Production of infectious RNA transcripts from full-length cDNA clones representing two subgroups of peanut stunt virus strains: mapping satellite RNA support to RNA1

Chung-Chi Hu, Margaret Sanger and Said A. Ghabrial

Department of Plant Pathology, University of Kentucky, Lexington, Kentucky 40546, USA

Full-length cDNA clones from which infectious transcripts could be generated were constructed from the genomic RNAs of two distinct strains of peanut stunt cucumovirus (PSV), PSV-ER and PSV-W. PSV-ER, a subgroup I strain, is known to support efficient replication of satellite RNA (satRNA) in infected plants, whereas PSV-W, a subgroup II strain, does not support satRNA replication. Although artificial reassortants (pseudorecombinants) of all possible combinations of infectious transcripts representing RNA1, RNA2 and RNA3 were infectious, only those having RNA1 from PSV-ER supported the replication of satRNA. These results demonstrate conclusively that support of PSV satRNA replication maps to RNA1. Comparisons of secondary structure predictions of the C-terminal helicase-like domain of the 1a proteins of four PSV strains belonging to two subgroups did not reveal any obvious differences between strains that differ in satRNA support. The complete nucleotide sequence of RNA1 from strains PSV-ER and PSV-W were determined and found to be 79% identical. Sequence comparison analysis of RNA1 sequences of cucumoviruses confirmed the placement of the PSV strains into two distinct subgroups.

Introduction

Peanut stunt virus (PSV) is a member of the genus *Cucumovirus* in the family *Bromoviridae* (Rybicki, 1995). Other members of the genus are tomato aspermy virus (TAV) and the type member, cucumber mosaic virus (CMV). Cucumoviruses have tripartite genomes of positive-strand RNAs, designated RNA1, 2 and 3 in order of decreasing size, that are packaged in isometric particles of about 28 nm in diameter. RNAs 1 and 2 encode the 1a and 2a proteins, respectively, which are required, along with host components, for replication. RNA3 is dicistronic and encodes a movement protein and a coat protein, which is expressed from subgenomic RNA4 (for a review, see Palukaitis *et al*., 1992). A small overlapping gene (2b), encoded by RNA2, was recently discovered in all cucumoviruses sequenced to date, and is probably expressed through the subgenomic mRNA RNA4A (Ding *et al*., 1994). Naturally occurring virions of CMV and PSV, but not TAV, may also package a fifth RNA, designated satellite RNA (satRNA), along with their genomic and subgenomic RNAs (Naidu *et al*., 1991, 1992; Roossinck *et al*., 1992). CMV and TAV, but not PSV, support the replication of CMV satRNAs. Only PSV supports the replication of PSV satRNAs (Kaper *et al*., 1978; Naidu *et al*., 1995).

The complete nucleotide (nt) sequences of the genomic RNAs of several strains of CMV have been reported, and CMV strains have been classified into two subgroups, I and II, based on percentage nt sequence identity (Palukaitis *et al*., 1992). Full-length cDNA clones, from which infectious transcripts of the genomic RNAs can be generated, are available for several CMV strains representing the two subgroups (Hayes & Buck, 1990; Rizzo & Palukaitis, 1990; Zhang *et al*., 1994; Tousignant *et al*., 1996). The complete nt sequences and the production of infectious RNAs from cDNA clones of the genomic RNAs of the V strain of TAV (Bernal *et al*., 1991; Moriones *et al*., 1991; Moreno *et al*., 1997) have also been reported. Strains of PSV have been classified into two major subgroups (I and II) based on serology and percentage nt sequence identity (Hu *et al*., 1997). Furthermore, evidence has been reported for the occurrence of natural reassortants

Author for correspondence: Said Ghabrial.
Fax +1 606 323 1961. e-mail saghab00@ukcc.uky.edu

The nucleotide sequence data reported in this paper have been deposited in GenBank under accession numbers U33145 (PSV-W RNA1) and U15728 (PSV-ER RNA1).

0001-5447 © 1998 SGM

between the two subgroups of PSV strains (Naidu *et al.*, 1995; Hu & Ghabrial, 1998). We have recently completed sequencing the genomes of two strains of PSV belonging to subgroups I and II (this study; Hu *et al.*, 1997; C.-C. Hu & S. A. Ghabrial, unpublished data). In addition, the complete nt sequence of a Japanese strain (PSV-J; a subgroup I strain) has previously been published (Karasawa *et al.*, 1991, 1992).

In this paper, we describe the production of infectious RNA transcripts from full-length cDNA clones of the genomic RNAs from two PSV strains, PSV-ER and PSV-W, representative of subgroup I and II strains, respectively. The infectious RNA transcripts were used to construct all possible reassortant viruses (pseudorecombinants), and evidence is presented that the ability to replicate satRNA maps to RNA1. The complete nt sequences of RNA1 from strains of PSV-ER and PSV-W, as well as comparisons of RNA1 sequences among cucumoviruses, are also reported.

Methods

Virus strains and purification. The sources of CMV strain Fny and PSV strains ER, W, J and B were as described previously (Naidu *et al.*, 1995; Hu *et al.*, 1997). All PSV strains were maintained in cowpea (*Vigna unguiculata* 'Blackeye'), and increased in Burley tobacco (*Nicotiana tabacum* cv. Ky 14) or in cowpea. Purification of PSV virions was as described by Ghabrial *et al.* (1977). CMV was increased in Burley tobacco and purified according to the procedure of Lot *et al.* (1972). Viruses were stored as purified RNAs in 75% ethanol at −80 °C.

Hybridization analysis of virion RNA. Viral RNAs were extracted from purified virions as described by Naidu *et al.* (1995), denatured with 6% formaldehyde, and separated by electrophoresis in 1·5% agarose gels made in TBE (89 mM Tris–borate buffer pH 8·3 containing 25 mM disodium EDTA). The RNAs were then transferred and fixed onto Hybond-N membrane (Amersham) according to the manufacturer's instructions. Procedures for preparation of ^{32}P-labelled nick-translated cloned probes, and conditions for pre-hybridization, hybridization, membrane washings and autoradiography, were as described by Hu *et al.* (1997).

