J. gen. Virol. (1989), **70**, 1297–1300. *Printed in Great Britain*

Key words: *RDV/genome segment 9/nucleotide sequence*

Nucleotide Sequence of Rice Dwarf Virus Genome Segment 9

By ICHIRO UYEDA,* HIROSHI KUDO, TOHRU TAKAHASHI, TERUO SANO, KAZUSATO OHSHIMA, TAKESHI MATSUMURA AND EISHIRO SHIKATA

Department of Botany, Faculty of Agriculture, Hokkaido University, Sapporo 060, Japan

(*Accepted 3 January 1989*)

SUMMARY

The complete nucleotide sequence of the phytoreovirus rice dwarf virus (RDV) genome segment 9 is presented. It consisted of 1305 nucleotides and had an open reading frame that codes for a putative polypeptide of 351 amino acids. M_r of the protein was calculated to be 38 598. The terminal nucleotides 5′ GGUAAA--GAU 3′ were the same as those of RDV genome segment 10. The fourth nucleotide from the 3′ end of an expected conserved sequence was a C rather than the U found in the previously sequenced genome segment 10. A structure similar to the segment-specific inverted repeat of wound tumour virus was also found in the terminal region of segment 9.

Rice dwarf virus (RDV) belongs to phytoreovirus subgroup 1 and has a genome comprising 12 segmented dsRNAs (Boccardo & Milne, 1984). The virus multiplies both in plant hosts and its insect vectors and has a virion-associated RNA polymerase (Kodama & Suzuki, 1973; Uyeda & Shikata, 1984). In order to elucidate the mechanisms operating in replication and transcription, it is essential to determine the structure of the genome. The nucleotide sequence of genome segment 10 has already been determined (Uyeda *et al.*, 1987; Omura *et al.*, 1988). In this paper, we present the complete nucleotide sequence of genome segment 9.

Rice dwarf virus was maintained in rice plants in a greenhouse by periodic transfers through an insect vector, the leafhopper *Nephotettix cincticeps*. The virus was purified as previously described (Uyeda & Shikata, 1982).

The viral RNA was extracted from the purified virus and cDNA cloning of genome segment 9 was done as previously described (Uyeda *et al.*, 1987). Two clones, pRD538 and pRD546, were made from the denatured genome dsRNA. They reacted positively with a ^{32}P-labelled genome segment 9 probe (Jordan & Dodds, 1983) by Southern blot hybridization (Southern, 1975). The sizes of the inserts of pRD538 and pRD546 were about 1350 and 1200 nucleotides respectively. A restriction endonuclease cleavage map of the overlapping inserts showed that these clones covered about 1400 nucleotides of genome segment 9.

In order to examine whether the cDNA clones cover the entire nucleotide sequence of genome segment 9, the 5′- and 3′-terminal regions were directly sequenced by a dideoxynucleotide chain termination method using reverse transcriptase (Seikagaku Kogyo) as described by Meshi *et al.* (1983). For 5′-terminal sequencing, a *Hin*dIII–*Acc*II restriction endonuclease cleavage fragment of the cDNA at nucleotides 96 to 71 was used as a primer and the transcript as a template. The transcript was prepared as described previously (Uyeda & Shikata, 1984). Analyses showed that the clone pRD545 contained the 5′-terminal region with a sequence identical to that derived from direct RNA sequencing. For sequencing the 3′-terminal region of the genome segment, a synthetic oligonucleotide, 5′ GAAGTTTTGACAGCGAA 3′ at nucleotides 1228 to 1244, was used as a primer and the genome dsRNA as a template (Fig. 1). The genome dsRNA was denatured in 90% DMSO for 30 min at 50 °C, quickly chilled in ice and precipitated by ethanol. The template was annealed with the primer at 40 °C for 6 h in a buffer described by Meshi *et al.*

