The chimeric nature of the genome of pea enation mosaic virus: the independent replication of RNA 2

Steven A. Demler, Deborah G. Rucker and Gustaaf A. de Zoeten*

Department of Botany and Plant Pathology, Michigan State University, East Lansing, Michigan 48824–1312, U.S.A.

The genome of pea enation mosaic virus (PEMV) consists of two plus-sense RNAs, both of which are required for mechanical transmission. RNA 1 (5706 nucleotides) has strong sequence similarity with members of the luteovirus group, a similarity that is also manifested in the symptomatology, cytopathology and vector transmission of this virus. RNA 2 (4253 nucleotides) is hypothesized to facilitate systemic invasion and mechanical transmission, attributes that distinguish PEMV from the phloem-limited luteoviruses. Sequence analysis of RNA 2 has demonstrated that PEMV is unique among multicomponent viruses in that it lacks 3'- and 5'-terminal homology between its genomic RNAs. Sequence analysis of RNA 2 has identified an open reading frame encoding a putative product of 65K that contains a series of polymerase-like motifs typical of viral RNA-dependent RNA polymerases. This protein sequence lacks homology with the polymerase encoded on RNA 1 of PEMV, instead being more closely affiliated with the polymerases of viruses related to the carmopol and tombusvirus groups. Inoculation of pea protoplasts with RNA transcripts derived from a full-length cDNA clone of RNA 2 has demonstrated that RNA 2 replicates autonomously in the absence of RNA 1, although comparable inoculation of whole plants failed to establish a systemic infection. There is no evidence that RNA 2 encodes structural proteins, suggesting that encapsidation functions are supplied in trans by RNA 1, comparable to the helper-dependent complexes occurring within the luteovirus group. These data suggest that the PEMV genome can be characterized as a symbiotic association of two taxonomically distinct viral RNAs cooperatively interacting in the establishment of a systemic virus infection.

Introduction

The type member of the monotypic enamovirus group is pea enation mosaic virus (PEMV). This virus is characterized by a genome of two positive-strand ssRNAs of Mr 1.9×10^6 and 1.4×10^6 encapsidated in separate, structurally distinct isometric particles. Although its bipartite genome and virion organization have dictated the establishment of a separate taxonomic group for PEMV, evidence suggests that PEMV may more accurately be considered a derivative of viruses in the luteovirus group. Analysis of the nucleotide sequence and genomic organization of RNA 1 has established strong similarity between PEMV and the potato leafroll (PLRV)–beet western yellows (BWYV) luteovirus subgroup, encompassing regions encoding the coat, aphid transmission and putative polymerase cassettes of these viruses (Demler & de Zoeten, 1991). The breadth of this relationship is also reflected in similarities in the symptomatology, cytopathology, virion composition and circulative aphid transmission of these two virus groups (summarized in Demler & de Zoeten, 1991). On the basis of these similarities, it appears that much of the basic biology associated with PEMV infection correlates with the luteovirus-like nature of RNA 1.

Despite these similarities, there are differences in both the basic biology and genomic organization of PEMV that distinguish this virus from those in the luteovirus group. Perhaps the most significant of these differences is the ability of PEMV to circumvent the phloem limitation of luteoviruses, coupled with its capacity for mechanical transmission. It has been hypothesized that these two traits and the occurrence of a second RNA are associated, and that RNA 2 provides functions facilitating systemic invasion. A second dissimilarity is the absence from PEMV RNA 1 of the highly conserved 17K to 19K protein-encoding reading frame nested within the coat protein gene of luteoviruses. This protein has been postulated to represent the 5'-terminal VpG common to both virus groups (Miller et al., 1988a; van der Wilk et al., 1989), and a recent report by Tacke et al. (1991) has speculated that it may have a role in virus replication, based on its affinity for single-stranded nucleic acids. Although the necessity and role of this...
protein have not been unequivocally determined, the conspicuous absence of a comparable open reading frame (ORF) in RNA 1 of PEMV raises the possibility that this function, if required, may be provided in trans by RNA 2.

To clarify the biological significance of RNA 2 in PEMV infection, we have undertaken an examination of the genomic strategy of this species. We demonstrate that RNA 2 represents a unique and self-replicating RNA species unrelated to the remainder of the PEMV genome. The data suggest that PEMV is not a conventional multicomponent virus, but instead represents what may be better characterized as a form of symbiosis between two unique viral RNAs of diverse taxonomic origins.

Methods

Virus isolates and purification. The aphid non-transmissible PEMV strain WSG was propagated under greenhouse conditions in Pisum sativum L. cv 8221 as described previously (Demler & de Zoeten, 1989). Purification of virus and viral RNAs were also as described (Demler & de Zoeten, 1989, 1991).

cDNA synthesis and cloning of PEMV RNA 2. The synthesis of an RNA 2 cDNA library and its insertion into pUC19 were based on the procedure of Gubler & Hoffman (1983) and were identical to the protocol described previously in the analysis of PEMV RNA 1 (Demler & de Zoeten, 1991).

Following sequence analysis, a second series of clones were generated by exploiting the presence of the two internal BamHI sites identified at positions 1406 and 3632. The synthetic oligonucleotide 5' AGAAGCGCCC 3' (nucleotides 124 to 140) with AMV reverse transcriptase as described by Allison (1989, 1991). The internal 2226 nucleotides of RNA 2 were derived from clone pS0 (nucleotides 4119 to 4135) serving as the primer on the negative-sense RNA strand. The sequence was analysed using the UWCGG software system (Devereux et al., 1984) as described previously (Demler & de Zoeten, 1991). Amino acid sequence analysis was performed using the programs COMPARE, DOTPLOT and GAP, and database comparisons were performed using the program TFASTA in conjunction with the GenBank (release 68) and EMBL (release 27) databases.

