Journal of General Virology (1990), **71**, 2217–2222. *Printed in Great Britain*

Nucleotide sequence of rice dwarf virus genome segment 4

Ichiro Uyeda,* Hiroshi Kudo, Naoki Yamada, Takeshi Matsumura and Eishiro Shikata

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

The complete nucleotide sequence of rice dwarf virus (RDV) genome segment 4 was determined. Genome segment 4 was 2468 nucleotides long and had a long open reading frame initiating at nucleotides 64 to 66 and terminating at 2245 to 2247. The deduced polypeptide contained 727 amino acid residues with an M_r of 79·8K. Amino acid sequences similar to a 'zinc-finger' and purine NTP-binding motifs were present in the deduced polypeptide. Considerable amino acid sequence homology was detected between genome segment 4 of RDV and wound tumor virus (WTV). One of the sequences similar to the 'zinc-finger' motif was present in a conserved region of the polypeptide of both viruses. However, the sequence similar to the purine NTP-binding motif was not present in the polypeptide encoded by genome segment 4 of WTV.

Introduction

Plant reovirus subgroup 1 consists of rice dwarf virus (RDV), wound tumor virus (WTV) and rice gall dwarf virus (Boccardo & Milne, 1984). The viruses in this group are transmitted by leafhoppers and have a genome made up of 12 segmented dsRNAs. The sizes of the RDV genome segments are between about 1000 and 4500 nucleotides (Uyeda *et al.*, 1990). The purified virions of the group exhibit RNA polymerase activity *in vitro* (Nuss & Peterson, 1981; Uyeda & Shikata, 1984; Yokoyama *et al.*, 1984).

During the last few years, the complete nucleotide sequences of several segments of RDV and WTV have been determined and information on the coding assignment of some genome segments has been obtained (Anzola *et al.*, 1987, 1989*a*, *b*; Asamizu *et al.*, 1985; Dall *et al.*, 1989; Uyeda *et al.*, 1987, 1989; Omura *et al.*, 1988, 1989; Fukumoto *et al.*, 1989; Nakashima *et al.*, 1990). The assignment of structural and non-structural viral polypeptides to all the genome segments has been proposed for WTV (Xu *et al.*, 1989). Comparison of the deduced amino acid sequences showed significant homologies between comparable segments of the two viruses. The amino acid sequence of a minor outer capsid protein encoded by genome segment 11 of WTV had homology to a protein encoded by RDV genome segment 9. The sequence of the WTV genome segment 10 encoding a non-structural polypeptide showed significant homology to that of RDV genome segment 10. The genome segments 7 and 8 of RDV were assigned to the 60K core protein and a major outer capsid polypeptide, respectively (Omura *et al.*, 1989; Nakashima *et al.*, 1990).

Little information is available on genome assignment and function of genome segment 4 of RDV. It has been reported that the ability to induce severe stunting in rice plants is associated with a mobility shift of genome segment 4 in PAGE (Kimura *et al.*, 1987).

In this paper, we have determined the complete nucleotide sequence of the RDV genome segment 4 and possible functions of the polypeptide encoded by the segment are discussed.

Methods

Viral RNA preparation. Purification of RDV, synthesis of viral transcripts *in vitro* from the purified virus and extraction of viral dsRNAs and ssRNA transcripts were done as previously described (Uyeda & Shikata, 1984).

Terminal sequencing of viral RNAs. The 5′ termini of genome dsRNAs and transcripts were labelled with ^{32}P by T4 polynucleotide kinase (Takara Shuzo Co.) after treating with calf intestine alkaline phosphatase (Boehringer Mannheim). The labelled genome segments of dsRNA were separated by 10% polyacrylamide gel electrophoresis (PAGE) according to Laemmli (1970), while the labelled transcripts were separated by 3% PAGE in Tris–boric acid–EDTA containing 8 M-urea as previously described (Uyeda & Shikata, 1984). Individual segments were eluted from the crushed gel (Smith, 1980) and partially degraded in 100 mM-NaOH for 60 min at 100 °C. The nucleotide sequence was determined by a wandering spot analysis (De Wachter & Fiers, 1972; Rensing & Schoenmakers, 1973). The alkali-treated RNAs were run in 7·4% PAGE gels with citric acid pH 3·5 containing 6 M-urea in the first dimension and in 20% PAGE gels in 50 mM-Tris, 50 mM-boric acid in the second dimension. The sequence was read after autoradiography of the gel at 6 °C for 2 to 4 days.