Cloning and sequencing of cDNA to RNA1. PSV RNA1 was purified from low-melting-point agarose following electrophoresis of unfractionated virion RNAs. cDNA libraries representing RNA1 from PSV-ER and PSV-W were synthesized using components of the Amersham cDNA synthesis kit. The cloning and sequencing strategies used were as previously described for PSV RNA3 (Hu *et al.*, 1997). The 5′-terminal sequences were determined by the dideoxy-chain termination method using a reverse transcriptase RNA sequencing kit method (USB). Internal primers complementary to the following nt positions were used to prime the reverse transcription sequencing of the RNA1 templates: 131–147 of PSV-ER, and 128–144 of PSV-W. The sequences were assembled using the GEL program in the IG-Suite (IntelliGenetics) and analysed using both the IG-Suite and the University of Wisconsin GCG package (UWGCG; Devereux *et al.*, 1984). The GAP, PILEUP and PRETTY programs in the UWGCG package (version 9.0) were used to align and compare nt and deduced amino acid (aa) sequence identities (and aa sequence similarities) among RNA1 sequences of several representative cucumoviruses. The lengths of the coding and noncoding regions in the cucumovirus RNA1 and GenBank accession numbers used for sequence comparisons are listed in Table 1.

Construction of full-length cDNA clones from which infectious transcripts can be generated. An RT–PCR approach was used to construct full-length cDNA clones of PSV-ER and PSV-W genomic RNAs. Three sense primers were designed based on the 5′-terminal nt sequences of the genomic RNAs of PSV-ER and PSV-W as follows.

P1f: 5′ TAATACGACTCACTATA**G**GTTTTATCACGAGCGTA-CGG 3′ (used for amplification of cDNAs to PSV-ER RNA 1 and PSV-W RNAs 1 and 2), containing the 5′-terminal 20 nt of PSV-ER RNA1 plus an extra G residue (bold) and a T7 RNA polymerase promoter (underlined).

P2f: 5′ TAATACGACTCACTATA**G**GTTTTACCAACCAGGAA-TCT 3′ (used with cDNA to RNA3 from either strain) containing PSV sequence identical to the 5′-terminal 20 nt of RNA3, with an extra G residue (bold) and T7 promoter (underlined).

Table 1. Lengths of the coding and noncoding regions in the RNA1s of the cucumoviruses included in this study

UTR, untranslated region; 1a, gene encoding the 1a protein.

Virus strain	Accession no.	Sub-group	5′ UTR	3′ UTR	Full-length	1a gene	Reference
PSV-ER	U15728	PSV-I	88	248	3357	3021	This study
PSV-J	D11126	PSV-I	90	247	3355	3018	Karasawa *et al.* (1992)
PSV-W	U33145	PSV-II	85	261	3355	3009	This study
CMV-Fny	D00356	CMV-I	94	281	3357	2982	Rizzo & Palukaitis (1989)
CMV-Ix	U20220	CMV-I	95	284	3361	2982	McGarvey *et al.* (1995)
CMV-Y	D12537	CMV-I	97	282	3361	2982	Kataoka *et al.* (1990)
CMV-Q	X02733	CMV-II	97	316	3389	2976	Rezaian *et al.* (1985)
TAV-V	D10044	Not applicable	94	334	3410	2982	Bernal *et al.* (1991)

P3f: 5′ GGGGAATTCATTTAGGTGACACTATAGTTTTATCA-AGAGC 3′ (used with cDNA to PSV-ER RNA2) containing PSV sequence identical to the 5′-terminal 14 nt of PSV-ER RNA2 with SP6 RNA polymerase promoter (underlined) and an engineered *Eco*RI recognition site (double-underlined).

Four antisense primers were designed based on the 3′-terminal nt sequences as follows.

P1r: 5′ AAGCTTGGTCTCCTATGGAAAC 3′ (used with cDNAs to PSV-ER RNAs 1 and 2), containing PSV sequence complementary to the 3′-terminal 17 nt of PSV-ER RNA1 with an engineered *Hin*dIII recognition site (double-underlined).

P2r: 5′ **GA**CTGCAGTGGTCTCCTATGG 3′ (used with cDNA to PSV-ER RNA3) containing PSV sequence complementary to the 3′-terminal 13 nt of RNA3 with an engineered *Pst*I recognition site (double-underlined) and two extra nucleotides for efficient digestion (bold).

P3r: 5′ **GA**CTGCAGTGGTCCCCATAAGG 3′ (used with cDNA to PSV-W RNA1), containing PSV sequence complementary to the 3′-terminal 14 nt of PSV-W RNA1, with an engineered *Pst*I site (double-underlined) and two extra nucleotides (bold).

P4r: 5′ **GA**CTGCAGTGGTCCCCTTTAA-GG 3′ (used with cDNAs to PSV-W RNAs 2 and 3), containing PSV sequence complementary to the 3′-terminal 15 nt of PSV-W RNAs 2 and 3, with an engineered *Pst*I site (double-underlined) and two extra nucleotides (bold).

First strand cDNA was synthesized with Superscript II reverse transcriptase (BRL) according to the manufacturer's instructions, using gel-purified individual genomic RNAs of PSV-ER or PSV-W as templates, and the reactions were primed with the appropriate reverse primers. PCR amplifications were then performed using the first-strand cDNAs as templates and the pertinent set of primers for each RNA to prime the amplification reactions for 30 cycles of 1 min at 94 °C, 1 min at 60 °C and 3·5 min at 72 °C, followed by a final extension at 72 °C for 10 min. The PCR products were then gel-purified, blunt-ended, digested with the appropriate restriction enzyme (*Hin*dIII or *Pst*I) and cloned into plasmid pUC119 previously digested with *Sma*I and *Hin*dIII or *Sma*I and *Pst*I. The recombinant plasmids containing cDNAs of PSV-ER genomic RNAs were designated pER1, pER2 and pER3. The corresponding plasmids derived from PSV-W were pW1, pW2 and pW3.