Construction of RNA 2 transcription vectors. A series of three cDNA clones and an additional polymerase chain reaction (PCR)-generated clone containing the phage T7 promoter sequence were used to assemble a full-length, transcriptionally active clone of RNA 2 (designated pPER2; Fig. 1). Specific restriction fragments were isolated in Nusieve GTG agarose (FMC), and the fidelity of all ligation segments originating from individual clones are identified by different line patterns; pertinent restriction sites are as indicated. Digestion with Smal followed by transcription with T7 RNA polymerase leads to the production of full length RNA 2 transcripts containing the native 5' and 3'-terminal nucleotides.

Template with the oligonucleotide 5' dGGTGCGTGAGAGCTTG (nucleotides 4119 to 4135) serving as the primer on the negative-sense RNA strand. The sequence was analysed using the UWCGG software system (Devereux et al., 1984) as described previously (Demler & de Zoeten, 1991). Amino acid sequence analysis was performed using the programs COMPARE, DOTPLOT and GAP, and database comparisons were performed using the program TFASTA in conjunction with the GenBank (release 68) and EMBL (release 27) databases.

 PCR/T7 promoter

Fig. 1. Assembly of a full-length cDNA clone of PEMV RNA 2 (pPER2) under the control of the bacteriophage T7 promoter. cDNA segments originating from individual clones are identified by different line patterns; pertinent restriction sites are as indicated. Digestion with Smal followed by transcription with T7 RNA polymerase leads to the production of full length RNA 2 transcripts containing the native 5' and 3'-terminal nucleotides.

Construction of RNA 2 transcription vectors. A series of three cDNA clones and an additional polymerase chain reaction (PCR)-generated clone containing the phage T7 promoter sequence were used to assemble a full-length, transcriptionally active clone of RNA 2 (designated pPER2; Fig. 1). Specific restriction fragments were isolated in Nusieve GTG agarose (FMC), and the fidelity of all ligation segments originating from individual clones are identified by different line patterns; pertinent restriction sites are as indicated. Digestion with Smal followed by transcription with T7 RNA polymerase leads to the production of full length RNA 2 transcripts containing the native 5' and 3'-terminal nucleotides.

Template with the oligonucleotide 5' dGGTGCGTGAGAGCTTG (nucleotides 4119 to 4135) serving as the primer on the negative-sense RNA strand. The sequence was analysed using the UWCGG software system (Devereux et al., 1984) as described previously (Demler & de Zoeten, 1991). Amino acid sequence analysis was performed using the programs COMPARE, DOTPLOT and GAP, and database comparisons were performed using the program TFASTA in conjunction with the GenBank (release 68) and EMBL (release 27) databases.

Construction of RNA 2 transcription vectors. A series of three cDNA clones and an additional polymerase chain reaction (PCR)-generated clone containing the phage T7 promoter sequence were used to assemble a full-length, transcriptionally active clone of RNA 2 (designated pPER2; Fig. 1). Specific restriction fragments were isolated in Nusieve GTG agarose (FMC), and the fidelity of all ligation segments originating from individual clones are identified by different line patterns; pertinent restriction sites are as indicated. Digestion with Smal followed by transcription with T7 RNA polymerase leads to the production of full length RNA 2 transcripts containing the native 5' and 3'-terminal nucleotides.

Template with the oligonucleotide 5' dGGTGCGTGAGAGCTTG (nucleotides 4119 to 4135) serving as the primer on the negative-sense RNA strand. The sequence was analysed using the UWCGG software system (Devereux et al., 1984) as described previously (Demler & de Zoeten, 1991). Amino acid sequence analysis was performed using the programs COMPARE, DOTPLOT and GAP, and database comparisons were performed using the program TFASTA in conjunction with the GenBank (release 68) and EMBL (release 27) databases.
extraction and ethanol precipitation, the PCR product was digested with EcoRI and BamHI (nucleotide 114), and ligated into pUC19. Following sequence confirmation, the PCR clone was digested with EcoRI and XbaI (nucleotide 88) and reintroduced into similarly digested p2152 (designated p2152b).

The final step in constructing pPER2 consisted of digesting p2152b with EcoRI, followed by partial digestion with BamHI and ligating the resulting 1428 nucleotide fragment into EcoRI 1 - BamHI-digested p80b. Digestion of pPER2 with SmaI followed by transcription with T7 RNA polymerase results in full-length RNA 2 transcripts containing native 5' and 3'-terminal nucleotides.

Preparation of RNA transcripts. Capped transcripts of pPER2 were prepared by suspending 5 µg of linearized plasmid in 250 µl of transcription buffer (40 mM-Tris–HCl pH 7.5, 6 mM-MgCl2, 10 mM-NaCl, 2 mM-steripinidine, 10 mM-DTT, 0.4 mM-AMP, -CTP and -UTP, 0.075 mM-GTP, 0.5 mM-G(5')ppp(5')G, 125 U RNAsin, 50 U T7 RNA polymerase (Promega), and incubating for 45 min at 37 °C. An additional 50 U T7 RNA polymerase was added and the GTP concentration increased to 0.16 mM for an additional 45 min incubation. The DNA template was destroyed with 5 U RQI DNase (Promega) (37 °C, 15 min), followed by two phenol-chloroform extractions and ethanol precipitation. Transcript size and integrity were evaluated on non-denaturing 1% agarose gels and quantified against known amounts of PENV RNAs.