The 5′-terminal nucleotides were determined by polyethylenimine (PEI)–cellulose chromatography of complete P1 nuclease digests of the labelled RNAs. The PEI–cellulose sheet was developed in 0·4 M-LiCl in a moist chamber and was autoradiographed.

cDNA cloning and sequencing. Oligo(A) was added to the viral transcript by poly(A) polymerase and oligo(dT)-primed cDNA synthesis was carried out according to the method of Gubler & Hoffman (1983). After homopolymeric tailing with oligo(dC), the cDNA was inserted at the *Pst*I site of pBR322 and *Escherichia coli* strain HB101 was transformed with the recombinant DNA by the $CaCl_2$ method (Maniatis *et al.*, 1982).

cDNA inserts were screened by colony hybridization using ^{32}P-labelled genome segment 4 dsRNA as a probe (Jordan & Dodds, 1983). Assignment of the cDNA inserts to genome segment 4 was confirmed by alkaline blotting analyses as described by Li *et al.* (1987). In brief, genome dsRNA was separated by 5% PAGE in Tris–acetate–EDTA buffer (Uyeda & Shikata, 1984) and transferred to a nylon membrane in 0·2 M-NaOH for 30 min at room temperature. Transferred RNA was hybridized with a recombinant cDNA probe. The recombinant cDNA was labelled with [α-^{32}P]dCTP by the Klenow fragment of DNA polymerase after digestion with *Hap*II (Maniatis *et al.*, 1982). Hybridization was at 52 °C overnight in 50% formamide, 5 × SSC, 50 mM-sodium phosphate pH 6·5, yeast tRNA (500 µg/ml), 0·1% SDS and 5 × Denhardt's solution. cDNAs were subcloned in M13mp18 or M13mp19 by the method of Messing (1983) and sequenced by the dideoxynucleotide chain-termination method (Sanger *et al.*, 1977) using the Klenow fragment or T7 DNA polymerase. When the subcloned cDNA was too long for the sequence to be read, the insert cDNA was deleted sequentially by the method of Henikoff (1984). Since cDNA clones did not cover the 3′-terminal region, the viral transcript RNA was sequenced directly by reverse transcription using a dideoxynucleotide chain-termination method (Uyeda *et al.*, 1989). A synthetic oligonucleotide complementary to 12 nucleotides of the 3′-terminal region was used as a primer.

Results

Terminal structure

The terminal nucleotides were determined by PEI–cellulose chromatography. When the genome dsRNA was analysed, radioactivity was detected predominantly in nucleotide A (Fig. 1). There was a little incorporation

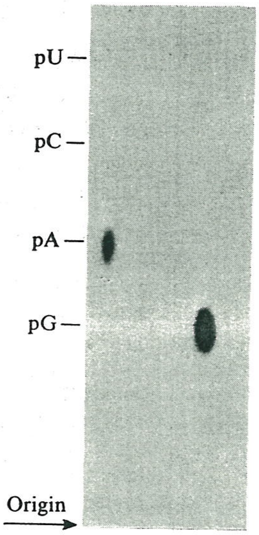

Fig. 1. Identification of terminal nucleotides by PEI–cellulose chromatography of 5′-terminal-labelled genome dsRNA (left lane) and transcripts (right lane) of RDV genome segment 4.

