Segment 5 of the rice dwarf virus genome encodes a protein highly conserved within the phytoreoviruses

Nagao Hayashi¹ and Yuzo Minobe²

¹National Agriculture Research Center and ²National Institute of Agrobiological Resources, Kannondai, Tsukuba, Ibaraki 305, Japan

The complete nucleotide sequence of segment 5 (S5) of the rice dwarf virus (RDV) genome was determined. RDV S5 is 2571 bp in length and has a single long open reading frame which encodes a polypeptide of 801 amino acids (Mr, 90495). When compared to the wound tumor virus genome S5, there was 56.9% and 52.8% similarity in the nucleotide and amino acid sequences, respectively. This high similarity suggests that the S5 proteins of these viruses are functionally similar.

Rice dwarf virus (RDV) (Reoviridae; Phytoreovirus subgroup), like other phytoreoviruses such as rice gall dwarf virus (RGDV) (Omura & Inoue, 1985) and wound tumour virus (WTV) (Boccardo & Milne, 1984), has particles that are composed of 12 segments of dsRNA (Fujii-Kawata et al., 1970; Kimura et al., 1987) and seven structural proteins (Kimura et al., 1987; Nakata et al., 1978). These proteins include those which have polymerase activity (Kodama & Suzuki, 1973) and those which may control the severity of the symptoms induced in infected plants (Kimura et al., 1987). Here, we report the complete nucleotide sequence of segment 5 of the RDV genome (RDV S5) and discuss the function of its predicted translational product. During the preparation of this manuscript, an incomplete sequence of RDV S5 was published (Suzuki et al., 1989).

Particles of RDV (kindly supplied by Dr T. Omura, National Agriculture Research Center, Tsukuba, Japan) were purified as described by Omura et al. (1982). Complementary DNA copies of RDV genome segments were synthesized and cloned into pBR322, essentially as described by Cashdollar et al. (1982). Transformants containing RDV-specific cDNA were selected by colony hybridization probes and plasmid DNA was extracted by the alkaline lysis method (Birnboim & Doly, 1979). Three recombinants were selected that had an insert of approximately 3.0 kb, which corresponds in size to the original full-length dsRNA (including the tails attached during the course of cloning). Nick-translation plasmid DNA from one of the three recombinants hybridized specifically with S5 in Northern blots (data not shown). The entire sequence of RDV S5 was determined by the dideoxynucleotide chain-termination method (Sanger et al., 1977) after subcloning of the appropriate restriction endonuclease fragments into M13mp18 or M13mp19 vectors (Messing et al., 1981). The entire sequence was confirmed by sequencing using 2'-deoxy-7-deazaguanosine triphosphate in place of dGTP (Mizusawa et al., 1986) and sequencing at least twice in each direction.

The complete nucleotide sequence of RDV S5 (Fig. 1) is 2571 bp long. The 3'-terminal sequence of UGAU 3' is the same as those of RDV S7 (Nakashima et al., 1990), S8 (Omura et al., 1989), S10 (Omura et al., 1988), S11 (Minobe, 1987) and S12 (Minobe, 1987) reported earlier. The 5'-terminal nucleotide sequence of 5' GGCAAA is the same as those of RDV S7 (Nakashima et al., 1990) and S8 (Omura et al., 1989). The three large clones were 5'-coterminal which suggests that the sequence is full length. Comparison with the sequence published by Suzuki et al. (1989) shows 47 transition and four transversion substitutions and six amino acid differences at positions 70, 190, 331, 371, 495 and 680. The length of the 5'-terminal non-coding region is 10 bp longer than that described by Suzuki et al. (1989), although the four 5'-terminal nucleotides and eight nucleotides adjacent to the first AUG were identical. This may be because we used a virus isolate different to that used by Suzuki et al. An inverted repeat was found near the termini, as shown in Fig. 2, as has also been observed in all segments of WTV (Anzola et al., 1987). A single long open reading frame (ORF) of 2403 nucleotides was found between position 27 and a UGA termination codon at position 2430 (Fig. 1). No other ORFs present in either the minus or plus-sense RNA strands exceeded 219 bp. The AUG at positions 27 to 29, the first potential initiation codon, has the sequence context AAAUGU; a at position -3 and a pyrimidine at position +4 are thought to be important as a ribosomal recognition sequence in
Fig. 2. Terminal sequence domains of RDV S5. The translational product of one large ORF is shown below the sequence. The terminal 5'-hexanucleotide and 3'-tetranucleotide sequences found in several other RDV genome segments are underlined. The asterisk indicates the termination codon.

Fig. 1. Nucleotide sequence of the plus-sense strand of RDV S5. The translational product of one large ORF is shown below the sequence. The terminal 5'-hexanucleotide and 3'-tetranucleotide sequences found in several other RDV genome segments are underlined. The asterisk indicates the termination codon.

eukaryotic mRNA (Kozak, 1987). Assuming that the first AUG is the initiator of protein synthesis, the S5 gene encodes a protein of 801 amino acids with an estimated Mr, 90495 (90-5K).

The S5 protein of RDV revealed strong identity to that of WTV throughout its amino acid sequence. The alignment of the S5 amino acid sequences of RDV and WTV is shown in Fig. 3. The nucleotide sequences and corresponding translational products are 56-9% and 52-2% identical, respectively, and the sequences are 83% similar if chemically similar amino acids are scored (Dayhoff et al., 1972). The alignment of the RDV and WTV translational products revealed that sequences in the C-terminal halves resemble each other more than do sequences in the N-terminal halves. This is largely because, firstly, the sequences in the N-terminal half
short communication

3083

Fig. 3. Alignment of the translational products of RDV S5 (top line) and WTV S5 (bottom line). Identical amino acids are indicated by colons; chemically similar amino acids, as described by Dayhoff et al. (1972), are indicated by periods.

Fig. 4. Alignment of the S5 segments of RDV and WTV showing possible frameshift deviation. Numbers refer to nucleotide and amino acid positions (see Fig. 1). Identical nucleotides are indicated by colons; dashes denote a nucleotide gap; codons are indicated by horizontal lines with the corresponding amino acid shown above or below these lines.

We would like to thank Dr Ilkka Havukkala for critical reading of the manuscript.

References


(Received 12 June 1990; Accepted 16 August 1990)