In vitro translation analysis. The translation products of 5' capped transcripts and viral RNAs were examined in both rabbit reticulocyte and wheatgerm systems (Promega and Boehringer-Mannheim, respectively) as recommended by the manufacturers. [35S]Methionine-labeled proteins were separated on 12% denaturing gels (Laemmli, 1970) and detected by fluorography.

Protoplast assays. Protoplasts of P. sativum cv. 8221 were prepared using an adaptation of the methods of Loesch-Fries & Hall (1980) & de Faria & de Zoeten (1986). Pea plants were grown under greenhouse conditions (20 °C, 16 h photoperiod) prior to harvest. Ten to 12 day (post-emergence) pea seedlings were dark conditioned for 24 to 48 h prior to protoplast isolation.

Portions (4 g) of fully expanded leaves were surface-sterilized by sequentially soaking for 1 min in 70% ethanol containing 150 p.p.m. Tween 20, 6 min in 10% household bleach containing 450 p.p.m. Tween 20, and rinsing thoroughly in five changes of sterile distilled water. The leaves were then sliced with a razor blade into 1 to 3 mm strips and digested in a 50 ml solution of 20 mg/ml Cellulysin (Calbiochem), 1 mg/ml Macerozyme R-10 (Yakult Honsha), 1 mg/ml BSA in 10% mannitol, pH 5.7. Digestion was performed for 1 h at 28 °C in a rotary water bath at 120 rotations/min. The digest was filtered through two layers of sterile cheesecloth and the protoplasts were concentrated by centrifugation for 3 min at 25 g in an IEC model HN-SII table top centrifuge equipped with a swinging bucket rotor. The protoplasts were washed three times by resuspension in 8 ml 10% mannitol followed by re centrifugation by centrifugation. The pelleted protoplasts were resuspended in 1 ml 10% mannitol, and the concentration and viability were assessed by fluorescence microscopy using the vital stain fluorescein diacetate. Approximate yields ranged from 1 x 108 to 2 x 108 protoplasts/g tissue with viability in excess of 90% at the time of inoculation, and in excess of 75% following a 24 h incubation.

Inoculation of protoplasts was performed using the polyethylene glycol/CaCl2 procedure described by Samac et al. (1983). Typically, 1 x 105 to 2 x 105 protoplasts and 1 to 5 µg RNA were used in each treatment. Following inoculation, the protoplasts were washed twice with 3 ml of 10% mannit0 and the final protoplast pellet was resuspended in 0.5 ml of protoplast medium (Loesch-Fries & Hall, 1980). Inoculated protoplasts were incubated for either 0 (control) or 24 h at room temperature under diffuse lighting. Following incubation, the protoplasts were concentrated by brief centrifugation, and the medium was replaced with 400 µl of RNA extraction buffer (100 mM-glycine, 100 mM-NaCl, 0.1 mM-EDTA, 1% SDS, 2.5 µg/ml bentonite pH 8.0; Loesch-Fries & Hall, 1980). The mixture was vortexed briefly, extracted twice with 400 µl of phenol-chloroform, and ethanol precipitated. Nucleic acids were concentrated by centrifugation and resuspended in 30 µl water.

Inoculation of plants. Seedlings of P. sativum L. cv 8221 and Nicotiana benthamiana were inoculated with 5 to 10 µl aliquots of viral transcripts or viral RNAs suspended in 10 mM-Tris–HCl pH 8.0, 1 mM-EDTA, 1% celite, 1% bentonite. Plants were harvested at 6 and 12 days post-inoculation and total RNA was isolated by the protocol of Silflow et al. (1979) as described by Demler & de Zoeten (1989).

Northern blot analysis. Protoplast nucleic acids were separated on 1% non-denaturing agarose gels in 0.5 x TAE buffer (20 mM-Tris-acetate, 10 mM-sodium acetate, 0.5 mM-EDTA pH 7.8) and electroblotted onto Hybond-N nylon transfer membranes (Amersham) in 0.5 x TAE buffer. Total RNAs isolated from whole plants were analysed on 1% glyoxal denaturing gels (Sambrook et al., 1989) and electroblotted in 0.5 x TAE as stated above. Prehybridization, hybridization and washing of membranes were as recommended by the manufacturer.

Preparation of 32P-labeled probes. RNA probes specific for the positive-sense strand of PENV RNA 2 were prepared by subcloning the EcoRI–SmaI fragment of pPER2 (nucleotides 1 to 4253) into the dual transcription vector pT7T3-18 (Pharmacia). Similarly, negative-sense probes specific for PENV RNA 1 were prepared by subcloning RNA 1 clone p216 (corresponding to ORF 2 of RNA 1, nucleotides 978 to 1766; Demler & de Zoeten, 1991) into the Pho site of pT7T3-18. T3 RNA polymerase transcription reactions were carried out with 1 µg of EcoRI-linearized template as described by Gilman (1989). Following DNase digestion of the template, the reaction mixture was diluted to 100 µl with 10 mM-Tris–HCl, 1 mM-EDTA pH 8.0 and residual nucleotides were removed by Sephadex G-50 column chromatography.