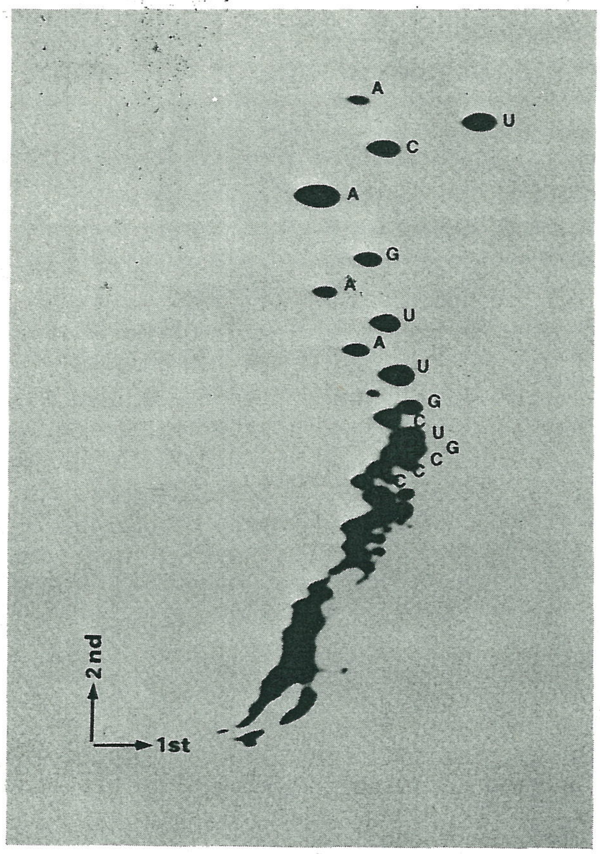

Fig. 2. Wandering spot analysis of the 5′-terminal-labelled genome dsRNA of RDV genome segment 4.

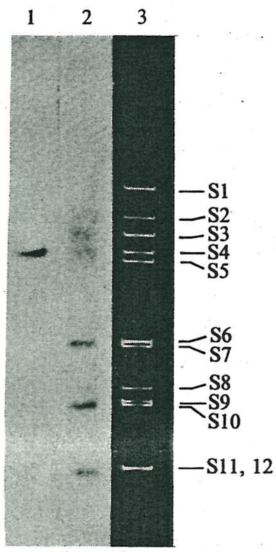

Fig. 3. Assignment of the cDNA clone pRD3A11 to RDV genome segment 4 by alkaline blotting analysis. Lane 1, total genome dsRNA was blotted, hybridized with ^{32}P-labelled cDNA clone pRD3A11 and autoradiographed. Lane 2, total genome dsRNA was labelled with ^{32}P, then blotted onto a nylon membrane and autoradiographed. Lane 3, total genome dsRNA was stained with ethidium bromide after electrophoresis.

into G but none at all in C and U. Incorporation of radioactive phosphate into the guanosine mononucleotide was less than 50% of that into the adenosine mononucleotide. However, when the transcript was analysed, the radioactive phosphate was exclusively incorporated into the guanosine mononucleotide (Fig. 1). Thus we concluded that the 5′ termini of the plus and minus strands were G and A, respectively.

Terminal sequences of genome segment 4 were determined by wandering spot analyses of the 5′ terminally labelled genome dsRNA and the transcript. The sequence of the transcript was determined to be 5′ NGUAAAUUGCUGC 3′. Analysis of the 5′ terminally labelled genome segment 4 dsRNA showed one pattern of sequence although a mixed pattern of spots derived from both plus and minus strands had been expected (Fig. 2). This single pattern was probably due to the uneven incorporation of the radioactive phosphate into the termini as was revealed by PEI–cellulose chromatography analyses of the terminal nucleotides. The sequence derived was that of the minus strand because the sequence complementary to the presumptive conserved 3′-terminal tetranucleotide (5′ UGAU 3′) was present and it was different from that obtained from the wandering analysis of the transcript (plus strand).

The expected conserved terminal sequences of RDV (Uyeda *et al.*, 1987, 1989; Omura *et al.*, 1988, 1989) were present in the genome segment 4 as 5′ GGUAAA---UGAU 3′.

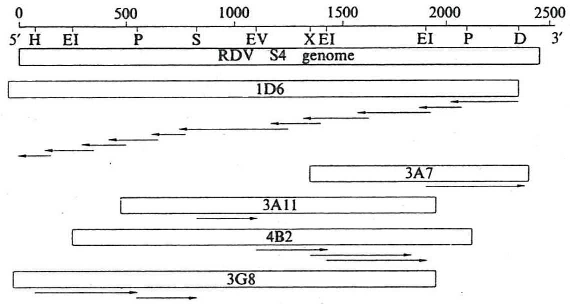

Fig. 4. Location of cDNA clones on RDV genome segment 4. Directions and extent of the sequences determined are shown as arrows below the respective cDNA clones. Numbers indicate the nucleotide position. Restriction cleavage sites are: *Hin*dIII, H; *Pvu*II, P; *Sma*I, S; *Eco*RV, EV; *Xba*I, X; *Eco*RI, EI; *Dra*I, D.

