



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Comparison of the 5' and 3' Termini of Tomato Ringspot Virus RNA1 and RNA2: Evidence for RNA Recombination¹

M. E. ROTT,* J. H. TREMAINE,† AND D. M. ROCHON†²

**Department of Plant Science, University of British Columbia, Vancouver, British Columbia V6H 2A2, Canada; and †Agriculture Canada Research Station, 6660 North West Marine Drive, Vancouver, British Columbia V6T 1X2, Canada*

Received February 4, 1991; accepted August 9, 1991

The sequences of the 5' terminal 1140 and 3' terminal 1546 nt of tomato ringspot virus (TomRSV) RNA1 have been determined. These sequences share a high degree of nucleotide sequence similarity with the previously determined TomRSV RNA2 sequence. Eighty-eight percent of the 5' terminal 907 nt of TomRSV RNA1 and RNA2 contain identical nucleotide residues; the first 459 nt are identical at all positions, whereas the next 447 nt are identical at only 75.8% of the nucleotide positions. The region of similarity includes not only the 5' nontranslated leader but also sequence probably encoding polyproteins. The 3' terminal 1533 nt of TomRSV RNA1 and RNA2 are identical and are noncoding. The sequences common to RNA1 and RNA2 account for almost 35% of the total genomic sequence. It is possible that the similar sequences at both ends of TomRSV RNA1 and RNA2 are a result of recombination between these two genomic RNA components. © 1991 Academic Press, Inc.

Recombination, involving the exchange of genetic information between genomic RNA molecules of RNA viruses, appears to play a number of important functions during the evolution and life cycle of RNA viruses (for reviews see 1, 2). The importance of nonhomologous RNA recombination in the generation of evolutionary diversity can be seen in the modularity of viral-encoded genes. In addition, many RNA viruses produce defective-interfering RNAs which may be a result of recombination. Homologous RNA recombination has been observed among the picornaviruses (3), coronaviruses (4), and bromoviruses (5) and is suspected to occur in the tobnaviruses (6). It has been suggested that homologous recombination in the picornaviruses and coronaviruses is important for repairing defective genomes. In addition, coronaviruses undergo site-specific recombination to express downstream genes from leader-primed subgenomic transcripts (see 7).

Tomato ringspot virus (TomRSV) is a member of the nepovirus group (8). Nepoviruses consist of 28-nm spherical particles composed of 60 copies of a single coat protein species and two separately encapsidated genomic RNA components. Nepoviruses share similarities in genomic structure and translational strategies with the plant como- and potyviruses as well as the animal picornaviruses (9). Previously, we reported that the 3' termini of TomRSV RNA1 and RNA2 share an

extended region of nucleotide sequence similarity, as determined by restriction enzyme cleavage maps and hybridization analysis (10). We report that extensive nucleotide sequence similarity also exists between the 5' termini of RNA1 and RNA2. The possibility that these repeated sequences may facilitate replication of TomRSV RNA, perhaps through recombination, will be discussed.

Two cDNA clones derived from TomRSV RNA1 (see Fig. 1) were used to determine the 5' and 3' terminal sequences of TomRSV RNA1. Clone J27, which has been previously described (10), was used to sequence the 3' terminal 1546 nt of TomRSV RNA1 and clone 25P6 was used to sequence the 5' terminal 1108 nt. 25P6 was obtained in essentially the same manner as J27 except that random priming was used for first-strand cDNA synthesis. Subcloning, sequencing, sequence assembly, and analysis were essentially as described previously (11).

Clone 25P6 was found to hybridize to both TomRSV RNA1 and RNA2 in Northern hybridization studies (12) (data not shown). However, the restriction enzyme map of 25P6 matched that of the RNA1-specific clone B54 (10) (Fig. 1) but was distinct from that of the RNA2-specific clone O35 (see 11). To confirm that 25P6 was derived from RNA1, the region 5' to the *Hind*III site of B54 was partially sequenced in one direction and was found to be identical to the corresponding region obtained from 25P6 (data not shown). The 5' terminal sequence not encoded by 25P6 was determined by dideoxynucleotide sequence analysis using TomRSV RNA as a template and a specific oligonucleo-

¹ Sequence data from this article have been deposited with the EMBL/Genbank Data Libraries under Accession Numbers M27935 and M73822.

