The proteins encoded by rice grassy stunt virus RNA5 and RNA6 are only distantly related to the corresponding proteins of other members of the genus Tenuivirus

Shigemitsu Toriyama,¹ Takao Kimishima² and Mami Takahashi¹

¹National Institute of Agro-Environmental Sciences, Kannondai 3, Tsukuba, Ibaraki 305, Japan
²Faculty of Agriculture, Ibaraki University, Ami-machi, Ibaraki 300-03, Japan

The genome of rice grassy stunt virus (RGSV) consists of six RNA segments. The nucleotide (nt) sequences of the two smallest segments, RNAs 5 and 6, were determined and found to comprise 2704 and 2584 nt, respectively. The 5′- and 3′-terminal sequences of both RNAs were identical over a length of 21 nt and could potentially form a panhandle-like structure due to intramolecular complementarity. Each RNA segment contained a virus (v) sense open reading frame (ORF) in the 5′-proximate region, and a virus complementary (vc) ORF in the 3′-proximate region, indicating an ambisense coding strategy. The protein encoded by the ORF on the vc strand of RNA5 was identified as the viral nucleocapsid protein (M, 35 927). The ORF on the v strand of RNA6 encoded a protein of M, 20581 which represented the major nonstructural protein, previously shown to be produced in RGSV-infected rice tissues. The predicted proteins encoded by RGSV RNAs 5 and 6 were only distantly similar in sequence to the four proteins encoded by RNAs 3 and 4 of other viruses belonging to the genus Tenuivirus. These low sequence similarities, together with the apparently distinct number of genome segments, set RGSV apart from the other tenuiviruses and indicate that it should be placed in a taxonomically separate genus.

Introduction

Rice grassy stunt virus (RGSV) causes a destructive disease of rice (Oryza sativa L.) in rice growing areas of southern and south-eastern Asia (Rivera et al., 1966; Iwasaki & Shinkai, 1979; Hibino, 1986). RGSV is circulatively and propagatively transmitted by the brown planthopper, Nilaparvata lugens (Stål) (Rivera et al., 1966), and two other species (Hibino, 1986), but transovarial transmission by N. lugens has not been detected. RGSV is a member of the genus Tenuivirus, which includes rice stripe virus (RSV), rice hoja blanca virus (RHBV) and maize stripe virus (MStV) (Toriyama, 1995; Toriyama & Tomaru, 1995).

RGSV particles are thread-like, 6–8 nm wide and mostly circular with a contour length of 200–2400 nm (Hibino et al., 1985). The nucleocapsid (NC) proteins have Mr values of 31 000 (Hibino et al., 1985), and 34 000 and 31 500 (Toriyama, 1985). The filamentous virus particles of RGSV contain RNA-dependent RNA polymerase activity similar to the filamentous RSV particles (Toriyama, 1986, 1987). In RSV- and MStV-infected tissues, large quantities of a noncapsid protein are produced (Kiso & Yamamoto, 1973; Gingery et al., 1981); this protein was subsequently termed the major nonstructural (NS) protein. Recently, a noncapsid viral protein has also been detected in RGSV-infected rice (Miranda & Koganezawa, 1995). All these characteristics of RGSV are common with other viruses of the genus Tenuivirus.

There are four different RNA species in RSV and RHBV (Toriyama, 1982; Ramirez et al., 1992) and five in MStV (Falk & Tsai, 1984). The 5′- and 3′-terminal sequences of about 18 nt are conserved and complementary in RNA3 and RNA4 of RSV, MStV and RHBV (Takahashi et al., 1990; Ramirez & Haenni, 1994). All RNAs of segments 2, 3 and 4 in tenuiviruses have an ambisense coding strategy, and RNA5 of MStV and RNA1 of RSV are negative-stranded (Huiet et al., 1993; Toriyama et al., 1994). The nucleotide sequences of the tenuiviruses share some features with the phleboviruses, which belong to the family Bunyaviridae (Kakutani et al., 1990; Takahashi et al., 1990, 1993; Zhu et al., 1991; Huiet et al., 1991; Ramirez & Haenni, 1994; Toriyama et al., 1994). In
particular, the terminal sequences of each segment and certain amino acid sequence motifs of the putative RNA polymerase are similar in tenuiviruses and phleboviruses (Toriyama et al., 1994).

