Journal of General Virology (1993), **74**, 1–14. *Printed in Great Britain*

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 carmo- 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 M_r $1{\cdot}9 \times 10^6$ and $1{\cdot}4 \times 10^6$ 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.*, 1988*a*; 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 *Bam*HI sites identified at positions 1406 and 3632. The synthetic oligonucleotide 5′ dCCTGTGAGCGCCTTGCCAG 3′ (nucleotides 3902 to 3884; nucleotide numbering throughout the text refers to that in Fig. 2) was used to prime first strand synthesis with avian myeloblastosis virus (AMV) reverse transcriptase (Sambrook *et al.*, 1989). The RNA template was removed by base hydrolysis (1 h, 65 °C), neutralized with 1 M-Tris–HCl pH 6·8, and extracted with phenol–chloroform. Following ethanol precipitation, the first strand cDNA was resuspended in T4 DNA polymerase buffer (33 mM-Tris–acetate pH 7·9, 66 mM-potassium acetate, 10 mM-magnesium acetate, 1 mM-dNTPs, 4 mM-DTT). The complementary DNA strand was synthesized by incubating with 3 units (U) of T4 DNA polymerase (30 min at 37 °C) using the oligonucleotide 5′ dGGATAGGGTTGTGGAG 3′ (nucleotides 1228 to 1243) as the second strand primer. The resulting ds cDNA was then purified, digested with *Bam*HI and inserted into *Bam*HI-digested pUC19.

Nucleotide sequence analysis. Sequence analysis was performed using the dideoxynucleotide sequencing technique of Sanger *et al.* (1977) as adapted for T7 DNA polymerase (USB, Tabor & Richardson, 1987). Sequence data for the entire sequence of RNA 2 were determined in both orientations from a minimum of three overlapping clones. Sequence data were derived from either ssDNA of M13mp18 and -mp19 subclones, or directly from pUC19 plasmids using specific oligonucleotide primers.

The 5′-terminal sequence of RNA 2 was confirmed by primer extension of the oligonucleotide primer 5′ dGCACACG-AGAAGCGCCC 3′ (nucleotides 124 to 140) with AMV reverse transcriptase as described by Allison *et al.* (1988). The 3′ terminus was confirmed similarly using the denatured replicative form of RNA 2 (prepared by the method of Morris & Dodds, 1979) as

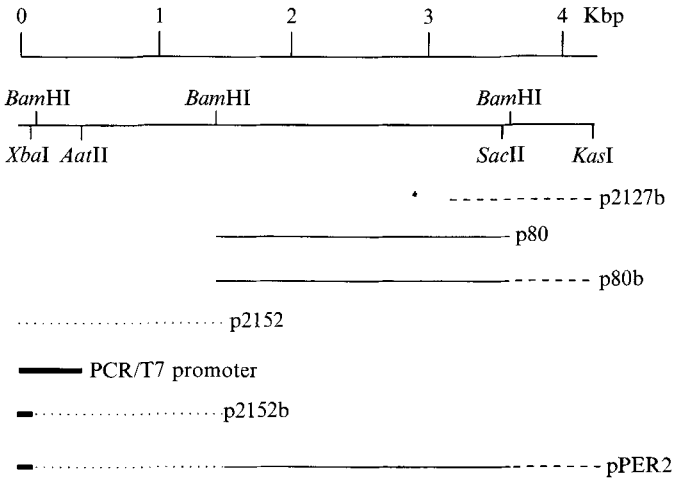


Fig. 1. Assembly of a full-length cDNA clone of PEMV RNA 2 (pPER2) under the control of the bacteriophage T7 promoter. cDNA segments originating from individual clones are identified by different line patterns; pertinent restriction sites are as indicated. Digestion with *Sma*I followed by transcription 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′ dGGTGCGTGGAGACTTGG 3′ (nucleotides 4119 to 4135) serving as the primer on the negative-sense RNA strand. The sequence was analysed using the UWGCG 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 junctions was confirmed by sequence analysis. The native 3′ terminus of RNA 2 was created by digesting clone p2127 (nucleotides 3241 to 4253) with *Kas*I (cuts at position 4248) and filling in the 5′ overhang using the Klenow fragment of DNA polymerase I. The entire cDNA fragment was then released by digestion with *Eco*RI (at the site contained in the vector polylinker) and ligated into pUC19 digested with *Kpn*I, treated with the Klenow fragment to eliminate the 3′ overhang, and then digested with *Eco*RI. This approach provided a 3′-terminal *Sma*I site that regenerates the native 3′ terminus of the RNA 2 transcript (designated p2127b).

The internal 2226 nucleotides of RNA 2 were derived from clone p80 (nucleotides 1406 to 3632). The 3′-terminal 676 nucleotides of p2127b were released by digestion with *Sac*II (3577) and *Pst*I (cuts in the polylinker), and ligated into *Sac*II–*Pst*I-digested p80 to create p80b, containing the 3′-terminal 2848 nucleotides of RNA 2.

The 5′-terminal regions of pPER2 were generated from clone p2152 (nucleotides 1 to 1438). The phage T7 promoter was attached by PCR amplification of the 5′-terminal *Pst*I–*Aat*II (nucleotide 470) fragment of p2152 using the primers 5′ dCGGAATTCT-AATACGACTCACTATAGGGTATTTATAGAG 3′ and 5′ dGAGTCGAACTTGGCGAG 3′ (nucleotides 382 to 366). The first primer contains the phage T7 promoter, preceded by an *Eco*RI site and followed by the initial 14 nucleotides of RNA 2. Following phenol

extraction and ethanol precipitation, the PCR product was digested with *Eco*RI and *Bam*HI (nucleotide 114), and ligated into pUC19. Following sequence confirmation, the PCR clone was digested with *Eco*RI and *Xba*I (nucleotide 88) and reintroduced into similarly digested p2152 (designated p2152b).

The final step in constructing pPER2 consisted of digesting p2152b with *Eco*RI, followed by partial digestion with *Bam*HI and ligating the resulting 1428 nucleotide fragment into *Eco*RI–*Bam*HI-digested p80b. Digestion of pPER2 with *Sma*I 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-$MgCl_2$, 10 mM-NaCl, 2 mM-spermidine, 10 mM-DTT, 0·4 mM-ATP, -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 RQ1 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 PEMV 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. [^{35}S]Methionine-labelled proteins were separated on 12% denaturing gels (Laemmli, 1970) and detected by fluorography.

