Arch Virol (1998) 143: 1535–1544

Archives of Virology

Nucleotide sequence of both genomic RNAs of a North American tobacco rattle virus isolate*

M. R. Sudarshana** and **P. H. Berger**

Division of Plant Pathology, Department of Plant, Soil and Entomological Sciences, University of Idaho, Moscow, Idaho, U.S.A.

Accepted March 21, 1998

Summary. The complete sequence of a North American tobacco rattle virus (TRV) isolate, 'Oregon yellow' (ORY), was determined from cDNA and RT-PCR clones derived from the two genomic RNAs of this isolate. The RNA-1 is 6790 bases and RNA-2 is 3261 bases. The sequence of TRV-ORY RNA-1 was similar to RNA-1 of TRV isolate SYM, and differs in 48 nucleotides. TRV-ORY RNA-1 was one base shorter than -SYM, and had 47 base substitutions resulting in 12 amino acid substitutions of which 4 were conservative. The RNA-2 of TRV-ORY was distinct from RNA-2 of other characterized TRV isolates and contained three open reading frames (ORFs) that could potentially code for proteins of MW 22.4 kDa, 37.6 kDa and 17.9 kDa. Based on the homology of the predicted amino acid sequence with those of other tobraviruses, ORF1 of RNA-2 encodes the coat protein (CP). The protein sequence of ORF2 had regions of limited similarity with those of ORF2 of two other TRV isolates and pea early browning tobravirus. The ORF3 was unique to TRV-ORY. Phylogenetic analysis of tobravirus CPs indicated that TRV-ORY was most closely related to pepper ringspot tobravirus and TRV-TCM. The relationship of tobravirus CPs to other rod-shaped tubular plant viruses is also discussed.

Introduction

In the US, TRV infection was reported in several crops [1, 4, 8, 30] and weeds [6,18] and in most cases the isolates were only partially characterized. Most plant

*The sequences reported in this paper have been deposited in the GenBank nucleotide sequence database as accession numbers AF034621 and AF034622.

**Present address: Department of Plant Pathology, University of California, Davis, CA 95616, U.S.A.

diseases caused by TRV are not of significant economic importance. However, incidence of corky ringspot syndrome (CRS) in potato tubers, caused by TRV, has increased in the Pacific Northwest, partly because of restrictions on the use of systemic nematicides. However, three variants of a TRV isolate from Oregon [17], and variable lengths of 'S' particles from California, Florida, and Oregon [13] indicate that TRV isolates present in the US are genetically diverse. Different pathotypes and serologically distinct isolates have also been encountered [6], but none of the isolates have been analyzed for genetic information. Based on nucleic acid hybridization, RNA-1 of 'Oregon yellow' (ORY) was similar to RNA-1 of TRV-SYM, but the RNA-2 was different from SYM RNA-2 [23].

Tobacco rattle virus (TRV) is the type member of the tobravirus group which includes nematode-transmitted, rigid rod-shaped, (+) sense plant RNA viruses with a bipartite genome enclosed in particles of unequal length [26]. Tobraviruses can be seedborne, but are primarily soilborne and transmitted by several species of Trichodorid nematodes. The long (L) particles of TRV contain RNA-1 (ca. 6.8 kb), which is capable of self-replication and systemic infection. The short (S) particles contain RNA-2 (ca. 1.9 to 3.9 kb) whose length is different in different isolates. Pea early browning virus (PEBV) and pepper ringspot virus (PepRSV) are the two other known members of the tobravirus group. Genomic RNAs of several European TRV isolates have been cloned and the complete nucleotide sequence of RNA-1 of the SYM isolate [12], and RNA-2 of the PLB [3], PSG [5], TCM [2], and PPK20 [14] isolates has been determined. About 2.2 kb of the 3′-end of PSG RNA-1 [5] and TCM [2] have also been sequenced. Nucleic acid hybridization studies [23, 27], and sequence homology of RNA-1 of TRV isolates PSG [5], SYM [12], and TCM, indicate that RNA-1 is highly conserved among TRV isolates. In contrast to RNA-1, RNA-2 was found to differ in length and homology [2].