Serological screening of viral proteins expressed in vitro. The three 3'-terminal ORFs of RNA 2 (corresponding to the 25K, 27K and 15K protein-encoding reading frames; Fig. 3) were expressed as β-galactosidase fusion proteins in vitro using the pEX vector system (Stanley & Luzio, 1984). Vectors expressing the overlapping 25K and 27K protein-encoding reading frames were constructed from an Asnl (nucleotide 2795)–BamHI (nucleotide 3632) fragment of clone p80 in which the S' Asnl overhang had been repaired with the Klenow fragment of DNA polymerase. This fragment was ligated into both pEX1 (27K protein-encoding reading frame) and pEX2 (25K protein-encoding reading frame) at the BamHI–BamHI polynucleotide sites. The 15K-encoding ORF was expressed from a His6II (nucleotide 3855)–Pho site (pUC19 polynucleotide) subclone of p2127 inserted into the SmaI–Pho site of pE2X. Constructions were confirmed by sequence analysis, and plasmids introduced into Escherichia coli strain N4830 containing the temperature-sensitive λ repressor.

Bacterial cultures were grown for 12 to 16 h at room temperature, followed by a 2 h heat induction at 42 °C. The cells were then pelleted, resuspended in 100 to 200 µl of HEMGN buffer (25 mM-HEPES pH 7.6, 100 mM-KCl, 0.1 mM-EDTA, 12.5 mM-MgCl2, 10% glycerol, 0.1% NP40) containing protease inhibitors (Hocy, 1990), and disrupted in a sonic bath at 0 °C for 5 min. The lysate was diluted with five volumes of 2 x Laemmli sample buffer and heated to 95 °C for 5 min; 20 µl aliquots were electrophoresed on 7.5% SDS-polyacrylamide gels (Laemmli, 1970). Proteins were then electroblotted onto nitrocellulose membranes and examined by Western blot analysis using rabbit anti-
Fig. 2. The nucleotide sequence and major translation products of PEMV RNA 2. Alternative nucleotides at specific positions are depicted above the sequence, alternative amino acids as lower case letters. The primary sequence depicted corresponds to the sequence of pPER2.
antiserum directed against PEMV virions in a double antibody sandwich
with goat anti-rabbit alkaline phosphatase conjugate (Sambrook et al.,
1989).

Electron microscopy. Pea protoplasts inoculated with either PEMV
RNAs 1 and 2 or with transcripts derived from pPER2, or mock-
inoculated negative controls were fixed 24 h after inoculation by over-
night dialysis at 4 °C against 0.5 % glutaraldehyde, 10 % mannitol, in 80
mM-sodium cacodylate buffer (pH 7.2). Protoplasts were concentrated
by centrifugation, washed with 80 mM-cacodylate buffer and post-fixed
for 2 h in 2 % osmium tetroxide in cacodylate buffer. The protoplasts
were dehydrated in a graded series of acetone to 70 % and stained with
a saturated uranyl acetate solution in 70 % acetone. The protoplasts
were then dehydrated to 100 % acetone and embedded in Spurr's
medium. Sections were prepared with a Reichert microtome, stained
with lead citrate and viewed on a JEOL-100C electron microscope.

Results and Discussion

Sequence analysis of PEMV RNA 2

Previous sequence analysis of RNA 1 of PEMV has
established a strong linkage with the PLRV-BWYV
subgroup of luteoviruses, particularly in regions en-
coding the putative viral polymerase and structural
proteins (Demler & de Zoeten, 1991). As an initial
approach toward defining the role of RNA 2 in the
infection process, we examined the nucleotide se-
equence for evidence of specific sequence elements and rela-
tionships that might identify the origin and function of this
RNA. Fig. 2 shows the entire 4253 nucleotide se-
quence of RNA 2, along with the predicted amino acid sequence
encoded by relevant ORFs. Also included are nucleotide
and amino acid variants found in alternative clones of
RNA 2. Fig. 3 depicts the size and organization of the
major RNA 2 ORFs. The 5' ORF begins following a
short 21 nucleotide non-coding region (Fig. 3) and
encodes a potential product of 33K. The second ORF
overlaps the final 37 codons of reading frame 1 (in a
different reading frame) and has the capacity to encode
a product of 65K. This putative 65K product is composed
of an 8K N-terminal region preceding the first meth-
ionine codon (shaded region, Fig. 3), followed by a 57K
C-terminal region terminating with an amber codon
(nucleotide 2558). Following the second ORF, a 221
nucleotide non-translated region precedes a third ORF
encoding a potential product of 25K. Superimposed (out
of frame) on this reading frame is an ORF encoding a
putative 27K product, beginning 17 nucleotides after the
reading frame 3 start codon and extending 76 nucleotides
beyond its termination codon. This is followed by a 293
nucleotide non-coding region, leading into a fifth ORF
potentially encoding a 15K product which lacks a
termination codon. Among the six clones examined in
this region, the 15K ORF was the predominant ar-
angement. Two variants were identified that had stop
codons nine and 59 codons prior to the 3' terminus.

