Cloning and sequence analysis of an infectious clone of *Citrus yellow mosaic virus* that can infect sweet orange via *Agrobacterium*-mediated inoculation

Qi Huang and John S. Hartung

USDA, Agriculture Research Service, Fruit Laboratory, Bldg 010A, BARC-West, 10300 Baltimore Avenue, Beltsville, MD 20705, USA

---

*Citrus yellow mosaic virus* (CYMV), a member of the family *Caulimoviridae*, genus *Badnavirus*, causes citrus mosaic disease, a disease that occurs commonly in India. The CYMV genome has been cloned and its complete nucleotide sequence determined. Its DNA genome is 7559 bp in length and contains six putative open reading frames (ORFs), all on the plus-strand of the genome and each capable of encoding proteins with a molecular mass of greater than 10 kDa. ORF 3, the largest ORF, encodes a putative polyprotein for functions involved in virus movement, assembly and replication. The other ORFs encode proteins whose exact functions are not completely understood. The genome also contains a plant tRNA<sup>val</sup>-binding site, which may serve as a primer for minus-strand DNA synthesis, in its intergenic region. Phylogenetic analysis of the badnaviruses revealed that CYMV is most closely related to *Cacao swollen shoot virus*. It was demonstrated that a construct containing 1–4 copies of the cloned CYMV genome could infect sweet orange via *Agrobacterium*-mediated inoculation.

---

**Introduction**

Citrus mosaic disease in India was first described by Dakshinamurti & Reddy (1975). It was later studied in detail by Ahlawat et al. (1984, 1996a, b). It is widely distributed and is a common and severe disease in India, especially in sweet orange (*Citrus sinensis* (L.) Osbeck) (Ahlawat et al., 1996a). The most characteristic symptoms of the disease are yellow mosaic of the leaves and yellow flecking along the veins (Ahlawat et al., 1996a). Trees affected by this disease not only produce significantly less fruit, but fruit from such trees also produce less juice and ascorbic acid (Ahlawat et al., 1996a). In severe cases, infected trees had to be abandoned because they were no longer productive (Ahlawat et al., 1996a).

Ahlawat et al. (1996a) provided a partial characterization of *Citrus yellow mosaic virus* (CYMV). The virus was transmitted by grafting and dodder to 14 citrus species and cultivars, including sweet orange, pummelo, Rangpur lime, Volkamer lemon and sour orange, but was not transmitted to Mexican lime. It was also mechanically transmitted to *Citrus decumana*, a native of India, Satgudi sweet orange and pummelo (Ahlawat et al., 1996b). The authors were unable to transmit CYMV using either aphids or mealybugs (Ahlawat et al., 1996a, b).

CYMV is a proposed member of the newly established family *Caulimoviridae*, genus *Badnavirus* (Pringle, 1998), based on its serological relationship with other badnaviruses and PCR amplification using degenerate primers designed from conserved badnavirus sequences (Ahlawat et al., 1996b).

Badnaviruses affect a wide range of tropical plant species, including economically important crops, such as banana, citrus, cacao, sugarcane, rice and yam. They are characterized by non-enveloped bacilliform particles (30 x 120–150 nm), which contain a circular dsDNA genome of 71–76 kb (Lockhart & Olszewsksi, 1999). The genomes of five badnaviruses, including *Commelina yellow mottle virus* (ComYMV) (Medberry et al., 1990), the type species of the genus, *Sugarcane bacilliform virus* (SCBV) (Bouhida et al., 1993), *Cacao swollen shoot virus* (CSSV) (Hagen et al., 1993), *Banana streak virus* (BSV) (Harper & Hull, 1998) and *Dioscorea bacilliform virus* (DBV) (Briddon et al., 1999) have been cloned and sequenced, but only the cloned genomes of ComYMV, SCBV and CSSV were shown to be infectious (Bouhida et al., 1993; Jacquot et al., 1999; Medberry et al., 1990). The sequence of *Rice tungro bacilliform virus* (RTBV), a former member of the genus *Badnavirus* but now the type species of the new ‘RTBV-like viruses’ genus within the family *Caulimoviridae* (Pringle, 1998), has also been published (Hay et al., 1991; Qu et al., 1991). The genomes of all badnaviruses have similar genome organization and contain three open reading frames (ORFs), ORFs 1–3, capable of
encoding proteins with a molecule mass greater than 10 kDa (except CSSV, which has two extra ORFs, designated X and Y, overlapping ORF 3). All of the ORFs are present on the plus-strand of the genome. ORFs 1 and 2 encode putative proteins of unknown function, although the C terminus of the ORF 2 product of CSSV has nucleic acid-binding activities for both dsDNA and ssRNA (Jacquot et al., 1996). ORF 3 encodes a polyprotein that is cleaved post-translationally by the viral aspartic protease to produce a virus movement protein, a coat protein, the aspartic protease itself and a replicase comprising reverse transcriptase and ribonuclease H (RNase H) (Bouhida et al., 1993; Hohn & Futterer, 1997).