In the absence of other reliable control measures, pathogen-derived resistance (PDR) (see [33]) in commercial potato varieties presents a reasonable approach to control. Previous efforts to obtain transgenic resistance have resulted in strain specificity in plants expressing TRV-TCM coat protein (CP) [32] or the replicase-derived sequence of PEBV [19]. In an effort to engineer transgenic resistance in potato to TRV isolate(s) in the Pacific Northwest, cDNA sequences derived from genomic RNAs of TRV-ORY were cloned and sequenced. Here we report the complete nucleotide sequence of RNA-1 and RNA-2 and analysis of the open reading frames relative to tobraviruses. To our knowledge, this is the first TRV isolate for which both RNAs have been completely sequenced.

Materials and methods

Virus source, purification, preparation of cDNA and sequencing of cDNA clones

TRV-ORY [17] was obtained from the American Type Culture Collection and maintained in *Nicotiana clevelandii* in a greenhouse. Virus was purified from systemically infected leaves of *N. clevelandii* 10–12 days after mechanical inoculation using the method of Robinson and Harrison [26], with the additional step of centrifugation of the extract through 8 ml of a 10%

sucrose cushion at 40,000 rpm for 2 h in a Beckman Ti70 rotor. The two virus zones were collected separately using an Isco Density Gradient Fractionator (Lincoln, NE), diluted 2.5 times (v/v) with 0.02 M Tris-HCl buffer, and the virus was pelleted by centrifuging at 40,000 rpm for 2 h in a Beckman Ti70 rotor. All operations were performed at 4 °C using ice cold buffers.

cDNA was prepared directly from 100 μg of virus particles and cloned as described previously by Wyatt et al. [34]. Clones containing virus cDNA were identified and sequenced using methods described elsewhere [16, 29]. Sequencing was by the dideoxy chain termination method using Sequenase 2.0 sequencing kit (US Biochemicals) and ^{35}S-labeled dATP. Nucleotide sequence was also obtained using a Licor (Lincoln, NE) Automated DNA Sequencer and SequiTherm Cycle Sequencing kit (Epicentre Technologies, Madison, WI), according to the manufacturer's instructions.

Determination of the 3′-end of RNA-1 and RNA-2

One mg of purified virus was suspended in 400 μl TAE buffer and incubated with 5 μl of 20% SDS and 5 μl of 5 mg/ml proteinase K at 37 °C for 30 min. After incubation, 25 μl of 20% SDS and 5μl of β-mercaptoethanol were added and the solution incubated at 60 °C for 5 min. RNA was extracted with an equal volume of phenol, followed by extraction with chloroform-isoamyl alcohol (24:1 V/V). The RNA was precipitated by ethanol and quantified using a spectrophotometer. One μg of RNA was incubated with 500 units of yeast poly(A) polymerase (US Biochem.) in a 10 μl reaction mixture for 30 min as per the manufacturer's instructions. Poly(A)-tailed RNA was used for cDNA preparation using oligo-d$(T)_{12–18}$ to prime first strand synthesis and cloned as previously described.

Based on the sequence obtained from overlapping cDNA clones, the remainder of RNA-1 and RNA-2 was obtained by RT-PCR. cDNA for RNA-1 was prepared in separate reactions using three reverse primers, complimentary to SYM RNA-1 bases 3817 to 3835, 5412 to 5429, and 6763 to 6779, respectively. The reverse primers were paired with forward primers, identical to bases 2535 to 2551, 3756 to 3783, and 4991 to 5009, respectively, of RNA-1. For PCR, 1 μl of the 20 μl first strand cDNA reaction was subjected to amplification in a 100 μl reaction mixture using 6 units of *UlTma* DNA polymerase according to the manufacturer's instructions (Perkin-Elmer Inc.). The nucleotide sequence of the ultimate 5′-ends of RNAs-1 and -2 was obtained by direct sequencing of RNA in chain termination reactions using an AMV reverse transcriptase based kit (US Biochem.) as per manufacturer's instructions.

Analysis of sequence data

Sequence analyses were performed at the Center for Visualization and Analysis of Molecular Sequences at Washington State University using the Genetics Computer Group (GCG) software. Sequences of RNA-1 and RNA-2 of other tobraviruses were obtained using GenBank and EMBL databases. Percent identity and similarity scores were obtained by pairwise alignment using GAP and multiple sequence alignment was done using PILEUP. The multiple sequence alignment was examined visually and analyzed by parsimony for phylogenetic inference using PAUP [31] on a Macintosh computer.

