Arch Virol (1997) 142: 1719–1726

Archives of Virology



# Rice ragged stunt oryzavirus genome segments S7 and S10 encode non-structural proteins of $M_r$ 68 025 (Pns7) and $M_r$ 32 364 (Pns10)*

Brief Report

**N. M. Upadhyaya**[1], **K. Ramm**[1] **J. A. Gellatly**[1], **Z. Li**[1], **W. Kositratana**[2], and **P. M. Waterhouse**[1]

[1] CSIRO Plant Industry, Canberra, Australia,
[2] Department of Plant Pathology, Kasetsart University, Nakhon Pathom, Thailand

Accepted February 14,1997

**Summary.** The nucleotide sequences of genome segments S7 and S10 of a Thai-isolate of rice ragged stunt virus (RRSV) were determined. The 1 938 bp S7 sequence contains a single large open reading frame (ORF) spanning nucleotides 20 to 1 843 that is predicted to encode a protein of $M_r$ 68 025. The 1 162 bp S10 sequence has a major ORF spanning nucleotides 142 to 1 032 that is predicted to encode a protein of $M_r$ 32 364. This S10 ORF is preceded by a small ORF (nt 20–55) which is probably a minicistron. Coupled in vitro transcription-translation from the two major ORFs gave protein products of the expected sizes. However, no protein was visualised from S10 when the small ORF sequence was included. Proteins were expressed in *Escherichia coli* from the full length ORF of S7 (P7) and from a segment of the S10 ORF (P10) fused to the ORF of glutathione S-transferase (GST). Neither fusion protein was recognised by polyclonal antibodies raised against RRSV particles. Furthermore, polyclonal antibodies raised against GST-P7 fusion protein did not recognise any virion structural polypeptides. These data strongly suggest that the proteins P7 and P10 do not form part of RRSV particle. This is further supported by observed sequence homology (though very weak) of predicted

* The nucleotide sequences reported in this paper have been submitted to the GenBank nucleotide sequence database and have been assigned the accession numbers U66712 (S10) and U66713 (S7).

RRSV P7 and P10 with those of rice dwarf virus (RDV) non-structural proteins Pns6 and Pns9, respectively.

*

Rice ragged stunt oryzavirus (RRSV) is a member of the family *Reoviridae* and the type species of the genus *Oryzavirus* [2, 6, 11]. It has distinctive icosahedral particles and is transmitted exclusively by the insect, brown planthopper (*Nilaparvata lugens*)[2]. Its genome is composed of 10 dsRNA segments each having the genus-specific conserved terminal nucleotide sequences of 5′ GAUAAA - - - and - - - GUGC 3′ [19]. RRSV particles are comprised of five major, highly immunoreactive structural proteins with estimated $M_r$ of 33K, 39K, 43K, 70K and 120K and at least five minor structural proteins ($M_r$ 49K, 60K, 76K, 90K and 94K). Three more proteins ($M_r$ 31K, 63K and 88K) have also been identified from in vitro translation of RRSV genomic dsRNA, [8, 9] and designated as non-structural proteins. However, until recently, it was not known which of these proteins are encoded by which segment. We and others have determined the nucleotide sequences of RRSV S5, S8 and S9 and demonstrated that these segments encode, respectively, a $M_r$ 90K minor structural protein [7], a $M_r$ 67K major structural protein which is further self-processed to $M_r$ 46K, 43K and 26K proteins [15], and a $M_r$ 38K major structural protein [14, 18]. In this paper we report the nucleotide sequences of RRSV genome segments S7 and S10 and show that they encode non-structural proteins of $M_r \sim 68$K and 32K, respectively.