Like caulimoviruses, RTBV, soybean chlorotic mottle-like viruses, *Cassava vein mosaic virus* (CsVMV) and *Petunia vein clearing virus* (PVCV), badnaviruses are considered to be plant pararetroviruses. The replication model proposed for them involves an initial synthesis of a larger than unit length terminally redundant transcript of the viral genomic DNA by the host DNA-dependent RNA polymerase (Hohn et al., 1985; Medberry et al., 1990; Pfeiffer & Hohn, 1983; Qu et al., 1991). This transcript serves both as a polycistronic mRNA for translation to produce viral proteins and as a template for reverse transcription to replicate the viral genome by the virus-encoded replicase (Gowda et al., 1989; Joshi, 1987). The host cytosolic initiator methionine tRNA (tRNA\(^{Met}\)) is thought to serve as the primer for the reverse transcriptase in the synthesis of minus-strand DNA (Hohn et al., 1985; Pfeiffer & Hohn, 1983). RNase H digests the RNA in the RNA/DNA hybrid, leaving one or more specific RNA fragments, possibly a purine-rich region, as might be the case in ComYMV leaving one or more specific RNA fragments, possibly a

---

**Methods**

**Source and maintenance of the virus isolate.** The CYMV isolate used in this study was kindly provided by Manjunath Keremane (University of Florida, FL, USA) as budwood from a naturally infected sweet orange tree from Andhra Pradesh State, India. The virus was multiplied in ‘Madame Vinous’ sweet orange seedlings by graft transmission and maintained in a greenhouse at 16–37 °C, with supplementary lighting to give a year-round photoperiod of 16 h. Plants were maintained in continuous quarantine under permits from the USDA–APHIS and from the Maryland Department of Agriculture.

**Virus purification and nucleic acid extraction.** CYMV was partially purified from infected sweet orange leaves by the method of B. Lockhart (University of Minnesota, MN, USA; personal communication) with modifications. A sample of 100 g of symptomatic leaves were ground in liquid nitrogen and extracted with 2 vol. of 500 mM potassium phosphate buffer, pH 7.5, containing 1 M urea, 5% polyvinylpyrrolidone (Mr 40000) and 0.5% Na\(_2\)SO\(_4\). The extract was filtered through eight layers of cheesecloth and Triton X-100 was added to the filtrate to a final concentration of 2% (v/v) and mixed thoroughly. The mixture was centrifuged at 12000 g for 15 min. The supernatant was then layered over a 5 ml cushion of 30% sucrose in 100 mM potassium phosphate, pH 7.2, and centrifuged for 70 min at 148000 g in a Beckman 50Ti rotor. The resulting pellet was resuspended in 100 mM potassium phosphate, pH 7.2, centrifuged twice in a microfuge to remove particulate matter and passed through a second 5 ml cushion of 30% sucrose at 148000 g for 70 min. The pellet resuspended in 100 mM potassium phosphate, pH 7.2, constituted the partially purified virus preparation. To isolate virion DNA, the partially purified virus sample was first treated with RNase A (100 µg/ml) and DNase (30 U/ml) in 10 mM phosphate buffer, pH 7.2, for 30 min at 37 °C. The reaction was stopped by the addition of EDTA to a final concentration of 1 mM. Virions were disrupted by adding Proteinase K to a final concentration of 0.5 mg/ml with 0.5% SDS and incubating for 30 min at 37 °C. DNA was purified by phenol–chloroform extraction followed by ethanol precipitation.

