Partial nucleotide sequence of poplar mosaic virus RNA confirms its classification as a carlavirus

Janey Henderson, M. J. Gibbs, M. L. Edwards, V. A. Clarke, K. A. Gardner and J. I. Cooper

NERC Institute of Virology and Environmental Microbiology, Mansfield Road, Oxford OX1 3SR and Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, U.K.

The nucleotide sequence of the 3' proximal 1328 nucleotides of poplar mosaic virus (PMV) was determined and shown to contain two large open reading frames (ORFs). The ORF nearer to the 3' terminus of the RNA is capable of encoding a polypeptide of 14K with a 'zinc-finger' motif, and is homologous to sequences in corresponding positions in five other carlaviruses. The other ORF encodes a protein of 36K which includes two sequences of amino acids identified in tryptic digests as virion capsid protein, and has amino acid sequences in common with both carlaviruses and potexviruses.

Poplar mosaic virus (PMV) is prevalent in poplar germplasm and commercial clones in which it has been associated with a 30 to 40% diminishment in growth (Cooper et al., 1986). The virus spreads naturally by unknown means, although experimentally it is mechanically transmissible (Cooper & Edwards, 1981; van der Meer, 1981). Largely on the basis of its ultrastructural appearance in infected foliage (Atkinson & Cooper, 1976; Brunt et al., 1976) and particle morphology, PMV, which has a modal length of 685 nm (Boccardo & Milne, 1976) or 675 nm (Biddle & Tinsley, 1971), has been grouped with the carlaviruses under the misnomer poplar latent virus (Matthews, 1979). The genome of one carlavirus [potato virus M (PVM)] has been completely sequenced (Zavriv et al., 1991) and partial sequences of four others [potato virus S (PVS) (MacKenzie et al., 1989), lily symptomless virus (LSV) (Memelink et al., 1990), Helenium virus S (HelVS) (Foster et al., 1990) and carnation latent virus (CLV) (Foster & Mills, 1991; Meehan & Mills, 1991)] have been published. In this communication we report the nucleotide sequence encoding two open reading frames (ORFs) at the 3' terminus of the ssRNA genome, confirming the classification of PMV as a carlavirus.

An isolate of PMV (ATCC PV275) was purified from infected tobacco plants (Nicotiana clevelandii) as described by Eweida et al. (1989). RNA extracted from PMV virions with the aid of proteinase K, SDS and sodium perchlorate was precipitated with ethanol and purified following adsorption to oligo(dT)-cellulose.

Using polyadenylated RNA and oligo(dT) for priming first-strand cDNA synthesis, a library of 3'-terminal clones was generated, blunt end-ligated into the vector pGEM-3Z (Promega) using EcoRI and NotI adaptors (Pharmacia), and transformed into Escherichia coli strain XL1-blue. The sizes of plasmid inserts in the resultant colonies were determined by gel electrophoresis after restriction enzyme digestion. A nested set of overlapping clones was prepared (using exonuclease III) from the largest cDNA inserts and sequenced using a U.S. Biochemicals Sequenase kit.

The nucleotide and derived amino acid sequences at the 3' terminus of PMV are shown in Fig. 1. There are two distinct ORFs: one capable of encoding a protein with an Mr of 35819 (36K) and the other a protein of Mr 14106 (14K). The N terminus of the viral capsid protein was blocked to Edman degradation, but tryptic digestion (Matsudaira, 1987; Aebersold et al., 1987) and automated Edman degradation using an Applied Biosystems 470A protein sequencer (Hewick et al., 1981) allowed the identification of short strings of amino acids (Fig. 1, underlined), indicating that the coat protein-coding region is in the 36K protein ORF. This allocation was confirmed when cDNA representing the predicted coat protein-coding region was transcribed using T7 polymerase and translated using rabbit reticulocyte lysate as recommended by the manufacturer (Amersham). Im-
Fig. 1. The DNA sequence corresponding to the 3'-terminal region of PMV-derived plus-strand RNA is shown below the derived amino acid sequences. Two ORFs are shown, one encoding a protein of 36K and the other a protein of 14K.

munoprecipitation with a polyclonal antiserum prepared against the virions of PMV yielded only one product, with an M₉ of 36K. This value agrees with that reported by Boccardo & Milne (1976) following their analysis of an Italian isolate of PMV, but falls outside the range (31K to 34K) listed by Francki (1980; Wetter & Milne, 1981).

There are several in-frame methionine codons, any of which may be the initiation site for the capsid protein. We have calculated the M₉ using the first AUG upstream to sequences in the genomes of carlaviruses and potexviruses using the programs GAP and PILEUP program from the GCG version 7.0 package (Devereux et al., 1984). The 36K ORF was homologous to the carlavirus capsid protein genes and, to a lesser extent, to the potexvirus coat protein genes. Pairwise comparisons revealed 30 to 36% identities between the PMV 36K ORF and the capsid protein-coding region of carlaviruses. Alignments showed that the amino acids conserved between the carlavirus, potexvirus and PMV coat protein sequences were not evenly distributed but concentrated in the C-terminal half.