The RNA genome of RGSV has been only partially characterized (Toriyama, 1985, 1987). In order to determine the coding characteristics of RGSV RNAs in detail, we are attempting to sequence the RGSV genome. In this paper, we report the complete nucleotide sequence of RNA segments 5 and 6, and compare them with the RNAs of other viruses, in particular other tenuivirus species.

**Methods**

- **Virus and RNAs.** Rice plant materials, infected with RGSV, were obtained from H. Koganezawa (formerly of the International Rice Research Institute, Philippines). RGSV was purified from infected leaves and stems by the method described for RSV purification (Toriyama, 1986). RNA was obtained from virus particles by SDS–phenol extraction (Toriyama, 1986). RGSV dsRNAs were separated by electrophoresis on a 1% agarose gel and each band was electro-eluted and ethanol-precipitated. The yield of RGSV from rice plant was 100–150-fold lower than that of RSV (Toriyama, 1987). RNA preparations contained less ssRNA, which formed broad bands during agarase gel electrophoresis, probably due to degradation. dsRNA was therefore preferred as template for cDNA synthesis.

- **NS protein.** The NS protein of RGSV was purified by a slight modification of the method of Miranda & Koganezawa (1995). RGSV-infected rice leaves (50 g) were homogenized in 150 ml 0.2 M K$_2$HPO$_4$, 0.1 M citric acid, pH 7.1 and the extract was centrifuged at 100000 g for 90 min. The supernatant fraction was mixed with an equal volume of 0.1 M citric acid and centrifuged. The pellet was dissolved in phosphate–citrate buffer, pH 7.1. After two cycles of precipitation and resuspension, the NS protein preparation was dissolved in 10 mM Tris–HCl, pH 7.5, and centrifuged at 10000 g for 10 min. The pellet containing the NS protein was dissolved in 10 mM Tris–HCl, pH 7.5, and then dialysed against distilled water and lyophilized.

- **Terminal sequence of RGSV RNAs.** Total RGSV RNA was 3’ end-labelled with 5’$[^{32}P]$$p$Cp and T4 RNA ligase, and electrophoresed on a 1% agarase gel. The bands of ss- and dsRNAs were excised and each RNA was electro-eluted and precipitated with ethanol. Terminal

---

**Fig. 1.** cDNA clones used to determine the nucleotide sequence of RGSV segment RNA5 (a) and RNA6 (b). cDNA clones synthesized by using non-denatured total RGSV RNA as template and synthetic primer P1; cDNA clones synthesized by using denatured dsRNA as template and primer mixture (P1 and a random hexamer); cDNA clones (with DNA inserts amplified by PCR); cDNA clones synthesized by primer extension using primer P2.
null
components of RGSV contained five distinct species of dsRNA (d2–d6) (Fig. 2b, lanes 1 and 2), and other not so well separated bands that contained ssRNA (s2–s6) (Fig. 2b, lanes 2 and 3). Banding of ssRNA species of RGSV on agarose gels was broad and the density of each species frequently differed with preparations, as shown in Fig. 2(b) (s2–s6, lanes 2 and 3). Generally, ssRNA species of RGSV seemed to be present in relatively low amounts when compared to RSV.

The largest dsRNA species (d1) (Fig. 2b, lane 3) was found in RGSV preparations recovered from the faint broad band (B3) shown in Fig. 2(a). This band was not always visible with RGSV preparations made from 80 g of rice leaf material. The band containing d1 after sucrose gradient centrifugation was therefore removed and the virus was recovered by centrifugation. The infectious portion of RGSV was found in the heavier fraction below the two distinct bands (Iwasaki et al., 1987). The terminal 16 nt of the RNA5 and RNA6 were complementary. Nucleotides of the RNA terminal sequences of RSV, MStV, RHBV and UUKV that are identical to RGSV RNA sequences are indicated by (·).

and dsRNAs and were also the same as the terminal nucleotide sequence of RSV and other tenuiviruses. Beyond nt 11, two lines of nucleotide spots (wandering spots) were found in dsRNAs 5 and 6 (Takahashi et al., 1990) and also in ssRNAs 5 and 6, suggesting that these ssRNAs might be a mixture of plus- and minus-sense RNAs. However, these sequences were partly ambiguous (data not shown). Terminal sequences obtained by direct sequencing and analysis of the cDNA sequence are shown in Fig. 3. The 5′- and 3′-terminal 21 nt were the same in RNA5 and RNA6. The 5′-terminal 17 nt and 3′-terminal 16 nt are complementary and could form pan-handle-like structures. These terminal sequences of RGSV RNAs are also very similar to those of other tenuiviruses and Uukuniemi virus, a phlebovirus (Fig. 3).