Recombinant plasmids containing sequences of S7 and S8 (28 clones) or S9 and S10 (12 clones) were identified from a library of cDNA clones of the RRSV genome [14] by hybridization using gel-purified, radiolabelled RRSV dsRNA (S7 and S8 migrate very closely in agarose gel electrophoresis, as do S9 and S10). These clones were sequenced (complete or partial) using universal M13 forward and reverse primers and sequence assembled using the programs of the University of Wisconsin Genetics Computer Group (GCG) [4] which produced four large contiguous sequences. Two of these have been previously identified and reported as S8 and S9 [14, 15]. Dot-blot hybridization of RRSV dsRNA segments (purified by high resolution polyacrylamide gel electrophoresis) with radiolabelled representative cDNA clones confirmed that a 1 926 bp (11 overlapping clones) contiguous sequence represents S7. Similarly, the remaining 1 156 bp (10 overlapping clones) contiguous sequence was shown to represent S10. Neither S7 nor S10 contiguous sequence had the previously reported [19] RRSV 5′ terminal conserved sequences but the S10 contiguous sequence had the previously reported 3′ terminal sequence of - - - AGGUGC. The nucleotide sequences of the S7 and S10 termini were therefore determined by amplifying these regions using the RACE method [5] and directly sequencing the RT-PCR products [15]. This increased the S7 sequence length to 1 938 bp and the S10 sequence to 1 162 bp and identified the previously reported terminal conserved sequences (5′ GAUAAAU - - - and - - - AGGUGC 3′).

The complete S7 and S10 nucleotide sequences together with the deduced amino acid sequences are presented in Fig. 1. A large ORF spans nt 20 to nt 1 843 in the S7 sequence and is predicted to encode a 608 aa polypeptide with a deduced molecular weight of 68 025. No other ORFs in either of the strands exceed 37 amino acids. The S10 possesses a long ORF which spans nt 142 to nt 1 032 and is capable of encoding a protein of 297 aa with a predicted molecular weight of 32 364. There is a short ORF (nt 20 to nt 55) preceding this major ORF. A similar coding arrangement with a short ORF preceding the major functional coding sequence has been found in the genome segment S5 of the recently characterised *Nilaparvata lugens* reovirus (NLRV) [10]. Such a short ORF (termed minicistron) in rice dwarf virus (RDV) segment S1 has been shown to down regulate the expression of the downstream major ORF [13].

In order to study the expression of S7 and S10 ORFs, cDNA sequences spanning these ORFs were reconstructed in the transcription vector pBluescript II $SK^{+/-}$ (Stratagene,USA) and coupled in vitro transcription-translation were performed with wheat germ extract using T3 polymerase (TNT T3 Coupled Wheat Germ Extract System). A near full-length S7-specific cDNA clone (pMNUS7_2a; nt 16 – nt 1 935) was reconstructed in pBluescript II $SK^{+/-}$ from three overlapping clones, RR405, RR511 and RR542, using internal restriction sites (*Bsp*EI and *Acc*I). A near full-length S10 sequence from a cDNA clone, RR499 (nt 74–1 138) was recloned into pBluescript II $SK^{+/-}$ as an *Eco*RI fragment (original cDNA clone had *Eco*RI adaptors at both ends). The transcription-translation from S7 ORF produced a $\sim$ 68K protein, P7 and that from S10 ORF produced a 32K protein, P10 (Fig. 2). Thus the predicted protein sizes of P7 and P10 are correct. However, a clone covering the entire S10 sequence, identical to the original cDNA sequence (produced by RT-PCR and nucleotides confirmed by sequencing) did not produce detectable levels of the 32K protein in a similar translation-translation experiment. This could have been because of active translation of the minicistron and consequent down regulation of transcription of the downstream major ORF.

Comparison of S7 and S10 nucleotide and deduced amino acid sequences with those in the GenBank and EMBL databases using BLAST or FASTA programs revealed no significant similarities. This was not surprising because RRSV genome segments S5, S8 and S9 also show no similarities with other reovirus sequences [7, 14, 15]. However, pair-wise similarity scores with RDV genome segments (obtained using GCG GAP program) suggest that RRSV S7 is analogous to RDV S6 and that RRSV S10 is analogous to RDV S9. Both RDV S6 and S10 encode non-structural proteins [13]. At the deduced amino acid level, the sequence similarity between RRSV P7 and RDV P6 is 50%, with 23% amino acid identity. RRSV P10 and RDV P9 have 49% similarity and 24% identity. Although these identities are very low, none of the other RDV segments showed identities of more than 20% with RRSV P7 or P10.