The motif (R/K)FA(G/A)FDXFXXVXXXAA which is conserved in the coat protein-coding regions of potexviruses and carlaviruses (AbouHaidar & Lai, 1989; MacKenzie et al., 1989; Rupasov et al., 1989) was recognized in the PMV sequence using the MOTIFS and Patterns (Devereux et al., 1984).

A 3'-terminal ORF capable of encoding a protein in the M₉ range 10K to 15K appears to be unique to carlaviruses. No homology was evident when this sequence was used to search for other viral sequences in the PMV sequence was used to search for other viral sequences in the genomes of five carlaviruses, i.e. PVM, PVS, CLV, PMV and PVM, has recently been presented (Haylor et al., 1990). Fig. 2 is an expanded alignment that includes the PMV data and shows that the 14K ORF is included in the PMV data and shows that the 14K ORF is homologous to the other sequences. Interestingly, each of the three polypeptides with a different Mr (12.6K, 16K, 10.8K) in the viral capsid protein-coding region of carlaviruses and potexviruses using the programs GAP and PILEUP from the GCG package version 7.0 (Devereux et al., 1984). The 36K ORF was homologous to the carlavirus capsid protein genes and, to a lesser extent, to the potexvirus capsid protein genes. Pairwise comparisons revealed 30 to 36% identities between the PMV 36K ORF and the capsid protein-coding region of carlaviruses. Alignments showed that the amino acids conserved between the carlavirus, potexvirus and PMV coat protein sequences were not evenly distributed but concentrated in the C-terminal half.

The motif (R/K)FA(G/A)FDXFXXVXXXAA which is conserved in the coat protein-coding regions of potexviruses and carlaviruses (AbouHaidar & Lai, 1989; MacKenzie et al., 1989; Rupasov et al., 1989) was recognized in the PMV sequence using the MOTIFS and Patterns (Devereux et al., 1984).

A 3'-terminal ORF capable of encoding a protein in the M₉ range 10K to 15K appears to be unique to carlaviruses. No homology was evident when this sequence was used to search for other viral sequences in the SWISSPROT and GenEMBL databases using the programs WORDSEARCH and T-FASTA (GCG).

An alignment of the predicted amino acid sequences encoded by the HelVS 12.6K ORF, LSV 16K ORF, PVS 10.7K ORF, CLV 11.6K ORF, PMV 14K ORF and PVM 10.8K ORF using the PILEUP program from the GCG version 7.0 package (Devereux et al., 1984). Identical amino acid residues are boxed, and termination codons are shown as asterisks. The putative zinc-finger domains are indicated by black triangles.

Fig. 2. Alignment of the predicted amino acid sequences encoded by the HeLVs 12.6K ORF, LSV 16K ORF, PVS 10.7K ORF, CLV 11.6K ORF, PMV 14K ORF and PVM 10.8K ORF using the PILEUP program from the GCG version 7.0 package (Devereux et al., 1984). Identical amino acid residues are boxed, and termination codons are shown as asterisks. The putative zinc-finger domains are indicated by black triangles.
there is a conserved motif conforming to the consensus sequence characterizing a putative nucleic acid-binding ‘zinc-finger’ (Klug & Rhodes, 1987).

The 11K protein encoded by the 3’-terminal ORF of PVM binds single- and double-stranded nucleic acids (Gramstat et al., 1990), but no other information about the function of this domain is available.

The program CLUSTAL V (Higgins & Sharp, 1989) was used to produce a classification of the 3’-terminal ORFs (Fig. 3); the result suggests that the putative 14K protein of PMV is more closely related to the corresponding proteins of PVM or CLV than to those of the other carlaviruses tested.

Pending more detailed information about the function of carlavirus genes in general, and those of PMV in particular, it is difficult to judge the significance of the substantial identities in the capsid protein-coding regions of carlaviruses and potexviruses. The vector relations of many of these viruses are unknown and there is little pattern to the antigenic cross-reactivities between members of these groups of plant pathogenic viruses. It is interesting to note that a classification based on properties of RNA-dependent RNA polymerases (Koonin, 1991) has placed a carlavirus (PVM) closer to apple chlorotic leafspot closterovirus than to a potexvirus (potato virus X). With regard to our data, this might imply that different parts of the PMV genome evolved from different stocks.

We thank Mrs D. MacCall for technical assistance, Mrs C. Howes for computer services and A. Willis of the MRC Immunochemical Unit for protein sequencing. The work was done under NERC contract F58/G13/05 in conjunction with the Department of Plant Sciences, University of Oxford, U.K.

References


(Received 3 February 1992; Accepted 3 April 1992)