Analysis of the 65K ORF

The most illuminating evidence concerning the origin
and function of RNA 2 came from comparisons of the
deduced amino acid sequence of the 65K protein-
encoding ORF with viral sequences in the GenBank
and EMBL databases. This amino acid sequence displayed
statistically significant identity with a number of viral
proteins belonging to the diantho-, tombus-, carmo-
neco- and barley yellow dwarf virus (BYDV)
subgroup of the luteoviruses, as well as with putative translation
products encoded by the BWYV ST9-associated RNA
(Table 1). In all cases, these proteins contain sequence
motifs associated with nucleic acid polymerase and
helicase-like functions, and are hypothesized to represent
the core viral RNA-dependent RNA polymerases (Fig.
4; Hodgman, 1988; Habili & Symons, 1989; Koonin,
1991). Based on the occurrence and arrangement of these
motifs, it has been proposed that this group of viruses
constitutes a separate taxonomic entity, different from
the PLRV-BWYV subgroup of the luteoviruses which
contains the NY-RPV isolate of BYDV and RNA 1 of
PEMV (Habili & Symons, 1989; Koonin, 1991; Demler
& de Zoeten, 1991; Vincent et al., 1991). Consistent with this model is the observation that this
65K product does not display homology with the
polymerase cassette of PEMV RNA 1, nor with its allied
luteoviruses. Thus, the bipartite genome of PEMV
appears to contain two polymerase-like proteins derived
from taxonomically distinct virus groups.

The out-of-frame overlap of the 33K and 65K protein-
encoding ORFs suggests that a frameshift event may be
involved in the expression of the 65K protein ORF,
comparable to that in the expression of the polymerase
cassettes described in PEMV RNA 1 and in the diantho-
and luteovirus groups. In vitro translation of full-length
transcripts of RNA 2 and of virion-derived RNA 2 did
not produce a 65K or 57K primary product (Fig. 5b,
lanes 7 and 8). Consistent with the frameshift hypothesis
is the occurrence of a minor 97K translation product,
which correlates with the 93K product predicted by a
frameshift mechanism. Supporting this observation,
transcripts truncated at the XhoI site (nucleotide 2056,
lane 5) shifted the migration of the 97K protein to 71K,
in good agreement with the size reduction anticipated by a
frameshift strategy (75K). Sequence comparisons of
the putative frameshift regions of PEMV RNA 1 and
RNA 2 failed to identify any regions of homology that
might implicate a common signal for such a frameshift
event. In contrast, the region immediately upstream of
the RNA 2 frame 1 amber stop codon (GGAUUUUU,
917) is comparable to putative frameshift signals identi-
fied in the short overlap regions of red clover necrotic
mosaic virus (RCNMV) and BYDV (GGAUUUUU
917).
and GGGUUUUU, respectively; Xiong & Lommel, 1989; Miller et al., 1988b). The five nucleotide uracil tract has been implicated in a $-1$ frameshift involving phenylalanine codons, and resembles similar events in a variety of retrovirus frameshifts (Jacks et al., 1988). Thus, the similarities evident in the primary sequence of the PEMV 65K protein and the BYDV and RCNMV polymerases also appear to be reflected in the translational expression of the reading frames.

**Analysis of the 3'-terminal 33K protein**

Despite the strong similarity evident in the 65K portion of the frameshift protein, the PEMV RNA 2 33K protein lacks statistically significant identity with either of its counterparts in the luteo- and dianthovirus subgroups, or with the N-terminal reading frames of the carmo-, necro- and tombusvirus groups. Additional database searches have failed to reveal amino acid homology with any other viral or non-viral sequences. Although the N-terminal component of the PEMV RNA 1 frameshift product displayed a clustering of hydrophobic residues indicative of membrane association, the RNA 2 33K product does not exhibit a comparable composition. At this time the role of this protein in infection is unclear.

The dominant product of RNA 2 in an *in vitro* rabbit reticulocyte translation system was a 42K protein, with some evidence for a minor 28K protein (Fig. 5b). The 42K product was also synthesized in a wheatgerm system, although we found no evidence of the minor 28K product (data not shown). This 42K primary product corresponds quite closely to a dominant 45K protein reported in previous translational analysis of PEMV RNA 2 (Gabriel & de Zoeten, 1984). In an attempt to identify the RNA regions encoding this product, we prepared a series of truncated transcripts to serve as templates for *in vitro* translation. Transcripts terminating immediately after the 33K protein reading frames ($AsnI$, nucleotide 1117; Fig. 5b, lanes 1 and 6) generated only the 45K protein, suggesting that this protein represents the ORF 1 product. Supporting this observation, transcripts terminating within the 33K protein reading frame (*AccI*, nucleotide 611; Fig. 5b, lane 3; *HincII*, nucleotide 874; Fig. 5b, lane 2) shifted the 42K protein to 31K and 40K respectively, closely approximating the size reduction anticipated from truncations in the 33K protein reading frame. The cause of the anomalous migration of this protein has not been determined, but the pattern is consistent for the translation products of both viral RNAs and *in vitro* transcripts.