To investigate whether RRSV P7 and P10 are structural or non-structural proteins, they were expressed as fusion proteins in bacteria and analysed using antibodies raised against RRSV particles. The near full-length S7 sequence

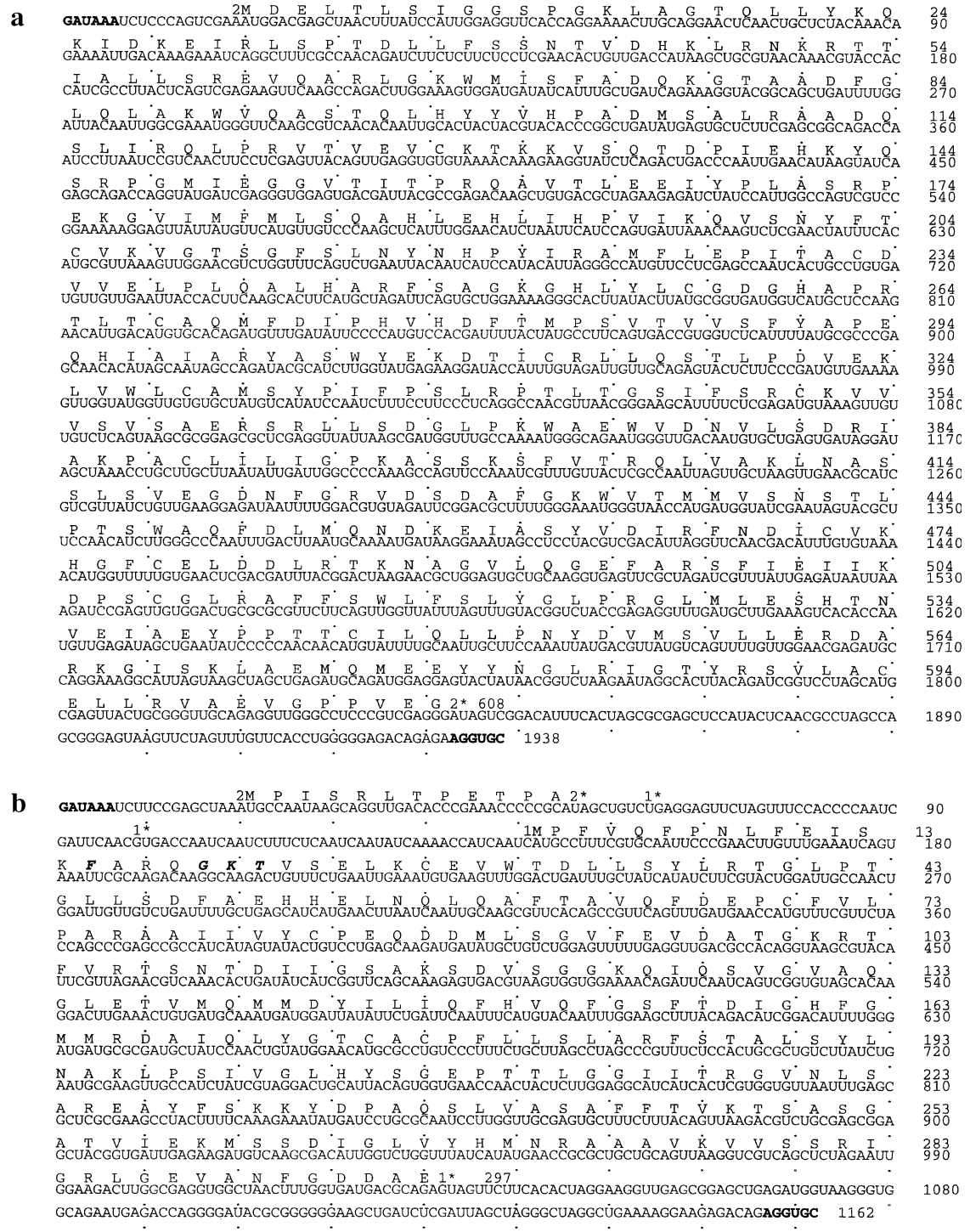


**Fig. 1.** Complete nucleotide and predicted amino acid sequences of RRSV genome segments S7 and S10. The nucleotide sequences of S7 (**a**) and S10 (**b**), determined from overlapping cDNA clones and endspecific cDNA clones, together with deduced amino acid sequences (indicated above the nucleotide sequence) are shown. Genus specific terminal sequences (bold letters) and the predicted translation termination codon (*) are indicated. The number preceding (*) indicates the reading frame. Possible purine nucleotide binding domain in S10 amino acid sequence is indicated (bold italics)