■ ***In vitro* transcription.** Capped RNA transcripts were synthesized *in vitro* from restriction enzyme-linearized plasmid cDNA clones of PSV strains ER and W genomic RNAs. Transcripts for PSV-W RNAs 1, 2 and 3, and ER RNA3, were prepared from *Pst*I-linearized plasmids pW1, pW2, pW3 and pER3 after the resultant 3′ overhangs were blunt-ended with DNA polymerase I Klenow fragment. Transcripts of PSV-ER RNAs 1 and 2 were synthesized from *Hin*dIII-linearized pERl and pER2, respectively. SatRNA transcripts were generated from *Xho*I-linearized pT7G2 (Naidu *et al.*, 1992). The 50 µl reaction mixture contained 40 mM Tris–HCl pH 7·5, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM dithiothreitol, 50 units RNasin (Promega), 0·5 mM each ATP, CTP and UTP (plus 0·1 mM GTP), 0·5 mM cap-analogue [$m^7G(5')ppp(5')G$; New England Biolabs], 1–5 µg template DNA and 50 units of T7 (for all templates except those derived from pER2), or SP6 (for the template derived from pER2) RNA polymerases (Stratagene). The satRNA transcript was synthesized as above but without cap analogue, and with 0·5 mM GTP. After incubation at 37–40 °C for 2 h, samples of 3–5 µl of the reaction mixture were analysed in a 1·0% agarose gel to examine the yield and integrity of the transcripts. The infectivity of the RNA transcripts was tested in tobacco, cowpea and bean (*Phaseolus vulgaris*).

■ ***In vitro* translation.** Purified virion RNAs or transcripts from full-length cDNA clones of PSV-ER and PSV-W genomic RNAs were translated in a rabbit reticulocyte lysate system (Promega). After 1 h of translation, the ^{35}S-labelled translation products were analysed in 12·5% SDS–polyacrylamide gels (Laemmli, 1970). The gels were fixed, dried and exposed to a Phosphor screen (Molecular Dynamics) for 5–16 h. The images were visualized by a PhosphorImager 445 SI system (Molecular Dynamics), and analysed with the ImageQuant 4.1 program.

■ **Preparation and electroporation of protoplasts.** Protoplasts were prepared from cultured tobacco suspension cells (*Nicotiana tabacum* cv. Xanthi nc) as described by Passmore *et al.* (1993). Electroporation was by procedures modified from Fromm *et al.* (1987) using 2×10^6 protoplasts suspended in 0·7 ml electroporation buffer (10 mM sodium HEPES, 120 mM KCl, 10 mM NaCl, 4 mM $CaCl_2$ and 200 mM mannitol pH 7·1) containing sheared salmon sperm DNA (428 µg/ml). Test RNAs were suspended in 20 µl water and contained approximately 1 µg of each genomic RNA transcript and approximately 0·25 µg satRNA. The protoplast suspension was combined with 20 µl of an aqueous solution containing test RNAs, and electroporated. A 'mock' inoculum of water without RNA was also tested. Electroporation was at 4 °C in 0·4 cm spaced electroporation cuvettes (Bio-Rad) using a Gene Pulser II electroporation system with Pulse Controller Plus and Capacitance Extender Plus modules set at 200 V and 700 µF. Immediately following electroporation, the cells were incubated on ice for 5–15 min and then diluted into 9·3 ml protoplast culture medium (Passmore *et al.*, 1993).

■ **Extraction and analysis of RNA from electroporated protoplasts.** Protoplast samples were collected from 1 ml aliquots of cultured protoplasts 3 days post-electroporation, quick-frozen and stored at −80 °C. Frozen cells were thawed by suspension in 500 µl 50 mM Tris–HCl pH 7·4, containing 1 mM EDTA, 10 mg/ml SDS and 100 µg/ml proteinase K. The mixture was incubated at 37 °C for 20 min and extracted once with 500 µl phenol and 500 µl chloroform. The aqueous phase was separated by centrifugation, and the nucleic acids were precipitated by addition of 30 µl 4 M sodium acetate pH 5·0, and 1 ml of 95% ethanol. The precipitated nucleic acids were collected by centrifugation, dried and suspended in 50 µl diethylpyrocarbonate-treated water. Aliquots of the nucleic acid solutions were denatured by mixing with 3 vols of denaturation solution (60% dimethyl sulfoxide, 9·5% glyoxal, 1·67 mM disodium EDTA, 33 mM HEPES pH 7·0, and containing 50 µg/ml ethidium bromide). The mixtures were incubated at 55 °C for 1 h, and samples were analysed by electrophoresis in 0·8% agarose gels in 20 mM sodium HEPES–1 mM disodium EDTA pH 7·0, at 9 V/cm. The gels were blotted onto Hybond-N^+ nylon membrane (Amersham) with 50 mM NaOH for 3–4 h, and rinsed in 6× SSC (1× SSC is 0·15 M NaCl, 15 mM citric acid pH 7·8). A hybridization probe specific to PSV satRNA was synthesized using plasmid pT7G2 (Naidu *et al.*, 1992) as template with the nick-translation system of Gibco/BRL (Life Technologies). Probes specific to PSV strains ER and W were similarly synthesized, but using plasmids pER3-3′ and pW3Δ1, respectively. Plasmid pER3-3′ contains the 3′-terminal 279 nt of PSV-ER RNA3, and plasmid pW3Δ1 was derived from pW3 by deletion of a 1586 bp fragment between the *Sal*I site at nt position 241 and the *Hpa*I site at nt position 1909 of the RNA3 cDNA insert. The membranes were hybridized at 50 °C according to the methods of Church & Gilbert (1984), and with post-hybridization washes according to Amasino (1986).

■ **Computer analysis of 1a protein secondary structure.** The computer programs PeptideStructure and PlotStructure (Jameson & Wolf, 1988) of the UWGCG package were used to predict the secondary structures of the 1a proteins of PSV-ER, PSV-J, PSV-W and PSV-B. The Garnier–Osguthorpe–Robson (GOR) algorithm was used (Garnier *et al.*, 1978).