**Table 1. Amino acid sequence comparisons of PEMV RNA 2 65K protein-encoding ORF product with selected viral polymerases***

<table>
<thead>
<tr>
<th>Virus†</th>
<th>Similarity (%)</th>
<th>Identity (%)</th>
<th>Quality‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tombus- TBSV</td>
<td>54</td>
<td>33</td>
<td>302 (45)</td>
</tr>
<tr>
<td>CNV</td>
<td>56</td>
<td>34</td>
<td>307 (37)</td>
</tr>
<tr>
<td>CyRSV</td>
<td>55</td>
<td>34</td>
<td>307 (33)</td>
</tr>
<tr>
<td>Carmo- CarMV</td>
<td>55</td>
<td>35</td>
<td>317 (50)</td>
</tr>
<tr>
<td>TCV</td>
<td>59</td>
<td>40</td>
<td>332 (48)</td>
</tr>
<tr>
<td>Necro- TNV-D</td>
<td>59</td>
<td>41</td>
<td>349 (43)</td>
</tr>
<tr>
<td>Diantho- RCNMV</td>
<td>51</td>
<td>30</td>
<td>290 (31)</td>
</tr>
<tr>
<td>Luteo- BYDV-PAV</td>
<td>54</td>
<td>34</td>
<td>290 (36)</td>
</tr>
<tr>
<td>BWVV-ST9</td>
<td>53</td>
<td>32</td>
<td>179 (40)</td>
</tr>
<tr>
<td>MCMV</td>
<td>55</td>
<td>35</td>
<td>328 (45)</td>
</tr>
</tbody>
</table>

* Comparisons were performed using the program GAP with comparison parameters defined as described previously (Demler & de Zoeten, 1991).
† Abbreviations: TBSV, tomato bushy stunt virus; CNV, cherry necrosis virus; CyRSV, cymbidium ringspot virus; CarMV, carnation mottle virus; TCV, turnip crinkle virus; TNV-D, tobacco necrosis virus strain D; MCMV, maize chlorotic mottle virus. The sequence of the ST9 isolate of BWYV is from Chin et al. (1989); that of TNV-D is from Coutts et al. (1991). All other sequences were derived from the GenBank and EMBL databases.
‡ Numbers in parentheses correspond to the number of standard deviations by which the quality exceeds the mean quality of 36 comparisons of randomized sequences.

![Fig. 3. Arrangement of ORFs of PEMV RNA 2. Open boxes indicate prominent ORFs; the grey box designates the 8K N-terminal extension preceding the start codon of the 57K protein-encoding ORF.](image-url)
perfecticosalhedral symmetry, and appear somewhat pleiotropic in electron microscopy (Gibbs et al., 1966; Hull & Lane, 1973; S. A. Demler & G. A. de Zoeten, unpublished results). The relative ratios of top to bottom component can also vary greatly between PEMV isolates, ranging from nearly undetectable levels of top component to isolates in which the level of top component is the dominant nucleoprotein species (Hull & Lane, 1973). In an analysis of the virion subunits of two aphid-transmissible isolates of PEMV, Hull (1977) detected the presence of three minor protein species, including one of 28K. In light of the described dual polymerase nature of the PEMV genome, it is relevant to determine whether these pleiotropic virions may be composed of a unique second structural subunit or of a mixture of subunits.

Several lines of evidence discount the possibility that the 3′-terminal RNA 2 proteins provide a supplementary component to the viral RNA 1 or 1a subunits (Gabriel, 1983; Demler & de Zoeten, 1991). We have found no evidence of minor species preserved in all six sequenced isolates, and lower case letters indicating conservation in two or more sequences. The designation of polymerase (POL) and helicase (HEL) motifs follows the classification scheme of Habib & Symons (1989).
with rabbit anti-PEMV antiserum demonstrated cross-reactivity with only the RNA 1-encoded product (Fig. 6). The viral antigen used to generate this antiserum was composed of approximately 25 to 30% top component, and therefore we would expect some level of recognition if the top component was composed primarily of a unique protein. However, this approach may lack adequate sensitivity for the detection of a trace capsid component. Finally, electron microscopy of sectioned protoplasts inoculated with RNA 2 of PEMV failed to provide unequivocal evidence of virions specific to RNA 2 infection, although there was evidence of virions localized in the nuclei of protoplasts co-infected with RNAs 1 and 2. The available evidence suggests that encapsidation functions are provided solely by the coat protein encoded by RNA 1.

Earlier we noted the absence from PEMV RNA 1 of the 17K to 19K nested reading frame present in all luteovirus coat proteins, and we hypothesized that this product may be provided in trans by RNA 2. Examination of the RNA 2-encoded proteins has not provided any evidence to support this contention. In direct amino acid sequence comparisons using the programs COMPARE, DOTPLOT and GAP, there was no evidence of statistically significant sequence homology, nor was there evidence of the acidic–basic residue charge distribution reported by Tacke et al. (1991) in the PLRV 17K protein. A notable structural characteristic of the PEMV 25K protein is a preponderance of basic residues in the N-terminal half of the molecule (38 of 126 residues), although we can ascribe no functional significance to this observation at this time.

Another instance of the nested overlap of reading frames occurs in the 3' non-structural proteins encoded by viruses of the tombusvirus group. Amino acid sequence comparisons did not identify statistically significant homology between viruses of these groups, or with any other sequences available in the databases. At this time, sequence analysis has not provided a clear picture of the function or role of these three PEMV proteins. It is also unclear at this time whether the unclosed 15K protein-encoding ORF represents a translationally active species.

**Analysis of non-coding regions**

The occurrence of dual, dissimilar polymerase cassettes raises the possibility that one or possibly both polymerases function autonomously or in concert in the replication of the PEMV genome. In marked contrast to other multicomponent viral genomes, comparisons of the 5' and 3' termini of RNA 1 and 2 has demonstrated an absence of sequence homology, supporting the hypothesis of independently acting polymerases. Consistent with the polymerase sequence similarity identified previously, the 5' termini of the plus and minus sense strands of RNA 2 end in three guanylate residues, a trait evident with both termini of CarMV as well as in the negative-sense strand of RCNMV, BYDV, TBSV, CyRSV, TCV and TNV-D. Although the combination of PEMV RNA 1 and 2 has been demonstrated to possess a covalently attached VpG of 17.5K (Reisman & de Zoeten, 1982), the dissimilarity between 5'- and 3'-terminal sequences suggests that we can no longer assume that both species are equivalent in this respect.