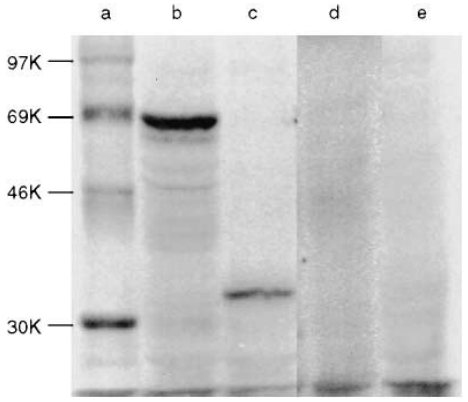


**Fig. 2.** In vitro transcription and translation from S7 and S10 coding region sequences. An autoradiograph (using PhosphorImager, Molecular Dynamics) of electrophoretically-separated (10% SDS-PAGE and blotted onto Hybond N) products of coupled in vitro transcription/translation of recombinant plasmids pMNUS7_2a containing S7 ORF (*b*), RR499 containing S10 ORF without minicistron (*c*), pMNUS10_1a containing S10 ORF with preceding minicistron (*d*) and BlueScript $SK^+$ plasmid vector (*e*) using the TNT Coupled Wheat Germ Extract System with T3 RNA ploymerase and $^{35}$S-methionine labelling. *a* contains $^{3}$H-labelled protein molecular weight markers

(pMNUS7_2a; nt 16-1 935) and a fragment of S10 sequence (clone RR217; nt 370–849) were cloned into the glutathione-S-transferase (GST) fusion protein expression vectors pGEX-1 and pGEX-2 [12], respectively. These fusion proteins and a fusion protein derived from RRSV S8 (encoding a major structural protein P8) [15] were expressed and analysed by immunoblot essentially as described previously [14]. The antiserum recognised the P8 fusion protein but not the P7 or P10 fusion proteins (Fig. 3). Furthermore, an

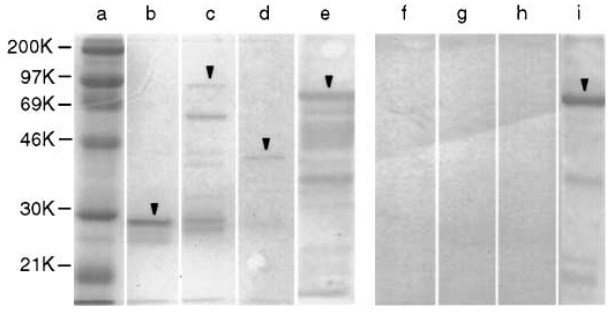


**Fig. 3.** Antigenicity of the GST-P7 and GST-truncated P10 fusion proteins. Electrophoretically-separated (10% SDS-PAGE) and stained (Coomassie blue R-250) protein molecular weight standards (*a*), GST alone (*b*), GST-P7 (*c*), GST-truncated P10 (*d*) and MBP-P8 [15] (*e*, + control) fusion proteins and Western blots of similarly separated GST alone (*f*), GST-P7 (*g*), GST-truncated P10 (*h*) and MBP-P8 (*i*) fusion proteins probed with alkaline phosphatase labelled anti-RRSV IgG and immunoreactive proteins detected by the use of BCIP and NBT as substrates. ▼ indicate respective proteins