**Cloning, subcloning and DNA sequencing.** We determined by endonuclease digestion and agarose gel electrophoresis that there was apparently only one *Kpn*I site within the CYMV genome. Thus, we ligated CYMV DNA, digested with *Kpn*I, into the *Kpn*I site of pBluescript KS\(^{\ast}\) (pBS) (Stratagene) to construct the genomic clones. Genomic subclones were produced by cloning the desired restriction fragments taken directly from the full-length CYMV clone into pBS. The subclones were then sequenced at the Interdisciplinary Center for Biotechnology Research (University of Florida, FL, USA).

**Analysis of sequence data.** The web-based software available at eBioinformatics (http://www.bionavigator.com) was used to analyse the DNA sequence, as well as to deduce and analyse amino acid sequence data. Sequence similarities between CYMV and other badnaviruses were evaluated using the Gap program in the GCG package (Genetics Computer Group) (Devereux et al., 1984). Database searches were carried out using FastA (Pearson & Lipman, 1988), tBlastn and BlastP (Altschul et al., 1997). Phylogenetic relationships among badnaviruses were estimated using the Phylogeny Analysis program, a macro created by eBioinformatics. Nucleotide or amino acid sequences were first aligned using CLUSTAL W (fast) (Thompson et al., 1994). The phylogenies were then estimated using the neighbour-joining method from a distance matrix generated by DNAdist or Protdist and unrooted phylogenetic trees were plotted by DrawTree using the PHYLIP software package (Felsenstein, 1989). To estimate the confidence placed in the phylogenetic trees, bootstrap sets of resampled alignments (1000 replicates) were generated by Seqboot and a consensus tree was obtained by Consense in the PHYLIP package of programs.

**Construction of an infectious clone.** A 4/4 genome length copy of CYMV was cloned into the binary vector pHBl101.2 (Clontech). The *BamH*I/*Kpn*I fragment of *pCYMV*, the full-length CYMV clone, was first cloned into pUC18 to create pBK. The *Sal*I/*Kpn*I fragment of pBK and the *Kpn*I/*Xba*I fragment of *pCYMV* were then ligated together with...
Infectious clone of CYMV

Infectious clone of CYMV

Fig. 1. Organization of the circular CYMV genome. The outer double circle represents the CYMV dsDNA genome. Arrows indicate the deduced ORFs 1–6. The KpnI and EcoRI restriction sites used to construct the genomic clones are shown. Numbers refer to their positions within the genome.

pBI101.2 to create pBICYMV, which was then transformed into Agrobacterium tumefaciens C58C1 by the freeze–thaw method (An et al., 1988). A. tumefaciens C58C1 transformants containing pBICYMV and pBI101.2 only were selected on LB plates containing kanamycin (50 µg/ml).

Agrobacterium-assisted inoculation. To introduce pBICYMV into plants, stems of 2-year-old sweet orange seedlings were wounded by three sets of 20 stem slashes using a disposable scalpel. Ten leaves of each plant were also wounded with a needle press inoculation tool. Both the scalpel and the needle press had been dipped into a saturated culture of A. tumefaciens grown at 28 °C in LB medium for 16 h. A total of nine plants was inoculated with C58C1(pBICYMV) and eight plants with C58C1(pBI101.2).

Confirmation of CYMV infection after Agrobacterium-assisted inoculation.

(i) PCR-based detection. We used PCR to determine if the CYMV sequence was present in inoculated plants. Total DNA from 0.2 g of inoculated plant leaves was extracted using a QiAamp Tissue kit (Qiagen) and one-fifth of the DNA preparation was used as a template. The two primers used were 5′ CGCAGGCAGAGAGCAAAA and 5′ CCAGATGGCAAAACATTT, corresponding to CYMV nt 4504–4521 and 5230–5213, respectively. PCR conditions used were 1 cycle of 95 °C for 5 min, 40 cycles of 95 °C for 1 min, 52 °C for 1 min, 72 °C for 1 min and 1 cycle of 72 °C for 10 min. PCR products were analysed by electrophoresis in 1% agarose gels stained with ethidium bromide.