² To whom requests for reprints should be addressed.

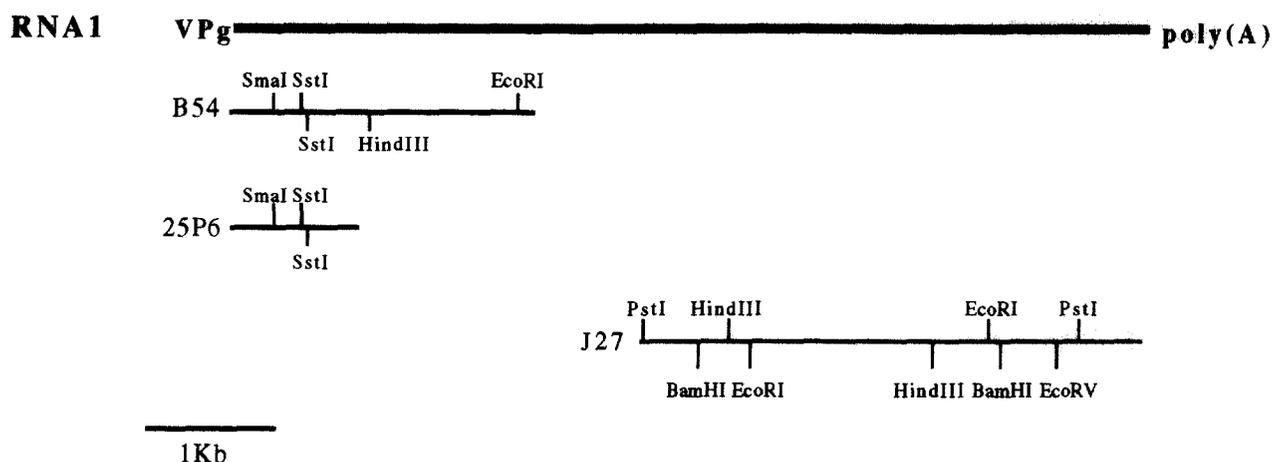


Fig. 1. Genomic location and restriction enzyme maps of clones used to determine the 5' and 3' terminal nucleotide sequences of TomRSV RNA1.

tide primer (5' GCCTTCGATGGCAACC 3') complementary to nucleotide positions 115–130 as described previously (11). The sequencing gel obtained from this experiment terminated with two strong stop points in all four lanes which likely correspond to the first and second nucleotides of TomRSV RNA1 (each denoted N in Fig. 2A).

Nucleotide sequence comparison of the 5' termini of TomRSV RNA1 and RNA2 revealed that within the first 907 nt, 88.8% of the nucleotide positions are identical. The first 459 nt of TomRSV RNA1 and RNA2 are identical. This region includes the 5' noncoding regions of RNA1 and RNA2 as well as two potential in-frame translation initiation sites at AUG₇₈ and AUG₄₄₁. Beginning from the first potential in-frame initiation site at AUG₇₈, the N-terminal regions of the TomRSV RNA1 and RNA2 polyproteins are identical for the first 132 amino acids, and of the next 145 amino acid residues 75.3% of the positions are identical (Fig. 3A). It is perhaps significant that the second in-frame initiation site at AUG₄₄₁ occurs shortly after the point where the homology between the RNA1 and the RNA2 polyproteins becomes less than perfect. As described in a previous paper (11), AUG₄₄₁ is in a better Kozak context for the initiation of translation than AUG₇₈ (13). Since it is unknown whether AUG₇₈ and/or AUG₄₄₁ act as initiation sites for translation, it cannot be said whether it is the conservation of amino acid or nucleotide sequence between AUG₇₈ and AUG₄₄₁ on RNA1 and RNA2 which is most significant. The deduced amino acid sequence of the regions beginning shortly after AUG₄₄₁ on RNA1 and RNA2 could be aligned with the deduced amino acid sequences encoded by the 5' terminal region of RNA1 of the nepoviruses tomato blackring (TBRV) and grapevine chrome mosaic (GCMV) (14, 15) (Fig. 3B). The fact that these regions of similarity are present only