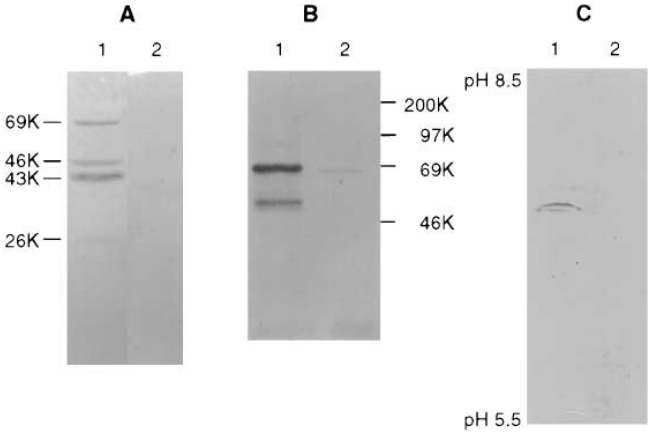


**Fig. 4.** Immunoblotting of RRSV proteins. Western blots (nitrocellulose membrane) of **A** RRSV particle proteins (separated by 10% SDS-PAGE) probed with alkaline phosphatase labelled, anti-P8 IgGs (*1*) or anti-P7 IgGs (*2*), **B** protein extracts (separated by 10% SDS-PAGE) from RRSV infected (*1*) and healthy (*2*) rice shoots probed with alkaline phosphatase labelled anti P7 fusion protein IgGs and **C** protein extracts from RRSV-infected (*1*) and healthy rice shoots (*2*), separated by isoelectric focussing (IEF, horizontal 5% polyacrylamide gel pH 5.5 to 8.5) and probed with alkaline phosphatase labelled anti-P7 fusion protein IgGs and immunoreactive proteins detected by the use of BCIP and NBT as substrates

antiserum raised against the P7 fusion protein did not recognise any proteins in an RRSV particle preparation (Fig. 4A). These data show that P7 and P10 are not major structural proteins and suggest that they are not present in the viral particles. In vitro translation and immunoprecipitation studies [8, 9] have shown that RRSV possesses three nonstructural proteins with $M_r$ of 88K, 63K and 31K. This is consistent with our data and suggests that the latter two are encoded by S7 and S10, respectively. The discrepancy in the size of S7 (63K/68K) is probably due to the inaccuracy of size determination by gel electrophoresis. This information coupled with the sequence similarity (although weak) of P7 and P10 with the non-structural proteins of RDV suggest that these proteins encoded by RRSV S7 and S10 are non-structural. We therefore term them Pns7 and Pns10.

The non-structural proteins encoded by animal reoviruses are known to function as RNA binding proteins, subunit proteins of microtubules, proteins associated with viral replication or as proteins associated with the release of virus particles from the cell [3, 16]. It therefore seems likely that Pns7 and Pns10 will each have one or more of the above mentioned functions. RRSV Pns 10 contains the amino acid triplet GKT bordered by hydrophobic residues which has been identified as a motif common in RNA binding proteins [1, 4] suggesting this protein has an RNA binding function. A zinc-finger motif is also usually present in such proteins [1, 4] but we could not identify this motif in Pns10.

Antibodies raised against the Pns7 fusion protein gave a strong signal against a ~ 68K protein present in immunoblots of proteins from RRSV infected plants (Fig. 4B), prepared as described previously [7]. This suggests that the protein is either highly immunoreactive or highly abundant. Anti-Pns7 antibodies also reacted with a protein of ~ 52K in infected plant extract, presumably a degradation product of the 68K protein. In contrast, the P7 antiserum produced a weak band of similar size in healthy rice plant extracts. When the immunoblot was repeated using P7 antiserum that had been cross-absorbed with GST-P7 fusion protein, the intensity of the 68K band in the infected material diminished while the intensity of the band in the healthy plant extracts remained the same (data not shown). This observation suggests that the 68K cross-reacting protein in the healthy plant extract is different from that in the infected plant extract. Furthermore, when prior to immunoblotting, protein extracts were subjected to isoelectric-focussing (IEF) (using a Ampholine PAGplate thin-layer gel of pH range 5.5 to 8.5, according to the manufacturer's instruction; Multiphor, LKB), no reacting band was observed in the healthy plant extract (Fig. 4C). The second cross-reacting protein in the infected plant extract (~ 52K) migrated closer to the 68K protein suggesting that it was a degraded form of the Pns7. It seems likely that the intensity of the Pns7 band in infected plants reflected an abundance of the protein in infected cells. This abundance suggests that Pns7 is a microtubule subunit since reoviral encoded microtubules are similarly abundant in infected cells whereas other reoviral NS proteins are found in low concentrations [16]. Nevertheless, further work is needed to confirm the tentative function assignments for these non-structural proteins.