Results

Analysis of RNA transcripts derived from PSV-ER and PSV-W full-length cDNA clones

Capped RNA transcripts synthesized *in vitro* by T7 or SP6 RNA polymerases from full-length cDNA clones of each of the genomic RNAs of PSV-ER and PSV-W were of the expected size when examined by agarose gel electrophoresis. Translation of the cDNA-derived RNA transcripts in a wheat germ cell-free system resulted in products indistinguishable in size from those produced with the corresponding virion RNAs (data not shown), thus indicating that the ORFs were complete. When transcripts representing PSV-ER genomic RNAs (500 µg/ml) were inoculated onto five tobacco seedlings, all plants showed chlorotic lesions on inoculated leaves approximately 6 days post-inoculation. Systemic symptoms were observed on two of the five plants. Inocula prepared from three plants from areas in the inoculated leaves containing chlorotic lesions were highly infectious: all 30 plants tested per inoculum (10 plants each of tobacco, cowpea and bean) became infected and showed typical systemic symptoms within 7 days of inoculation. Similar results were obtained with transcripts from strain PSV-W.

PSV SatRNA support maps to RNA1

The eight different combinations of *in vitro* synthesized RNA transcripts representing PSV-ER and PSV-W (parental sets and all possible reassortants) were tested for their ability to replicate/accumulate in tobacco protoplasts and to support the replication of satRNA. All eight combinations were infectious and the predicted RNA species accumulated in tobacco protoplasts, yielding the expected hybridization pattern for each RNA combination (Fig. 1*a*, *b*). Efficient replication and accumulation of satRNA occurred only with transcript sets containing PSV-ER RNA1, as indicated by the strong hybridization observed in RNA blots probed with a satRNA-specific probe (Fig. 1*c*). No or negligible amounts of satRNA-specific hybridization signals were detected with extracts from protoplast electroporated with any transcript sets containing PSV-W RNA1 (Fig. 1*c*). However, the origin of the trace amount of hybridization detected with some transcript sets (see for example lane [W1, ER2, ER3], Fig. 1*c*) is unknown. It is unlikely that this trace amount of hybridization is due to residual inoculum because it was not detected in control protoplasts electroporated with satRNA alone or when representative samples of selected protoplast treatments were processed immediately following electroporation. It remains to be determined whether such reassortants might support very low levels of satRNA replication in protoplasts.

To determine whether satRNA accumulation may be detected *in planta* in cases where satRNA was not evident in protoplasts (transcript sets containing PSV-W RNA1), extracts from infected protoplasts (7 days post-electroporation) were rub-inoculated onto tobacco seedlings and the plants were monitored for symptom development and viral/satellite RNA content. Symptoms were observed on all inoculated leaves 4 days post-inoculation. Total RNA was extracted from inoculated leaves 18 days post-inoculation and analysed by Northern blot hybridization using PSV strain-specific probes and satRNA-specific probes. The results were the same as those obtained with protoplasts; satRNA was detected only in samples inoculated with progeny of transcript sets containing PSV-ER RNA1, but not with those containing PSV-W RNA1 (data not shown). Unlike some of the protoplast samples where trace amounts of apparently satRNA-specific hybridization signals were detected, none of the plant samples containing PSV-W RNA1 exhibited any satRNA-specific hybridization signals, even in overexposed autoradiograms.

