Journal of General Virology (1996), 77, 139–146. *Printed in Great Britain*

Complete nucleotide sequence of the *Nilaparvata lugens* reovirus: a putative member of the genus *Fijivirus*

Nobuhiko Nakashima,* Mika Koizumi, Hirofumi Watanabe and Hiroaki Noda

National Institute of Sericultural and Entomological Science, Owashi, Tsukuba, Ibaraki 305, Japan

The nucleotide sequences of all genome segments of the *Nilaparvata lugens* reovirus (NLRV), which is found in the brown planthopper *Nilaparvata lugens*, have been determined and some genes have been assigned to structural and functional proteins. The genome of NLRV consists of 28699 nucleotides and contains at least 11 large open reading frames (ORFs). The genome of NLRV is the largest among viruses of the family *Reoviridae* reported to date. The deduced amino acid sequence of genome segment S1 contained the major motifs of RNA polymerase and that of S7 had the purine NTP-binding motif. Based on the molecular masses of the deduced proteins and the particle structure of NLRV, segments S1, S3 and S7 were assigned to the 160, 140 and 75 kDa proteins, respectively, that are located in the inner core. It was deduced that S2 codes for the 135 kDa protein (B spike), which is located on the surface of the inner core. Most reported ORFs of rice black streaked dwarf virus (RBSDV), which shares many properties with NLRV, had similarities with the corresponding ORFs of NLRV. An exception was S7 ORF2, which is found in RBSDV but not NLRV and may therefore be involved in multiplication of RBSDV in rice plants. These results and our previous observations indicate that NLRV should be classified in the genus *Fijivirus*.

Introduction

Planthoppers and leafhoppers are vectors of phytopathogenic viruses which are classified into three genera, *Phytoreovirus*, *Fijivirus* and *Oryzavirus*, in the family *Reoviridae* (Murphy *et al.*, 1995). The brown planthopper, *Nilaparvata lugens*, is one of the most important pests of the rice plant because of its sucking damage and its ability to transmit plant-pathogenic viruses (Sōgawa, 1982; Hibino, 1989). We found a virus, *Nilaparvata lugens* reovirus (NLRV), in a healthy colony of the brown planthopper. NLRV has 10 segmented dsRNAs, termed S1 to S10 based on their electrophoretic mobility in polyacrylamide gels (Noda *et al.*, 1991*a*). The terminal nucleotide sequences of each genome segment of NLRV are similar to maize rough dwarf virus (MRDV) and rice black streaked dwarf virus (RBSDV), which are members of the genus *Fijivirus* (Noda *et al.*, 1994). All fijiviruses described so far are phytopathogenic. However, NLRV does not multiply in rice plants, although the virus is transmitted from a viruliferous planthopper to a non-viruliferous one through the rice plant (Nakashima & Noda, 1995). Some non-phytopathogenic reoviruses are also found in leafhoppers and planthoppers, such as leafhopper A virus (Boccardo *et al.*, 1980; Ofori & Francki, 1985) and *Peregrinus maidis* virus (Falk *et al.*, 1988). Non-pathogenic reoviruses in planthoppers and leafhoppers can help to explain the origin of plant pathogenic reoviruses because of their common characteristics such as genome organization, life-cycle and morphology (Nuss & Dall, 1990; Nault & Ammar, 1989).

In the genus *Phytoreovirus*, many nucleotide sequences of genome segments have been reported: all 12 genome segments of rice dwarf virus (RDV; Uyeda *et al.*, 1995), S4 to S12 of wound tumor virus (Nuss & Dall, 1990) and S3 (Takahashi *et al.*, 1994), S8 and S10 (Noda *et al.*, 1991*b*) of rice gall dwarf virus (RGDV). On the other hand, in the genus *Fijivirus*, only S6 of MRDV (Marzachi *et al.*, 1991) and S7, S8 (Azuhata *et al.*, 1993) and S10 (Uyeda *et al.*, 1990) of RBSDV have been fully sequenced. We previously reported the nucleotide sequences of genome segments S8 (Nakashima & Noda, 1994) and S10 (Noda *et al.*, 1994) of NLRV; their deduced amino acid sequences were similar to those of RBSDV S10 and MRDV S6, respectively. However,

---

* Author for correspondence. Fax +81 298 38 6028.
e-mail nakaji@ss.nises.affrc.go.jp

The nucleotide sequence date reported in this paper for NLRV S1–S7 and S9 will appear in the GSDB, DDBJ, EMBL and NCBI databases with the accession numbers D49693–49700, respectively.

amino acid sequence similarity between NLRV and fijiviruses (about 18%) was much lower than that between MRDV and RBSDV (more than 85%; Azuhata *et al.*, 1993).