Protoplast assays. Protoplasts of *P. sativum* L. 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 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 reconcentration 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 1×10^6 to 2×10^6 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/$CaCl_2$ procedure described by Samac *et al.* (1983). Typically, 1×10^5 to 2×10^5 protoplasts and 1 to 5 μg RNA were used in each treatment. Following inoculation, the protoplasts were washed twice with 3 ml of 10% mannitol 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 mg/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.

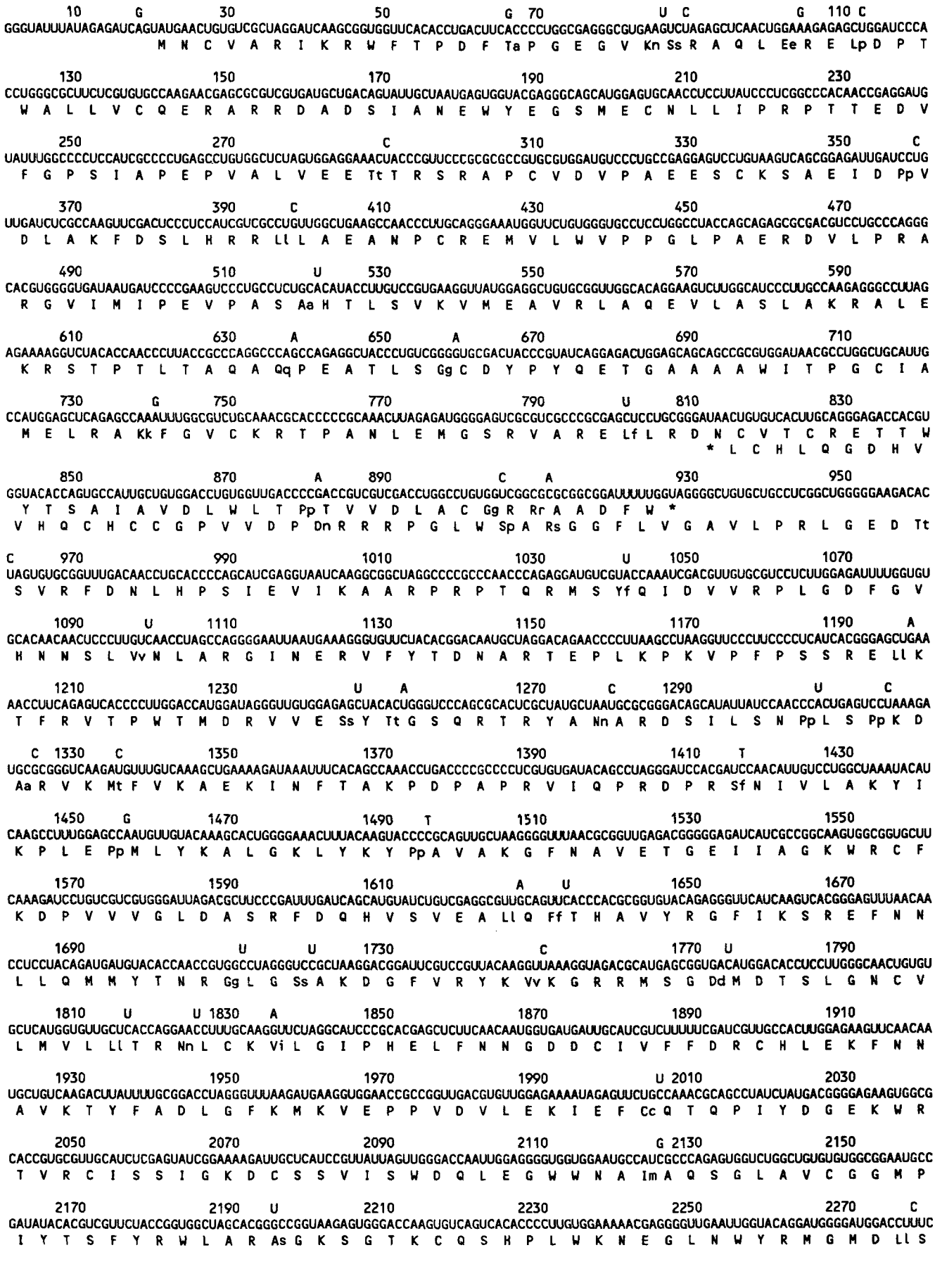
Inoculation of plants. Seedlings of *P. sativum* L. cv 8221 and *Nicotiana benthamiana* were inoculated with 5 to 10 μg aliquots of viral transcripts or viral RNAs suspended in 10 mM-Tris–HCl pH 8·0, 1 mM-EDTA, 1% celite, 1% bentonite. Plants were harvested at 6 and 12 days post-inoculation and total 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 × 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 × TAE buffer. Total RNAs isolated from whole plants were analysed on 1% glyoxal denaturing gels (Sambrook *et al.*, 1989) and electroblotted in 0·5 × TAE as stated above. Prehybridization, hybridization and washing of membranes were as recommended by the manufacturer.

Preparation of ^{32}P-labelled probes. RNA probes specific for the positive-sense strand of PEMV RNA 2 were prepared by subcloning the *Eco*RI–*Sma*I fragment of pPER2 (nucleotides 1 to 4253) into the dual transcription vector pT7T3-18 (Pharmacia). Similarly, negative-sense probes specific for PEMV 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 *Pst*I site of pT7T3-18. T3 RNA polymerase transcription reactions were carried out with 1 μg of *Eco*RI-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 *Asn*I (nucleotide 2795)–*Bam*HI (nucleotide 3632) fragment of clone p80 in which the 5′ *Asn*I 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 *Sma*I–*Bam*HI polylinker sites. The 15K-encoding ORF was expressed from a *Hin*cII (nucleotide 3855)–*Pst*I (pUC19 polylinker) subclone of p2127 inserted into the *Sma*I–*Pst*I site of pEX2. 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-$MgCl_2$, 10% glycerol, 0·1% NP40) containing protease inhibitors (Hoey, 1990), and disrupted in a sonic bath at 0 °C for 5 min. The lysate was diluted with five volumes of 2 × 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