The absence of the 27K, 25K and 15K products from the in vitro translation profile suggests that generation of subgenomic mRNAs is a likely expression strategy for
the corresponding downstream ORFs. The sequences of the two internal non-coding regions spanning nucleotides 2560 to 2782 and 3550 to 3723 were unique and had no homology with comparable regions in the tombus-carmo-like viruses.

A curious phenomenon present in the first RNA 2 intergenic region is the occurrence of duplicated sequence blocks:

RNA 2
2590 gguCaCACCUGAAGuGAcACGGGuAgacauc AAGGGAAGCGGGGAgucc acc
2642 aacCeCACCUGAAuGACAGGGGcAaa......AAGGGAAGCGGgGcAe 2684

RNA 1
cggCcCACCa UGAAuGACAu GGgu
3830 u AAGGGAAGCGGGGcuAc
3868

An examination of the intergenic region of PEMV RNA 1 preceding the coat protein cistron also revealed a similar pair of these sequence blocks. We cannot confirm a functional role for these sequences, but it is noteworthy that they precede internal reading frames hypothesized to be expressed by subgenomic messengers. Until the promoter signals and termini of the putative RNA 1 and RNA 2 subgenomic messengers are defined, we cannot confirm a role for these blocks as promoter-related elements.

**Infectivity studies**

Dissimilarity in the 5' and 3' termini of PEMV RNAs 1 and 2 and the presence of an apparent second polymerase cassette in RNA 2 suggest that RNA 2 may be replicatively competent in the absence of RNA 1. To test this hypothesis, we inoculated pea protoplasts, pea seedlings and seedlings of the non-leguminous host *N. benthamiana* with full-length *in vitro* transcripts of RNA type PEMV RNAs 1 and 2 in this protoplast system. The RNA analysed in lanes 7 and 8 was isolated from protoplasts inoculated with full-length RNA 2 transcripts containing the native 3' and 5' termini. As is evident in these lanes, full-length transcripts derived from pPER2 demonstrate replicative competence in the absence of RNA 1.

To assess the possibility that the positive signals evident in lanes 1, 7 and 8 were due to the survival of the input inoculum, two sets of controls were used to confirm the *de novo* replication of RNAs 1 and 2. Lane 6 represents a non-replicating negative control consisting of RNA isolated from protoplasts inoculated with transcripts lacking the 3'-terminal 227 nucleotides of RNA 2 (*Eag* I, nucleotide 4026). Although we cannot rule out the possibility that this input inoculum is significantly less stable than its full-length counterparts, the absence of hybridization in this lane suggests that positive signals are not due to the survival of residual inoculum.
Analysis of pea enation mosaic virus RNA 2

Supporting this conclusion, lanes 3 to 5 contain 0 h controls consisting of protoplast RNAs isolated immediately after inoculation (and washing) with viral RNA (lane 3), with EagI-truncated transcripts (lane 4) or with full-length RNA 2 transcripts (lane 5). Relative to the 0 h controls, there is a clear enhancement of signal in lanes 1, 7 and 8, consistent with de novo replication of RNAs 1 and 2. Thus, these analyses support the conclusion that RNA 2 is replicatively competent in pea protoplasts in the absence of PEMV RNA 1.

In contrast to protoplast assays, we found no evidence of systemic invasion in pea or *N. benthamiana* seedlings inoculated with RNA 2 transcripts. Neither host displayed virus-induced symptoms, nor could we detect evidence of RNA 2 by Northern blot analysis of uninoculated leaves (Fig. 8). This is consistent with previous infectivity studies of electrophoretically fractionated viral RNAs (Hull & Lane, 1973). Thus, although transcripts of RNA 2 replicate in a pea protoplast system, they are not competent in the establishment of systemic infection in these hosts.

**Electron microscopy**

For each of the experimental treatments, a total of 40 viable protoplasts were examined for evidence of cytopathological ultrastructure. Of the protoplasts inoculated with PEMV RNAs 1 and 2, 60% demonstrated the typical cytopathic aberrations associated with PEMV infection (de Zoeten *et al.*, 1972). This was most evident in the presence of membrane-bound vesicles in the cytosol and associated with the perinuclear space, the presence of virions within the nucleus, and a general disorganization of the nuclear membrane. In contrast, mock-inoculated protoplasts and protoplasts inoculated with RNA 2 transcripts were indistinguishable by electron microscopic analysis. Although Northern blot analysis confirmed active replication of RNA 2 in these protoplasts (Fig. 7), and assuming levels of infection comparable to the RNA 1 and 2 combination, the RNA 2-inoculated protoplasts did not display the cytopathic structures typical of PEMV infection. They also did not display any cytopathology common to members of the tombus- and carmo-like viruses. There was also no unequivocal evidence of virions in these tissues. There was some limited evidence of vesiculation in RNA 2-infected protoplasts, although similar structures were also evident in mock-inoculated controls, a probable reflection of cell membrane and wall repair processes.