Comparison of nucleotide and amino acid sequences of non-pathogenic and plant-pathogenic reoviruses is necessary for elucidating the functions of proteins such as those involved in multiplication in plants. This paper reports the nucleotide sequences of genome segments S1–S7 and S9 of NLRV, and the assignment of structural proteins to each segment. Morphological properties of NLRV are also described because A and B spikes, which have been observed in fijiviruses (Milne & Lovisolo, 1977), have not been confirmed in NLRV.

Methods

cDNA library and sequencing. The cDNA library of NLRV genome segments and the screening method used to identify the clones containing each genome segment were reported previously (Noda *et al.*, 1994). A total of 153 clones containing each genome segment were identified (S1, 15 clones; S2, 36; S3, 39; S4, 24; S5, 9; S6, 11; S7, 13; S9, 6). Deletion mutants were derived from the cDNA clones of each genome segment and sequenced as described (Noda *et al.*, 1994). A part of the S1 cDNA (residues 2640 to 3394) was absent in the library and only one clone containing residues 2434 to 3427 of the cDNA from S5 was available. Therefore, DNAs corresponding to these regions were amplified by RT–PCR from genomic dsRNA using synthesized primers based on the flanking sequences of the regions, according to the method reported previously (Nakashima & Noda, 1995). Each genome segment was entirely sequenced in both directions using at least two clones.

Computer analysis of the sequenced data. The obtained sequence data were processed to search for open reading frames (ORFs) with the software GENETYX (Software Development). To search for homologous sequences in databases, the determined nucleotide sequences of each genome segment were divided into 240 nucleotide blocks that overlapped by 60 nucleotides, and their deduced amino acid sequences were divided into 180 residue blocks that overlapped by 60 amino acids. These blocks were then used for a similarity search analysis against the GenBank and PIR databases using the FASTA program (Pearson, 1990) installed in the DDBJ database.

Morphological observations of NLRV. An *N. lugens* from the viruliferous colony was dissected and the salivary glands removed. The dissected glands were homogenized in a few drops of 0·1 M-sodium phosphate buffer pH 7·0. The homogenate was then mounted on a grid and stained with 2% uranyl acetate for observation by an electron microscope.

Expression of a truncated S3 gene in Spodoptera frugiperda (Sf) 9 cells using a baculovirus vector. A truncated ORF of segment S3 was expressed in Sf9 cells using a baculovirus expression system (MaxBac Complete System Plus; Invitrogen) to obtain antiserum against the S3 product and to assign the product to a structural protein of NLRV. The inserted DNA to be expressed was amplified with RT–PCR using synthesized primers based on the nucleotide sequence of S3 from residues 15 to 35 (forward) and from 1045 to 1025 (reverse). *Pst*I and *Hin*dIII recognition sequences were added at the 5′ terminus of the forward and reverse primers, respectively, in order to insert the truncated ORF into a transfer vector, pBlueBacHis. The sequences of the primers were 5′ TTCTGCAGATGCATAAACAAAGCTCAAGC 3′ (forward) and 5′ TTAAGCTTGCGATCACCGT**TTA**AACACTG 3′ (reverse). An artificial termination codon introduced in the reverse primer is shown in bold, and recognition sequences of the restriction enzymes are underlined. The transfer vector was designed to produce a metal-binding fusion protein to facilitate purification of the expressed protein using an immobilized metal affinity chromatography column. Production and purification of the expressed truncated protein were conducted according to the manufacturer's recommendations.

Immunodetection of the truncated S3 product. The separated S3 product was injected into a mouse and antiserum against the protein was obtained. Purified NLRV particles were electrophoresed in a 10% SDS–polyacrylamide gel and proteins in the gel were transferred to a PVDF membrane using a semi-dry blotting apparatus according to the method of Hirano & Watanabe (1990) with some modifications. To improve blotting efficiency of the high molecular mass proteins of NLRV, SDS (0·05% w/v) and 2-mercaptoethanol (0·08% v/v) were added to the blotting buffers and the methanol concentration was decreased to 10%. A part of the blotted membrane was stained with Coomassie brilliant blue R-250 (CBB) to determine the electroblotting efficiency. The rest of the blotted membrane was treated with 500-fold-diluted antiserum directed against the truncated S3 product and proteins that reacted with the antiserum were detected with peroxidase-labelled goat anti-mouse IgG (Bio-Rad).