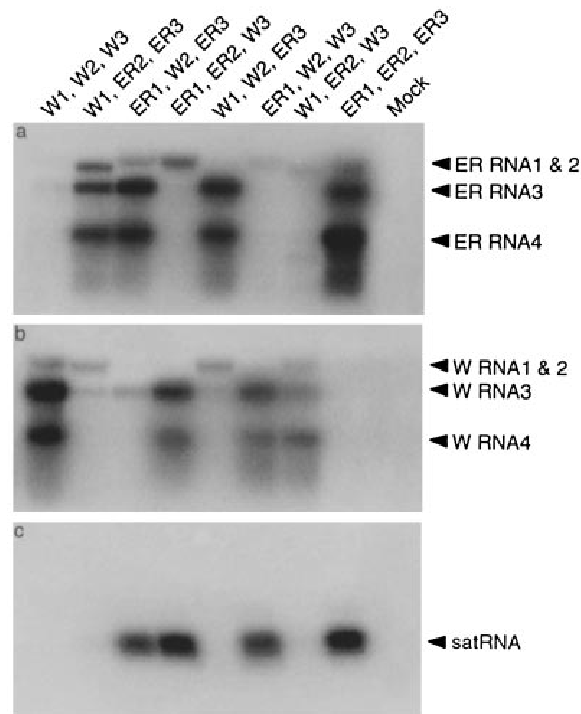

Fig. 1. Accumulation of PSV RNA and satRNA in protoplasts inoculated with satRNA and each of eight different combinations of *in vitro* synthesized RNA transcripts representing PSV-ER and PSV-W (parental sets and all possible reassortants). Tobacco protoplasts were electroporated with approximately 1 µg each of *in vitro* synthesized capped transcripts of RNA1, RNA2 and RNA3, derived from cDNA clones (plasmids pW1, pW2, pW3, pER1, pER2 and pER3) and approximately 0·25 µg *in vitro* synthesized satRNA transcribed from cDNA clone pG2 (first eight lanes in *a*–*c*), or were electroporated without transcript RNAs (lane labelled 'Mock', *a*–*c*). Nucleic acids extracted from protoplasts 3 days post-electroporation were denatured, analysed by electrophoresis through 0·8% agarose gel, and blotted to nylon membrane. The membranes were then hybridized with ^{32}P-labelled, nick-translated probes derived from pER3-3′ (specific to PSV-ER) for (*a*), pW3Δ1 (with specificity to PSV-W) for (*b*) or pG2 (specific to PSV satRNA) for (*c*). Test combinations of genomic RNAs are indicated above each lane. Migration positions of RNAs 1, 2, 3 and 4 are indicated by arrowheads on the right. Images were produced using a Molecular Dynamics PhosphorImager 445SI system and analysed with the ImageQuant 4.1 program.

```
PSV-ER   CAGACCTAAGCTAAGTTGGTAGTTATATTACTCTAAACTACCT.AGTCACTAAACAG...TTCTACTGGTGAACGGGT
PSV-W    TCCTCCACTTCTAAGTTGGCAGTTATATTATTCTAAACTGTCTGGGTCACTAAACAGAGTTCTCTCTGGTGAACGGGT
CONS     -----C----CTAAGTTGG-AGTTATATTA-TCTAAACT--CT--GTCACTAAACAG---T-TT-CTGGTGAACGGGT
                     ▼
PSV-ER   TGTCCATCCAGCTAACGGCTAAAATGGTCAGTCTTACCTCTGAGGTAAGCCGAGTTTCTTACAAGAACTCGAGGTACT
PSV-W    TGTCCATCCAGCTGACGGCTAAAATGGTCAGTCTCATCCTAAAGGTGAGCC AGGTTCTTACAAGAACCTGAGGTACT
CONS     TGTCCATCCAGCT-ACGGCTAAAATGGTCAGTCT-A-C----AGGT-AGCC-AG-TTCTTACAAG-AC--GAGGTACT

PSV-ER   TTATATCATCTACTCCTGGTGATTGTGAAGAGCCGTGTACAATGTACAC.ATCCGGCTCGGGTACCCCCTCTTCGG..
PSV-W    TTATACTATCCACCCCTGATAGTTGTGAAGAGCTGCATACAATGTATGCGGTTCACGTCGGGTACCCCCTCATAAGTG
CONS     TTATA--ATC-AC-CCTG-T--TTGTGAAGAGC-G--TACAATGTA--C--T-----TCGGGTACCCCCTC-T--G--

PSV-ER   .......AGGGTTTCCATAGGAGACCA  3357
PSV-W    CAGCTTATGAGGTCCTTATGGGGACCA  3355
CONS     --------G-G-T-C----GG-GACCA
```

Fig. 2. Comparison of the nt sequences at the 3′ UTR of PSV RNA1 from strains ER (subgroup I) and W (subgroup II). The highly conserved 40 nt domain present in the 3′ UTR of all cucumovirus RNA sequences is underlined and the position of the only known variant nt is indicated with an arrowhead. The alignment shown was derived from a multiple cucumovirus RNA1 sequence alignment generated by the PILEUP program. In the consensus sequence 'CONS', positions of identity for the two sequences are indicated by the nt symbol. The 3′ UTR for PSV-ER RNA1 starts at nt position 3110 and terminates at position 3357. The corresponding nt positions for PSV-W are 3095 and 3355.

Comparative nucleotide sequence analysis of RNA1 from PSV strains in subgroups I and II

The complete nt sequences of RNA1 of PSV-W and PSV-ER have been determined and deposited in GenBank. RNA1 sequences of PSV-ER and PSV-W are 3357 and 3355 nt in length, and each has a single large ORF of 3021 and 3009 nt, respectively (Table 1), which encodes a component of the replicase complex, the 1a protein. The N-terminal region of the PSV 1a proteins (like other members of the alphavirus-like superfamily) contains a putative methyltransferase domain with three conserved sequence motifs. These conserved motifs are strictly associated with superfamily 1 (SF1) helicases and supergroup 3 RNA-dependent RNA polymerases (Koonin & Dolja, 1993). The C-terminal region contains a putative helicase domain with seven conserved motifs, designated I, Ia and II–VI, arranged in a collinear manner (Koonin & Dolja, 1993; Kadaré & Haenni, 1997). These seven conserved motifs are shared by members of SF1 helicases (Koonin & Dolja, 1993). The nt sequences of the 3′ untranslated regions (UTR) of RNA1 from the two strains are compared in Fig. 2. They both contain the highly conserved 40 nt region (underlined) found in all cucumoviruses sequenced to date (McGarvey *et al*., 1995), with only one variant nt, A in PSV-ER and G for PSV-W (Fig. 2, arrowhead).

Hybridization analysis of RNA1 from two subgroups of PSV strains

Northern blots of total RNAs from purified virions of PSV strains ER, J (subgroup I), W and B (subgroup II) were probed with a cloned cDNA probe to PSV-W RNA1. Whereas the PSV-W RNA1 probe hybridized strongly with RNA1 from PSV strains W and B, it did not react with RNA1 from PSV strains ER and J (data not shown). These results confirmed that strain B belongs to PSV subgroup II strains (Hu *et al*., 1997).

Comparative analysis of RNA1 sequences of cucumoviruses

The RNA1 sequences of PSV strains were compared to those of several representative strains of CMV and TAV using the GAP program in the UWGCG suite of programs (Fig. 3). Whereas the percentage nt sequence identity of full-length RNA1 or 1a gene between subgroup I strains (PSV-ER or PSV-J) and PSV-W (subgroup II) is about 79%, the percentage identity between strains in the same subgroup (PSV-ER and PSV-J) is 91% (Fig. 3*a*). All three PSV strains showed similar percentage sequence identities to CMV and TAV strains, ranging from 65·3 to 68·4%, indicating that PSV strains are approximately equidistant to CMV and TAV. Similar results were observed for CMV RNA1 with strains in the same subgroup (CMV-Fny and CMV-Y, subgroup I) showing 96·8% sequence identity, whereas strains belonging to different subgroups exhibited approximately 76% identity. All CMV strains are also approximately equidistant to PSV and TAV strains.

