNA B was first amplified with the primer SGMV10 (5′ AGC GAATTC GGC GAA TTG GAG CTC 3′) and the reverse primer of pBluescript KS(−). With primer SGMV10 an additional EcoRI site (underlined) was added, allowing transfer of the partial DNA B repeat as a HindIII–EcoRI fragment into pBin19. Clones were mobilized into Agrobacterium tumefaciens LBA4404 by triparental mating (Stanley et al., 1990).

**Inoculation of plants and characterization of viral DNA forms.** Cloned genomic components were introduced into plants by agroinoculation (Stanley et al., 1990). Total cellular nucleic acids were extracted 14 days post-inoculation from systemically infected leaves as described by Frischmuth & Stanley (1991). Samples were analysed either by agarose gel electrophoresis or by the PCR method. Samples containing 5 µg total nucleic acids were

---

**Fig. 1.** (a) Maps of the SiGMV-Co genomic A and B components. The solid arrows define the position of the open reading frames (C for complementary and V for virus sense). The common region (CR) between DNAs A and B is indicated by a shaded box. The positions of selected restriction endonuclease sites are indicated. The calculated molecular masses (MW) of putative proteins encoded by DNA A and B ORFs are shown below the maps. (b) Nucleotide sequence comparison of the common region between DNA A (A) and B (B). Nucleotide mismatches are indicated by asterisks and an inverted repeat sequence able to form a hairpin–loop structure by arrows above the sequence.
analysed on agarose gels in 0.5× TBE buffer and nucleic acids were transferred to nylon membranes. Viral DNA was detected by hybridization with digoxigenin-labelled probes prepared by random priming according to the manufacturer’s recommendations (Boehringer Mannheim). DNA components were specifically detected using an EcoRV fragment of DNA A and an SstI fragment of DNA B.

For PCR analysis primers SGCP1 (5′ TTT ATT TCA TTA TCG 3′) and SGCP2 (5′ CAA ATG GCC TAA GCG CGA 3′) were used. This primer pair allows a specific amplification of the coat protein gene of SiGMV-Co. PCR fragments were analysed on agarose gels in 0.5× TBE.

Results and Discussion

The increasing incidence of geminiviruses threatens the production of tomatoes, beans and peppers in Latin America (Simone et al., 1990; Hidayat et al., 1993; Torres-Pacheco et al., 1993; Brown et al., 1995). A number of bipartite geminiviruses from infected tomatoes, beans and peppers have been analysed (Morales et al., 1990; Gilbertson et al., 1991; Abouzid et al., 1992; Hidayat et al., 1993; Torres-Pacheco et al., 1993; Paplomatas et al., 1994).

*S. rhombifolia* grows as a common weed in Central America and plants are often infected with whitefly-transmitted viruses (Bird, 1958). To confirm the presence of a geminivirus in ornamental *S. rhombifolia*, plants were collected in Costa Rica and a bipartite geminivirus was isolated. The nucleotide sequences of the full-length genomic components A and B were determined. The sequence data are available in the EMBL/GenBank under accession no. X99550 (DNA A) and X99551 (DNA B). The sequences are numbered beginning at the first nucleotide of the common region (CR) between DNAs A and B (Fig. 1b). The genomic structure of SiGMV-Co, deduced from the nucleotide sequences, resembles that of other bipartite geminiviruses (Fig. 1a). DNA A encodes five and DNA B for two putative genes (Fig. 1a). The sequence of the CR between DNAs A and B is almost identical over

<table>
<thead>
<tr>
<th>Percentage identity to SiGMV-Co</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>AbMV</td>
</tr>
<tr>
<td>ACMV</td>
</tr>
<tr>
<td>BDMV</td>
</tr>
<tr>
<td>BGMV</td>
</tr>
<tr>
<td>PHV</td>
</tr>
<tr>
<td>PYMV</td>
</tr>
<tr>
<td>SqLCV</td>
</tr>
<tr>
<td>TGMV</td>
</tr>
<tr>
<td>ToMoV</td>
</tr>
<tr>
<td>TCrV</td>
</tr>
</tbody>
</table>

Table 1. Comparison of amino acid sequences encoded by putative genes of SiGMV-Co with 10 other bipartite geminiviruses

The computer program PROSIS was used to compare sequences.

* Sources of sequences are given in Fig. 2.
† The AV1 and BV1 gene sequences for TCrV are from Paplomatas et al. (1994).
‡ Amino acid sequence of BV1 of clone pBMx9B (Paplomatas et al., 1994).
§ Amino acid sequence of BV1 of clone pTMX44B (Paplomatas et al., 1994).
Table 2. Host range of SiGMV-Co based upon agroinoculation experiments.