Results

General features of the NLRV genome

The length, GC content and coding capacity of each genome segment of NLRV are summarized in Table 1. The largest genome segment, S1, had 4391 nucleotides and the smallest, S10, had 1430. The mean GC content in the whole sequence of NLRV was 34·8%. The most GC-rich segment was S5 (39·0%) and the segment with the lowest GC content was S2 (33·1%). The coding capacity of NLRV ORFs was 8916 amino acids. The non-coding regions at the 5′ and 3′ termini of each genome segment totalled 623 and 1239 nucleotides, respectively.

Reoviruses usually have one ORF in a genome segment. A large ORF was found in all segments except S9 which had two non-overlapped ORFs. The first ORF in S9 was located at nucleotides 53 to 925 and the second at 982 to 1602. MRDV S6 (Marzachi *et al.*, 1991) and RBSDV S7 (Azuhata *et al.*, 1993) also have two non-overlapped ORFs. In addition to the ORFs listed in Table 1, six small ORFs were found in the plus strands and 19 in the minus strands using a program for ORF prediction (Fickett, 1982). The longest possible polypeptides among them had 177 and 145 amino acids, respectively.

Flanking nucleotide sequences of the initiation codons

Flanking sequences of the initiation codons of large ORFs are shown in Fig. 1. Two AUG codons were found at the 5′ terminus of S8 as described (Nakashima & Noda, 1994). ORFs 7, 8 and 9-1 had a suitable consensus sequence for translational initiation by eukaryotic ribosomes (Kozak, 1984, 1986); they had A and G at the −3

Table 1. *Features of the NLRV genome segments and their deduced proteins*

Genome segment	Length (bp)	GC content (%)	Non-coding region (5′/3′)	Number of codons	Molecular mass of deduced protein (kDa)	pI*	Molecular mass in SDS–PAGE (kDa)	Locations†	Remark	Accession number
S1	4391	33·5	20/42	1442	165·9	8·66	160	Core	RNA polymerase	D49693
S2	3732	33·1	21/111	1199	136·6	7·87	135	Outer shell	B spike	D49694
S3	3753	34·2	14/67	1223	138·5	5·17	140	Core	Major core capsid	D49695
S4	3560	33·8	75/86	1132	130·0	6·97	–	–	–	D49696
S5	3427	39·0	201/307	972	106·4	5·00	–	–	–	D49697
S6	2970	36·8	149/328	830	95·1	6·45	–	–	–	D49698
S7	1994	34·1	40/64	629	73·5	8·95	75	Core	NTP binding	D49699
S8	1802	35·3	6/107	562	62·4	6·11	65	Outer shell	Major outer capsid	D26127
S9	1640	33·2	52/56/38‡	290	33·0	7·45	–	–	Non-structural	D49700
				206	23·6	5·87	–	–	Non-structural	
S10	1430	35·2	45/89	431	49·4	5·89	–	–	Non-structural	D14691

* Isoelectric point.

† Modified from Noda *et al.* (1991*a*; Fig. 4).