10 G 30 50 G 70 U C G 110 C
GGGUAUUUAUAGAGAUCAGUAUGAACUGUGUCGCUAGGAUCAAGCGGUGGUUCACACCUGACUUCACCCCUGGCGAGGGCGUGAAGUCUAGAGCUCAACUGGAAAGAGAGCUGGAUCCCA
M N C V A R I K R W F T P D F Ta P G E G V Kn Ss R A Q L Ee R E Lp D P T
130 150 170 190 210 230
CCUGGGCGCUUCUCGUGUGCCAAGAACGAGCGCGUCGUGAUGCUGACAGUAUUGCUAAUGAGUGGUACGAGGGCAGCAUGGAGUGCAACCUCCUUAUCCCUCGGCCCACAACCGAGGAUG
W A L L V C Q E R A R R D A D S I A N E W Y E G S M E C N L L I P R P T T E D V
250 270 C 310 330 350 C
UAUUUGGCCCCUCCAUCGCCCCUGAGCCUGUGGCUCUAGUGGAGGAAACUACCCGUUCCCGCGCGCCGUGCGUGGAUGUCCCUGCCGAGGAGUCCUGUAAGUCAGCGGAGAUUGAUCCUG
F G P S I A P E P V A L V E E Tt T R S R A P C V D V P A E E S C K S A E I D Pp V
370 390 C 410 430 450 470
UUGAUCUCGCCAAGUUCGACUCCCUCCAUCGUCGCCUGUUGGCUGAAGCCAACCCUUGCAGGGAAAUGGUUCUGUGGGUGCCUCCUGGCCUACCAGCAGAGCGCGACGUCCUGCCCAGGG
D L A K F D S L H R R Ll L A E A N P C R E M V L W V P P G L P A E R D V L P R A
490 510 U 530 550 570 590
CACGUGGGGUGAUAAUGAUCCCCGAAGUCCCUGCCUCUGCACAUACCUUGUCCGUGAAGGUUAUGGAGGCUGUGCGGUUGGCACAGGAAGUCUUGGCAUCCCUUGCCAAGAGGGCCUUAG
R G V I M I P E V P A S Aa H T L S V K V M E A V R L A Q E V L A S L A K R A L E
610 630 A 650 A 670 690 710
AGAAAAGGUCUACACCAACCCUUACCGCCCAGGCCCAGCCAGAGGCUACCCUGUCGGGGUGCGACUACCCGUAUCAGGAGACUGGAGCAGCAGCCGCGUGGAUAACGCCUGGCUGCAUUG
K R S T P T L T A Q A Qq P E A T L S Gg C D Y P Y Q E T G A A A A W I T P G C I A
730 G 750 770 790 U 810 830
CCAUGGAGCUCAGAGCCAAAUUUGGCGUCUGCAAACGCACCCCCGCAAACUUAGAGAUGGGGAGUCGCGUCGCCCGCGAGCUCCUGCGGGAUAACUGUGUCACUUGCAGGGAGACCACGU
M E L R A Kk F G V C K R T P A N L E M G S R V A R E Lf L R D N C V T C R E T T W
* L C H L Q G D H V
850 870 A 890 C A 930 950
GGUACACCAGUGCCAUUGCUGUGGACCUGUGGUUGACCCCGACCGUCGUCGACCUGGCCUGUGGUCGGCGCGCGGCGGAUUUUUGGUAGGGGCUGUGCUGCCUCGGCUGGGGGAAGACAC
Y T S A I A V D L W L T Pp T V V D L A C Gg R Rr A A D F W *
V H Q C H C C G P V V D P Dn R R R P G L W Sp A Rs G G F L V G A V L P R L G E D Tt
C 970 990 1010 1030 U 1050 1070
UAGUGUGCGGUUUGACAACCUGCACCCCAGCAUCGAGGUAAUCAAGGCGGCUAGGCCCCGCCCAACCCAGAGGAUGUCGUACCAAAUCGACGUUGUGCGUCCUCUUGGAGAUUUUGGUGU
S V R F D N L H P S I E V I K A A R P R P T Q R M S Yf Q I D V V R P L G D F G V
1090 U 1110 1130 1150 1170 1190 A
GCACAACAACUCCCUUGUCAACCUAGCCAGGGGAAUUAAUGAAAGGGUGUUCUACACGGACAAUGCUAGGACAGAACCCCUUAAGCCUAAGGUUCCCUUCCCCUCAUCACGGGAGCUGAA
H N N S L Vv N L A R G I N E R V F Y T D N A R T E P L K P K V P F P S S R E Ll K
1210 1230 U A 1270 C 1290 U C
AACCUUCAGAGUCACCCCUUGGACCAUGGAUAGGGUUGUGGAGAGCUACACUGGGUCCCAGCGCACUCGCUAUGCUAAUGCGCGGGACAGCAUAUUAUCCAACCCACUGAGUCCUAAAGA
T F R V T P W T M D R V V E Ss Y Tt G S Q R T R Y A Nn A R D S I L S N Pp L S Pp K D
C 1330 C 1350 1370 1390 1410 T 1430
UGCGCGGGUCAAGAUGUUUGUCAAAGCUGAAAAGAUAAAUUUCACAGCCAAACCUGACCCCGCCCCUCGUGUGAUACAGCCUAGGGAUCCACGAUCCAACAUUGUCCUGGCUAAAUACAU
Aa R V K Mt F V K A E K I N F T A K P D P A P R V I Q P R D P R Sf N I V L A K Y I
1450 G 1470 1490 T 1510 1530 1550
CAAGCCUUUGGAGCCAAUGUUGUACAAAGCACUGGGGAAACUUUACAAGUACCCCGCAGUUGCUAAGGGGUUUAACGCGGUUGAGACGGGGGAGAUCAUCGCCGGCAAGUGGCGGUGCUU
K P L E Pp M L Y K A L G K L Y K Y Pp A V A K G F N A V E T G E I I A G K W R C F
1570 1590 1610 A U 1650 1670
CAAAGAUCCUGUCGUCGUGGGAUUAGACGCUUCCCGAUUUGAUCAGCAUGUAUCUGUCGAGGCGUUGCAGUUCACCCACGCGGUGUACAGAGGGUUCAUCAAGUCACGGGAGUUUAACAA
K D P V V V G L D A S R F D Q H V S V E A Ll Q Ff T H A V Y R G F I K S R E F N N
1690 U U 1730 C 1770 U 1790
CCUCCUACAGAUGAUGUACACCAACCGUGGCCUAGGGUCCGCUAAGGACGGAUUCGUCCGUUACAAGGUUAAAGGUAGACGCAUGAGCGGUGACAUGGACACCUCCUUGGGCAACUGUGU
L L Q M M Y T N R Gg L G Ss A K D G F V R Y K Vv K G R R M S G Dd M D T S L G N C V
1810 U U 1830 A 1850 1870 1890 1910
GCUCAUGGUGUUGCUCACCAGGAACCUUUGCAAGGUUCUAGGCAUCCCGCACGAGCUCUUCAACAAUGGUGAUGAUUGCAUCGUCUUUUUCGAUCGUUGCCACUUGGAGAAGUUCAACAA
L M V L Ll T R Nn L C K Vi L G I P H E L F N N G D D C I V F F D R C H L E K F N N
1930 1950 1970 1990 U 2010 2030
UGCUGUCAAGACUUAUUUUGCGGACCUAGGGUUUAAGAUGAAGGUGGAACCGCCGGUUGACGUGUUGGAGAAAAUAGAGUUCUGCCAAACGCAGCCUAUCUAUGACGGGGAGAAGUGGCG
A V K T Y F A D L G F K M K V E P P V D V L E K I E F Cc Q T Q P I Y D G E K W R
2050 2070 2090 2110 G 2130 2150
CACCGUGCGUUGCAUCUCGAGUAUCGGAAAAGAUUGCUCAUCCGUUAUUAGUUGGGACCAAUUGGAGGGGUGGUGGAAUGCCAUCGCCCAGAGUGGUCUGGCUGUGUGUGGCGGAAUGCC
T V R C I S S I G K D C S S V I S W D Q L E G W W N A Im A Q S G L A V C G G M P
2170 2190 U 2210 2230 2250 2270 C
GAUAUACACGUCGUUCUACCGGUGGCUAGCACGGGCCGGUAAGAGUGGGACCAAGUGUCAGUCACACCCCUUGUGGAAAAACGAGGGGUUGAAUUGGUACAGGAUGGGGAUGGACCUUUC
I Y T S F Y R W L A R As G K S G T K C Q S H P L W K N E G L N W Y R M G M D Ll S