cDNA cloning and sequencing

Several clones from the cDNA clone library have been selected and sequenced. All the clones cross-reacted in Southern blot hybridization (Southern, 1975) and also

```
5' GGUAAAUUGC UGCCGUUGCU GUCAACCAAC GUAAAACCAG CGAUUAACUA CAGUGUAUUC 60
AGGAUGAACC AAUCUCGAAG CUUUGUUACG GGGAGAGGAC GCGACUUGUC UCGGACUCCU 120
UCGGCUUUAA GCUCGAAUUC GGAAACUCCA GGAUCAAUGA GUUCACCAUC UGAAGGGAAG 180
ACGAACGCUU GGGUUAACUC GGCGUACGUC AGCAACUUUC CUGCGCUGGG CCAAUCGCAG 240
GGGUUGCCAU CCCAUAAAUG UUCAGCUUUG GCUCUGCGUA GUUCACAGAC UACGUACAUU 300
AUAAACUUUC CACGCCAGCA CUGGAACAUC AUGACUUUUC CAAAUCAGUC UGAGGCGAUU 360
CUGGCUAAUG UUGCCUCAUA UGCAAAAGAC CUUGACGGAA AGAACUCCUU CGCCGUAUCC 420
GAUACGCUCA AGAUAGCCUU GGUUCUGUCG CCUACGGAGA AGAGACUGUU CCGGAAUGAU 480
ACUUUAAGCC ACCUCGCUGA UGUUCACAUG UACGUGUUGG ACCCAGCUGA GGCGGAAGGG 540
AAGAAUUUGU CUGACUCGGA GACUGUUUAC GUCUACGUGA CCCCACCCAA UUUGACUGAU 600
GUGAAGCCGA CGACCGUUGU GCUGACUGAG UGUGCCGCUA ACGCUAAGUC AGCAAAUGAU 660
CUGAGGCAAU ACAUCGUCAC CCAACUAAGG AAGAUGCCCU CCCUUCCAUU UGGGUGUACU 720
ACUUAUGCCC CAGGUUUUUU GUCUGACGGU GUGUGCAAGG AACAUCCUAA CUUGUUCACU 780
UCUGAGGAGC UGGGCGCCAA AAUCAAGGUG CUCACGAAAU UGCUCAUUCG UUGUGCUACU 840
UCCAUGUCUC AGGACGGAUC UAACGCGUUC UGUCCUAAGC AUCCCAAAGU GAAGAUUGUG 900
CACGAAUCGA ACGCCACAUC CUACGUCCUG UUCAAUCGUC CGAACGGGAU GGUGGCUACG 960
AAUCUAAUUC UCUCAGAUUU GCCCGAUGAU GAUUGUCCAA CGUGCUGGAU ACUUAAGUUA 1020
GCGAUUAGCG AAGCUAGGUU CUAUGCUCUU GACGGUCAUC ACCGCUGUCG CUCUGGAAUA 1080
AUAACAUCCA GUGUCUUCAG AUAUCUCGCU UCCAUAGUCA UCAGAGUAAG GAUGGAUUCC 1140
GUGCUCGCUC CCUCUGACGC GUCGUCAACU GAUCACGCCG CCCUCGUCAA UAUGAUGUGC 1200
GGCAUUAUCC AAAACACACC UGCGAUGCGA CAGCUCGGCA UAUCGACGGG UAGUGAGAAA 1260
GUGAACAACA GAAGCAUGAG AGUCAUCAUC AUGCAAGAGA AUGCCGACAG GGCUACGCAA 1320
AUGACAGCUC UGUAUCAUCU GUUUCUAGAC UACUUUGGAG CCCUCAAUGG UUGGGGGUUC 1380
UACUUCUGCU CGCUGACGUC UCUCUAUGGU GAAUUCCAUG GUUUCAGCGU CGGAUUUAGC 1440
GGAGAAAUAA CUCAUGUGAA CGUUGCGAGC GUGAUUGCUA AGAAUUGGGA CACGCAGUCU 1500
GGCAUCGACA ACAUUUUGGA AUUUAAGACU AUCACUAUUC CCGUCCACAA CGGGGACAUA 1560
GUGUGCAUGG UUGAGCGUAC UUUGGCUGAG UCAUUCGAAG UCGUAAUGAA UGAGCACUUU 1620
AAUGGUGCUA GCACCAUAAA AGUGCGUAGA AAUGGUGGAG AUUCGCGAUU CAACUUCACC 1680
AUUUCCAAUC CUCGUGACGC CUUCUUACUG CUGCAAAAGG CGGUAGCUGA UGGAGGCAUU 1740
UUGCAAAAAA UCUUGUGCAG GGCUAUGCUU AAGGCUAUCA CUAGCCAUGC UCUACGAGCU 1800
GAUAGGGAGG UACAGGAUGU UUCCUUCAGC UUCGUAUUAA AAAUGAGACU CAACGCGGUU 1860
AACAAGAGCG ACUCGAAGUC GUCAGAAUUG GCGCACACGG CUCAGAUGAA UUCCCUACCC 1920
GUGUUUCUAG CUUCGACGCC CUUUACCAUG CAGCUAGGCA CUCUUCGAGA UGCCUUGCUG 1980
AAAAAGACUG AAAAUGUCAC GGUCAUCAAC AUGGCCCGAA CUACUGAAGA GGUGUCGAAG 2040
GACGCCCUUC GAGAAAUAUU GAAAAGUAUC GGGGGUAGUU CCAUGACCCU CGAAGACACA 2100
GCUGAGCCAG UAUCGGACGC UGAGUCAAUC CCUGAUCCAC CCCCUAGAUC CUGGGCCUCU 2160
GAAGAUGAAG CUGUAAACUC ACCACAGAUU UACAGCAGCA GAAGGAAGGC CAGGAAGGCU 2220
AGGGCAGCUA GCAAGCUGUC GAAGUAAUUA UCUGUUCUUA UAUUUAGUAG UUGUUAAUGU 2280
CGUAGACUUA GUAACUAGUG CUGUAUAUUA UGCUUUUUGU GUUUAAAAUC AGAUCUGCCC 2340
CCCUGGCGAC GCACCCAAGG UGAUGGCAGA CGCGCGAAGA UUGCAAAUGC CAUUCCGUCU 2400
CGAACCUAUA CCGAGUUAGA GGUAAGAGUA CCUGAUCCGU AGUGGAUCGU CAGGGCAGCA 2460
UAUCUGAU 3'
```