‡ Two large non-overlapped ORFs were observed in S9.

```
                                              (S1)AGUCUUUCAACACAUAACCGAUGAAACCAAUAACUAUUCAG
                                                                       M  K  P  I  T  I  Q

                                            (S2)AGUUCGUGAACCAAGGAUUAAAUGUCGCUCGAAACGAAAGUU
                                                                       M  S  L  E  T  K  V

                                                    (S3)AGUUUGUUGUGAUCAUGCAUAAACAAAGCUCAAGC
                                                                       M  H  K  Q  S  S  S

(S4)AGUUU-----AUAAGAUUACUUAAUGCCUUCGUAUUUGAAUAUAAUAGAAAUACUGCUCAAUAAACAUGACAGAAUAUAAACCAAUU
                                                                       M  T  E  Y  K  P  I
                            M  P  S  Y  L  N  I  I  E  I  L  L  N  K  H  D  R  I  *(18aa)

(S5)AGUAUGCUU---UAUUGAUAA --CCCUUUAAACGAUUUAAGGAUAUCUCUUCACGAUUUUAUUCAAUGGCUUUUAAAUUUGACGCU
                                                                       M  A  F  K  F  D  A
        M  L ---  Y  *(31aa)

(S6)AGUUCGAAA----------------------UAAAUAUUAAAAAUAUUGUUCAUAAAUAAAUAAAAAUGUCAGCAUCAUCAUCUACG
                                                                       M  S  A  S  S  S  T

                          (S7)AGUCAAUACCAGACAGAAAGCAGAUCUUACGUAUUUAACAAUGGAACAUCUCGAAACAAUC
                                                                       M  E  H  L  E  T  I

                                                      (S8)AGUUUGAUGACCAUGGAUCAGAAACUUAAAAUA
                                                                 M  T  M  D  Q  K  L  K  I

              (S9)AGUUUCUAAACCACUAGAACUAGCGAAUUCAAGGAAAAGAUUACAAUUCACUAUGGAGAGCAUCGGAAGAUUC
                                                               (ORF 1) M  E  S  I  G  R  F
           UAAUUUAUAUUUUCAUGUUUCUUAUAUGUAUAUUUAUGUUCCUUGUUUAACGUAUCAUCAUGACUACACGAAUAAAAUUA
           *(ORF 1 stop)                                       (ORF 2) M  T  T  R  I  K  L
                          M  F  L  I  C  I  F  M  F  L  V  *(11aa)

                    (S10)AGUUCGAAUCACAACGAUAUAUGGAAGGUUAAGAUUUGCACAAAAAUGACUUUGCGACGAGCAACG
                                                                       M  T  L  R  R  A  T
                                              M  E  G  *(3aa)
```

Fig. 1. Flanking nucleotide sequences of the initiation codons of NLRV ORFs. The predicted initiation codons and residues located at the −3 and +4 positions are shown in bold. Start and stop codons of minicistronic ORFs observed in these regions are underlined (S4, S5, S8, S9 and S10). The predicted amino acid sequences are shown under the nucleotide sequence.

and +4 positions of their AUG codons, respectively. Other ORFs deviated from the consensus sequence. The largest ORFs started from the first AUG codon in S1, S2, S3, S6, S7, S8 and S9, but started from the second in S4, S5 and S10. The reading frames starting from the first AUG codon of S5 and S10 had termination codons before the second AUG codon, which started the larger ORFs. In S4, a termination codon in-frame with the first AUG codon emerged nine nucleotides from the second AUG codon. Although two AUG codons were found between the termination codon of ORF9-1 and the AUG codon of ORF9-2, these codons do not seem to work because there was a stop codon (UAA) before the start codon of ORF9-2.

Computer analysis

The amino acid sequence coded by NLRV S1 from residues 792 to 1028 showed 18·7% identity to a 227 amino acid sequence in S1 of bluetongue virus serotype 10 (BTV-10; Roy *et al.*, 1990) using the LFASTA program (data not shown). This region included the GDD motif, which is well known as a core motif of RNA-dependent RNA polymerase (Kamer & Argos,

```
(i)                            -         A          B                         C          D
(ii)                           -         1          2                         3          -
(iii)                          I         IV         V                         VI         -
                            (*) * *      * *  *     **   *      (*)           ***        *
NLRV                 (646) IDRRGRII 60 DMSGMDAH  90 SGLFATSGQHTMFL 21 NYVMGDDIF 34 YSKYS (894)
RDV                  (643) AWRPVRPI 73 DCSSWDQT  76 SGRLDTFFMNSVQN 22 FQVAGDDAI 33 PQKTV (890)
Rotavirus SA11       (455) PGRRTRII 57 DVSQWDSS  63 SGEKQTKAANSIAN 21 IRVDGDDNY 30 KVKAL (669)
Reovirus serotype 3  (521) VQRRPRSI 56 DISACDAS  89 SGSTATSTEHTANN 33 YVCQGDDGL 30 GWKYD (772)
BTV-10               (515) PIKATRTI 72 DYSEYDTH 119 SGENSTLIANSMHN 23 EQYVGDDTL 32 PSKTM (804)
```