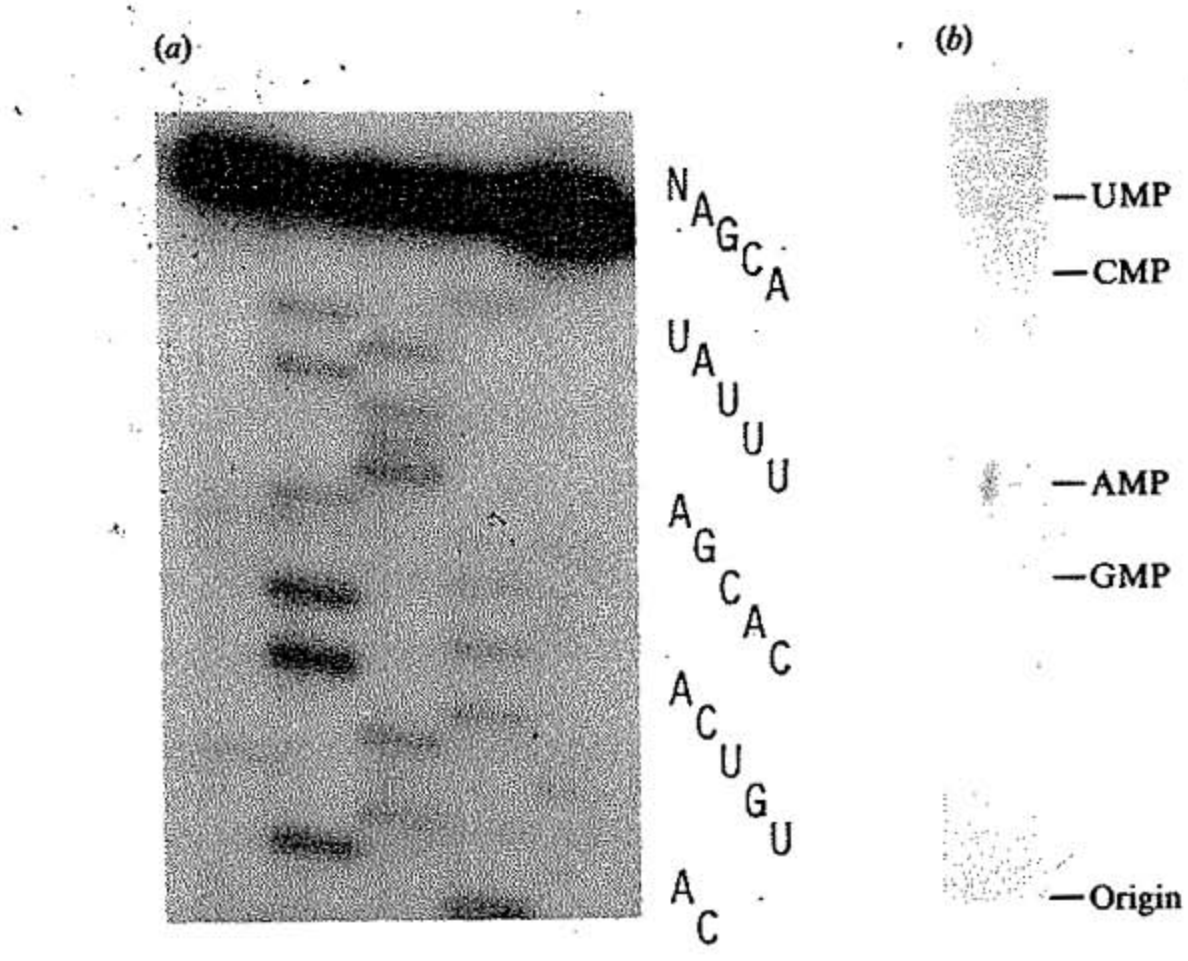

Fig. 1. Analyses of the 3′-terminal region of RDV genome segment 9. (*a*) Direct RNA sequencing of the 3′-terminal region of the genome segment 9 by a dideoxynucleotide chain termination method. (*b*) PEI–cellulose thin-layer chromatography of genome segment 9 terminally labelled by polynucleotide kinase and then completely digested with nuclease P1. Positions of the mononucleotides are shown on the right-hand side of the chromatograph.

(1983). It was found that the cDNA clone pRD538 lacked 11 nucleotides of the 3′-terminal region.

The nucleotide at the 3′ terminus of the (+) strand was determined by identifying that of the 5′ terminus of the (−) strand. The 5′ termini of the genome dsRNAs were labelled with [γ-^{32}P]ATP by T4 polynucleotide kinase (Takara Shuzo). Genome segment 9 was separated on a 10% polyacrylamide gel according to Laemmli (1970), eluted from the gel as described by Smith (1980), and the labelled genome segment was completely digested with nuclease P1. The digested RNA was separated on PEI-cellulose in 0·4 M-LiCl using the four mononucleotides as standards. Autoradiography after chromatography showed that the radioactive phosphate was predominantly incorporated into the adenosine mononucleotide. Incorporation into GMP was about 50% of that in AMP (Fig. 1). Direct RNA sequencing by two-dimensional electrophoresis (De Wachter & Fiers, 1972; Rensing & Schoenmakers, 1973) indicated that the radioactivity was predominantly incorporated into the 5′ terminus of the (−) strand (unpublished). Thus we concluded that the 3′-terminal nucleotide residue was U, complementary to the 5′ terminus of the (−) strand.