(ii) Detection of virus particles. Immunosorbent electron microscopy was used to detect virus particles in leaves of inoculated plants by the method of Lockhart et al. (1992), modified as follows. Carbon-coated Formvar grids (SPI) were coated for 15 min with 25 µl of SCBV antiserum (0.1 mg/ml) and then rinsed with 30 drops of 10 mM sodium phosphate, pH 7.2. The grids were then floated for 2 h on 50 µl of a partially purified virus sample prepared from 1.25 g of leaves, as described above. Finally, the grids were rinsed with 20 drops of the same buffer and incubated for 5 min in 25 µl of 2% sodium phosphotungstate, pH 7.2.

(iii) Symptoms. Inoculated plants were observed at regular intervals for yellow mosaic symptoms in the leaves. Symptoms were compared

---

**Table 1. Protein-coding regions of the CYMV genome**

<table>
<thead>
<tr>
<th>ORF</th>
<th>Starting nucleotide</th>
<th>Ending nucleotide*</th>
<th>No. of amino acids</th>
<th>Calculated molecular mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>231</td>
<td>662</td>
<td>143</td>
<td>16549</td>
</tr>
<tr>
<td>2</td>
<td>659</td>
<td>1072</td>
<td>137</td>
<td>15394</td>
</tr>
<tr>
<td>3</td>
<td>1069</td>
<td>7020</td>
<td>1983</td>
<td>225931</td>
</tr>
<tr>
<td>4</td>
<td>1973</td>
<td>2512</td>
<td>179</td>
<td>20627</td>
</tr>
<tr>
<td>5</td>
<td>3977</td>
<td>4264</td>
<td>95</td>
<td>10814</td>
</tr>
<tr>
<td>6</td>
<td>6594</td>
<td>7058</td>
<td>154</td>
<td>17491</td>
</tr>
</tbody>
</table>

* Last base of the stop codon.
to those occurring on plants infected by graft-inoculation in the same greenhouse as well as to healthy plants and plants inoculated with the binary vector only.

Results

Cloning of the CYMV genome

To clone the full-length genome of CYMV, we first determined that the restriction endonuclease KpnI had a unique site within the genome (Fig. 1). We then digested the viral DNA to completion with KpnI and cloned it into pBS. One clone, pCYMV, which contained the entire CYMV genome was chosen for further study. To exclude the possibility that two closely spaced KpnI sites existed within the genome, a 1–4 kb EcoRI fragment containing the KpnI site (Fig. 1) was cloned separately and sequenced to verify that only a single KpnI restriction site existed within the CYMV genome.

Nucleotide sequence of the CYMV genome

The complete nucleotide sequence of both strands of the insert in pCYMV was determined and submitted to GenBank (accession no. AF347695). The genome consists of 7559 bp and has a G+C content of 43.6%. A long sequence of 731 bp that does not code for a protein exists within the CYMV genome. We identified a region in this intergenic sequence with 16 of 18 nucleotides complementary to the consensus sequence of plant tRNA\textsuperscript{met}. This region is also highly homologous to a sequence that is conserved among all of the badnaviruses sequenced so far (Fig. 2). As with other published Caulimoviridae sequences, the numbering of the CYMV sequence starts at the first nucleotide of this putative tRNA\textsuperscript{met}-binding site. A potential TATA box (Boeke & Corces, 1989) at nt 7360–7366, with the sequence TATATAA, as in BSV, Soybean chlorotic mottle virus (SbCMV) (Hasegawa et al., 1989) and PVCV (Richard-Poggeler & Shepherd, 1997), was found within the intergenic region at a position expected for a transcriptional promoter for a possible larger than unit length transcript of the viral genome. Upstream of the TATA box were CACAAT and TGACG sequences at nt 7303–7308 and 7317–7321, similar to that found upstream of the TATA box in the 35S promoter of Cauliflower mosaic virus (CaMV) (Odell et al., 1985). The GT motif (GTGGA/TA/TA/T), one of the animal transcriptional enhancer cores and also found upstream of the TATA box in the CaMV 35S promoter, was not present upstream of the putative CYMV promoter. This motif was also absent in the SbCMV promoter (Hasegawa et al., 1989). Downstream of the putative CYMV promoter was a possible polyadenylation signal (AATAAA) (Boeke & Corces, 1989; Joshi, 1987) at nt 7478–7483, similar to that seen in ComYMV, BSV (Nigerian isolate) (Geering et al., 2000) and PVCV.