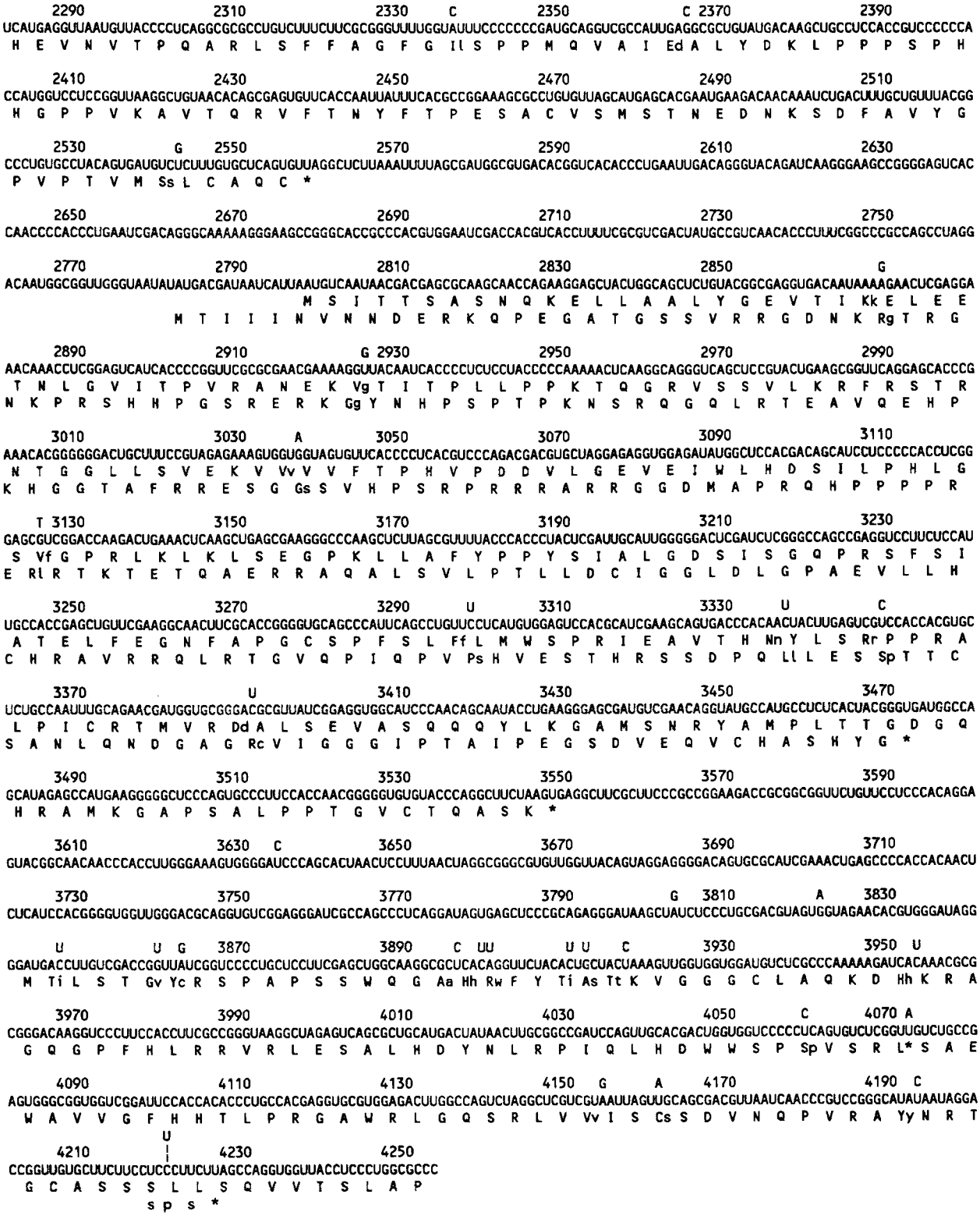


Fig. 2. The nucleotide sequence and major translation products of PEMV RNA 2. Alternative nucleotides at specific positions are depicted above the sequence, alternative amino acids as lower case letters. The primary sequence depicted corresponds to the sequence of pPER2.

antisera 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 overnight 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 encoding 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 sequence for evidence of specific sequence elements and relationships that might identify the origin and function of this RNA. Fig. 2 shows the entire 4253 nucleotide sequence 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 methionine 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 arrangement. 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-, necro- and barley yellow dwarf virus subgroup (BYDV) 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 also 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. 5*b*, 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 *Xho*I 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 identified in the short overlap regions of red clover necrotic mosaic virus (RCNMV) and BYDV (GGAUUUUU

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

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

* 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.* (1992), 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.