Results and discussion

cDNA library and sequence of RNA-1 and RNA-2 of TRV-ORY

The complete nucleotide sequence of the genomic RNAs of a North American TRV isolate was obtained from several clones, generated by different strategies. Sequencing of overlapping clones compiled together with the sequence of the

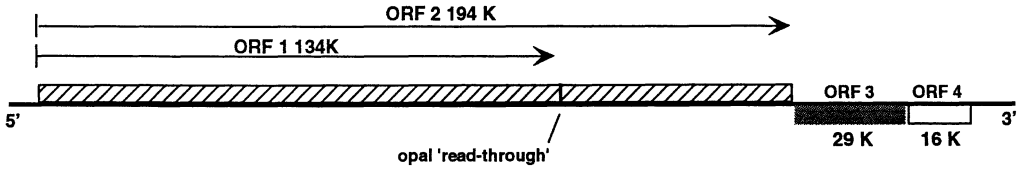

Fig. 1. The genome organization of TRV ORY-RNA1 (6790 nucleotides) (GenBank accession no. AF034622). Predicted open reading frames are represented by boxes

5′-end of RNA-1, determined by direct sequencing of RNA, indicated that the RNA-1 of TRV-ORY was 6790 bases long and contained four presumed open reading frames (ORFs) (Fig. 1). In the first attempt to obtain complete cDNA clones, a strategy that was successfully used for several plant viruses [34] was used, and cDNA clones corresponding only to the 5′-ends of RNA-1 and RNA-2 were obtained. Limited success with TRV may have been due to poor destabilization of the virions, because Mg^{2+} ions, present in the virus purification and cDNA synthesis buffers, are known to stabilize the virion structure of TRV [28]. Completed sequences of several overlapping cDNA clones and any remaining portions of RNA-2, cloned from RT-PCR, indicated that the RNA-2 was 3261 bases long and contained three non-overlapping open reading frames. The 3′-end of RNA-2 had homology to RNA-1 in the region beginning at base 2696, and this region included part of the 16 kDa ORF of RNA-1.

Comparison of RNA-1 sequence of TRV-ORY with RNA-1 of other tobraviruses

The RNA-1 of TRV-ORY is 6790 bases and is organized in the same manner as TRV-SYM, indicating that these isolates are strains of the same virus. A total of forty eight base changes were observed between TRV-SYM and -ORY. TRV-ORY RNA-1 was one base shorter than TRV-SYM RNA-1 and this difference was due to the absence of G^{137} of SYM RNA-1. Two nucleotide changes occurred in the 5′-noncoding region (NCR) and three in the 3′-NCR. Nineteen nucleotide changes were in the 134K ORF and of these sixteen were silent substitutions in the third position. Similarly, fifteen nucleotide changes were in the readthrough portion of the 194K ORF and of these twelve were also silent substitutions. Of the six amino acid substitutions in the 194K ORF, three resulted in non-conserved amino acid substitutions. Within the 29K and 16K ORFs, there were 9 nucleotide substitutions, resulting in two or four amino acid substitutions, respectively. Only one of the amino acid substitutions in the 16K ORF was conservative.

Genome organizations of RNA-1 of TRV-ORY was also similar to PEBV and PepRSV. At the nucleotide level, identities ranged from 58% to 65% between TRV-ORY, -SYM, PEBV and PepRSV. Homologies between PEBV and PepRSV are greater between each other than with TRV, at both nucleotide and amino acid levels (Table 1). Putative zinc finger motifs, previously reported in ORF4 of PEBV-SP5, PepRSV-CAM, and TRV isolates PSG, SYM, and TCM [20], were also found to be conserved in TRV-ORY. The predicted protein sequence of TRV-

Table 1. Identity (%) at nucleotide level of tobravirus RNA1 sequences and identity and similarity of their protein products