Fig. 2. Conserved amino acid sequences specific for RNA-dependent RNA polymerases of reoviruses. The conserved motifs presented by (i) Poch *et al.* (1989), (ii) Bruenn (1991) and (iii) Koonin (1992) are shown at the top. Asterisks show highly conserved amino acids and asterisks in parentheses show mostly conserved amino acids. Numbers in parentheses on the left and right indicate the starting and finishing positions of the aligned sequences. Numbers between motifs indicate the number of residues that occur between the motifs. The sequences of RDV (Suzuki *et al.*, 1992), rotavirus SA11 (Mitchell & Both, 1990), reovirus serotype 3 (Weiner & Joklik, 1989) and BTV-10 (Roy *et al.*, 1988) were used for comparison.

```
               370       380       390       400       410       420
NLRV S7   YTMRIGRPIATSFSTFSRERMAENFKDAVRERMPIWVVANKGSGKTVLRKELEE--IGFN
          .....  .  .....    :. . .. . .:   . ...::: ::. .   :.:    .
RBSDV S8  FKLNLLETGLSPLDDVVTIRLKDLLR-VFKEGQDVQIIGNKGVGKSEIGVMLAERYPHLL
          320       330       340        350       360       370
                  430       440          450       460       470
NLRV S7   VIDSDAYGWFVSKVAYIHKEVNQN---TLSTIELSQDHINRLVSEVLEEDRGISYFNVIM
          :.:::.::.:. ..  .  .. .:   ....  :...   . .:. ...... .    ..
RBSDV S8  VVDSDDYGRFLVMLLNLVPSLFKNFDFEINEELLTEEIYFQAMSDFIKAKQTEDVSLQTI
           380       390       400       410       420       430
          480       490       500       510       520       530
NLRV S7   YSLTKQKRDINKVLTFPYVTHRWNDTLTTFGKVFTEIADSPAIGYPRFIEGC--MEY--F
          .. ...   . ....   ...  . .:..:..:: .:..:  :::..:. ..  . :  :
RBSDV S8  FEYVMEGIILPNTVESGETNE--EVILDIFNRVFHSIQGSKFIGYRKFMIEYSRLMYTNF
           440       450         460       470       480       490
              540         550        560       570        580
NLRV S7   RVRETLHKPIIF--FAHTESELSRISI-RHFSCVIYNSIDSRS-ILSKRADPVVELMLHA
          . ..: :    .  .. .  .:. :.. . ..... : . ... . ..... ... .::
RBSDV S8  NKSQTCHFVHSYCELSFVPHSLAYITLYSSYNSAVLNVVPRNGQVECSSSKMMANTLLHH
             500       510       520       530       540       550
        590       600       610
NLRV S7   YYYMNQVAIDKIPFSLFRQWLGMSN
          .: .   ... .:  ::. ..:...
RBSDV S8  FYERYTSNMNPTPVFLFSYFFGLTK
             560       570
```

Fig. 3. Comparison of the C termini of the deduced amino acid sequences coded by NLRV S7 and RBSDV S8 (Azuhata *et al.*, 1993). The sequences were aligned by the LFASTA program (Pearson, 1990). Purine NTP-binding motifs (Golbalenya & Koonin, 1989) are shown in bold type.

1984), indicating that NLRV S1 codes for this enzyme (Fig. 2). In the aligned residues, an asparagine (N) residue in motif V (Koonin, 1992) was substituted with a leucine (L) residue. Three nucleotide substitutions were required to achieve this.

The C terminus of the deduced amino acid sequence of NLRV S7 showed 16·2% identity in a 265 amino acid overlap with the deduced sequence of RBSDV S8 (Azuhata *et al.*, 1993) (Fig. 3). The N-terminal regions of the two segments were not aligned by the LFASTA program although their hydrophobicity plots were similar (data not shown). By a motif search analysis using the program PROSITE (release 11.0), the purine NTP-binding motif A/GXXXXGKS/T (Golbalenya & Koonin, 1989) was found at residues 401 to 408 in the predicted amino acid sequence of NLRV S7. RBSDV S8 also possesses the same predicted motif at residues 355 to 362.

With regard to the nucleotide sequences and the deduced amino acid sequences coded by S2, S3, S4, S5, S6 and S9 of NLRV, no other remarkable similarities to the sequences deposited in the databases were found.