Restriction endonuclease cleavage fragments of the cDNAs were subcloned in M13 phage mp18 or 19 (Messing, 1983) and sequenced by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977) using a sequencing kit purchased from Takara Shuzo. When subcloned cDNAs were too long for the sequence to be read, they were deleted sequentially for sequencing using exonuclease III (Takara Shuzo) and mung bean nuclease (Takara Shuzo) as described by Henikoff (1984). When it was compared with the sequence of RDV genome segment 10 (Uyeda *et al.*, 1987; Omura *et al.*, 1988), hexanucleotide 5′ GGUAAA--- at the 5′ terminus were found to be conserved, and the first four, GGUA, were also as found in wound tumour virus (Asamizu *et al.*, 1985). Tetranucleotide ---UGAU 3′ of wound tumour virus was conserved in all 12 segments (Asamizu *et al.*, 1985) and the sequence is the same as that of RDV genome segment 10. However, segment 9 was found to have ---CGAU 3′. That is, the fourth base from the 3′ terminus is C instead of U. Whether this change is peculiar to our isolate, which has been maintained for at least 20 years in our greenhouse, is not known at the moment.

```
        10          20          30          40          50          60
GGUAAAAAUCGUGUGUCCUCCGUGAUGGGUAAGCUCCAAGAUGGAAUCGCCAUCAAGCGG
                              M   G   K   L   Q   D   G   I   A   I   K   R
        70          80          90          100         110         120
AUCAACGACGCGAUUACCACUUUCAAGAAUUACAAGCUUGGUGAACUGGAACAGGGCGGC
I   N   D   A   I   T   T   F   K   N   Y   K   L   G   E   L   E   Q   G   G
        130         140         150         160         170         180
UCAAUGGCCAUCAACACAUUGAGUAACGUCCGCGCCCAUGUUGGGCUGGCUUGGCCGGCC
S   M   A   I   N   T   L   S   N   V   R   A   H   V   G   L   A   W   P   A
        190         200         210         220         230         240
AUCUUGCGAAAUUGUUUGAUACACACUUCAUCCCAUCUUGGGUUCAUGAAGUUUAUGAUU
I   L   R   N   C   L   I   H   T   S   S   H   L   G   F   M   K   F   M   I
        250         260         270         280         290         300
GAUAUUGCUACUACCUGGAAAGUUGGUGCUUUCACCCUUCUGGGCAGCGUCGGUGACGAA
D   I   A   T   T   W   K   V   G   A   F   T   L   L   G   S   V   G   D   E
        310         320         330         340         350         360
GAUCCUUUCACUGACGUUGACUUGAUUUACACUAAGACCUGCUUGCAUUUGGGUCUUAAA
D   P   F   T   D   V   D   L   I   Y   T   K   T   C   L   H   L   G   L   K
        370         380         390         400         410         420
GACAAUGAUUUUCUGCAAUUUCCAGAAGAGUUUGCCUAUGAGGCGAAUUCUUUUCUAGAA
D   N   D   F   L   Q   F   P   E   E   F   A   Y   E   A   N   S   F   L   E
        430         440         450         460         470         480
GCGCAGUCGAUGAAUGCUAGGGUGGACAUGCUCACUGGUGUCCACAAUAUUGAAGAUAAA
A   Q   S   M   N   A   R   V   D   M   L   T   G   V   H   N   I   E   D   K
        490         500         510         520         530         540
UAUGUCUUUAGAAUAGAGUCUAUAUCUAAGUUUUUGAAAGCUUACUAUACUGCUUCAGAA
Y   V   F   R   I   E   S   I   S   K   F   L   K   A   Y   Y   T   A   S   E
        550         560         570         580         590         600
GACGUUGCUUACUUGACUGGAUUUAUAAAGCCUGACGGCUCUAAGGAGUCAAUCUUGAGU
D   V   A   Y   L   T   G   F   I   K   P   D   G   S   K   E   S   I   L   S
        610         620         630         640         650         660
GCCGAACUCUUGAAAGCGCAGGUCACAUCCGAGGUGCUACGCGUGCGUAAUUUAAUUACC
A   E   L   L   K   A   Q   V   T   S   E   V   L   R   V   R   N   L   I   T
        670         680         690         700         710         720
ACCAAGAUUCAGCAGUACAUUAAUUUGUACGAAGAUUCGCAGUUACCGCAUUUUCGGCGA
T   K   I   Q   Q   Y   I   N   L   Y   E   D   S   Q   L   P   H   F   R   R
        730         740         750         760         770         780
GCGGCUUUGUCCUACACUCAGGAUUGGGAUGUUGAUGGCGGUGUGCCCGCUGCACUCCCA
A   A   L   S   Y   T   Q   D   W   D   V   D   G   G   V   P   A   A   L   P
        790         800         810         820         830         840
CAGCCUGAUACAACUGACGAUGAAAGUCCCGUCACUAAGCCUGGCGCUAGUGCACCAACA
Q   P   D   T   T   D   D   E   S   P   V   T   K   P   G   A   S   A   P   T
        850         860         870         880         890         900
GUGAGUAAAGGUGCUGAUCAGCCAGAAGACGAGGAGAUAAUACAUAAAAAGGUGGAUGCU
V   S   K   G   A   D   Q   P   E   D   E   E   I   I   H   K   K   V   D   A
        910         920         930         940         950         960
UCGAAAGAUGCUCCACCUAAGGCAGUUUCUUCUGGAAAUGUAAGCGCCAGAGGUAUUCCU
        970         980         990         1000        1010        1020
GCCUUUUUAGAAGAUGAUAUGAGUGAGAUGGACGCACCUGAUGGCUUCCAUGAUUACUUA
A   F   L   E   D   D   M   S   E   M   D   A   P   D   G   F   H   D   Y   L
        1030        1040        1050        1060        1070        1080
ACGAGGGAACAUGAGAACAACUUCGACUUGGCGCAGUUGGGACUCGCACCCUCAGUCUGA
T   R   E   H   E   N   N   F   D   L   A   Q   L   G   L   A   P   S   V   *
        1090        1100        1110        1120        1130        1140
CUUUACGCUGGAGUAGAUGAAUGCCUCACCAACAUUUGUUACUCCAUAAUGAUGUUAUCA
        1150        1160        1170        1180        1190        1200
UGUCUGCAUGGAUUCUUCAGAAUGUUGAUUCGUGUAUUAGGUUGCUUGUAAACAACGGAG
        1210        1220        1230        1240        1250        1260
UGGUGAUCGGUAGAUCCGUCUGGCGUGGAAGUUUUGACAGCGAACCUGUGUCCUAACGCG
        1270        1280        1290        1300
GAUGCGGGAGGGUCAUUAACUUCCAUGUCACACGAUUUAUACGAU
```