Comparison of percentage identities of deduced aa sequences of 1a proteins of cucumoviruses (Fig. 3*b*, below diagonal) revealed similar results to those of nt sequence comparisons. However, the aa sequences of 1a proteins appeared to be highly conserved, as the percentage similarities among 1a proteins from different cucumovirus species are all above 75% (Fig. 3*b*, above diagonal). Unlike CMV, the 5′ UTR is highly conserved among PSV strains regardless of subgroup classification. On the other hand, the 3′ UTR of RNA1 (Fig. 2)

(a)

RNA1

		PSV-ER	PSV-J	PSV-W	CMV-Fny	CMV-Y	CMV-Q	TAV-V
	PSV-ER		91.0	79.2	67.2	67.0	68.5	65.8
	PSV-J	90.5		79.0	67.1	66.9	68.4	65.3
1a gene	PSV-W	79.3	79.3		67.3	66.8	67.5	65.9
	CMV-Fny	67.7	67.7	68.2		96.8	76.6	68.4
	CMV-Y	67.5	67.5	67.9	96.8		76.1	67.9
	CMV-Q	69.2	68.9	68.6	77.5	76.9		68.4
	TAV-V	66.4	66.3	66.8	68.3	68.0	68.1	

(b)

1a Protein

	PSV-ER	PSV-J	PSV-W	CMV-Fny	CMV-Y	CMV-Q	TAV-V
PSV-ER		97.5	90.9	79.7	78.5	79.1	75.8
PSV-J	96.7		89.9	79.5	78.3	79.1	76.0
PSV-W	87.1	87.5		78.9	77.1	78.9	75.1
CMV-Fny	73.8	74.0	72.7		98.1	89.5	78.9
CMV-Y	72.5	72.7	71.7	97.9		87.6	77.1
CMV-Q	73.8	74.0	73.2	85.2	83.6		79.6
TAV-V	70.1	70.3	69.1	73.5	71.7	73.2	

(c)

3′ UTR

		PSV-ER	PSV-J	PSV-W	CMV-Fny	CMV-Y	CMV-Q	TAV-V
	PSV-ER		96.7	74.4	61.0	61.0	63.1	55.6
	PSV-J	89.7		73.2	60.6	57.4	61.5	54.9
5′ UTR	PSV-W	89.4	90.4		58.2	56.2	55.2	55.5
	CMV-Fny	60.2	66.6	62.3		97.1	65.1	65.8
	CMV-Y	61.3	66.6	62.3	94.6		63.8	63.0
	CMV-Q	59.0	64.4	62.3	80.8	82.4		71.0
	TAV-V	65.9	66.6	63.5	69.8	70.2	69.1	

Fig. 3. Percentage nt sequence and deduced aa identities of RNA1 between PSV strains and other cucumoviruses. (*a*) Full-length RNA1, above diagonal; 1a gene, below diagonal. (*b*) Deduced aa sequence similarity (above diagonal) and sequence identity (below diagonal) of 1a protein. (*c*) 3′ UTR above diagonal, 5′ UTR below diagonal. Values are the identity (similarity) scores as calculated by the GAP program in the UWGCG package.

is the least conserved among PSV strains as well as among CMV strains (Fig. 3*c*, above diagonal).

Discussion

With the availability of the complete nt sequences of two PSV subgroup I strains, PSV-J and PSV-ER, and one subgroup II strain, PSV-W (this study; Karasawa *et al.*, 1991, 1992; Hu *et al.*, 1997; Hu & Ghabrial, 1998), and with the ability to generate infectious RNA transcripts from full-length cDNA clones of PSV-ER and PSV-W RNAs (this study), it is now possible to map PSV gene functions more precisely. Using infectious RNA transcripts representing the genomic RNAs of PSV-ER and PSV-W, we constructed all possible reassortants (pseudorecombinants) between the two strains and demonstrated conclusively that support of PSV satRNA replication maps to RNA1. Results of sequence comparison analysis (Fig. 3), as well as the finding that exchanging RNA1 or RNA2 between strains PSV-ER and PSV-W resulted in viable reassortants, support the classification of these two strains into two distinct subgroups instead of two different virus species (Hu *et al.*, 1997).

Two strains of CMV, CMV-Sny and CMV-Ix, have previously been reported to be defective in supporting the replication of certain CMV satRNAs (Roossinck & Palukaitis, 1991; McGarvey *et al.*, 1995; Tousignant *et al.*, 1996). Reassortment studies involving these defective strains as well as those with reassortants derived from CMV and TAV (Moriones *et al.*, 1994) indicated that CMV satRNA replication support mapped to RNA1. Recent molecular characterization of PSV strain BV-15 (Hu & Ghabrial, 1998), a natural reassortant between PSV subgroup I and II strains, also suggests that the ability to replicate PSV satRNA maps to RNA1. PSV-W differs from the two CMV strains (CMV-Sny and CMV-Ix) in that it does not replicate satRNA in any of the host species thus far tested, regardless of the satRNA variant involved (Naidu *et al.*, 1995). CMV-Sny, on the other hand, supports the replication of certain satRNAs in solanaceous plant species but not in cucurbits (Roossinck & Palukaitis, 1991), whereas CMV-Ix supports the replication of certain satRNAs (T- and GP-satRNAs) but not others (D-satRNA), in the same host species (Tousignant *et al.*, 1996). It is thus possible that more than one mechanism may be responsible for the failure of certain cucumoviruses to replicate satRNA.

Fig. 4. Alignment of the aa sequences of the C-terminal region of cucumovirus 1a proteins containing conserved motifs V and VI of the helicase-like domain typical of superfamily 1 helicases (Koonin & Dolja, 1993). The sequence motifs V and VI present in the 1a proteins from nine cucumoviruses are boxed. The position of the Cys residue downstream of motif VI (implicated in satRNA replication in CMV-Fny) is boxed. The conserved motifs were derived from an alignment generated by the PILEUP program and the consensus sequence was produced using the PRETTY program with plurality of 7. The aa positions in the C-terminal region of PSV-ER 1a protein which are included in the sequence comparisons are indicated at the top.

PSV RNA1 encodes the 1a protein which contains two conserved domains: the N-terminal domain has sequence similarity to the Sindbis virus nsP1 methyltransferase protein, which is believed to function in capping viral RNAs (Ahlquist *et al*., 1985); the C-terminal domain has similarity to viral and cellular helicases and, like other alphavirus-like (+) strand viruses, belongs to SF1 helicases (Koonin & Dolja, 1993). Of the seven conserved motifs shared by members of SF1, motifs I, II, V and VI are the best conserved and of these, motif VI is the one implicated in ATP hydrolysis and RNA unwinding (Kadaré & Haenni, 1997). Because of the highly structured nature of satRNAs, it is reasonable to propose that the deficiency in satRNA replication may be attributed to impairment in the helicase activity as a consequence of alteration in the aa sequence of motif VI and flanking sequences. The report by Roossinck *et al*. (1997), that the loss of satRNA support in strain Sny of CMV maps to a single aa downstream from motif VI at the C-terminal region of the 1a protein, appears to be consistent with this proposal (see below).