at the N-termini of the TBRV and GCMV RNA1-encoded polyproteins but are present at the N-termini of both TomRSV RNA1- and RNA2-encoded polyproteins suggests that a large portion of coding and noncoding sequences at the 5' terminus of TomRSV RNA1 have been duplicated and are now present at the 5' termini of both RNA1 and RNA2. The function of this coding region in RNA1 is unknown, however, it has been suggested that it may have a role in proteolytic processing of the viral-encoded polyprotein (15). It is interesting that *in vitro* translation studies of cherry leafroll virus (CLR) RNA2 (16), another nepovirus with a large RNA2 component (17), resulted in proteolytic processing of the RNA2 polyprotein in the absence of RNA1-encoded protease. It is possible that a subgroup of nepoviruses with large RNA2 components, which would include TomRSV and CLR (18), may encode another protease on RNA2 which is involved in proteolytic processing. However, it is possible that proteolytic processing of the CLR RNA2 polyprotein may not be due to a specific viral-encoded protease.

The 3' noncoding regions of TomRSV RNA1 and RNA2 are almost identical for 1533 nt (excluding the 3' poly(A) tail sequence) with only three nucleotide differences at positions 703, 720, and 770 as shown in Fig. 2B. These sequences are preceded by 13 and 17 nt of noncoding sequence which are unique to RNA1 and RNA2, respectively (UAAAUCCUCUUUG and UAAGUUGGCUUCCUGAA, underlined nucleotides indicate stop codons for the large ORFs of RNA1 and RNA2, respectively). When we first reported the sequence similarity at the 3' termini of TomRSV RNA1 and RNA2 (10), we proposed that extensive 3' terminal identity between RNA1 and RNA2 may be characteristic for other nepoviruses with large RNA2 components. This has been confirmed for the nepovirus CLR (19) which

(A)

```

*****
RNA1 1  MSSICFAGGNHARLPSKAAYYRAISDRELDREGRFPCGCLAQYTVQAPPPAKTQEKAVGRSADLQ
RNA2 1  MSSICFAGGNHARLPSKAAYYRAISDRELDREGRFPCGCLAQYTVQAPPPAKTQEKAVGRSADLQ

*****
RNA1 66  KGNVAPLKKQRCDVVAVSGPPPELVYPARVQHRLDQPSKGPLAVPSAKQTSTAMEVVLSAEE
RNA2 66  KGNVAPLKKQRCDVVAVSGPPPELVYPARVQHRLDQPSKGPLAVPSAKQTSTAMEVVLSAEE

** ***** * * ***** * ** ***** * ***** * **
RNA1 131  AALTAPWLLCSYKSGVSSPPPMQTRQQFAAIKRRLVQKQQI IRELIRARKAAKYAFAARKKA
RNA2 131  AAITAPWLLRPCK-GEAPPPPLTQRQQFAALKRRLAVKQQI IREHIRARKAAKYAAIAKAKKA

** ** * ***** ***** ***** ***** ** *****
RNA1 196  AAVAAQKARAEAPRLAAQKAAIAKILRDRQLVSLPPPPPSAARLAAEAEELASKSASLQRLKAFH
RNA2 195  AALAAVKAQAEAPRLAAQKAAISKILRDRDVAALPPPPPSAARLAAEAEELASKAESLRRLLKAFK

***** ** * * **
RNA1 261  RANRVRPVLNNSFPSP
RNA2 260  TFSRVRPALNTSFPSP
    
```

(B)

```

* ^ * * ^ ^* ^ ^ * * ^ *
TomRSV RNA1 138  LLCYKSGVSSPPPMQTRQQFAAIKRRLVQKQQI IREL---IRARKAAKYAFAARKKAAAV
TomRSV RNA2 138  LLRPCK-GEAPPPPLTQRQQFAALKRRLAVKQQI IREH---IRARKAAKYAAIAKAKKAAAL
TBRV RNA1 189  KLNKAKALGEAHRSAVARAQAKAEVLEFEPSPQIQRALEAQIFADRLSRKYAALTARVRAKRA
GCMV RNA1 189  KLTKANALGAAHRSAVATAQAKAEVLEFEPSPAHIQI AVKAHIFAEKLSRKYADLTAQVRARRA

** ^ * ^ ^ * ^ ^ ^ ^ ^ * ^
TomRSV RNA1 200  AAQKARAEAPRLAAQKAAIAKILRDRQLVSLPPPPPSAARLAAEAEELASKSASLQRLKAF
TomRSV RNA2 199  AAVKAAQAEAPRLAAQKAAISKILRDRDVAALPPPPPSAARLAAEAEELASKAESLRRLLKAF
TBRV RNA1 255  AARELREKELFLETQDLLNAPLLPMEKVGIERKY-RKVRPTGSNVTSTPKPNVLENLCPF
GCMV RNA1 255  AARDLRAKEIYLEIVDLLGAPLLSIPQIKIKGKYLR--RSVAEVEVPHTRNPMaelvpy
    
```