Tobravirus	TRV-SYM	PEBV-SP5	PepRV-CAM
RNA1 nucleotide sequence			
TRV-ORY	99.3	59.8	58.4
PEBV-SP5			64.6
Amino acid sequence[a]			
	a) 194K		
TRV-ORY	99.6/99.7	59.1/67.5	58.1/67.5
PEBV-SP5			68.1/75.6
	b) 29K		
TRV-ORY	99.2/99.2	55.7/67.2	58.5/67.6
PEBV-SP5			67.7/75.5
	c) 16K[b]		
TRV-ORY	97.2/97.2	26.9/36.1	27.5/38.5
PEBV-SP5			64.8/72.2

[a]Percent identity/percent similarity
[b]This protein is 12 kDa in PEBV and PepRSV

ORY and -SYM ORF4 has greater similarity to PEBV and PepRSV than to other characterized TRV isolates. To date, no biological function has been attributed for the protein product of ORF4 [3, 11].

Analysis of the sequence and genome organization of RNA-2

Despite sharing features such as similar virus architecture and transmission by Trichodorid nematodes, tobravirus RNA-2 sequences vary not only in length, but also in genome organization (Fig. 2). Regions of consensus sequence, complementary to the putative subgenomic RNA promoters in (−) RNA of tobraviruses [9], could be found upstream of each predicted ORF of TRV-ORY RNA-2. However, the putative promoter region proximal to ORF3 was closer to the expected translational start of this ORF than reported for other tobraviruses (data not shown).

Computer predictions indicated three open reading frames in TRV-ORY RNA-2, potentially coding for proteins of 22.4 kDa, 37.6 kDa, and 17.9 kDa (Fig. 2), as well as a reading frame corresponding to part of the 16 kDa cistron from the 3′-end of RNA-1 (Fig. 2). The primary sequence of TRV-ORY CP was predicted by homology to the predicted translation product of ORF1. This sequence corresponded to previous analyses of the coat protein of other TRV isolates [7, 21, 30]. Attempts to characterize the CP by direct protein sequencing of purified virions were unsuccessful, which suggests N-terminal blockage of the CP [22]. The putative CP of TRV-ORY has two AUG codons at the 5′-end of this ORF separated by three other codons. The translational context around the two codons, -CGCG*AUG*AC- and -CGGT*AUG*TA-, respectively, were different from those found in plant genes of abundant proteins [15]; we presume that the first AUG is

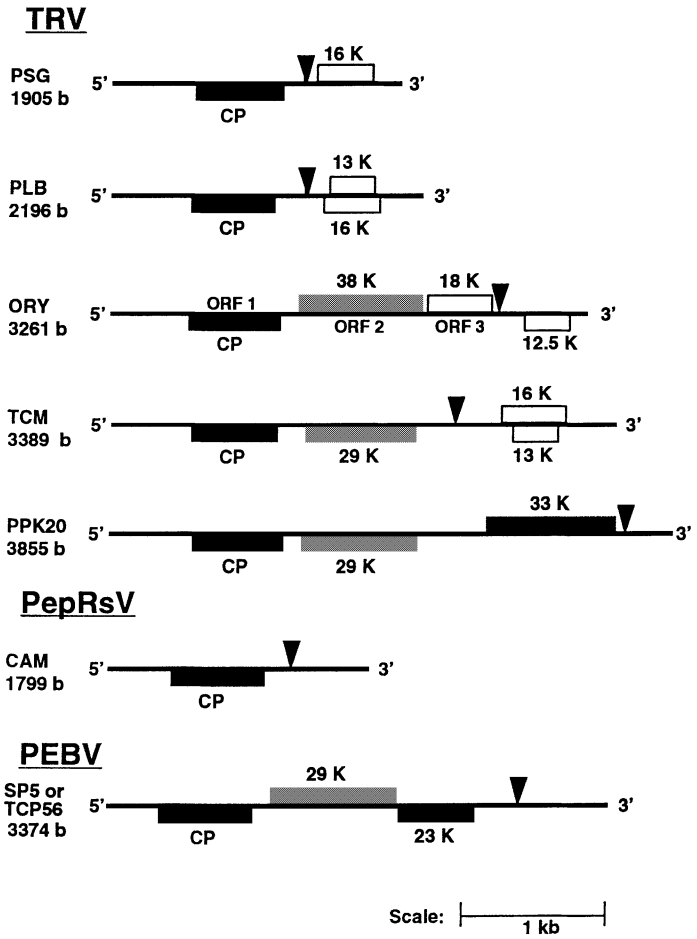

Fig. 2. Tobravirus RNA-2 genome organizations, including TRV-ORY (3261 nucleotides) (GenBank accession no. AF034621). Downward arrow indicates start of sequence identical to the 3′-end of RNA-1. Different shadings correspond to different regions of homology