Sequence of RNA5. The complete nucleotide sequence of RGSV RNA5 expressed as virus (v) sense DNA is shown in Fig. 4. RNA5 comprises 2704 nt with a base composition of 32.8% A, 32.2% U, 19.0% C and 15.9% G. No significant similarities were detected in database searches. The sequence contains a long ORF in the 5′-proximal region of virus complementary (vc) RNA5 from nt 1558–2535 in the virus sense. The predicted translation product contains 325 amino acids with an M, of 35927 (35-9K). An additional ORF found at the 5′ end of vRNA5 encodes a protein of 191 amino acids with an M, of 21586 (21-6K) (Figs 4 and 6). Between these two ORFs there is a long intergenic region of 878 nt (nt 680–1557 from the 5′ end) in which oligo(A)- and oligo(U)-rich sequences frequently appear. Other smaller ORFs were found in a different region and in the same frame as vRNA5 (M, 12000), and as a result of a 1 frameshift of vcRNA5 (M, 17000), but no similarities were found with any proteins in the databases.

A database search with the sequence of the putative product of ORF 35-9K revealed a weak but significant similarity (21–22% identity) with the NC protein of other tenuiviruses. The amino acid sequence of the 21-6K protein was 19–23% identical with the NS3 protein of other tenuiviruses (see Fig. 7a).
Rice grassy stunt virus RNA5 and RNA6

Fig. 4. Nucleotide sequence of RGSV RNA5 presented as vDNA sequence and the amino acid (single letter) sequence encoded by the two large ORFs. The amino acid sequence encoded by vRNA is printed above the DNA sequence; that encoded by vcRNA is printed below. Asterisks indicate termination codons. The amino acids of the RGSV NC protein that were identified by Edman degradation analysis are boxed.
Fig. 5. Nucleotide sequence of RGSV RNA6 presented as vDNA sequence and the amino acid (single letter) sequence encoded by the two large ORFs. Refer to Fig. 4 for further details. The amino acids of the RGSV NS protein that were identified by Edman degradation analysis are boxed.
Sequence of RNA6. Fig. 5 shows the entire nucleotide sequence of RGSV RNA6 expressed as vDNA. RNA6 consists of 2584 nt with a base composition of 32.4% A, 32.9% T, 17.5% C and 17.2% G. No matches were found in database searches, although the 5'- and 3'-terminal sequences were conserved among tenuiviruses. One ORF was found on vRNA6 and one on vcRNA6. The ORF on vRNA6 encodes a protein of 178 amino acids with an Mr of 20581, whereas the ORF on vcRNA6 encodes a protein of 325 amino acids with an Mr of 36421 (Figs 5 and 6). Between these two ORFs there is a long intergenic region of 913 nt (nt 615–1527 from the 5' end) which is rich in oligo(A) and oligo(U) sequences. Another possible ORF resulting from -1 frameshift was found on vcRNA6, but the putative protein product (Mr 11904) did not match any protein in the databases.

Database searches with the sequence of the putative product of ORF 20±6K revealed a weak but significant similarity (26–31% identity) with the major NS proteins of other tenuiviruses, which are encoded on vRNA4. Another ORF on vcRNA6 also showed similarity to proteins encoded by vcRNA4 of other tenuiviruses; the amino acid identity was 27–28% (Fig. 7b).

Amino acid sequence of NC and NS proteins. The sequence of the N-terminal 21 amino acids of the NC protein was determined and found to be identical with the N-terminal amino acid sequence of the 35-9K protein deduced from the ORF on vcRNA5 (Fig. 4). To determine the internal sequence, the NC protein was digested with S. aureus V8 protease. The NC protein was fragmented into three major peptides and one minor peptide. The 18 amino acid sequence of the smallest of the major fragments was determined. This internal sequence was identical to amino acids 262–280 encoded by the 35-9K ORF on vcRNA5. The other two fragments had the N-terminal sequence of non-digested NC protein. Thus, the NC protein of RGSV is encoded by vcRNA5 (Fig. 4).

