The nucleotide sequence of a satellite RNA associated with strawberry latent ringspot virus

S. Kreiah, J. I. Cooper and G. Strunk

NERC Institute of Virology and Environmental Microbiology, Mansfield Road, Oxford OX1 3SR, U.K. and
Max-Planck Institute for Bio-Physical Chemistry, Nikolausberg, 3400 Goettingen, Germany

The nucleotide sequence of a satellite RNA associated with a strawberry isolate (H) of strawberry latent ringspot nepovirus (SLRSV) was determined from cDNA copies and the 5' end sequence was deduced from directly sequenced virion RNA. At the 3' end a poly(A) sequence was identified. A long open reading frame encoding a polypeptide of 331 amino acids (M r 36 488) was determined. Sequence comparisons showed that SLRSV satellite RNA has no extensive homology with other sequences in the GenEmbl and Swiss-Prot databases.

Strawberry latent ringspot (SLRSV) is provisionally classified in the nepovirus genus of the proposed family Comoviridae (Martelli, 1992). The virus is transmitted by longidorid nematodes (Lister, 1964) and has a bipartite positive sense ssRNA genome (Mayo et al., 1974) contained within isometric particles.

Although extra-genomic RNAs have not been reported in comoviruses or fabaviruses, several nepoviruses encapsidate RNAs which are satellites (satRNAs) as defined by Murant & Mayo (1982). Satellite RNAs have been classified according to their physical and messenger properties. Two types (B and D) of nepovirus satRNAs have been recognized (Martelli, 1992). The Hampshire strawberry isolate H of SLRSV (SLRSV-H) contains a type B satRNA of approx. 0.4 x 10^6 M r (Mayo et al., 1974; Gallitelli et al., 1982). No satRNAs were found in other strawberry isolates such as in isolate J (Mayo et al., 1974) or in isolates from peach or olive (Gallitelli et al., 1982). The encapsidated satRNA in SLRSV-H is message sense and, like the genomic RNAs, is 3' polyadenylated and has a 5' VPg (Mayo et al., 1979, 1982). In this paper we report the nucleotide sequence of this satRNA and compare the data with those of other satRNAs.

The SLRSV-H (culture T39; obtained from the Scottish Crop Research Institute) was propagated as single lesions in inoculated leaves of Chenopodium murale before multiplication in, and virion purification from, systemically invaded leaves of Chenopodium quinoa. Butanol–chloroform was used for virus purification as described by Harrison & Nixon (1960). RNA was extracted from virions as described by Massalski & Cooper (1986). Virion-derived RNA (three size classes) was used as a template for cDNA synthesis with Moloney murine leukaemia reverse transcriptase and oligo(dT)_{12-18} (cDNA synthesis kit; Pharmacia). The Klenow fragment of DNA polymerase was used to create blunt-ended cDNA before EcoRI adaptors were added and the resulting construct was ligated into the plasmid pT7T3 18U (Pharmacia) and used to transform Escherichia coli XL-1Blue (Stratagene). All subsequent molecular manipulations were as described by Sambrook et al. (1989). Recombinants, selected after extensive restriction enzyme analysis, were used as nick-translated [α-32P]dATP-labelled probes in Northern hybridization against virion-derived RNA. A satRNA clone of 1088 nucleotides (nt) was inserted in both orientations into pT7T3 18U and used to generate nested deletion libraries by controlled digestion with exonuclease III. Single-stranded DNA was generated with the helper phage M13KO7 (Pharmacia): dsDNA for sequencing was produced as described by Chen & Seeberg (1985). The sequences were determined by dideoxynucleotide chain termination (Sanger et al., 1977) using either the 'universal' sequencing primer or the M13 reverse sequencing primer (Pharmacia) and the modified DNA polymerase Sequenase (U.S. Biochemical Corporation). A synthetic oligodeoxynucleotide complementary to residues 86 to 112 was synthesized by the phosphotriester method with an Applied Biosystem 380B DNA synthesizer. After deprotection at 55 °C, the primer was purified by the

The nucleotide sequence data reported in this paper will appear in the EMBL database under the accession number X69826.
satRNA sequence had an in-frame AUG codon at positions 41 to 43, according well with the ribosome recognition sequence identified by Kozak (1986; i.e. there was a G residue at position +4 and an A at position −3). The termination codon (UAG) was located between 1034 to 1036 nt, 82 bases from the 3'-terminal run of 12 A residues. In this SLRSV satRNA sequence, the ORF was preceded by 40 nt and, in this respect, resembled the satRNAs of tomato black ring virus (TBRV) strains C, E and G (Hemmer et al., 1987). In the satRNA of the lilac isolate of arabis mosaic virus (ArMV-L; Liu et al., 1990) the corresponding upstream sequence was 14 nt long. In its length, the 3'-non-coding region in the SLRSV satRNA sequence closely resembled those of the satRNAs of TBRV strains E and C (Hemmer et al., 1987), although it was longer than the equivalent region in the satRNA of ArMV (7 nt; Liu et al., 1990). The DOT PLOT, PILEUP and CLUSTAL comparison program (Staden, 1984, 1986) and GCG search programs (e.g. BLAST, FASTA, T-FASTA, WORDSEARCH; Devereux et al., 1984) did not reveal any significant homology (except at the 5' end), either at the nucleic or amino acid level, between the SLRSV satRNA sequence and satRNAs from the following: TBRV (Meyer et al., 1984), ArMV-L (Liu et al., 1990), grapevine fan leaf nepovirus (Fuchs et al., 1989), chicoory yellow mottle nepovirus (Rubino et al., 1990), or the type D satRNAs from ArMV (Kaper et al., 1988), or tobacco ringspot nepovirus (Buzayan et al., 1986) or the cucumber mosaic cucumovirus (type C; Collmer et al., 1983). We have determined the complete nucleotide sequence in cDNA clones representing the RNA 2 of SLRSV-H. Using these data, we confirmed our Northern hybridization results and showed that the base sequence in the satRNA was different from that of RNA 2 [except for the poly(A) region and the 5' end].

The theoretical coding capacity of the ORF in the satRNA from SLRSV is for 331 amino acids, equivalent to a protein of $M_r$ 36488. The predicted protein is rich in basic amino acids (e.g. 30 arginine and 20 lysine residues) which may neutralize the net negative charge on RNA (Rice et al., 1985). The C-terminal region of the predicted protein is hydrophobic and might anchor it in membranes (Coia et al., 1988). We had suspected that SLRSV-H and ArMV-L (which are transmitted by the same nematode species and often co-exist in the same plants) might share the same satRNAs but now know that these satRNAs are distinct. Similarities in size and structure of type B satellites suggest that, irrespective of the virus with which they are now associated, they all shared a common progenitor.

Fig. 1. The RNA sequence corresponding to virion-derived RNA 3 from SLRSV and (below) the deduced amino acid sequence.

method of Sawadogo & Van Dyke (1991), and used with reverse transcriptase (Life Sciences) to sequence the 5' end of RNA 3 by primer extension. In addition, RNA sequencing was performed using fluorescent dideoxynucleotide terminators, the sequence upstream of the primer was confirmed and the 5'-terminal six bases were identified. No evidence of sequence heterogeneity was observed. In common with many other RNAs that have VPg, the 5'-most nucleotide in the RNA 3 of SLRSV is a U (Fig. 1). Furthermore, the terminal base sequence UUGAAAAGAA GAA fits with the consensus recognized by Fuchs et al. (1989).

The longest open reading frame (ORF) of the putative
Department of the Environment. The senior author was funded by NERC and the Syrian Government.

References


(Received 12 November 1992; Accepted 12 January 1993)