Roossinck *et al*. (1997) presented evidence that a change from Cys to Ser at position 978 of the 1a protein is responsible for the loss of satRNA support function in CMV-Sny, and attributed this to a significant change in the predicted secondary structure of the C-terminal region of the protein. To ascertain whether a similar mechanism underlies the inability of PSV-W to support satRNA, the secondary structure of the C-terminal region of PSV-W 1a protein was determined and compared to that of other PSV strains known to support satRNA. Results of computer-assisted protein secondary structure prediction, using the GOR algorithm (as used by Roossinck *et al*., 1997), revealed that a 13 aa α-helix at the C-terminal region of the 1a proteins from subgroup I strains PSV-ER and PSV-J was shortened to 7 aa followed downstream by a 4 aa β-sheet region in the two subgroup II strains, PSV-W and PSV-B (data not shown). These differences in secondary structures apparently have no bearing on support of satRNA replication, since the predicted secondary structure of the 1a protein from PSV-B, a subgroup II strain that supports satRNA replication (Militão *et al*., 1998), is identical to that of PSV-W. Thus the proposal that changes in the secondary structure of the C-terminal region of the 1a protein are responsible for the loss in ability to replicate satRNA may not be applicable to PSV satRNA.

The 1a proteins from all PSV and CMV strains sequenced to date (with the apparent exception of CMV-Sny) contain a Cys residue (implicated in satRNA support) at a comparable position (boxed, Fig. 4) regardless of whether or not they support satRNA replication. The aa sequence within the helicase motifs V and VI, as well as the sequences flanking these two motifs, is highly conserved among the alphavirus-like supergroup (members of SF1 helicases). Of the 45 aa contained in this region, 43 are identical in at least seven of the nine cucumovirus sequences shown (Fig. 4). The present data on aa sequence and predicted secondary structure of the C-terminal region of 1a protein of PSV-W thus do not allow us to identify specific differences that may be responsible for its inability to support the replication of satRNA.

The interaction of PSV with satRNA presents a less complicated model system to study satRNA replication, because of the apparent lack of specificity in regard to host species and satRNA variant (Naidu *et al*., 1995). We have recently cloned and sequenced RNA1 from strain PSV-B which belongs to subgroup II of PSV strains (C.-C. Hu and others, unpublished data). Because PSV-B is closely related to PSV-W (about 95% nt sequence identity) and, unlike PSV-W, is capable of replicating satRNA (Militão *et al*., 1998), it should provide valuable material for mutational and molecular genetic analysis to map the RNA1 sequences required for satRNA replication.

This research was funded by grants from The University of Kentucky Tobacco and Health Research Institute (5-41124 and 5-41147) and is published as Journal Series Paper 97-14-181 of the Kentucky Agricultural Experiment Station.

References

Ahlquist, P., Strauss, E. G., Rice, M., Strauss, J. H., Haseloff, J. & Zimmern, D. (1985). Sindbis virus proteins nsP1 and nsP2 contain homology to nonstructural proteins from several RNA plant viruses. *Journal of Virology* **53**, 536–542.

Amasino, R. M. (1986). Acceleration of nucleic acid hybridization rate by polyethylene glycol. *Analytical Biochemistry* **152**, 304–307.

Bernal, J. J., Moriones, E. & García-Arenal, F. (1991). Evolutionary relationships in the cucumoviruses: nucleotide sequence of tomato aspermy virus RNA 1. *Journal of General Virology* **72**, 2191–2195.

Church, G. M. & Gilbert, W. (1984). Genomic sequencing. *Proceedings of the National Academy of Sciences, USA* **81**, 1991–1995.

Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of nucleotide sequence analysis programs for the VAX. *Nucleic Acids Research* **12**, 387–395.

Ding, S. W., Anderson, B. J., Haase, H. R. & Symons, R. H. (1994). New overlapping gene encoded by the cucumber mosaic virus genome. *Virology* **198**, 593–601.

Fromm, M. E., Callis, J., Taylor, L. P. & Walbot, V. (1987). Electroporation of DNA and RNA into plant protoplasts. *Methods in Enzymology* **153**, 351–366.

Garnier, J., Osguthorpe, D. J. & Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *Journal of Molecular Biology* **120**, 97–120.

Ghabrial, S. A., Pickard, C. M. & Stuckey, R. E. (1977). Identification and distribution of virus diseases of soybean in Kentucky. *Plant Disease Reporter* **61**, 690–694.

Hayes, R. J. & Buck, K. W. (1990). Infectious cucumber mosaic virus RNA transcribed *in vitro* from clones obtained from cDNA amplified using the polymerase chain reaction. *Journal of General Virology* **71**, 2503–2508.

Hu, C.-C. & Ghabrial, S. A. (1998). Molecular evidence that strain BV-15 of peanut stunt cucumovirus is a reassortant between subgroup I and II strains. *Phytopathology* **88**, 92–97.

Hu, C.-C., Aboul-Ata, A. E., Naidu, R. A. & Ghabrial, S. A. (1997). Evidence for the occurrence of two distinct subgroups of peanut stunt cucumovirus strains: molecular characterization of RNA3. *Journal of General Virology* **78**, 929–939.

Jameson, B. A. & Wolf, H. (1988). The antigenic index: a novel algorithm for predicting antigenic determinants. *Computer Applications in the Biosciences* **4**, 181–186.

Kadaré, G. & Haenni, A.-L. (1997). Virus-encoded RNA helicases. *Journal of Virology* **71**, 2583–2590.