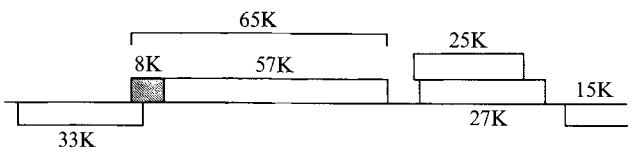


Fig. 3. Arrangement of ORFs of PEMV RNA 2. Open boxes indicate prominent ORFs; the grey box designates the 8K N-terminal extension preceding the start codon of the 57K protein-encoding ORF.

and GGGUUUUU, respectively; Xiong & Lommel, 1989; Miller *et al.*, 1988*b*). 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 5′-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. 5*b*). 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 (*Asn*I, nucleotide 1117; Fig. 5*b*, 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 (*Acc*I, nucleotide 611; Fig. 5*b*, lane 3; *Hin*cII, nucleotide 874; Fig. 5*b*, 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.

Analysis of the 3′-terminal reading frames.

The 3′-terminal position and arrangement of the remaining RNA 2 ORFs, and particularly the nested overlap of the 25K and 27K protein reading frames, are similar to the genomic organization of the coat protein gene of the luteovirus group and to the 3′ non-structural protein genes of tombusviruses. Although the dominant structural subunits of PEMV are encoded by RNA 1 (Hull & Lane, 1973; Demler & de Zoeten, 1991), there are irregularities in PEMV virion structure and composition that hypothetically could result from the presence of additional or alternative structural subunits. For example, virions forming the top component (containing RNA 2) are distinct from those in the bottom component in that they do not conform to

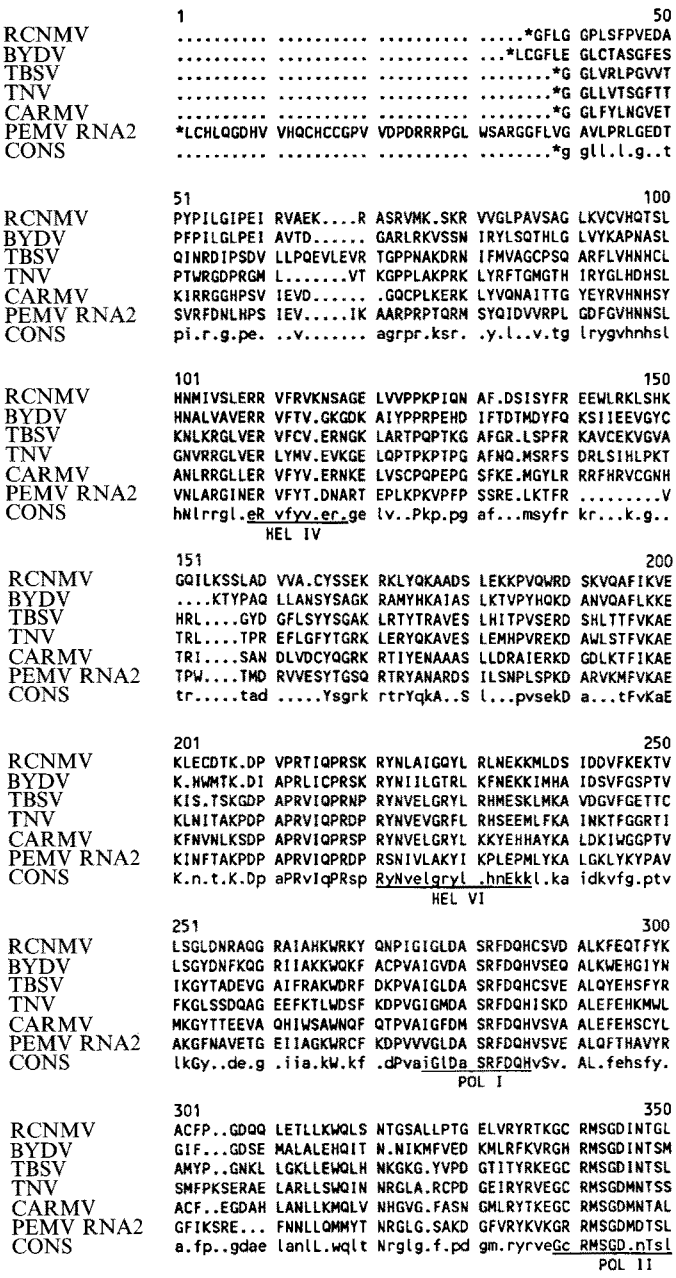


Fig. 4. Amino acid sequence alignment of the putative RNA-dependent RNA polymerases of selected members of the diantho-, luteo-, tombus-, necro- and carmovirus groups with that of PEMV RNA 2. CONS designates the consensus sequence, with capital letters indicating amino acids conserved in all six sequences, and lower case letters indicating those conserved in two or more sequences. The designation of polymerase (POL) and helicase (HEL) motifs follows the classification scheme of Habili & Symons (1989).

perfect icosahedral 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 inquire 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 structural function. First, analysis of viral coat protein by SDS–PAGE has consistently identified a single protein of 21K, corresponding to the product of the fourth reading frame of RNA 1 (Gabriel, 1983; Demler & de Zoeten, 1991). We have found no evidence of minor capsid species of a size appropriate to the 25K, 27K or 15K products in our isolates. Western blot screening of β-galactosidase fusion proteins expressing the 27K, 25K and 15K proteins of RNA 2 and the RNA 1 coat protein