Analysis of coding regions

The CYMV sequence contains six putative ORFs, each capable of encoding proteins with a molecular mass of greater
than 10 kDa and all of which are present on the plus-strand of the genome (Fig. 1 and Table 1). An ATGA motif at nt 659–662 is used for the stop codon of ORF 1 (TGA) and the start codon of ORF 2 (ATG), with a –1 nucleotide shift. The same sequence motif at nt 1069–1072 was also used for the termination codon of ORF 2 and the start codon of ORF 3. ORF 4 is located within ORF 3 near its N terminus. ORF 5 is located approximately in the middle of ORF 3 and ORF 6 overlaps the C-terminal region of ORF 3. Sequence comparisons between CYMV and other badnaviruses and RTBV at both the nucleotide and amino acid levels are summarized in Table 2. In general, CYMV has limited, but recognizable, similarities to other badnaviruses. The putative proteins of ORFs 1–3 have homology to those of other badnaviruses and ranges from 34 to 67%. ORF 3 of CYMV is more similar to that of other badnaviruses than ORFs 1 and 2. It contains domains, identified previously by Bouhida et al. (1993), that are conserved, although weakly, among badnaviruses and RTBV. These include a movement protein domain (Fig. 3), a cysteine-rich, zinc finger-like RNA-binding domain (CXC\textsubscript{a}CX\textsubscript{b}HX\textsubscript{c}C,\textsubscript{d})
corresponding to aa 927–942 in CYMV) and a second cysteine-rich region (CX_CX1CX2CX3CX4C, corresponding to aa 1045–1072). ORF 3 also contains domains homologous to those of the aspartic protease, reverse transcriptase and RNase H, which are highly conserved among all plant pararetroviruses (Fig. 3). No proteins were found to be similar to the putative products of ORFs 4 and 5 by either BlastP or tBlastn searches.

The predicted product of ORF 6 of CYMV is homologous with the predicted product of ORF Y of CSSV, with 36% similarity and 30% identity. We did not detect any α-helical coiled-coil structure in the products of ORFs 4–6 of CYMV by the CoilScan program in the GCG package. Such a structure allows protein subunit assembly and is conserved in the ORF III and II products of caulimoviruses and badnaviruses, respectively (Leclerc et al., 1998). Using the Gap program in the GCG package, we directly compared the sequences of ORFs 4–6 of CYMV to the ORF sequences with unknown functions in other members of the Caulimoviridae family, including ORFs 1b, III, VII and VIII of SbCMV (Hasegawa et al., 1989) and ORFs II, IV and V of CsVMV (Calvert et al., 1995). We found that only the ORF V product of CYMV has limited similarity to the products of ORFs IV and V of CsVMV (23 and 21% identities, respectively), as well as to the ORF VIII product of SbCMV (21% identity).

**Phylogenetic relationships among plant badnaviruses**

A phylogenetic analysis of the relationship of CYMV with other badnaviruses and RTBV is presented in Fig. 4. The phylogenetic tree derived from the complete nucleotide sequence showed that CYMV is more closely related to other badnaviruses than to RTBV. In addition, among the badnaviruses sequenced to date, CYMV has the closest relationship with CSSV (100% of bootstrap replicates), the only other badnavirus that contains more than three ORFs. These two viruses are more similar to DBV than to SCBV (Fig. 4a). A similar tree was obtained when the deduced amino acid sequences of ORF 3 were used to do the analysis (Fig. 4b).

**The cloned CYMV genome is infectious in sweet orange**

To determine whether the cloned CYMV genome is infectious in sweet orange, we constructed pBICYMV, a clone containing 1–4 copies of the CYMV genome (Fig. 5), and introduced it into sweet orange by *A. tumefaciens*-assisted...
inoculation. At 5 months after inoculation, we detected a PCR product of 726 bp in all nine plants inoculated with A. tumefaciens carrying pBICYMV (Fig. 6, lane 3). We confirmed the PCR product was the predicted CYMV product by sequencing. Bacilliform virus particles (30 × 130 nm) were observed in leaves distal from the inoculation site of the inoculated plants by immunosorbent electron microscopy (Fig. 6b). These particles were the same size those reported previously for CYMV (Ahlawat et al., 1996b). We also observed yellow mosaic symptoms in three of nine CYMV-positive plants (Fig. 6c, centre). No CYMV PCR product, viroids or disease symptoms were observed in the eight control plants inoculated with A. tumefaciens containing the binary vector pBI101.2 alone.