<table>
<thead>
<tr>
<th>Plant species</th>
<th>Infectivity (infected/inoculated)</th>
<th>Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nicotiana benthamiana D.</td>
<td>9/9</td>
<td>Severe stunting, leaf curling</td>
</tr>
<tr>
<td>Nicotiana tabacum L. cv. Xanthi</td>
<td>6/6</td>
<td>Severe stunting, leaf curling</td>
</tr>
<tr>
<td>Nicotiana tabacum L. cv. Samsun</td>
<td>6/6</td>
<td>Severe stunting, leaf curling</td>
</tr>
<tr>
<td>Datura stramonium L.</td>
<td>4/6</td>
<td>Severe stunting, leaf curling, mosaic</td>
</tr>
<tr>
<td>Lycopersicon esculentum M.</td>
<td>2/6</td>
<td>Stunting, leaf curling</td>
</tr>
<tr>
<td>Phaseolus vulgaris L.</td>
<td>6/6</td>
<td>Severe stunting, leaf curling</td>
</tr>
<tr>
<td>Sida rhombifolia L.</td>
<td>3/20</td>
<td>Initial stunting and leaf curling*, mosaic</td>
</tr>
<tr>
<td>Malva parviflora L.</td>
<td>9/9</td>
<td>Severe stunting, leaf curling, mosaic</td>
</tr>
</tbody>
</table>

* Following agroinoculation of S. rhombifolia, plants were initially stunted and leaves showed curling and mosaic. The stunting and leaf curling symptoms disappeared 4 weeks later and only the mosaic on leaves remained.

approximately 180 nucleotides (Fig. 1b). The CR contains a sequence which is capable of producing a hairpin–loop structure (Fig. 1b). This structure is identical to that found in other geminiviruses (Lazarowitz, 1992).

Comparison of the sequences of SiGMV-Co DNAs A and B with those of other bipartite geminiviruses revealed a close relationship between geminiviruses from the New World and less homology with ACMV from Africa. Phylogenetic analysis of the CR sequences showed that SiGMV-Co clusters with bean dwarf mosaic virus (BDMV) isolated from beans in Colombia, tomato mottle virus (ToMoV; Abouzid et al., 1992) isolated from infected tomato plants in Florida (USA) and Abutilon mosaic virus (AbMV; Frischmuth et al., 1990) (Fig. 2a). The infected Abutilon plants were probably from the West Indies (Regel, 1875).

A comparison of amino acid sequences encoded by single ORFs confirmed the close relationship between SiGMV-Co and BDMV (Table 1). The amino acid sequences of the coat protein and the BV1 protein of tomato leaf crumple virus (TLCrV; Paplomatas et al., 1994) were also included in this analysis and for all the viruses examined, these two TLCrV sequences showed the greatest identity to their counterparts in SiGMV-Co (Table 1). In the multiple alignment analysis of these amino acid sequences, SiGMV-Co clusters with BDMV, AbMV, ToMoV and TLCrV (Fig. 2a). In contrast, in the CR analysis (Fig. 2a) TLCrV clusters with pepper huasteco virus (PHV; Torres-Pacheco et al., 1993) and bean golden mosaic virus (BGMV; Howarth et al., 1985). This difference for TLCrV may reflect its origin because PHV as well as TLCrV was isolated in Mexico (Torres-Pacheco et al., 1993; Paplomatas et al., 1994). A noticeable feature of the amino acid sequence analysis was the low identity between AC4 in SiGMV-Co compared to the other viruses studied (Table 1). This diversity might be due to the fact that no function of this gene has been

Fig. 3. PCR analysis of SiGMV-Co agroinoculated Nicotiana benthamiana, N. tabacum cv. Samsun, Datura stramonium, Lycopersicon esculentum, Phaseolus vulgaris, Sida rhombifolia, Malva parviflora and uninfected N. benthamiana (negative control). SiGMV-Co DNA A was used as template for PCR (positive control). The molecular mass standard was EcoRI/HindIII digested λ-DNA (marker). Sizes are indicated in bp.
found in planta (Elmer et al., 1988; Etessami et al., 1991), although protoplast experiments have indicated that AC4 may be involved in the regulation of AC1 (Grönig et al., 1994). Taken together, results show that SiGMV-Co is most closely related to bipartite geminiviruses from Central America.

Utilizing the agroinoculation method we were able to infect various plant species with SiGMV-Co (Table 2). The general symptoms in all infected plant species were stunting of the plant and leaf curling. Some plant species, e.g. *Datura stramonium*, *Malva parviflora* and *S. rhombifolia*, also exhibited yellow-green mosaic on infected leaves (Table 2). Symptomatic plants were analysed by Southern blotting (data not shown) and PCR (Fig. 3). Viral DNA was detected in all infected plants (Fig. 3). With the exception of *S. rhombifolia*, all other plants were inhibited in their growth for several weeks post-inoculation. The first symptoms of infection (stunting, leaf curling and mosaic) on the leaves of inoculated *S. rhombifolia* plants resembled those on other plant species. However, approximately 4 weeks later the growth inhibition and leaf curling symptoms disappeared and only the mosaic remained (Table 2). After this period of time symptoms of agroinoculated plants were indistinguishable from those of the original *S. rhombifolia* plants.

We have shown that SiGMV-Co infects a number of plant species, including tomato and bean. Therefore, infected *S. rhombifolia* plants might serve as a natural host reservoir for geminiviruses from which crop plants could be infected. That ornamental *Sida* or *Abutilon* plants might serve as natural host reservoirs for geminiviruses has also been proposed by Gilbertson et al. (1993). *S. rhombifolia* is a common weed in Central America and close monitoring of *S. rhombifolia* plants for infection could lead to a reduced incidence of geminiviruses in crops.

SiGMV-Co clones were held according the Gentechnik Gesetz (Licence no. 76–14/8829.02/Uni.S.01.01–5 and 76–14/8829.02/Uni.S.01.04–2). The work was supported by a DFG grant to T.F. (Fr. 1122/1–1).

References


Received 8 November 1996; Accepted 24 February 1997