Sequence of RNA 2 of a nematode-transmissible isolate of tobacco rattle virus

Carmen Hernández,† Alex Mathis, Derek J. F. Brown and John F. Bol

1 Institute of Molecular Plant Sciences, Gorlaeus Laboratories, PO Box 9502, 2300 RA Leiden, The Netherlands and 2 Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK

Tobacco rattle virus (TRV) isolate PPK20 is transmitted by Paratrichodorus pachydermus nematodes. The factor(s) determining vector transmissibility has been shown to be located on TRV RNA 2. Sequence determination revealed that PPK20 RNA 2 contains three open reading frames encoding the coat protein (cp) and proteins with molecular masses of 29.4 and 32.8 kDa. The 29.4 and 32.8 kDa protein-coding genes showed no significant sequence similarity to any other known tobravirus gene. A full-length cDNA of PPK20 RNA 2 cloned between the 35S promoter and nos terminator infected plants when co-inoculated with PPK20 RNA 1. Deletions in the reading frames of the 29.4 and 32.8 kDa proteins revealed that these sequences are dispensable for replication of PPK20 RNA 2 in plants. Subgenomic RNAs for translation of cp and the putative 29.4 and 32.8 kDa proteins were detected in infected leaves. The possible role of PPK20 RNA 2 non-structural genes in TRV vector transmission is discussed.

Tobravirus isolates may be characterized as belonging to one of three subgroups represented by tobacco rattle virus (TRV; type member), pea early-browning virus (PEBV) and pepper ringspot virus. The tobravirus bipartite genome consists of positive-sense ssRNAs with RNA 1 being highly similar between isolates from the same subgroup. As it encodes the replicase and the movement proteins, and a small protein of unknown function, RNA 1 can replicate and spread in plants on its own (Hamilton et al., 1987; MacFarlane et al., 1989). In contrast, the RNA 2 is highly variable in size and nucleotide sequence and in all instances the 5'-proximal gene encodes the coat protein (cp; Bergh et al., 1985; Cornelissen et al., 1986; Angenent et al., 1986, 1989; Goulden et al., 1990). The variability in length of RNA 2 is due both to the presence of additional genes of unknown function on the larger molecules and a variably sized 3'-terminal region homologous to that of the corresponding RNA 1.

Tobraviruses are naturally transmitted by root-feeding Trichodorus and Paratrichodorus nematodes and serologically distinct isolates (serotypes) have been shown to be specifically transmitted by different nematode species (Brown et al., 1989; Ploeg et al., 1989). The determinants on the tobravirus genome of this apparent vector–virus specificity have not been elucidated. However, studies with pseudorecombinant viruses generated by combining RNA 1 and RNA 2 from the nematode-transmissible TRV isolate PPK20 and the non-transmissible TRV isolate PLB showed that nematode transmissibility was associated with RNA 2 (Ploeg et al., 1993). Whilst these observations suggest that the cp gene might play a role in determining the virus–nematode interaction, additional genes on RNA 2 could also be involved.

Recently, sequence information from the genome of a nematode-transmissible isolate of PEBV has been obtained and the involvement of an RNA 2-encoded 29 kDa protein in the transmission process has been suggested (MacFarlane & Brown, 1995). Sequence data from other vector-transmitted tobravirus isolates are required for comparison to determine which virus gene products may be involved. We have determined the complete sequence of RNA 2 of TRV isolate PPK20, originally transmitted by a single P. pachydermus (Ploeg et al., 1992). In addition, we constructed an infectious cauliflower mosaic virus (CaMV) 35S promoter-driven clone of PPK20 RNA 2 and two deletion mutants which were tested for their ability to replicate in plants.

A cDNA library was created using a Stratagene Lambda ZAP cDNA Synthesis Kit. First-strand cDNA
elementary to the T-terminal sequence of all TRV RNAs (5' dACTCGAGGGCGTAATAACGC 3'). The primer contain an was synthesized on purified PPK20 RNA with primer TRV3 and TRV5 (5' dCTCTAGATAAAACATT- GCACC 3'). The 3'-terminal 15 bases of primer TRV3 (5' dACTCGAGGGCGTAATAACGC 3'). The primer contain an XhoI restriction site. A full-length identical to all known TRV RNA 2 molecules; the 5'—terminal 7 nucleotides of the primer contain an XhoI restriction site. A full-length subgenomic RNAs are arrowed.