## Acknowledgements

Financial support from the Rockefeller Foundation is gratefully acknowledged. Thanks to Drs. Mark Gibbs, Petra Schünmann and Frank van de Loo for critical reading of the manuscript. RRSV infected materials were imported into Australia under Australian quarantine permit (28 546, BM 1 243).

## References

1. Berg JM (1986) Potential metal-binding domains in nucleic acid binding proteins. Science 232: 485–494
2. Boccardo G, Milne RG (1984) Plant reovirus group. CMI/AAB Descriptions of Plant Viruses, No. 294
3. Cross RK, Fields BN (1976) Reovirus-specific polypeptides: analysis using discontinuous gel electrophoresis. J Virol 19: 162–173
4. Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12: 387–396
5. Frohman MA, Dush MK, Martin GR (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci USA 85: 8 998–9 002

6. Holmes IH, Boccardo G, Estes MK, Furuichi MK, Hoshino Y, Joklik WK, McCrae M, Mertens PPC, Milne RG, Shikata E, Winton JR (1994) Family *Reoviridae*. In: Murphy FA, Fauquet CM, Bishop DHL, Ghabrial SA, Jarvis AW, Martelli GP, Mayo MA, Summers MD (eds) Virus Taxonomy. Classification and Nomenclature of Viruses. Sixth Report of the International Committee on Taxonomy of Viruses. Springer, Wien New York, pp 208–239 (Arch Virol [Suppl] 10)
7. Li Z, Upadhyaya NM, Kositratana W, Gibbs A, Waterhouse PM (1996) Genome segment 5 of rice ragged stunt oryzavirus encodes a virion protein. J Gen Virol 77: 3 155–3 160
8. Lu HH, Gong Z, Cao T (1988) Studies on the RNA polymerase activity associated with rice ragged stunt virus. Sci Sinica Ser B 31: 573–575
9. Lu H, Gong Z, Cao T (1990) Studies on the genomic coding assignments of rice ragged stunt virus. Chin J Virol 6: 167–195
10. Nakashima N, Kaizumi M, Watanabe H, Noda H (1996) Complete sequence of the *Nilaparvata lugens* reovirus: a putative member of the genus *Fijivirus*. J Gen Virol 77: 139–146
11. Nuss DL, Dall DJ (1990) Structural and functional properties of plant reovirus genomes. Adv Virus Res 38: 249–306
12. Smith DB, Johnson KS (1986) Single-step purification of polypeptides expressed in *Escherichia coli* as fusion with glutathione S-transferase. Gene 67: 31–40
13. Suzuki N (1995) Molecular analysis of the rice dwarf virus genome. Semin Virol 6: 89–95
14. Upadhyaya NM, Yang M, Kositratana W, Ghosh A, Waterhouse PM (1995) Molecular analysis of rice ragged stunt oryzavirus segment S9 and sequence conservation among isolates from Thailand and India. Arch Virol 140: 1945–1956
15. Upadhyaya NM, Zinkowsky E, Li Z, Kositratana W, Waterhouse PM (1996) The $M_r$ 43K major capsid protein of rice ragged stunt oryzavirus is a post-translationally processed product of a $M_r$ 67,348 polypeptide encoded by genome segment 8. Arch Virol 141: 1 689–1 701
16. Urakawa T, Roy P (1988) Bluetongue virus tubules made in insect cells by recombinant baculovirus: expression of NS1 gene of bluetongue virus serotype 10. J Gen Virol 62: 3 919–3 927
17. Uyeda I, Kudo H, Yamada N, Matsumura T, Shikata E (1990) Nucleotide sequence of rice dwarf virus genome segment 4. J Gen Virol 71: 2 217–2 222
18. Uyeda I, Suda N, Lee SY, Yan J, Hataya T, Kimura I, Shikata E (1995) Rice ragged stunt *Oryzavirus* genome segment 9 encodes a 38 600 $M_r$ structural protein. J Gen Virol 76: 975–978
19. Yan J, Kudo H, Uyeda I, Lee S, Shikata E (1992) Conserved terminal sequences of rice ragged stunt virus genomic RNA. J Gen Virol 73: 785–789

Authors' address: Dr. N. M. Upadhyaya, CSIRO Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia.

Received November 11, 1996