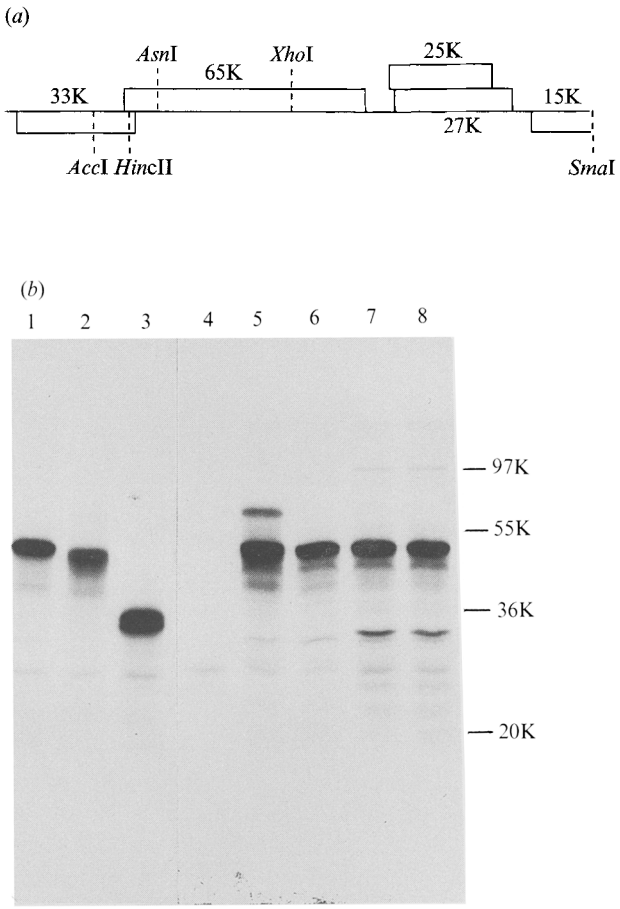


Fig 5. (*a*) Position of restriction sites used in the production of capped RNA 2 transcripts translated *in vitro*. (*b*) Autoradiogram of ^{35}S-labelled *in vitro* translation products directed by PEMV viral and transcript RNAs in rabbit reticulocyte lysate. Lane 1, transcripts from *Asn*I-digested pPER2; lane 2, transcripts from *Hinc*II-digested pPER2; lane 3, transcripts from *Acc*I-digested pPER2; lane 4, negative control (lysate without exogenous RNA added); lane 5, transcripts from *Xho*I-digested pPER2; lane 6, transcripts from *Asn*I-digested pPER2; lane 7, full-length RNA 2 transcripts; lane 8, electrophoretically fractionated PEMV virion RNA 2. Translation products were separated on a 12% SDS–polyacrylamide gel and detected by fluorography. M_r values relative to unlabelled markers are indicated to the right.

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 a virion 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

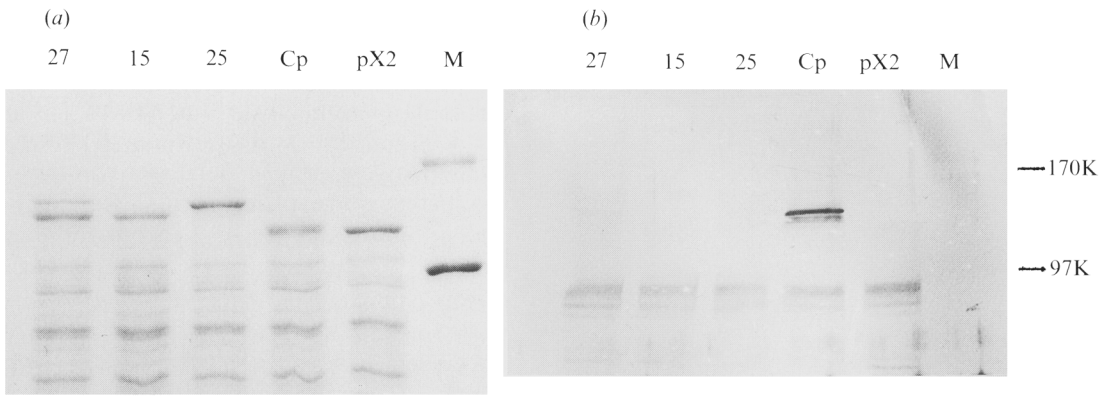


Fig. 6. Western blot of the 27K, 15K and 25K proteins (lanes 27, 15 and 25) of RNA 2 and the N-terminal 7K of the RNA 1-encoded coat protein (lane Cp) expressed in the pEX vector system. Fusion products were separated by 7·5% SDS–PAGE and blotted onto nitrocellulose membranes. (*a*) An SDS–polyacrylamide gel stained with Coomassie blue R-250. (*b*) Comparable Western blot probed with rabbit anti-PEMV antiserum and detected with goat anti-rabbit alkaline phosphatase conjugate. Lane pX2 is a negative control consisting of the pEX2-encoded fusion product lacking an exogenous insert. Lane M, M_r markers, the sizes of which are indicated to the right.

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 gguCaCACCCUGAAUuGACAGGGuAcagauc AAGGGAAGCCGGGgAgucacc

2642 aacCcCACCCUGAAUcGACAGGGcAaa.....AAGGGAAGCCGGGcAc 2684

RNA 1

3830 u AGGGAAGCCGGcuAc

cggCcCACCa UGAAUuGACAu GGuu 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 2. Fig. 7(*a*) shows a Northern blot of total RNA isolated from pea protoplasts hybridized with an RNA 2-specific probe, and Fig. 7(*b*) displays a replicate blot hybridized with a probe specific to the ORF 2-encoding region of PEMV RNA 1. Lane 2 shows mock-inoculated protoplasts, confirming the uninfected status of the protoplast source. Lane 1 shows protoplasts inoculated with virion RNA, confirming the infectivity of the combined wild-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.