Fig. 1. Nucleotide sequence of RNA 2 from TRV isolate PPK20. The amino acid sequences are deduced from the ORFs coding for the cp, 29.4 kDa protein and 32.8 kDa protein. The putative 5' termini of the subgenomic RNAs are arrowed.

was synthesized on purified PPK20 RNA with primer TRV3 (5' dACTCGAGGGCGTAATAACGC 3'). The 3'-terminal 14 nucleotides of this primer are complimentary to the 3'-terminal sequence of all TRV RNAs sequenced to date; the Y-terminal 7 nucleotides of the primer contain an XhoI restriction site. A full-length PPK20 RNA 2 cDNA was obtained by performing PCR on the double-stranded cDNA library using primers TRV3 and TRV5 (5' dCTCTAGATAAAACATT-GCACCC 3'). The 3'-terminal 15 bases of primer TRV5 are the same as the 5'-terminal sequence, which is identical to all known TRV RNA 2 molecules; the 5'-terminal 7 nucleotides contain an XhoI restriction site. The cDNA fragment was isolated from an agarose gel.
made blunt and ligated into the EcoRV site of the pBluescript SK(−) vector. The nucleotide sequence of the PPK20 RNA 2 cDNA was determined on both strands using a series of overlapping subclones generated by exonuclease III unidirectional digestion with the Nested Deletion Kit from Pharmacia. Small stretches of the viral cDNA which could not be reached with these subclones were sequenced by priming internally on the cDNA with proper oligonucleotides.

Fig. 1 shows the complete sequence of the 3856 nucleotides of PPK20 RNA 2. A schematic representation of the ORFs found in this molecule is given in Fig. 2 with those of other sequenced tobravirus RNA 2 molecules provided for comparison.

The first ORF starts at the AUG located at nucleotides 557–559 and extends to a UAG codon at nucleotides 1169–1171. This 5′-proximal reading frame is the \( \text{cp} \) gene, as the encoded protein of 204 amino acids, with a molecular mass of 22330 Da, is homologous to the \( \text{cp} \) of other known tobraviruses.

The second ORF starts with an AUG at nucleotides 1340–1342 and terminates with a UAG codon at nucleotides 2111–2113, encoding a potential protein of 257 amino acids with a molecular mass of 29403 Da. The third ORF begins with an AUG codon at nucleotide 2588–2590, extending to a UGA codon at nucleotides 3464–3466. This ORF encodes a potential protein of 293 amino acids with a molecular mass of 32839 Da. A database search did not reveal any homology between the predicted products from ORF2 and ORF3 and any previously described tobravirus (or other) protein.

In addition to the ORFs indicated in Fig. 1, an ORF potentially able to encode a product of 57 amino acids, with a molecular mass of 6511 Da, is present in the region between ORFs 2 and 3 (positions 2231–2442). However, examination of the codon usage of this small ORF suggests that it is not translated.

The non-coding regions of PPK20 RNA 2 are represented by a leader sequence of 556 nucleotides, two intercistronic regions of 166 and 474 nucleotides (assuming no expression of the small ORF found in the molecule) and a 3′-terminal sequence of 390 nucleotides. A sequence comparison revealed an extensive similarity between the last 271 nucleotides of this molecule and the equivalent region in TRV RNA 1 molecules. When the sequence of RNA 2 of TRV isolate PPK20 was compared with those of TRV isolates TCM, PLB, and PSG, and PEBV isolate SP5, significant similarities were found in the 5′ region of 1327 nucleotides (including the leader sequence, the \( \text{cp} \) gene and about 150 nucleotides downstream from this gene), a 3′ non-coding region of 271 nucleotides and an intercistronic sequence of 160 nucleotides (positions 2419 to 2579). This intercistronic region shares 70.6% similarity with a corresponding region located in a similar position in TCM RNA 2 (positions 2065 to 2224).

To obtain an infectious clone, a full-length cDNA of PPK20 RNA 2 was inserted between the CaMV 35S promoter and the terminator sequence from the nopaline synthase gene (\( \text{Tnos} \)); the resulting construct was called pCaK20-2T7. Plasmid pCa35J was kindly provided by the Kirin Brewery Company (Japan). This plasmid contains the CaMV 35S promoter with a \( \text{StuI} \) restriction site at the position used by \( \text{RNA polymerase II} \) to initiate transcription (Yamaya et al., 1988). The \( \text{Tnos} \) sequence was inserted as a \( \text{SphI} \)-EcoRI fragment derived from plasmid pCa32T (kindly provided by L. Neeleman). Construct pCaK20-2T7 contains four non-viral nucleotides (5′ CTAG 3′) immediately downstream of the transcription start site.

Restriction sites in plasmid pCaK20-2T7 were used to generate two deletion mutants: ABX with a deletion between nucleotides 1202–2096, created by digesting pCaK20-2T7 with \( \text{BglII} \) and \( \text{XbaI} \), filling and religating; and \( \text{AXN} \), generated by digestion of pCaK20-2T7 with \( \text{XbaI} \) and \( \text{NruI} \) (positions 2096–3067 in PPK20 RNA 2) and subsequent religation of the plasmid.