Fig. 5. Complete nucleotide sequence of RDV genome segment 4. The putative segment-specific inverted repeat is underlined. Initiation and termination codons are boxed.

```
1          MNQSRSFVTGRGRDLSRTPSALSSNSETPGSMSSPSEGKTNAWVNSAYYSNFPA
1   MKQVAGFIGMSQKNKGIQQNQWHSGPPQGLLSGQTKAEQGTSSQQAGVNQGENSKSGAVV

55  LGQSQGLPSHKCSALALRSSQTTYIINFPRQHWNIMTFPNQSEAILANVASYAKDLDGKN
                 .*.   **..  *** ** * ....  .* *  .......
61  QRKMPMRRDVNSAAQRQHVRRLVYTPYFPNETWNISTFKNGAKDVERSVISTLNNVAENR

115 SFAVSDTLKIALVLSPTEKRLFRNDTLSHLADVHMYVLDPAEAEGKNLSDSETVYYVYV-TP
    .. ... .. . .....  ... . .... *  .*. * .*. . .* . . * *.
121 VMDNCTS-RVIFEMTQIQFESLPDIIRNEFTRVGDDALKWAVPEDLKSADLDHMMV-VKL

175 PNLTDVKPTTVVL-TECAANAKSANDLRQYIVTQL-RKMPSLPFGCTTYAPGFLSD--GV
    .. ... ***... ..*.. **  ...  .. .** *  ... *  . .*.... .. ..
179 STEGTIYPTTLIFPGGCSGMAK-LKSVYSFLESQLERIVTPTPSVSLKYVTSWAEHLFDL

231 [CKEH]PNLFTSEELGAKIKVLTKLLIRCATSMSQDGSNAF[CPKH]PKVKIVH-ESNATSYVL
    *...  *..*.. .   *.*.  ..   ...... ..  * .**.*  . .........
238 CSGQ--LINSQN-ERVDKLLGYMIWDIEKAITLTDQVIA[CYNH]PEVVLRRLGASDIACAV

290 FNRPNGMVATNLILSDLPDDD[CPTC]WILKLAISEARFYALDG[HHRC]RSGIITSSVFRYLA
    ..  .  .   *.* **.  * *.*.  *.**.* .. .. .  *  .. . .*. .**.**..
295 LAGESVVKLTRLALSKSPVDG[CSCC]RILELILNLPSRKPNDKVPQVPLDILFASVYRYVS

350 SIVIRVRMDSVLAPSDASSTDHAALVNMMCGIIQNTPAMRQLGISTGSEKVNNRSMRVII
    .. .   ... ...*.  *****.   .. .** *. ..* .*....*.   ...***..
355 AMCMGRVLNGRIDASGIQSTDHATASIKLNDIIVNDLELRSMGVDKTSSFRGTQSMRAFY

410 MQENADRATQMTALYHLFLDYFGALNGWGFYFCSLTSL--YGEFHGFSVGFSGEITHVNV
    ..** . .. .. .   * . .** *. ***   * .   * ..* . .*..  .* ...
415 VPENLA-GSILDRINVLVMRHFGILHMWGFNGVVLQNQEGYCDYH-IITGLNHLTTITHT

468 ASVIAKNWDTQSGIDNILEFKTITIPVHNGDIVCMVERTLAESFEVVMNEHFNGASTIKV
    .*..* .*.*.* .***.*.*. *.*. ..... ..*..* * .. ....... . ...*
473 NSMVAVHWGTESRMDNIFEIKARTLPTASETMITLIENALKEQLTSIVKDNLRKGVSFSV

528 RRNGGDSRFNFTISNPRDAFLLLQKAVADGGILQKILCRAMLKAITSHALRADREVQDVS
    .** .****.*....... ** *.. . ..  . ..*. *. *.*.... . .*.. .*.
533 KRNINDSRFGFETNSSPAIFLKLRDMLKRAKPFSDLLSLALSKVIKKENAMIQRSITTVE

588 FSFVLKMRLNPVNKSDSKSSELAHTAQMNSLPV--FLASTPFTMQLGTLRDALLKKTENV
    . ..**..  ...  *  .   ..*. .***.  **  .. .   .   *. .*.*.*
593 VAVAIKMKVYGLDEYVSLMKVEKKEAESGSLPLQEFLKLKSNAAGAQSSTVAVKMKEEEV

646 TVIN--MARTTEEVSKDALREILKSIGGSSMTLEDTAEPVSDAESIPDPPPRSWASEDEA
    . .  . ..*  *..**.*.   . .........  .  . .**.
653 NSKAYCLISETIVVNMDAVRSACGVVQSENLVIKSELSGPELSESVTSGLMELLGRNAGP

704 VNSPQIYSSRRKARKARAASKLSK

713 SKSWADQVEEAENEEEKQKE
```