The N-terminal amino acid of the NS protein was blocked. The NS protein was fragmented into three peptides, one major and two minor, by V8 protease digestion. The N termini of the two minor peptides were also blocked and therefore were presumably derived from the N terminus of the NS protein. The N-terminal 13 residues of the major fragment were identical to the sequence deduced for residues 57–69 of the 20±6K protein in Fig. 5. Thus, the NS protein specific to RGSV infection is encoded on vRNA6.

Discussion

The 5'- and 3'-terminal sequences of RGSV RNA5 and RNA6 are very similar to those of other tenuiviruses and are complementary to each other for 17 and 16 terminal nucleotides, respectively. Furthermore, the presence of ORFs on both the v- and vc strands indicates that these RNA segments have an ambisense coding strategy (Fig. 6). Both segments have a long intergenic region rich in oligo(A) and oligo(U) sequences and it is likely that some regions can be folded into secondary structures. Also, it was shown that the NC protein of RGSV is encoded by vcRNA5 and the NS protein by vRNA6. The Mr values of the putative NC protein and NS protein (NS6) were almost identical to those measured by mass spectrum analysis, Mr 35782 and Mr 20722, respectively (H. Hirano, unpublished data). Thus, RGSV RNAs 5 and 6 correspond to RNAs 3 and 4, respectively, of other tenuiviruses. However, the sequence identity of the RGSV NC protein to those of RSV, MSV and RHBV was only about 21% over 230 amino acids, whereas the identity among the NC proteins of these latter three tenuiviruses was higher (48–65%) over a longer stretch of 310 amino acids (Fig. 7a).
The predicted protein encoded by vRNA5 of RGSV also shares significant sequence identity with the NS proteins encoded by vRNA3 of RSV, MStV and RHBV (Fig. 7a).

The NS protein encoded by RGSV RNA6 (NS6) is similar to the major NS protein encoded by vRNA4 of other tenuiviruses (Fig. 7b). For this protein only a low sequence similarity of 26–31% was found between RGSV and the other tenuiviruses, although Miranda & Koganezawa (1995) reported a positive serological cross-reaction for this protein in the case of RGSV and RSV. On the contrary, the similarity between the major NS proteins of other tenuiviruses was strikingly high, 61–74% (Fig. 7b).

The terminal nucleotide sequences are well conserved between tenuiviruses and phleboviruses of the family Bunyaviridae. In particular, the RNA-dependent RNA polymerase encoded by RSV RNA1 is homologous to the L protein of phleboviruses: 39–3% identity over 569 amino acids of the putative RNA polymerase domain (Toriyama et al., 1994). The NC protein of RGSV has 22–8% identity over 184 amino acids with the NC proteins of Punta Toro phlebovirus (Ihara et al., 1984), and 20–4% identity over 221 amino acids with that of Toscana phlebovirus (Giorghi et al., 1991). This is the same level of identity as now found between the NC proteins of RGSV and the other tenuiviruses. Thus, RGSV appears only distantly related to the other tenuiviruses as far as the NC proteins and major NS proteins are concerned. The number of RNA segments in tenuiviruses was reported to vary from four in RSV and RHBV to five in MStV, whereas six RNA segments have now been found in RGSV. Recently, we confirmed that RNA3 and RNA4 of RGSV do not represent subgenomic RNAs but genuine genomic RNA segments, since cDNA from RNA3 and RNA4 did not hybridize with the other RNA genomic RNA species (S. Toriyama, T. Kimishima, M. Takahashi & T. Shimizu, unpublished data).

The biological properties of RGSV seem to differ from those of other tenuiviruses; for example RGSV induces yellowing and severe stunting, whereas other tenuiviruses induce typical striate symptoms and blight but not marked stunting. So far, transovarial transmission has not been shown for RGSV in their planthopper vectors. Taken together, these properties suggest that it might be more appropriate for RGSV to be excluded from the genus Tenuivirus and classified in a new separate genus.

We thank Dr. M.A. Mayo (Scottish Crop Research Institute) for critical reading of the manuscript and helpful advice. We also thank Dr. H. Koganezawa (formerly of IRRI) for the gift of RGSV-infected rice material, Professor H. Hirano (Yokohama City University, Kihara Institute for Biological Research) for help with protein analysis, Dr. K. Akatsu (Ibaraki University) for his interest and encouragement and T. Shimizu (Ibaraki University) for his help with some of the experiments. This work was supported in part by Grant-in-Aid from the Ministry of Agriculture, Forestry and Fisheries [Biocosmos Program 97-L-C-1(1)].

References


Received 23 January 1997; Accepted 11 April 1997