A and B spikes of NLRV

The members of the genus *Fijivirus* typically have two kinds of spikes on the virus capsid, A spikes on the surface of the outer capsid and B spikes on the inner

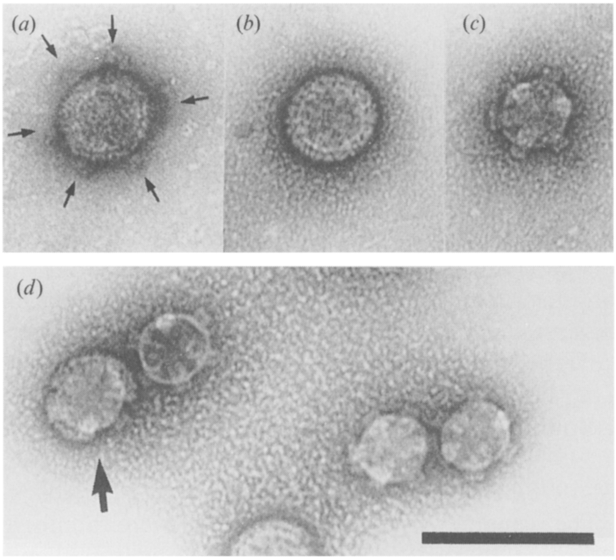

Fig. 4. Structure of NLRV particles. Particles were stained with 2% uranyl acetate. (*a*) An NLRV particle with attached A spikes (arrows) in the homogenate of the salivary glands of viruliferous *N. lugens*. (*b*) A purified NLRV particle obtained by the method of Noda *et al.* (1991*a*). (*c*) An inner core particle with attached B spikes. (*d*) NLRV particles with an area (arrow) in which the outer capsid proteins are partially stripped away. Bar represents 125 nm.

capsid (Milne & Lovisolo, 1977). In our earlier study of NLRV (Noda *et al.*, 1991*a*), we observed the double-shelled structure of the particle, but the A and B spikes were not detected.

Fig. 4 shows the particle morphology of NLRV. A spikes were observed when NLRV particles were prepared directly from the salivary glands of a viruliferous planthopper (Fig. 4*a*), but were not seen in the purified NLRV particles (Fig. 4*b*). A spikes seem to come off easily from the outer capsid of NLRV during purification from planthoppers. Repeated CsCl ultracentrifugations stripped particles of their outer capsid proteins (Fig. 4*d*). B spikes on the inner core were recognized in these particles (Fig. 4*c*). In an earlier study of NLRV, we treated purified NLRV particles with 1·9 M-$MgCl_2$ solution to remove the outer capsid (Noda *et al.*, 1991*a*), but B spikes also seemed to be removed because of the high stringency of the solution.

Blotting analysis of the structural proteins of NLRV and immunodetection of the S3 product

The result of immunodetection of the S3 product is shown in Fig. 5. NLRV has seven structural proteins, three major (140, 135 and 65 kDa), three intermediate (160, 110 and 75 kDa) and one minor (120 kDa) (Noda *et al.*, 1991*a*). In this experiment, to separate two close bands corresponding to the 140 and 135 kDa proteins, the amount of sample loaded on to the gel was reduced so that the 110 kDa protein was scarce and the 120 kDa protein was not detected at all (Fig. 5, lane 1). Blotting efficiency differed among the structural proteins of NLRV. The 140 and 65 kDa proteins were transferred to the membrane easily (Fig. 5, lane 3), but the 160, 75 and 135 kDa proteins remained in the gel after the electroblotting procedure (Fig. 5, lane 2). The truncated ORF coded by S3 was expressed in Sf9 cells and isolated with metal-binding affinity chromatography. A protein having the expected molecular mass (42·4 kDa) was then observed on an SDS–polyacrylamide gel (data not shown). Mouse antiserum raised against the 42·4 kDa protein reacted with the 140 kDa protein of NLRV (Fig. 5, lane 4) and with the 42·4 kDa protein in the lysate of Sf9 cells that had been infected with the recombinant baculovirus (Fig. 5, lane 5). These results indicate that S3 codes for the 140 kDa protein.

Assignment of the structural proteins of NLRV

Morphological observations of the NLRV particles suggested that at least four major structural proteins are required to assemble them, namely the A and B spike proteins and the outer and inner capsid proteins (Fig. 4). As shown in Fig. 4(*b*), A spikes came away from purified NLRV particles and therefore should not appear in SDS–PAGE profiles as a major structural protein. The purified NLRV particle has three major structural proteins of 140, 135 and 65 kDa (Noda *et al.*, 1991*a*). The 65 kDa protein was the major outer capsid protein and was coded by S8 (Nakashima & Noda, 1994); the 140 kDa protein was a major inner core structural protein and was coded by S3 (Fig. 5, lane 4). The 135 kDa protein is apparently the B spike protein. This conclusion is in accordance with previous observations. When the inner core and the outer shell of purified NLRV particles were separated by treatment with a high concentration of $MgCl_2$, the inner core particles did not retain B spikes (Noda *et al.*, 1991*a*; Fig. 2*b*). The 135 kDa protein was partly recovered from the outer shell fraction but not from the inner core (Noda *et al.*, 1991*a*; Fig. 4*a*, lane 3).