the authentic start codon based on alignment with other Tobravirus coat proteins. In vitro translation of RNA-2 resulted in synthesis of three proteins corresponding to the 5′-proximal three ORFs, and no protein corresponding to the putative open reading frame derived from RNA-1 was observed (data not shown). Three ORFs in RNA-2 are also found in PEBV-SP5 [9], -TCP56, and TRV-PPK20 [14] (Fig. 2). The RNA-2 sequences of TRV-PSG, -PLB, and -TCM RNA-2 included in Fig. 2, were also compared previously by Angenant et al. [3]. The protein products of ORF2 of tobraviruses, while of low homology at the nucleotide level, have regions of similarity across different tobraviruses at the amino acid level (data not shown). The predicted protein sequence of RNA-2 ORF2 had limited regions of similarity with PEBV-SP5, TRV-TCM, and -PPK20. There was no indication of the possible function of ORF2, based on TFASTA searches. Results of BLAST searches suggested that protein products of ORF2 and ORF3 shared domains of unrelated membrane-binding proteins (data not shown). The ORF3 of RNA-2

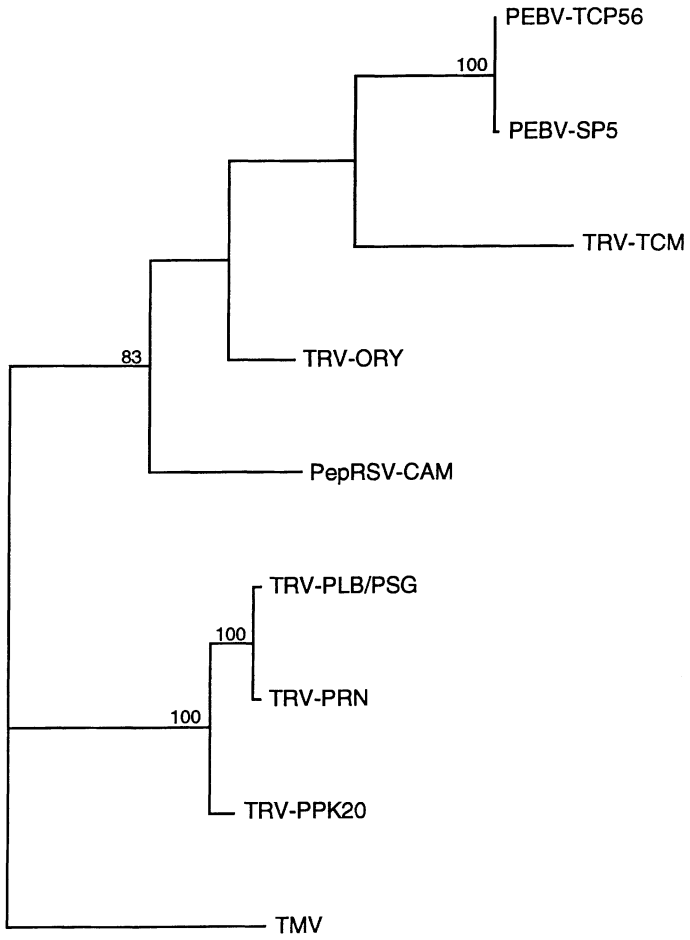

Fig. 3. Phylogenetic inference by maximum parsimony analysis of coat protein amino acid sequences of Tobraviruses. Branch lengths are proportional to distance, while vertical position can be arbitrary. Numbers on some branches are the percentage support of branching based on bootstrap analysis. Bootstrap percentages less than 50% are not shown. This is a rooted tree with tobacco mosaic virus (TMV) as the outgroup

of TRV-PPK20 was larger than ORF3 of TRV-ORY, but shared no homology with ORF3 of TRV-ORY, and thus the function of these open reading frames is uncertain.