Fig. 6. Comparison between the amino acid sequences of putative proteins encoded by genome segment 4 of RDV (upper) and WTV (lower). * and . denote identical and equivalent amino acids, respectively. Possible 'zinc-finger' motifs are boxed. The putative 'A' site of the purine NTP-binding domain of RDV is underlined.

hybridized specifically to genome segment 4 in alkaline Northern blot hybridization (Fig. 3). The location of the cDNA clones pRD3A7, 3A11, 4B2, 3G8 and 1D6 togther with the direction and extent of determination of the sequence are shown in Fig. 4. Most portions of the sequence were determined by reading from both directions using two different cDNA clones. Since the 3′-terminal portion of the genome was not covered by the cDNAs, the sequence was determined by reverse transcription of the viral transcript. The primer for direct RNA sequencing was designed based on the 3′-terminal sequence determined by the wandering spot analysis. One clone, pRD1D6, contained the complete sequence of the 5′-terminal region.

Nucleotide and amino acid sequences

The genome segment 4 was 2468 nucleotides long and encoded a protein of 727 amino acid residues with M_r of

```
Consensus 1    -!!!-g--g-gKs-----!--
                           t

RDV S4    28   TPGSMSSPSEGKTNAWVNSAY
                *   ..** ****  .  *
B19 NS1   318  TLWFYGPPSTGKTNLAMAIAK
                  ..  ...* ****
BPV NS1   307  STLFYGPASTGKTNLAKAICH


Consensus 2    --!!----G-GKt--!---!-
                           s

RDV S4    28   TPGSMSSPSEGKTNAWVNSAY
                 *    .*.** **
SDV NS1   68   TFQIVSPPSAGKNFFIETVLA
```

Fig. 7. Comparison between the 'A' site of the purine NTP-binding domain of RDV and of viruses in the *Parvoviridae*. Numbers on the left indicate the position of the amino acid from the N terminus. The sequences are taken from Gorbalenya & Koonin (1989). !, * and . denote hydrophobic residues, identical and equivalent amino acids, respectively. B19, BPV and SDV are the human parvovirus B19, bovine parvovirus and a densovirus from *Bombyx mori*, respectively.