Fig. 2. Complete nucleotide sequence and predicted amino acid sequence of RDV genome segment 9.

The complete nucleotide sequence of genome segment 9 is presented in Fig. 2. It consisted of 1305 nucleotides and the longest predicted open reading frame started with AUG at nucleotides 25 to 27 and terminated with UGA at nucleotides 1078 to 1080. The initiation codon had the strong and consensus eukaryotic initiator context GXXAUGG (Kozak, 1987). The M_r of the

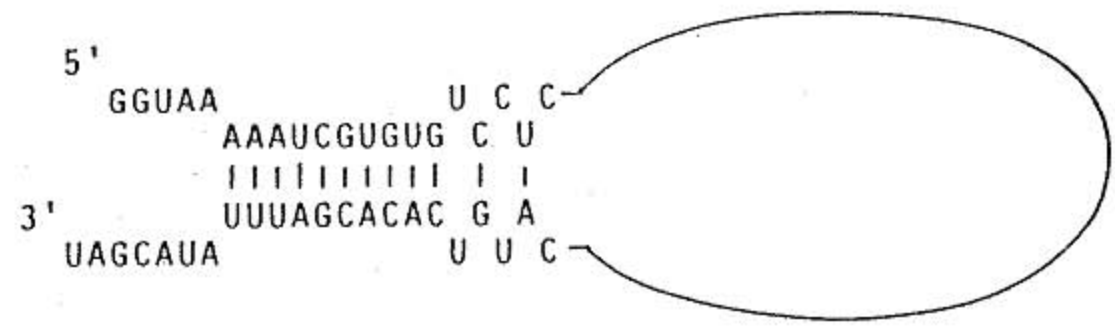

Fig. 3. Inverted repeat and a possible secondary structure found in the terminal region of RDV genome segment 9.

protein was calculated to be 38598. The predicted gene product of segment 9 is probably a non-structural protein, because its M_r is lower than the smallest structural protein of about 45000 (Nakata *et al.*, 1978; Matsuoka *et al.*, 1985).