Fig. 3. (A) Alignment of the N-terminal regions of the TomRSV RNA1 and RNA2 polyproteins. Asterisks indicate amino acids common to both sequences. Numbering to the left of each line refers to amino acid positions in TomRSV RNA1 and RNA2 polyproteins. (B) Alignment of the N-terminal regions of TomRSV RNA1 and RNA2, and TBRV and GCMV RNA1 polyproteins. Asterisks indicate amino acids common to all four sequences, while a carat (^) indicates that three of four amino acids at that position are identical. Numbering to the left of each line refers to amino acid position in each nepovirus polyprotein.

shows 3' noncoding sequence identity between RNA1 and RNA2 for a length which is similar to that found in TomRSV. However, very little sequence similarity is detectable between the 3' noncoding regions of TomRSV and CLRV (19). The 3' noncoding regions of the two RNA components for several other nepoviruses are also identical but much shorter (less than 300 nt) (14, 15). Extensive sequence similarity at the 3' termini has also been reported for members of the tobnavirus group (20, 21) and includes both potential coding sequences and noncoding sequences.

In summary, the total amount of duplicated sequences between TomRSV RNA1 and RNA2 as well as within RNA2 (see 11) accounts for almost 35% of the total genomic sequence. The extensive amount of nucleotide sequence identity at the 5' and 3' termini of TomRSV RNA1 and RNA2 may be required for recognition by a highly selective replicase. It is also possible that RNA recombination is responsible for maintaining nucleotide sequence identity at the 5' and 3' termini of RNA1 and RNA2. RNA recombination has been postulated to explain the duplication of 820 nt at the 3' ter-

Fig. 2. Nucleotide sequence and deduced amino acid sequence of the 5'(A) and 3'(B) regions of TomRSV RNA1. (A) The first two nucleotides which could not be determined from the sequencing gel are each represented by an N. The second in-frame AUG which is a potential initiation site for translation is underlined. Nucleotides are numbered on the left beginning at the first N. Amino acids are numbered on the right and begin at the first M. Numbering in (B) begins at the termination codon (UAA) for the RNA1-encoded long ORF as determined from clone J27. Three nucleotide substitutions in the 3' noncoding region of RNA2 (11) compared to RNA1 are shown below the sequence of RNA1.

mini of RNA1 and RNA2 of the tobavirus tobacco rattle virus (TRV) strain PLB (6). However, after several passages of a pseudorecombinant consisting of RNA1 from TRV strain TCM and RNA2 from strain PLB, which differ at their 3' termini in 39 of 820 nt, a recombinant RNA molecule was not detected. Not only was RNA recombination in this system not detected experimentally, but the viability of a pseudorecombinant consisting of heterologous 3' termini in RNA1 and RNA2 suggests that in this system precise 3' terminal sequences in RNA1 and RNA2 are not essential for replication. It has been suggested that the high frequency of homologous recombination in the animal picornaviruses is important for removing deleterious mutations introduced by the poor fidelity of the RNA replicase (22). Since it has been suggested that the picornavirus replicative machinery may not function efficiently in *trans*, nondefective genes, located on different RNA molecules which possess errors in genes involved in replication, can only be utilized after genetic recombination with an RNA molecule which encodes functional replicative genes (23, 24). It is possible that, in TomRSV, replication begins in *cis* with RNA1 and that *trans* replication of RNA2 occurs only following disassociation and reassociation of the initial negative-strand transcript with the corresponding region in RNA2. A similar mechanism involving recombination could account for the sequence conservation observed between the 5' termini of RNA1 and RNA2. Such a mechanism has recently been proposed for leader-primed generation of subgenomic RNAs in coronaviruses (see 7). The size of the duplicated sequences in TomRSV may be the minimum required to facilitate efficient replication through RNA recombination between RNA1 and RNA2. Alternatively, the entire length may not be required for recombination but may serve other important functions in addition to a postulated role in recombination. We are planning further experiments to determine the biological significance of the repeated sequences between TomRSV RNA1 and RNA2 and whether they are in fact involved in RNA recombination during replication.