Discussion

We cloned and sequenced the genome of CYMV and demonstrated that the cloned genome was infectious. We conclude that CYMV is indeed a member of the Badnavirus genus based on the following features of its genome. (i) It contains a putative tRNA\textsuperscript{met}-binding site that is highly conserved among all members of the Caulimoviridae family sequenced to date. This site was found to be indispensable for virus viability (Dixon & Hohn, 1984). It is believed that a tRNA binds to the tRNA-binding site on the virus transcript and primes the synthesis of the minus-strand DNA in all retroviruses and most pararetroviruses (Boeke & Corces, 1989; Hull & Covey, 1983; Pfeiffer & Hohn, 1983). The existence of this tRNA\textsuperscript{met}-binding site suggests a similar function in CYMV. (ii) The size and physical arrangement of CYMV ORFs 1–3 are similar to those of other badnaviruses as well as RTBV. In the CYMV genome, ORF 1 overlaps ORF 2 using the motif ATGA, as observed previously in CSSV, BSV and RTBV, with ATG acting as the start codon of the downstream ORF and TGA acting as the stop codon of the upstream ORF. The same motif was also used to overlap ORFs 2 and 3 in CYMV, similar to that seen in SCBV and RTBV. (iii) The predicted product of ORF 1 of CYMV is similar only to that of other badnaviruses and RTBV, although no highly conserved domains were identified among them. The same holds true for ORF 2. The products of ORFs 1 and 2 of ComYMV were shown to be virion-associated (Cheng et al., 1996). ORF 1 has been proposed to be involved in mealybug transmission (Joshi, 1987) and ORF 2 has been proposed to be a nucleic acid-binding protein (Jacquot et al., 1996). The exact functions of these proteins, however, are still obscure. (iv) Like other badnaviruses and RTBV, the predicted ORF 3 product of CYMV is very large (224 kDa) and is possibly a polyprotein encoding for functions involved in virus movement, assembly and replication. It contains conserved domains typical of a viral coat protein, as characterized by the cysteine-rich, zinc finger-like RNA-binding region CXX\textsubscript{4}CX\textsubscript{2}HX\textsubscript{2}C (Berg, 1990; Bouhida et al., 1993), an aspartic protease, reverse transcriptase and RNase H (Hohn & Futterer, 1997). Such an arrangement of domains is analogous to the ‘gag–pol’ ORFs in most of the retro- and pararetroviruses (Rothnie et al., 1994). Laco & Beachy (1994) demonstrated that insect cells expressing the RTBV genome containing the putative aspartic protease and replicate domains synthesized and processed the polyprotein, with the resulting accumulation of polypeptides possessing reverse transcriptase, DNA polymerase and RNase H activities. The product of ORF 3 of CYMV also shares homology, although weak, with a virus movement protein domain (Bouhida et al., 1993; Hagen et al., 1993). The second cysteine-rich domain CX\textsubscript{4}CX\textsubscript{4}CX\textsubscript{4}CX\textsubscript{2}C is also present in the product of ORF 3 of CYMV. Kano et al. (1992) first identified this region to be conserved among other badnaviruses and RTBV, but not among other retroelements. This region may be uniquely shared by badnaviruses and RTBV, although little is known about its function.

CYMV differs from other badnaviruses in that it contains three additional, relatively small ORFs (4–6 with predicted translation products of 21, 11 and 15 kDa, respectively). ORF 4 of CYMV and ORF X of CSSV are similarly located within their respective ORF 3 and both encode very basic proteins (pl = 11.2 and 11.1, respectively), but no sequence homology was detected between the predicted translation products. ORF 5 of CYMV overlaps the middle of ORF 3 and encodes a highly basic protein with a pl of 11.1. It has no counterpart
ORF in either other badnaviruses or RTBV and has no sequence similarity to any proteins by BlastP and tBlastn searches. Direct sequence comparison using the Gap program, however, revealed that it has limited similarity to ORFs 4 and 5 of CsVMV and ORF 8 of SbCMV. ORF 6 of CYMV overlaps the C-terminal region of ORF 3, as does ORF Y of CSSV, and they share 36% amino acid similarity. It remains unclear, without the construction of mutants and a test of their effects, whether the three additional ORFs are fortuitous or correspond to real viral genes that maybe unique to CYMV. This is especially the case with ORF 4. Without any experimental evidence, it is also hard to predict how these three additional ORFs are translated, if they ever are. Leaky scanning by the ribosome, splicing of the primary transcript, as found for the expression of RTBV ORFs (Futterer et al., 1994, 1997), or the use of subgenomic RNAs are among the possible mechanisms for the expression of ORFs 4–6 of CYMV.