In a pairwise alignment of the predicted CP sequences of tobraviruses, TRV-ORY had the highest levels of identity and similarity with PepRSV-CAM followed by PEBV-SP5 (Table 2). A phylogenetic tree was constructed based on the maximum parsimony of CP amino acid sequences of tobraviruses and, as an outgroup, tobacco mosaic virus. The tree (Fig. 3), in general agreement with pair-wise comparisons (Table 2), indicated that tobravirus coat proteins formed two phylogenetic clusters. TRV-ORY was part of a cluster that includes PepRSV, TRV-TCM, and PEBV (Fig. 3). TRV isolates PPK20, PSG and PRN, all from The Netherlands, are phylogenetically distinct from other tobravirus isolates and appear to form a subgroup.

Table 2. Percent identity/similarity determined by pair-wise alignment of primary sequences of the coat proteins of six TRV isolates and one each of PEBV and PepRSV isolates

	PepRSV-CAM	PEBV-SP5[a]	TRV-TCM	TRV-PPK20	TRV-PSG[b]
TRV-ORY	69.9/79.6	59.1/67.7	44.9/61.7	46.7/62.6	44.1/60.5
PepRSV-CAM		52.9/64.8	44.3/61.1	46.8/63.2	46.6/64.1
PEBV-SP5[a]			45.0/60.5	38.4/54.7	35.4/53.4
TRV-TCM				38.6/53.5	36.3/52.5
TRV-PPK20					88.2/94.1

[a] PEBV-SP5 and -TCP56 differ by one amino acid
[b] TRV-PSG and -PLB are identical

The RNA-1 of tobravirus isolates studied to date is highly conserved, while RNA-2 is highly variable. There is evidence for natural genetic recombination among TRV isolates, as appears to be the case of TRV-PSG and -PLB, which have identical CPs but different 3′-ends [3]. Recombination may be common in the tobraviruses [10, 24, 25]. The observed diversity of tobravirus RNA-2, combined with phylogenetic analysis of their CP's, indicates that taxonomic determinations must utilize multiple criteria.

Acknowledgements

This work was supported in part by a grant from the Idaho Potato Commission. We gratefully appreciate the assistance of Steve Thompson of the VADMS Center, Washington State University, Pullman, WA. Finally, we are grateful for the assistance and critical comments of Dr. P. J. Shiel in carrying out this work and to Drs. M. Wiese and W. Chun for a critical review of this manuscript.

References

1. Allen TC (1963) A strain of tobacco rattle virus from Oregon grown potatoes. Plant Dis Rep 47: 920–923
2. Angenant GC, Linthorst HJM, van Belkum AF, Cornelissen BJC, Bol JF (1986) RNA 2 of tobacco rattle virus strain TCM encodes an unexpected gene. Nucleic Acids Res 14: 4 673– 4 682
3. Angenent GC, Posthumus E, Brederode FT, Bol JF (1989) Genome structure of tobacco rattle virus strain PLB: Further evidence on the occurrence of RNA recombination among tobraviruses. Virology 171: 271–274
4. Corbett MK (1967) Aster ringspot virus as a strain of tobacco rattle virus. Phytopathology 57: 198–202
5. Cornelissen BJC, Linthorst HJ, B'rederode HJM, Bol JF (1986) Analysis of the genome structure of tobacco rattle virus PSG. Nucleic Acids Res 14: 2 157–2 169
6. Davis JR, Allen TC (1975) Weed hosts of the tobacco rattle virus in Idaho. Am Pot J 52: 1–8
7. Ghabrial SA, Lister RM (1973) Coat protein and symptom specification in tobacco rattle virus. Virology 52: 1–12
8. Gold AH, Grogan R, Bardin R (1963) An apparent strain of tobacco rattle virus associated with a yellow spotting of lettuce. Phytopathology 53: 1 139