Comparison of the calculated molecular masses of the predicted proteins coded by the genome segments of NLRV indicated that S1 codes for the largest protein (160 kDa) because no other genome segment had such a large coding capacity (Table 1). S7 seems to code for the 75 kDa protein because it is the only segment that had a coding capacity close to 75 kDa. The 160 and 75 kDa proteins remained in the polyacrylamide gel (Fig. 5, lane 2) after the electroblotting procedure. Basic proteins are rarely electroblotted from a polyacrylamide gel to a membrane because blotting buffers are basic (Van Seuningen & Davril, 1990). Indeed, the isoelectric points

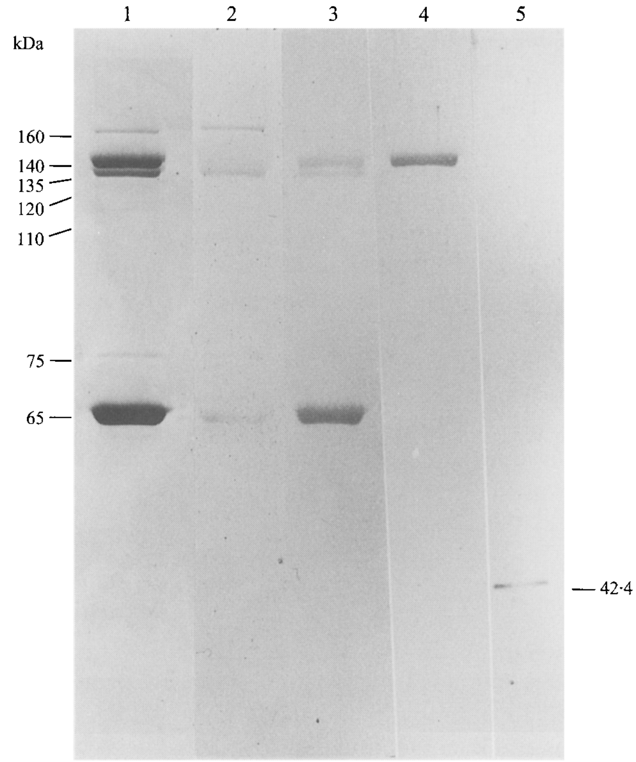

Fig. 5. Immunodetection of the structural protein coded by genome segment S3 of NLRV. Structural proteins of NLRV were separated by 10% SDS–PAGE and electroblotted onto a PVDF membrane for immunodetection of S3 product using antiserum directed against the S3 truncated ORF product. Lane 1, purified NLRV particles; lane 2, proteins remaining in the gel after electroblotting, stained with CBB; lane 3, electroblotted proteins on PVDF membrane stained with CBB; lane 4, immunodetection of S3 product (140 kDa) on the membrane; lane 5, immunodetection of the truncated S3 product (42·4 kDa) in Sf9 insect cells that had been infected with a recombinant baculovirus.

of the deduced proteins coded by S1 and S7 (8·66 and 8·95, respectively) were highly basic (Table 1).

The second largest protein (140 kDa) was coded by S3 (Fig. 5, lane 4). Of the remaining genome segments, only S2 had the capacity to code for the 135 kDa protein, the B spike protein. The 135 kDa protein did not electroblot well onto the membrane compared to the 140 kDa protein (Fig. 5, lanes 2 and 3). Again, the isoelectric point of the deduced protein coded by S2 was basic. These observations suggest that S2 codes for the 135 kDa protein.

The predicted proteins coded by NLRV S9 and S10 seemed to be non-structural proteins because S9 and S10 do not have coding capacities that exceed the smallest structural protein, 65 kDa (Table 1). We were unable to assign two structural proteins, the 120 and 110 kDa proteins, to coding segments because the molecular masses deduced from predicted amino acid sequences and those estimated from SDS–PAGE did not correspond well. However, one of the segments in S4, S5 and S6 must code for the A spike protein and this could be the 120 kDa protein or a protein that did not appear in SDS–PAGE of purified particles.

