Nucleotide Sequence of Segment S10 of the Rice Dwarf Virus Genome

By TOSHIHIRO OMURA,* YUZO MINOBE and TSUNEKO TSUCHIZAKI
National Agriculture Research Center and 1National Institute of Agrobiological Resources,
Tsukuba Science City, Yatabe, Ibaraki 305, Japan

(Accepted 7 October 1987)

SUMMARY

DNA complementary to the tenth largest (S10) of the 12 genome segments of rice dwarf virus (RDV) was cloned and its sequence was determined. This genome segment is 1319 nucleotides in length and has a single long open reading frame extending for 1059 nucleotides from the first AUG triplet (residues 27 to 29). The predicted translation product comprises 352 amino acids and has a mol. wt. of 39094. RDV transcripts synthesized in vitro have the same polarity as the plus strand of the genome. Terminal sequences were (+) 5' GGUA---UGAU 3' and (−) 3' CCAU---ACUA 5' which are similar to those of wound tumour virus RNA.

Particles of rice dwarf reovirus (RDV), which is persistently transmitted by Nephotettix species (Boccardo & Milne, 1984), are composed of 12 genome segments of dsRNA (Omura & Inoue, 1985) and seven major proteins (Nakata et al., 1978). These proteins, which are encoded in viral genomic RNA, include those which have polymerase activity (Kodama & Suzuki, 1973) and those which control the severity of the symptoms induced in infected plants (Kimura et al., 1987). The determination of the nucleotide sequence of the genomic RNA is expected to provide a great deal of information on the relationship between the genome structure and the functions of translational products, some of which may be expected to control the virus pathogenicity and its transmissibility by insect vectors. However, among the plant reoviruses the only nucleotide sequence known is that of the smallest segment (S12) of wound tumour virus (WTV) (Asamizu et al., 1985).

This paper describes the sequence of cDNA to the tenth largest dsRNA (S10) segment which was selected from the cDNA library of RDV genome segments (Minobe et al., 1986) cloned into pBR322 as described by Cashdollar et al. (1982).

Of the 11 clones that hybridized specifically with S10 dsRNA labelled with [32P]pCp, five had an insert of approximately 1550 bp of cDNA. These were considered to correspond to the full length of the original dsRNA (including the tail attached during the course of the cloning) because this size was the same as that deduced from the molecular weight of the original dsRNA. When one of the five clones was nick-translated using [32P]dCTP and annealed with RDV dsRNA species transferred onto a Zeta-Probe sheet (Bio-Rad) (Bodkin & Knudson, 1985), only the S10 segment was labelled (data not shown). Then, the cDNA insert was subdivided by restriction endonuclease cleavage and Bal 31 digestion (Fig. 1) and subcloned into M13 mpl 8 and mpl9 as described by Messing et al. (1981). The DNA sequence was determined by the dideoxy chain-termination reaction (Sanger et al., 1977) (M13 sequencing kit; Takara, Kyoto, Japan). All the parts of the cDNA were sequenced at least twice in each direction (Fig. 1). For the determination of the sequence of the 5' and 3' ends, cDNA was made by extending the primers shown in Fig. 1 and sequenced. All the sequence was confirmed by sequencing using 2'-deoxy-7-deazaguanosine triphosphate in place of dGTP (Mizusawa et al., 1986).

The nucleotide sequence of segment S10 of the cloned genome is shown in Fig. 2. The nucleotides at each end agree with those obtained by direct analysis of the dsRNA of RDV (Miura, 1981) confirming that the clone was full length. The cloned DNA of S10 RNA contains
1319 bp with one long open reading frame which starts from the first AUG triplet (residues 27 to 29) and extends for 1059 nucleotides. Three contiguous termination codons were detected at the third triplet downstream from the first stop codon (Fig. 2). The initiation codon has the sequence context AACAUUGG, regarded as an initiation sequence of high efficiency (Kozak, 1981). The open reading frame has the coding potential for a polypeptide composed of 352 amino acids giving a calculated mol. wt. of 39094 (39-1K). However, no major component protein of RDV is this size (Nakata et al., 1978). The other reading frames including those in the opposite polarity strand contained numerous stop codons and none coded for more than 59 amino acids.

To determine the polarity of transcription, RDV transcripts synthesized in vitro were hybridized to the individual complementary full-length single-stranded cDNA copies of segment S10 cloned into phage M13. Transcription in vitro was carried out using [32P]UTP essentially as described for cytoplasmic polyhedrosis virus by Smith & Furuichi (1980), except for the pH (8.5) and the omission of S-adenosyl methionine from the reaction mixture (Kodama & Suzuki, 1973). After incubation of the reaction mixture for 4 h virus particles and bentonite were removed by centrifugation (96000 g for 30 min) and the solution was passed through a Sepharose 2B column to remove unreacted materials. Dot hybridization was performed according to the method described by Maniatis et al. (1982). The labelled transcripts hybridized to cDNA containing ---TACC 3' but not to that with the sequence ---TGAT at the 3' end (Fig. 3). This result suggests that the sequence 5' GGUA---UGAU 3' (Fig. 2), which is the strand of coding (positive) polarity, has the same polarity as the RDV transcripts synthesized in vitro. A similar result was observed with transcripts of WTV RNA (Asamizu et al., 1985).

The 5' terminal structures of in vitro transcripts of RDV and WTV RNA are 5' m7GpppAm--- (Shimotohno & Miura, 1975; Rhodes et al., 1977), rather than 5' m7GpppGm--- predicted from the genome sequences of RDV (Fig. 2) and WTV (Asamizu et al., 1985). It is not clear why transcripts that are complementary copies of the genome negative strand (Fig. 3) differ from the positive strand at the 5' terminus. However, as Asamizu et al. (1985) pointed out the transcripts were made by transcription in vitro and further studies are needed to see whether this difference is found in vivo.

The 5' and 3' ends of RDV S10 are the same as the terminal sequences of each genome segment of WTV (Asamizu et al., 1985), that is: (+) 5' GGUA---UGAU 3' and (−) 3' CCAU---ACUA 5'. RDV and WTV belong to the same Phytoreovirus subgroup (Boccardo & Milne, 1984), though they have quite distinct host ranges and are geographically separated. Their common terminal sequences may correspond to the polymerase recognition site, as suggested by Shatkin & Kozak (1982). Although the possibility that the 39-1K polypeptide is included in the virus particles as an undetectable minor protein cannot be ruled out, it is presumably a non-
Short communication

Fig. 2. Nucleotide sequence of the plus-sense strand of segment of S10 of RDV (shown as DNA) and the amino acid sequence of its predicted translation product. A possible initiation codon is surrounded by a box. In-phase termination codons are underlined.
structural protein. The polypeptide may therefore have a biological function such as polymerase activity, as reported for sigma non-structural protein of reovirus (Gomatos et al., 1980; Richardson & Furuichi, 1983). However, a comparison of the 39.1K polypeptide sequence with data from the protein sequence library prepared by the National Biomedical Research Foundation, Washington, D.C., U.S.A. using the FASTP (Lipman & Pearson, 1985) and SEQHP (Kanehisa, 1982) programs resulted in no strong homologies with a range of proteins including polymerases.

We would like to thank Drs H. Kano, H. Mizuno and Y. Kawakami for their valuable suggestions.

REFERENCES


(Received 19 June 1987)