9. Goulden MG, Lomonossoff GP, Davies JW, Wood KR (1990) The complete nucleotide sequence of PEBV RNA2 reveals the presence of a novel open reading frame and provides insights into the structure of tobraviral subgenomic promoters. Nucleic Acids Res 18: 4 507–4 512
10. Goulden MG, Lomonossoff GP, Wood KR, Davies JW (1991) A model for the generation of tobacco rattle virus (TRV) anomalous isolates: pea early browning virus RNA-2 acquires TRV sequences from both RNA-1 and RNA-2. J Gen Virol 72: 1 751–1 754
11. Guilford PJ, Ziegler-Graff V, Baulcombe D (1991) Mutation and replacement of the 16-kDa protein gene in RNA-1 of tobacco rattle virus. Virology 182: 607– 614
12. Hamilton WDO, Boccara M, Robinson DJ, Baulcombe DC (1987) The complete nucleotide sequence of tobacco rattle virus RNA-1. J Gen Virol 68: 2 563–2 575
13. Harrison BD, Woods RD (1966) Serotypes and particle dimensions of tobacco rattle viruses from Europe and America. Virology 28: 610 – 620
14. Hernández C, Mathis A, Brown DJF, Bol JF (1995) Sequence of RNA 2 of a nematode-transmissible isolate of tobacco rattle virus. J Gen Virol 76: 2 847–2 851
15. Joshi CP (1987) An inspection of the domain between putative TATA box and translation start site in 79 plant genes. Nucleic Acids Res 15: 6 643–6 653
16. Le Gouill C, Dery CV (1991) A rapid procedure for the screening of recombinant plasmids. Nucleic Acids Res 19: 6 655
17. Lister RM, Bracker CE (1969) Defectiveness and dependence in three related strains of tobacco rattle virus. Virology 37: 262–275
18. Locatelli EA, Allen TC, Koespell PA, Appleby AP (1978) Diagnosis of tobacco rattle viruses in weed and rotation crops in potato fields. Am Pot J 55: 249–257
19. MacFarlane SA, Davies JW (1992) Plants transformed with a region of the 201-kiloldalton replicase gene from pea early browning virus RNA1 are resistant to virus infection. Proc Natl Acad Sci USA 89: 5 829–5 833
20. MacFarlane SA, Taylor SC, King DI, Hughes G, Davies JW (1989) Pea early browning virus RNA1 encodes four polypeptides including a putative zinc-finger protein. Nucleic Acids Res 17: 2 245–2 260
21. Miki T, Okada Y (1970) Comparative studies on some strains of tobacco rattle virus. Virology 42: 993–998
22. Offord RE, Harris JI (1965) The protein sub-unit of tobacco rattle virus. In: Proc. 2nd FEBS Meeting, 1965
23. Robinson DJ (1983) RNA species of tobacco rattle virus strains and their nucleotide sequence relationships. J Gen Virol 64: 657– 665
24. Robinson DJ (1994) Sequences at the ends of RNA-2 of I6, a recombinant tobravirus. Arch Virol [Suppl] 9: 245–251
25. Robinson DJ, Hamilton WDO, Harrison BD, Baulcombe DC (1987) Two anomalous tobravirus isolates: evidence for RNA recombination in nature. J Gen Virol 68: 2 551–2 561
26. Robinson DJ, Harrison BD (1989) Tobacco rattle virus. AAB Descriptions of Plant Viruses, No 346
27. Robinson DJ, Harrison BD (1985) Unequal variation in the genome parts of tobraviruses and evidence for the existence of three separate viruses. J Gen Virol 66: 171–176
28. Robinson J, Raschke JH (1977) Inactivation of tobacco rattle virus by EDTA, and the role of divalent metal ions in the stability of the virus. J Gen Virol 34: 547–559
29. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, New York
30. Semancik JS (1966) Purification and properties of tobacco rattle virus from pepper in California. Phytopathology 56: 1 190 –1 193

31. Swofford DL (1993) PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1. Illinois Natural History Survey, Champaign, IL
32. Van Dun CMP, Bol JF (1988) Transgenic potato plants accumulating tobacco rattle virus coat protein resist infection with tobacco rattle virus and pea early browning virus. Virology 167: 649–652
33. Wilson T (1993) Strategies to protect crop plants against viruses: Pathogen-derived resistance blossoms. Proc Natl Acad Sci USA 90: 3134–3141
34. Wyatt SD, Druffel K, Berger PH (1993) In vitro destabilization of plant viruses and cDNA synthesis. J Virol Methods 44: 211–220

Authors' address: Dr. P. H. Berger, Division of Plant Pathology, Department of Plant, Soil and Entomological Sciences, University of Idaho, Moscow, ID83844-2339, U.S.A.