Kaper, J. M., Tousignant, M. E., Diaz-Ruiz, J. R. & Tolin, S. A. (1978). Peanut stunt virus-associated RNA 5: second tripartite genome virus with an associated satellite-like replicating RNA. *Virology* **88**, 166–170.

Karasawa, A., Nakaho, K., Kakutani, T., Minobe, Y. & Ehara, Y. (1991). Nucleotide sequence of RNA 3 of peanut stunt cucumovirus. *Virology* **185**, 464–467.

Karasawa, A., Nakaho, K., Kakutani, T., Minobe, Y. & Ehara, Y. (1992). Nucleotide sequence analyses of peanut stunt cucumovirus RNAs 1 and 2. *Journal of General Virology* **73**, 701–707.

Kataoka, J., Masuta, C. & Takanami, Y. (1990). Complete nucleotide sequence of RNA1 of cucumber mosaic virus Y strain and evolutionary relationships among genome RNAs of the virus strains. *Annals of the Phytopathological Society of Japan* **56**, 501–507.

Koonin, E. V. & Dolja, V. V. (1993). Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Critical Reviews in Biochemistry and Molecular Biology* **28**, 375–430.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Lot, H., Marrou, J., Quiot, J. B. & Esvan, C. (1972). Contribution – l'étude de virus de la mosaïque du concombre (CMV). II. Méthode de purification rapide du virus. *Annales de Phytopathologie* **4**, 25–38.

McGarvey, P., Tousignant, M., Geletka, L., Cellini, F. & Kaper, J. M. (1995). The complete sequence of a cucumber mosaic virus from *Ixora* that is deficient in the replication of satellite RNAs. *Journal of General Virology* **76**, 2257–2270.

Militão, V., Moreno, I., Rodríguez-Cerezo, E. & García-Arenal, F. (1998). Differential interactions among isolates of peanut stunt cucumovirus and its satellite RNA. *Journal of General Virology* **79**, 177–184.

Moreno, I. M., Bernal, J. J., deBlas, B. G., Rodríguez-Cerezo, E. & García-Arenal, F. (1997). The expression level of the 3a movement protein determines differences in severity of symptoms between two strains of tomato aspermy cucumovirus. *Molecular Plant–Microbe Interactions* **10**, 171–179.

Moriones, E., Roossinck, M. J. & García-Arenal, F. (1991). Nucleotide sequence of tomato aspermy virus RNA 2. *Journal of General Virology* **72**, 779–783.

Moriones, E., Diaz, I., Fernandez-Cuartero, B., Fraile, A., Burgyan, J. & García Arenal, F. (1994). Mapping helper virus functions for cucumber mosaic virus satellite RNA with pseudorecombinants derived from cucumber mosaic and tomato aspermy viruses. *Virology* **205**, 574–577.

Naidu, R. A., Collins, G. B. & Ghabrial, S. A. (1991). Symptom-modulating properties of peanut stunt virus satellite RNA sequence variants. *Molecular Plant–Microbe Interactions* **4**, 268–275.

Naidu, R. A., Collins, G. B. & Ghabrial, S. A. (1992). Peanut stunt virus satellite RNA: analysis of sequences that affect symptom attenuation in tobacco. *Virology* **189**, 668–677.

Naidu, R. A., Hu, C.-C., Pennington, R. E. & Ghabrial, S. A. (1995). Differentiation of eastern and western strains of peanut stunt cucumovirus based on satellite RNA support and nucleotide sequence homology. *Phytopathology* **85**, 502–507.

Palukaitis, P., Roossinck, M. J., Dietzgen, R. G. & Francki, R. I. B. (1992). Cucumber mosaic virus. *Advances in Virus Research* **41**, 281–348.

Passmore, B., Sanger, M., Chin, L.-S., Falk, B. W. & Bruening, G. (1993). Beet western yellows virus-associated RNA: an independently replicating RNA that stimulates virus accumulation. *Proceedings of the National Academy of Sciences, USA* **90**, 10168–10172.

Rezaian, M. A., Williams, R. H. V. & Symons, R. H. (1985). Nucleotide sequence of cucumber mosaic virus RNA 1. Presence of a sequence complementary to part of the viral satellite RNA and homologies with other viral RNAs. *European Journal of Biochemistry* **150**, 331–339.

Rizzo, T. M. & Palukaitis, P. (1989). Nucleotide sequence and evolutionary relationships of cucumber mosaic virus (CMV) strains: CMV RNA 1. *Journal of General Virology* **70**, 1–11.

Rizzo, T. M. & Palukaitis, P. (1990). Construction of full-length cDNA clones of cucumber mosaic virus RNAs 1, 2 and 3: generation of infectious RNA transcripts. *Molecular and General Genetics* **222**, 249–256.

Roossinck, M. J. & Palukaitis, P. (1991). Differential replication in zucchini squash of a cucumber mosaic virus satellite RNA maps to RNA 1 of the helper virus. *Virology* **181**, 371–373.

Roossinck, M. J., Sleat, D. & Palukaitis, P. (1992). Satellite RNAs of plant viruses: structures and biological effects. *Microbiological Reviews* **56**, 265–279.

Roossinck, M. J., Kaplan, I. & Palukaitis, P. (1997). Support of a

cucumber mosaic virus satellite RNA maps to a single amino acid proximal to the helicase domain of the helper virus. *Journal of Virology* **71**, 608–612.

Rybicki, E. P. (1995). The *Bromoviridae*. In *Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses*, pp. 450–457. Edited by F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo & M. D. Summers. Vienna & New York: Springer-Verlag.

Tousignant, M. E., McGarvey, P. M., Piazzolla, P. & Kaper, J. M. (1996). Cucumber mosaic virus from *Ixora*: infectious RNA transcripts confirm deficiency in satellite support and unusual symptomatology. *Research in Virology* **147**, 363–373.

Zhang, L., Hanada, K. & Palukaitis, P. (1994). Mapping local and systemic symptom determinants of cucumber mosaic cucumovirus in tobacco. *Journal of General Virology* **75**, 3185–3191.

Received 16 January 1998; Accepted 2 April 1998