Before inoculating plants, the different plasmids were digested with \( \text{BglII} \), which cleaves at positions 1380 bp upstream of the 35S promoter and 150 bp downstream of \( \text{Tnos} \). Ten \( \mu \)g of the corresponding digested plasmid and 150 \( \mu \)g of total RNA from tobacco leaves infected with only RNA 1 of isolate PPK20 were combined to inoculate three half-leaves of two \( \text{Nicotiana tabacum} \) (cv. Samsun NN) plants. Samples were collected 3 days after inoculation and total RNA was extracted and analysed by Northern blot hybridization using radioactive probes as described previously (Cornelissen et al., 1986).
Northern blot analysis revealed that the full-length clone pCaK20-2T7 and the deletion mutants ΔBX and ΔXXN generate RNA 2 molecules which co-replicate efficiently in combination with PPK20 RNA 1 (Fig. 3). Infection of healthy plants using homogenates from the primary infected material was successful for virus reconstituted from the different clones, indicating proper encapsidation of the viral RNAs; this was confirmed by electron microscopy. In Fig. 3(c) the blot was probed with a cDNA corresponding to a region of the 32.8 kDa protein-encoding ORF in PPK20 RNA 2. In addition to RNA 2 and RNA 2a (the putative subgenomic cp mRNA of 3300 nucleotides in length), this blot revealed the accumulation of a subgenomic RNA with an estimated length of 1300 nucleotides in plants inoculated either with the wild-type isolate or with the combination of PPK20-RNA 1 plus pCaK20-2T7 or ΔBX (Fig. 3c, lanes 2 to 8). The 5' end of this subgenomic RNA would map approximately 20 nucleotides upstream of the 32.8 kDa protein cistron. Also, in these plants mRNA potentially involved in the expression of the PPK20 RNA 2-encoded 29.4 kDa protein was not detected. The putative subgenomic messenger for the expression of the 32.8 kDa protein-encoding ORF was not present in plants inoculated with PPK20 RNA 1 plus ΔXXN (Fig. 3c, lanes 9 to 11) as this construct has a large deletion which includes the postulated promoter for the generation of this mRNA. However, in this infected material a new subgenomic RNA, with an estimated size of 1500 nucleotides, was detected (Fig. 3c, lanes 9 to 11). This RNA is probably involved in the translation of the 29.4 kDa cistron as a probe derived from this gene also detected this RNA (data not shown). Competition for the factors involved in the generation of the mRNAs may result in the activation of the upstream subgenomic promoter when the downstream one is deleted.

The specific relationship found between tobravirus serotypes and their associated trichodorid nematode vector species (Ploeg et al., 1992) suggests that the virus particle protein plays a primary role in the interaction with the vector. Studies using nuclear magnetic resonance have revealed that the C-terminus is a flexible segment of the virus molecule which can protrude from the particle surface and probably reacts directly with the specific sites of virus attachment within the nematode (Mayo et al., 1993). The cp amino acid sequence of isolate PPK20 is about 90% identical to those of the non-nematode-transmissible isolates PLB and PRN. A comparison of the cp sequences revealed that most of the changes occur at the C-termini of the proteins (data not shown). These changes may account for the difference in nematode transmissibility.

Another possibility is that the 29.4 and/or 32.8 kDa protein-encoding genes present in PPK20 RNA 2 are involved in vector transmission. Significant similarities are not present between the potential products of these genes and those of any other known tobravirus product, including the proteins of 29.1 kDa encoded by strain TCM RNA 2 and 29.6 and 23 kDa encoded by PEBV RNA 2. However, some of these proteins could be
functionally equivalent. The PEBV 29-6 kDa protein has been detected in infected plants, indicating that this gene is expressed (Johnsen et al., 1991) and recently its participation in vector transmission has been proposed (MacFarlane & Brown, 1995). In our study the 29-4 and 32-8 kDa protein cistrons have been shown not to be required for replication of PPK20 RNA 2 in mechanically inoculated tobacco plants. However, these two genes may each contribute characteristics for the ecological fitness of the virus under natural conditions which, to a large extent, are a function of the vector–virus–plant interaction. A requirement for another gene product, besides the cp, has been reported for vector transmission in other virus groups such as potyviruses (Pirone & Thornbury, 1984), caulimoviruses (Armour et al., 1983) and luteoviruses (Brault et al., 1995).

Vector transmission studies have revealed that the virus progeny derived from PPK20 RNA 1 plus pCaK20-2T7 are transmitted as efficiently as the wild-type isolate (C. Hernández, A. Mathis, D. Brown & J. Bol, unpublished results). The availability of this PPK20 RNA 2 cDNA clone permits an analysis of the genetic determinants in this genome segment involved in the transmission process. We are currently investigating the transmissibility of the virus progeny derived from PPK20 RNA 1 plus ÐBX or ÐXN. Other PPK20 RNA 2 mutants are being made and will be tested to obtain a detailed map of the genetic determinants of vector transmissibility.

We thank F. T. Brederode for the preparation of TRV PPK20 RNA 1. C.H. is a recipient of fellowships from the Consejo Superior de Investigaciones Científicas (Spain) and from the European Community. A.M. was supported by a grant from the Swiss National Foundation.

References


(Received 28 March 1995: Accepted 3 July 1995)