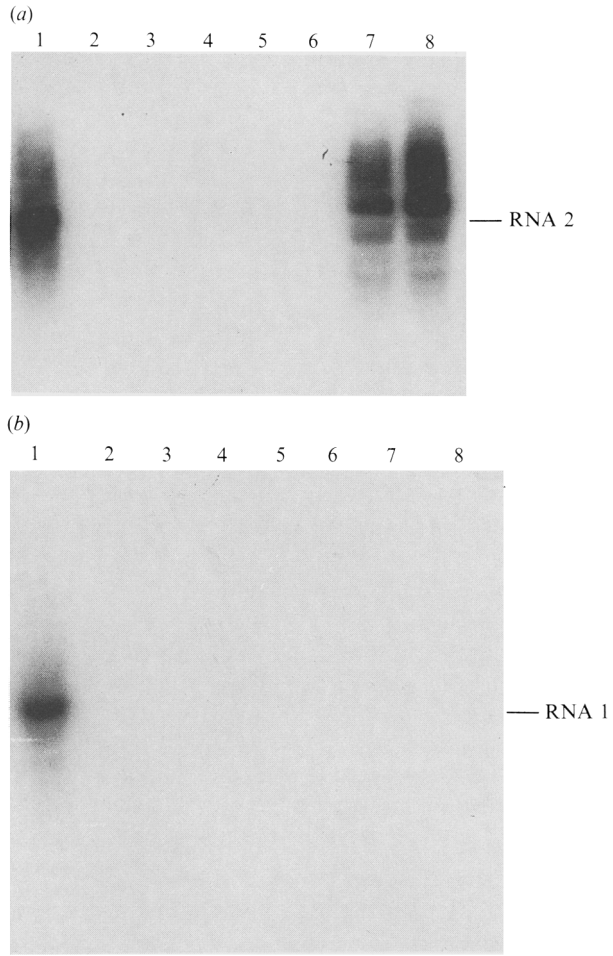


Fig. 7. Autoradiogram of a Northern blot of total RNA isolated from inoculated *P. sativum* protoplasts. (*a*) A blot hybridized with a probe specific for PEMV RNA 2 sequences; (*b*) a blot hybridized with a probe specific for the region encoding ORF 2 of PEMV RNA 1. Lanes 1, total RNA isolated from protoplasts inoculated with PEMV virion RNAs 1 and 2, 24 h post-inoculation; lanes 2, total RNA isolated from mock-inoculated protoplasts, 24 h post-inoculation; lanes 3, total RNA isolated from protoplasts inoculated with PEMV virion RNAs, 0 h post-inoculation; lanes 4, total RNA isolated from protoplasts inoculated with transcripts linearized with *Eag*I, lacking the 3′-terminal 227 nucleotides of RNA 2, 0 h post-inoculation; lanes 5, total RNA isolated from protoplasts inoculated with full-length RNA 2 transcripts, 0 h post-inoculation; lanes 6, total RNA isolated from protoplasts inoculated with transcripts linearized with *Eag*I, lacking the final 227 nucleotides of RNA 2, 24 h post-inoculation; lanes 7 and 8, duplicate samples of total RNAs isolated from protoplasts inoculated with full-length RNA 2 transcripts, 24 h post-inoculation. The RNA in lanes 1 was diluted threefold prior to gel electrophoresis.

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 *Eag*I-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

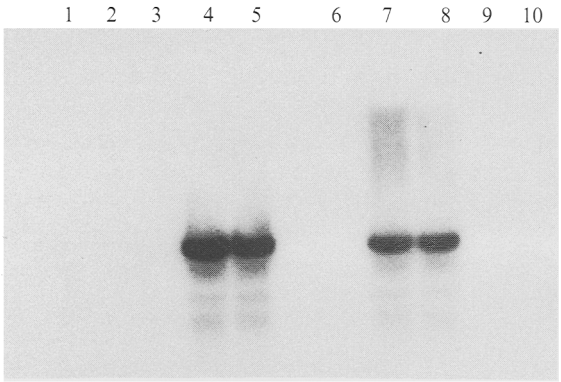


Fig. 8. Autoradiogram of a Northern blot of total RNAs isolated from the uninoculated leaves of seedlings of *P. sativum* or *N. benthamiana* harvested 6 (lanes 2, 4, 7 and 9) or 12 (lanes 1, 3, 5, 8 and 10) days post-inoculation. The blot was hybridized with a ^{32}P-labelled RNA transcript probe encompassing the entire sequence of PEMV RNA 2. Lanes 1 to 5 show total RNAs from *P. sativum* seedlings; lanes 6 to 10 show total RNAs from *N. benthamiana* seedlings. Lane 1, mock-inoculated negative control; lane 2, inoculation with RNA 2 transcripts; lane 3, inoculation with RNA 2 transcripts; lane 4, inoculation with PEMV RNAs 1 and 2; lane 5, inoculation with PEMV RNAs 1 and 2; lane 6, mock-inoculation; lane 7, inoculation with PEMV RNAs 1 and 2; lane 8, inoculation with PEMV RNAs 1 and 2; lane 9, inoculation with RNA 2 transcripts; lane 10, inoculation with RNA 2 transcripts.

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 although a discrete replication complex has been identified in PEMV-infected tissue (Powell *et al.*, 1977; Powell & de Zoeten, 1977; de Zoeten *et al.*, 1972, 1976), we can no longer presume that the replication of RNA 1 and RNA 2 are spatially linked. In this regard, the distinct ultrastructure associated with PEMV infection provides a method for the specific localization of individual viral gene products, and should complement these infectivity studies.

Despite 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.*, 1979*a*; Reddy *et al.*, 1985; Cockbain *et al.*, 1986). Similarly, Falk *et al.* (1979*b*) reported limited mechanical transmission of BWYV from plants infected with the lettuce speckles complex, which was interpreted as a result of the enhanced movement or replication of BWYV 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 BWYV 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.*, 1979*b*) 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

ALLISON, R. F., JANDA, M. & AHLQUIST, P. (1988). Infectious *in vitro* transcripts from cowpea chlorotic mottle virus cDNA clones and exchange of individual RNA components with brome mosaic virus. *Journal of Virology* **62**, 3581–3588.

CHIN, L. S., FOSTER, J. & FALK, B. W. (1992). The beet western yellows virus ST9-associated RNA shares structural and nucleotide sequence homology with carmo- and tombusviruses. *Virology* (in press).

COCKBAIN, A. J., JONES, P. & WOODS, R. D. (1986). Transmission characteristics and some other properties of bean yellow vein-banding virus, and its association with pea enation mosaic virus. *Annals of Applied Biology* **108**, 59–69.

COUTTS, R. H. A., RIGDEN, J. E., SLABAS, A. R., LOMONOSSOFF, G. P. & WISE, P. J. (1991). The complete nucleotide sequence of tobacco necrosis virus strain D. *Journal of General Virology* **72**, 1521–1529.

DE FARIA, J. C. & DE ZOETEN, G. A. (1986). Improved isolation method of pea (*Pisum sativum* L.) mesophyll protoplasts. *Archives of Biology Technology* **29**, 707–710.

DEMLER, S. A. & DE ZOETEN, G. A. (1989). Characterization of a satellite RNA associated with pea enation mosaic virus. *Journal of General Virology* **70**, 1075–1084.

DEMLER, S. A. & DE ZOETEN, G. A. (1991). The nucleotide sequence and luteovirus-like nature of RNA 1 of an aphid non-transmissible strain of pea enation mosaic virus. *Journal of General Virology* **72**, 1819–1834.

DEVEREUX, J., HAEBERLI, P. & SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* **12**, 387–395.

DE ZOETEN, G. A., GAARD, G. & DIEZ, F. B. (1972). Nuclear vesiculation associated with pea enation mosaic virus-infected plant tissue. *Virology* **48**, 638–647.

DE ZOETEN, G. A., POWELL, C. A., GAARD, G. & GERMAN, T. L. (1976). *In situ* localization of pea enation mosaic virus double-stranded ribonucleic acid. *Virology* **70**, 459–469.

FALK, B. W. & DUFFUS, J. E. (1981). Epidemiology of helper-dependent persistent aphid transmitted virus complexes. In *Plant Diseases and Vectors*, pp. 161–179. Edited by K. Maramorosch & K. F. Harris. New York: Academic Press.

FALK, B. W. & DUFFUS, J. E. (1984). Identification of small single- and double-stranded RNAs associated with severe symptoms in beet western yellows virus-infected *Capsella bursa-pastoris*. *Phytopathology* **74**, 1224–1229.