ACKNOWLEDGMENTS

The technical assistance of L. Lee is gratefully acknowledged. This work was partially supported by an NSERC operating grant awarded to J. H. Tremaine.

REFERENCES

1. KING, A. M. Q., ORTLEPP, S. A., NEWMAN, J. W. I., and McCAHON, D., In "The Molecular Biology of the Positive Strand RNA Viruses" (D. J. Rowlands, M. A. Mayo, and B. W. J. Mahy, Eds.), pp. 129-152. Academic Press, London, 1987.
2. KING, A. M. Q., In "RNA Genetics" (E. Domingo, J. J. Holland, and P. Ahlquist, Eds.), pp. 149-165. CRC Press, Boca Raton, FL, 1988.
3. HIRST, G. K., *Cold Spring Harbor Symp. Quant. Biol.* **27**, 303-308 (1962).
4. MAKINO, S., KECK, J. G., STOHLMAN, S. A., and LAI, M. M. C., *J. Virol.* **57**, 729-737 (1986).
5. BUJARSKI, J. J., and KAESBERG, P., *Nature* **321**, 528-531 (1986).
6. ANGENENT, G. C., POSTHUMUS, E., BREDERODE, F. T., and BOL, J. F., *Virology* **171**, 271-274 (1989).
7. LAI, M. M. C., *Annu. Rev. Microbiol.* **44**, 303-333 (1990).
8. HARRISON, B. D., and MURANT, A. F., In "Descriptions of Plant Viruses," No. 185. Commonwealth Mycological Institute and Association of Applied Biologists, Surrey, England, 1977.
9. GOLDBACH, R., *Microbiol. Sci.* **4**, 197-202 (1987).
10. ROTT, M. E., ROCHON, D. M., and TREMAINE, J. H., *J. Gen. Virol.* **69**, 745-750 (1988).
11. ROTT, M. E., TREMAINE, J. H., and ROCHON, D. M., *J. Gen. Virol.* **72**, 1505-1514 (1991).
12. VRATI, S., MANN, D. A., and REED, K. C., *Mol. Biol. Rep.* **1**, 1-4 (1987).
13. KOZAK, M., *Cell* **44**, 283-292 (1986).
14. GREIF, C., HEMMER, O., and FRITSCH, C., *J. Gen. Virol.* **69**, 1517-1529 (1988).
15. LEGALL, O., CANDRESSE, T., BRAULT, V., and DUNEZ, J., *Nucleic Acids Res.* **17**, 7795-7807 (1989).
16. PONZ, F., ROWHANI, A., MIRCETICH, S. M., and BRUENING, G., *Virology* **160**, 183-190 (1987).
17. MURANT, A. F., TAYLOR, M., DUNCAN, G. H., and RASCHKÈ, J. H., *J. Gen. Virol.* **53**, 321-332 (1981).
18. MARTELLI, G. O., In "Nematode Vectors of Plant Viruses" (Lamberti, Taylor, and Seinhorst, Eds.). Plenum, London/New York, 1975.
19. BORJA, M. J., ROWHANI, A., ALEXANDER, D., FRIAS, A., BRUENING, G., and PONZ, F., *VIIIth Int. Cong. Virol.* (1990). [Abstract No. P86-007]
20. BERGH, S. T., KOZIEL, M. G., HUANG, S., THOMAS, R. A., GILLEY, D. P., and SIEGEL, A., *Nucleic Acids Res.* **13**, 8507-8518 (1985).
21. ANGENENT, G. C., LINTHORST, H. J. M., BELKUM, A. F., VAN CORNELISSEN, B. J. C., and BOL, J. F., *Nucleic Acids Res.* **14**, 4673-4682 (1986).
22. STEINHAEUER, D. A., and HOLLAND, J. J., *Annu. Rev. Microbiol.* **41**, 409-433 (1987).
23. BERNSTEIN, H. D., SARNOV, P., and BALTIMORE, D., *J. Virol.* **60**, 1040-1049 (1986).
24. KIRKEGAARD, K., and BALTIMORE, D., *Cell* **47**, 433-443 (1986).