79780 (Fig. 5 and 6). A putative segment-specific inverted repeat was present at positions 8 to 14 and 2454 to 2460 as has been found in the other genome segments (Fig. 5) (Uyeda *et al.*, 1989; Omura *et al.*, 1989).

Sequences similar to the 'zinc-finger' motif were present in the polypeptide from residues 231 to 335. The $CX_2CX_{17}HX_2C$ sequence from residues 311 to 335 fitted best to the consensus sequences $CX_{2-4}CX_{2-15}\dot{a}X_{2-4}a$ where a may be either cysteine or histidine (Berg, 1986).

Furthermore a sequence similar to the 'A' site of consensus purine NTP-binding patterns (Gorbalenya & Koonin, 1989) was present. The sequence present in the RDV polypeptide was particularly similar to those of the NS1 protein of viruses in the *Parvoviridae* (Fig. 7). Although the sequence did not perfectly match the two proposed consensus patterns, deviations appeared to be within the range of those of other purine NTP-binding protein sequences listed in Fig. 7.

Sequence comparison with WTV genome segment 4

The 5′ non-coding regions of the two viruses have the same length with the initiation codon at nucleotides 64 to 66. Both viruses have the strong initiation context of 5′ ANNAUGA proposed by Kozak (1987). There were no sequence homologies in this region. The 3′ non-coding regions of the two viruses differ both in length and sequence, and no sequence homology could be detected.

Significant homology in amino acid sequence was detected in several regions of the two genes except for the N- and C-terminal regions (Fig. 6). Secondary structure predictions made by the program DNASIS based on a method of Chou & Fasman (1978) also showed similarity to each other except for N- and C-terminal regions (data not shown).

One of the sequences similar to the 'zinc-finger' motif present in the RDV polypeptide was conserved in the two viruses at comparable positions. These are the $CX_2HX_{37}CX_2C$ of RDV at residues 269 to 314 and $CX_2HX_{38}CX_2C$ of WTV at residues 274 to 319. A sequence similar to the purine NTP-binding consensus found in the RDV polypeptide was not present in the WTV polypeptide.

Discussion

We have determined the nucleotide sequence of RDV genome segment 4 and the amino acid sequence of an encoded polyeptide was deduced. The amino acid sequence showed significant homology to that of the WTV genome segment 4 indicating that both polypeptides have the same functional role. There was no similarity in the nucleotide sequence of the terminal non-coding regions of the two viruses. However, the presumptive segment-specific inverted repeat was present in the RDV genome segment 4.

Recent analyses of polypeptides interacting with nucleic acids have revealed two motifs. One is called the 'zinc-finger' of nucleic acid-binding proteins (Berg, 1986), and the other is found among nucleotide-binding proteins (Dever *et al.*, 1987; Gorbalenya & Koonin, 1989). We found both motifs in the polypeptide encoded by RDV genome segment 4. In particular, the 'zinc-finger' motif is likely to be functional, since it was also present in WTV genome segment 4 at a comparable position in a highly conserved region. Although the sequence $CX_2CX_{17}HX_2C$ at residues 311 to 335 fits best to the consensus proposed by Berg (1986) in terms of the number of amino acid residues present in a loop, the conserved one at residues 269 to 314 ($X_2HX_{37}CX_2C$) has a longer loop. The length of the loop appears to be more variable than those of the proposed consensus. For example, the recently characterized 'zinc-finger' protein of poly(ADP–ribose) polymerase has the sequence $CX_2CX_{30}HX_2C$ (Mazen *et al.*, 1989). Whether these putative assignments are actually functional requires further analysis.

We thank the Research Center of Molecular Genetics, Hokkaido University for use of computer facilities. This research was supported by Grant-in-aid for Scientific Research No. 01560045, Grant-in-aid for Co-operative Research No. 63304013, and a grant from the Akiyama Foundation.

References

ANZOLA, J. V., ZHENGKAI, X., ASAMIZU, T. A. & 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.

ANZOLA, J. V., DALL, D. J., XU, Z. & NUSS, D. L. (1989*a*). Complete nucleotide sequence of wound tumor virus genomic segments encoding nonstructural polypeptides. *Virology* **171**, 222–228.