A structure similar to the segment-specific inverted repeat of wound tumour virus (Anzola *et al.*, 1987) is also found in genome segment 9 (Fig. 3).

This research was supported by Grant-in-Aid for Scientific Research No. 62480042, Grant-in-Aid for Co-operative Research No. 63304013 from the Ministry of Education, Science and Culture, Japan, and a grant from the Akiyama Foundation. We thank the Research Center for Molecular Genetics, Hokkaido University for the use of their computer facilities.

REFERENCES

ANZOLA, J. V., XU, X., ASAMIZU, T. & NUSS, D. L. (1987). Segment-specific inverted repeats found adjacent to conserved terminal sequences in wound tumor virus genome and defective interfering RNAs. *Proceedings of the National Academy of Sciences, U.S.A.* **84**, 8301–8305.

ASAMIZU, T., SUMMERS, D., MOTIKA, M. B., ANZOLA, J. V. & NUSS, D. L. (1985). Molecular cloning and characterization of the genome of wound tumor virus: a tumor-inducing plant reovirus. *Virology* **144**, 398–409.

BOCCARDO, G. & MILNE, R. G. (1984). Plant reovirus group. *CMI/AAB Descriptions of Plant Viruses*, no. 294.

DE WACHTER, R. & FIERS, W. (1972). Preparative two-dimensional polyacrylamide gel electrophoresis of ^{32}P-labeled RNA. *Analytical Biochemistry* **49**, 184–197.

HENIKOFF, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28**, 351–359.

JORDAN, R. L. & DODDS, J. A. (1983). Hybridization of 5′-end-labelled RNA to plant viral RNA in agarose and acrylamide gels. *Plant Molecular Biology Reporter* **1**, 31–37.

KODAMA, T. & SUZUKI, N. (1973). RNA polymerase activity in purified rice dwarf virus. *Annals of the Phytopathological Society of Japan* **39**, 251–258.

KOZAK, M. (1987). An analysis of 5′-noncoding sequence from 699 vertebrate messenger RNAs. *Nucleic Acids Research* **15**, 8125–8132.

LAEMMLI, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, London* **227**, 680–685.

MATSUOKA, M., MINOBE, Y. & OMURA, T. (1985). Reaction of antiserum against SDS-dissociated rice dwarf virus and a polypeptide of rice gall dwarf virus. *Phytopathology* **75**, 1125–1127.

MESHI, T., ISHIKAWA, M., TAKAMATSU, N., OHNO, T. & OKADA, Y. (1983). The 5′-terminal sequence of TMV RNA; question on the polymorphism found in vulgare strain. *FEBS Letters* **162**, 282–285.

MESSING, J. (1983). New M13 vectors for cloning. *Methods in Enzymology* **101**, 20–77.

NAKATA, M., FUKUNAGA, K. & SUZUKI, N. (1978). Polypeptide components of rice dwarf virus. *Annals of the Phytopathological Society of Japan* **44**, 288–296.

OMURA, T., MINOBE, Y. & TSUCHIZAKI, T. (1988). Nucleotide sequence of segment S10 of the rice dwarf virus genome. *Journal of General Virology* **69**, 227–231.

RENSING, U. F. E. & SCHOENMAKERS, J. G. G. (1973). A sequence of 50 nucleotides from coliphage R17 RNA. *European Journal of Biochemistry* **33**, 8–18.

SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, U.S.A.* **74**, 5463–5467.

SMITH, H. O. (1980). Recovery of DNA from gels. *Methods in Enzymology* **65**, 371–380.

SOUTHERN, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503–517.

UYEDA, I. & SHIKATA, E. (1982). Ultrastructure of rice dwarf virus. *Annals of the Phytopathological Society of Japan* **48**, 295–300.

UYEDA, I. & SHIKATA, E. (1984). Characterization of RNAs synthesized by the virion-associated transcriptase of rice dwarf virus in vitro. *Virus Research* **1**, 527–532.

UYEDA, I., MATSUMURA, T., SANO, T., OSHIMA, K. & SHIKATA, E. (1987). Nucleotide sequence of rice dwarf virus genome segment 10. *Proceedings of the Japan Academy* **63**, 227–230.

(Received 10 October 1988)