Previously, Ahlawat et al. (1996a) reported that CYMV was related serologically to CSSV, BSV, ComYMV, DBV and SCBV, but had the closest serological relationship to SCBV. Our phylogenetic analysis based on complete nucleotide sequences, however, revealed that CYMV is most closely related to CSSV and DBV rather than to SCBV. Our result is supported by a separate analysis using the deduced amino acid sequences from ORF 3, which is the most conserved ORF among all of the badnaviruses and which includes the viral coat protein sequence. A recent phylogenetic study by Geering et al. (2000) based on the conserved region of the RNase H domain also showed that CSSV is more similar to DBV than to SCBV, indirectly supporting our conclusion. We do not know why there is a discrepancy between the phylogenetic relationship determined by serological methods and the one based both on the total genome sequence and on the amino acid sequence of ORF 3. It is possible that the viral coat protein or some other regions, such as an epitope of the coat protein, evolve at a different rate (Hearne et al., 1990). The coat protein sequence is not very well conserved among badnaviruses, except for the cysteine-rich, zinc finger-like RNA-binding region with the consensus sequence CX_CX_CX_CHX_C. It is also unclear where the coat protein sequence begins and ends in ORF 3, so it is difficult to determine whether the phylogenetic relationship based on the coat protein sequence agrees with the serological results. On the other hand, the serological study was carried out non-reciprocally and was based solely on immunosorbert electron microscopy. Because this method is more qualitative than quantitative in nature, additional quantitative serological tests are needed to determine precisely the serological relationship between CYMV and other badnaviruses. We also found that CYMV is more similar to other badnaviruses than to RTBV, reflecting the fact that badnaviruses and RTBV are also biologically distinct (Harper & Hull, 1998; Hibino, 1983; Hull, 1996).

Agrobacterium-mediated inoculation has been used successfully to deliver cloned viral genomes into mono- and dicotyledons and to induce virus infection (Grimsley & Bisaro, 1987; Grimsley et al., 1986, 1987). For plant pararetroviruses, clones used for such inoculations require a greater than genome length insert that gives rise to a larger than unit length terminally redundant transcript. This is important for virus infection (Boughida et al., 1993; Dasgupta et al., 1991; Jacquot et al., 1999; Medberry et al., 1990). Using a similar approach, we demonstrated that the cloned CYMV genome was infectious in sweet orange. However, we noticed that symptom development is somewhat slower in our plants, as it took 5 months or longer for the symptoms to develop. When CYMV is transmitted by grafting, disease symptoms become apparent 4 months after transmission. This may reflect high virus titres in infected buds used for graft-inoculation. When CYMV was transmitted by mealybugs, approximately 6 months were required for disease symptoms to develop (J. Hartung, unpublished data). Disease symptoms also developed slowly when ComYMV was inoculated to its host plant Commelina diffusa via A. tumefaciens (Medberry et al., 1990). Our successful delivery of an infectious CYMV clone to sweet orange will facilitate further molecular studies of this virus, by making possible the introduction of mutated viral genes into host plants to test the effect of mutations in planta.

The sequence information obtained from this study will be used to design primers specific to CYMV for quarantine purposes. It may also be useful to develop this dsDNA virus into a practical tool for plant transformation and to control the expression of foreign genes in citrus and/or other dicotyledonous woody plants.

We thank B. Lockhart, University of Minnesota, for his generous gift of the SCBV antiserum and for helpful discussions and M. Roose, University of California-Riverside, for kindly providing Agrobacterium tumefaciens C58C1. We also thank Peggy Dieneldt and Cristina Gouin for technical assistance.

References


Received 17 April 2001; Accepted 29 June 2001