**Conclusions**

Previous analyses of the interactions occurring between PEMV, its host and its aphid vector have established strong parallels with the luteovirus group, an analogy that is supported by similarities in the genomic organization and sequence of PEMV RNA 1 (Demler & de Zoeten, 1991). However, there is a clear difference between these two groups, evident in the adaptation of PEMV to the establishment of a systemic infection, its...
supplementary capacity to be mechanically transmissible, and in discrete differences in its genomic organization. We have hypothesized that these non-luteovirus-like characteristics are associated with the acquisition of RNA 2, and the goal of this study was to characterize the functional and evolutionary significance of RNA 2 in PEMV infection.

The data presented in this study support the hypothesis that RNA 2 represents the recruitment by a luteo-like virus of accessory functions derived from an unrelated virus. This is most effectively demonstrated in the conspicuous absence of sequence homology to RNA 1, particularly in the 3'- and 5'-terminal regions, a trait atypical among multicomponent viruses. This dissimilarity is emphasized by the presence of separately encoded polymerase-like units, which are associated with different virus groups. The demonstration in this study of the autonomous replication of RNA 2 in protoplasts, coupled with comparable evidence of the replicative independence of RNA 1 (unpublished results), further illustrates that the replicative functions of RNA 2 are independent from those of RNA 1. It follows that the replicative autonomy of RNA 2, there does appear to be some element of reliance on RNA 1 for functions relating to encapsidation, vector transmission and systemic invasion. Several lines of evidence have pointed to an absence of RNA 2-encoded structural functions, suggesting that encapsidation and vector transmission functions are provided by structural components encoded by RNA 1. Inoculation of seedlings with RNA 2 transcripts demonstrated that mechanical transmission does not occur in the absence of RNA 1, corroborating previous evidence with virion-derived RNA 2 (Hull & Lane, 1973). Reciprocal infectivity studies have also established the incapacity of RNA 1 to be transmitted mechanically (Hull & Lane, 1973). Thus, in the hosts analysed, mechanical passage appears to be dependent on a complementary interaction between the RNA species.

In many respects, the chimeric nature of the PEMV genome bears a striking resemblance to the dependent transmission phenomena evident with PEMV and bean yellow vein banding virus (Cockbain et al., 1986), and in a number of disease complexes associated with other members of the luteovirus group (summarized in Falk & Duffus, 1981). At the centre of these associations is the reliance of the dependent virus on encapsidation and vector transmission functions provided by the luteovirus helper. Although the dependent viruses have not been extensively characterized at the molecular level, physical characterization suggests that several of these viruses are also coat protein-deficient entities (Murant et al., 1969, 1973; Falk et al., 1979a; Reddy et al., 1985; Cockbain et al., 1986). Similarly, Falk et al. (1979b) reported limited mechanical transmission of BWV from plants infected with the lettuce speckles complex, which was interpreted as a result of the enhanced movement or replication of BWV into non-phloem tissues. All of these traits parallel the association between PEMV RNA 1 and RNA 2.

An apparent discrepancy in this pattern is that unlike PEMV RNA 2 (and the BWV ST9 RNA 2; Falk & Duffus, 1984), the dependent viruses in these complexes can be transmitted mechanically. Although we found no evidence of systemic infection by RNA 2 alone in P. sativum and N. benthamiana, we cannot assume that the host range of the individual PEMV genomic components will parallel that of the combined genome. Host range studies (Falk et al., 1979b) on the lettuce speckles complex have indicated that the host ranges of the dependent and helper virus do not absolutely coincide. Thus, the choice of host in our systemic infection assay may not accurately reveal the systemic
competence of RNA 2. A second discrepancy between PEMV and the helper dependent complexes is that in the dependent complexes the two entities are separable and can generate independent infections. In PEMV, the inability to dissociate RNAs 1 and 2 is supported by infectivity studies of PEMV based on mechanical inoculation (Hull & Lane, 1973), and by data presented in this study. However, based on the luteovirus-like aspects of RNA 1, we cannot extrapolate the previous mechanical transmission data to demonstrate that RNA 1 is ineffective in establishing an independent infection. We have alluded to the possibility that in suitable hosts RNA 2 may also be autonomously infective. Thus, the relationship between the RNAs of PEMV may yet be reduced to a classical example of a dependent transmission phenomenon. We are currently assessing this possibility by performing comparable infectivity studies with infectious transcripts of RNA 1 and by expanding the host range analysis of RNA 2.

An alternative interpretation of the seemingly inseparable nature of the two PEMV RNAs is that this interaction has evolved into a more permanent association. We have already mentioned the short sequence homology present in the internal non-coding regions of each RNA, as well as a number of other differences from the classical dependent-helper association. We cannot rule out the possibility that protein or RNA derived from either RNA 1 or 2 may interact in trans with its counterpart. A clear emphasis in future research will be to uncover the contributions of each RNA to this symbiotic association and to assess the cohesion of this relationship. As evident in this study, the relationship between the RNA species in the dependent transmission complexes and the genomic RNAs of PEMV represents a gradient of dependence and complementation. It appears that these cases may represent transitional phases toward the evolution of a more permanent association of dissimilar viral RNAs.

The authors thank P. Traynor and R. Allison for assistance in developing the pea protoplast system and for critical review of the manuscript. D. Fulbright is also acknowledged for helpful comments on the manuscript. We are also grateful to L. S. Chin, J. Foster and B. W. Falk, University of California, Davis, Ca., U.S.A. for providing sequence data and a preprint of their manuscript prior to publication. This work was supported by the National Science Foundation grant number DCB-9105334.

References


(Received 5 May 1992; Accepted 17 September 1992)