ANZOLA, J. V., XU, Z. & NUSS, D. L. (1989*b*). Complete nucleotide sequence of wound tumor virus genomic segment S7. *Nucleic Acids Research* **17**, 3300.

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.

BERG, J. M. (1986). Potential metal-binding domains in nucleic acid binding proteins. *Science* **232**, 485–484.

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

CHOU, P. Y. & FASMAN, G. D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. *Advances in Enzymology* **47**, 45–148.

DALL, D. J., ANZOLA, J. V., XU, Z. & NUSS, D. (1989). Complete nucleotide sequence of wound tumor virus genomic segment S11. *Nucleic Acids Research* **17**, 3599.

DEVER, T. E., GLYNIAS, M. J. & MERRICK, W. C. (1987). GTP-binding domain: three consensus sequence elements with distinct spacing. *Proceedings of the National Academy of Sciences, U.S.A.* **84**, 1814–1818.

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

FUKUMOTO, F., OMURA, T. & MINOBE, Y. (1989). Nucleotide sequence of segment S9 of the rice dwarf virus genome. *Archives of Virology* **107**, 135–139.

GORNALENYA, A. E. & KOONIN, E. V. (1989). Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Research* **17**, 8413–8440.

GUBLER, U. & HOFFMAN, B. J. A. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263–269.

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.

KIMURA, I., MINOBE, Y. & OMURA, T. (1987). Changes in a nucleic acid and a protein component of rice dwarf virus particles associated with an increase in symptom severity. *Journal of General Virology* **68**, 3211–3215.

KOZAK, M. (1987). An analysis of 5′ non-coding sequences 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.

LI, J. K. K., PARKER, B. & KOWLIK, T. (1987). Rapid alkaline blot-transfer of viral dsRNAs. *Analytical Biochemistry* **163**, 210–218.

MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning: A Laboratory Manual.* New York: Cold Spring Harbor Laboratory.

MAZEN, A., DE MURCIA, J. M., MOLENETE, M., SIMONIN, F., GRADWOHL, G., POIRIER, G. & DE MURCIA, G. (1989). Poly(ADP–ribose) polymerase: a novel finger protein. *Nucleic Acids Research* **17**, 4689–4698.

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

NAKASHIMA, K., KAKUTANI, T. & MINOBE, Y. (1990). Sequence analysis and product assignment of segment 7 of the rice dwarf virus genome. *Journal of General Virology* **71**, 725–729.

NUSS, D. L. & PETERSON, A. J. (1981). Resolution and genome assignment of mRNA transcripts synthesized *in vitro* by wound tumor virus. *Virology* **114**, 399–404.

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.

OMURA, T., ISHIKAWA, K., HIRANO, H., UGAKI, M., MINOBE, Y., TSUCHIZAKI, T. & KATO, H. (1989). The outer capsid protein of rice dwarf virus is encoded by genome segment S8. *Journal of General Virology* **70**, 2759–2764.

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. (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., OHSHIMA, K. & SHIKATA, E. (1987). Nucleotide sequence of rice dwarf virus genome segment 10. *Proceedings of the Japan Academy* **63**, 227–230.

UYEDA, I., KUDO, H., TAKAHASHI, T., SANO, T., OHSHIMA, K., MATSUMURA, T. & SHIKATA, E. (1989). Nucleotide sequence of rice dwarf virus genome segment 9. *Journal of General Virology* **70**, 1297–1300.

UYEDA, I., LEE, S. Y., LENG, G. W., AO, G. M., DENG, Z. P., LIANG, X. S., OHSHIMA, K. & SHIKATA, E. (1990). Size of dsRNA genome segments of rice dwarf and rice ragged stunt viruses determined by electron microscopy. *Journal of Electron Microscopy* **39**, 54–58.

XU, Z., ANZOLA, J. B. & NUSS, D. L. (1989). Assignment of wound tumor virus nonstructural polypeptides to cognate dsRNA genome segments by *in vitro* expression of tailored full-length cDNA clones. *Virology* **168**, 73–78.

YOKOYAMA, M., NOZU, Y., HASHIMOTO, J. & OMURA, T. (1984). *In vitro* transcription by RNA polymerase associated with rice gall dwarf virus. *Journal of General Virology* **65**, 533–538.

(*Received 30 March 1990; Accepted 3 July 1990*)