FALK, B. W., MORRIS, T. J. & DUFFUS, J. E. (1979*a*). Unstable infectivity and sedimentable ds-RNA associated with lettuce speckles mottle virus. *Virology* **96**, 239–248.

FALK, B. W., DUFFUS, J. E. & MORRIS, T. J. (1979*b*). Transmission, host range, and serological properties of the viruses that cause lettuce speckles disease. *Phytopathology* **69**, 612–617.

GABRIEL, C. J. (1983). *Investigations into the* in vitro *translation and aphid transmission of pea enation mosaic virus*. Ph.D. thesis, University of Wisconsin-Madison.

GABRIEL, C. J. & DE ZOETEN, G. A. (1984). The *in vitro* translation of pea enation mosaic virus. *Virology* **139**, 223–230.

GIBBS, A. J., HARRISON, B. D. & WOODS, R. D. (1966). Purification of pea enation mosaic virus. *Virology* **29**, 348–351.

GILMAN, M. (1989). Ribonuclease protection assay. In *Current Protocols in Molecular Biology*, pp. 4.7.1–4.7.8. Edited by F. M. Ausubel, R. Brent, R. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith & K. Struhl. New York: John Wiley.

GUBLER, U. & HOFFMAN, B. J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263–269.

HABILI, N. & SYMONS, R. H. (1989). Evolutionary relationship between luteoviruses and other RNA plant viruses based on sequence motifs in their putative RNA polymerases and nucleic acid helicases. *Nucleic Acids Research* **17**, 9543–9555.

HODGMAN, T. C. (1988). A new family of replicative proteins. *Nature, London* **333**, 22–23.

HOEY, T. (1990). Expression and purification of *lacZ* and *trpE* fusion proteins. In *Current Protocols in Molecular Biology*, pp. 16.5.1–16.5.6. Edited by F. M. Ausubel, R. Brent, R. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith & K. Struhl. New York: John Wiley.

HULL, R. (1977). Particle differences related to aphid-transmissibility of a plant virus. *Journal of General Virology* **34**, 183–187.

HULL, R. & LANE, L. C. (1973). The unusual nature of the components of a strain of pea enation mosaic virus. *Virology* **55**, 1–13.

JACKS, T., MADHANI, H. D., MASIARZ, F. R. & VARMUS, H. E. (1988). Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region. *Cell* **55**, 447–458.

KOONIN, E. V. (1991). The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *Journal of General Virology* **72**, 2197–2206.

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

LOESCH-FRIES, L. S. & HALL, T. C. (1980). Synthesis, accumulation and encapsidation of individual brome mosaic virus RNA components in barley protoplasts. *Journal of General Virology* **47**, 323–332.

MILLER, W. A., WATERHOUSE, P. M., KORTT, A. A. & GERLACH, W. L. (1988*a*). Sequence and identification of the barley yellow dwarf virus coat protein gene. *Virology* **165**, 306–309.

MILLER, W. A., WATERHOUSE, P. M. & GERLACH, W. L. (1988*b*). Sequence and organization of barley yellow dwarf virus genomic RNA. *Nucleic Acids Research* **16**, 6097–6111.

MORRIS, T. J. & DODDS, J. A. (1979). Isolation and analysis of double-stranded RNA from virus-infected plant and fungal tissue. *Phytopathology* **69**, 854–858.

MURANT, A. F., GOOLD, R. A., ROBERTS, I. M. & CATHRO, J. (1969). Carrot mottle – a persistent aphid-borne virus with unusual properties and particles. *Journal of General Virology* **4**, 329–341.

MURANT, A. F., ROBERTS, I. M. & GOOLD, R. A. (1973). Cytopathological changes and extractable infectivity in *Nicotiana clevelandii* leaves infected with carrot mottle virus. *Journal of General Virology* **21**, 269–283.

POWELL, C. A. & DE ZOETEN, G. A. (1977). Replication of pea enation mosaic virus RNA in isolated pea nuclei. *Proceedings of the National Academy of Sciences, U.S.A.* **74**, 2919–2922.

POWELL, C. A., DE ZOETEN, G. A. & GAARD, G. (1977). The localization of pea enation mosaic virus-induced RNA-dependent RNA polymerase in infected peas. *Virology* **78**, 135–143.

REDDY, D. V. R., MURANT, A. F. RASCHKE, J. H. & MAYO, M. A. (1985). Properties and partial purification of infective material from plants containing groundnut rosette virus. *Annals of Applied Biology* **107**, 65–78.

REISMAN, D. & DE ZOETEN, G. A. (1982). A covalently linked protein at the 5′-ends of the genomic RNAs of pea enation mosaic virus. *Journal of General Virology* **62**, 187–190.

SAMAC, D. H., NELSON, S. E. & LOESCH-FRIES, L. S. (1983). Virus protein synthesis in alfalfa mosaic virus infected alfalfa protoplasts. *Virology* **131**, 455–462.

SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edn. New York: Cold Spring Harbor Laboratory.

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

SILFLOW, C. D., HAMMET, J. R. & KEY, J. L. (1979). Sequence complexity of polyadenylated ribonucleic acid from soybean suspension culture cells. *Biochemistry* **18**, 2725–2731.

STANLEY, K. K. & LUZIO, J. P. (1984). Construction of a new family of high efficiency bacterial expression vectors: identification of cDNA clones coding for human liver proteins. *EMBO Journal* **3**, 1429–1434.

TABOR, S. & RICHARDSON, C. C. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proceedings of the National Academy of Sciences, U.S.A.* **84**, 4767–4771.

TACKE, E., PRÜFER, D., SCHMITZ, J. & ROHDE, W. (1991). The potato leafroll luteovirus 17K protein is a single-stranded nucleic acid-binding protein. *Journal of General Virology* **72**, 2035–2038.

VAN DER WILK, F., HUISMAN, M. J., CORNELISSEN, B. J. C., HUTTINGA, H. & GOLDBACH, R. (1989). Nucleotide sequence and organization of potato leafroll virus genomic RNA. *FEBS Letters* **245**, 51–56.

VINCENT, J. R., LISTER, R. M. & LARKINS, B. A. (1991). Nucleotide sequence analysis and genomic organization of the NY-RPV isolate of barley yellow dwarf virus. *Journal of General Virology* **72**, 2347–2355.

XIONG, Z. & LOMMEL, S. A. (1989). The complete nucleotide sequence and genome organization of red clover necrotic mosaic virus RNA-1. *Virology* **171**, 543–554.

(*Received 5 May 1992; Accepted 17 September 1992*)