Discussion

The total number of nucleotides in the NLRV genome was 28699, which is the largest number reported so far for a member of the family *Reoviridae*. For comparison, the sizes in nucleotides of the genomes of other reoviruses are: reovirus type 3 (genus *Orthoreovirus*), 23549 (Weiner *et al.*, 1989); rotavirus SA11 (*Rotavirus*), 18555 (Mitchell & Both, 1990); BTV-10 (*Orbivirus*), 19218 (Fukusho *et al.*, 1989); and RDV (*Phytoreovirus*), 25749 (Uyeda *et al.*, 1995).

The genome segments of NLRV were numbered according to their mobilities in 10% polyacrylamide gels, with higher numbers corresponding to higher mobilities. Segments S4 and S5 do not separate under these conditions (Noda *et al.*, 1994). The relatively low mobility of S5 in polyacrylamide gels was assumed to be due to its high GC content. Segment S2 was shorter than S3 by 21 bp. This reversed mobility, however, cannot be due to GC content because their GC contents were almost the same. There may be some additional factors affecting the relative mobilities of dsRNAs in polyacrylamide gels, as indicated by Fukumoto (1992).

MRDV S6 has nucleotide sequence variation at a level of 1·18%, and one reason for this is that the source viruses were collected from field-harvested plants (Marzachi *et al.*, 1991). In the case of NLRV, the average percentage of heterogeneous nucleotides in all segments except S5 was 0·09%. This value seems to be due to false incorporation of nucleotides by the reverse transcriptase used for cDNA synthesis. NLRV S5 differed from the other segments in that it had a nucleotide heterogeneity of 2·5%. Although this high rate was probably partly due to false incorporation, the reason why S5 alone had so many heterogeneous residues is unknown because cDNA synthesis of all segments was conducted concurrently from the total genome of NLRV (Noda *et al.*, 1994). The viruliferous *N. lugens* were collected from a field in 1987 and were then reared in our laboratory. The virus for cDNA cloning was purified in 1990. It is unlikely that the heterogeneity in S5 could have arisen in this 3 year period, since the rate of nucleic acid substitution, estimated at $2{\cdot}2 \times 10^{-3}$ substitutions/site/year in BTV (Kowalik & Li, 1991) is very low. Thus, NLRV S5 appears to have already been heterogeneous when the viruliferous insects were collected.

Consensus motifs of RNA-dependent RNA polymerases have been shown in various kinds of viruses by Poch *et al.* (1989), Bruenn (1991) and Koonin (1992).

Motifs A, B and C defined by Poch *et al.* (1989), motifs 1, 2 and 3 from Bruenn (1991) and motifs IV, V and VI from Koonin (1992) are strongly conserved in reoviruses (Fig. 2). Motif I defined by Koonin (1992) and motif D from Poch *et al.* (1989) also seem to be conserved, although these two motifs were not noted by Suzuki *et al.* (1992). Motifs 4–8 in the work of Bruenn (1991) and motif VII defined by Koonin (1992) were difficult to identify in reoviruses. Koonin (1992) suggested that the enzymes of reoviruses are monophyletic. However, alignment of their deduced amino acid sequences was difficult because, except for the conserved motif regions shown in Fig. 2, they are not well conserved. To align the sequences of the enzymes and to elucidate the phylogenetic relations of the members of the family *Reoviridae*, the sequence data of other groups (e.g. oryzaviruses, coltiviruses, cypoviruses and aquareoviruses) is required.

NLRV S3 coded for the major structural core protein. The equivalent proteins of RDV and RGDV, which belong to the genus *Phytoreovirus*, are both encoded by segment S3 (Suzuki *et al.*, 1990; Takahashi *et al.*, 1994). However, the deduced amino acid sequence coded by NLRV S3 was not homologous to those of RDV S3 and RGDV S3. Hopper-borne reoviruses are classified into three genera, *Phytoreovirus*, *Fijivirus* and *Oryzavirus*. Phytoreoviruses and fijiviruses have a double-shelled structure (Shikata, 1989; Murphy *et al.*, 1995). Members of the genus *Fijivirus* and NLRV both have A and B spikes, whereas phytoreoviruses have no obvious spikes. This morphological difference between the members of the genera *Phytoreovirus* and *Fijivirus* (or NLRV) seems to account for the low similarity of the deduced amino acid sequence of their structural proteins.

NLRV has many properties similar to those of RBSDV, including morphology, terminal nucleotide sequences of genome segments (Noda *et al.*, 1994), deduced amino acid sequences (Nakashima & Noda, 1994), manner of transmission through rice plants and multiplication ability in the small brown planthopper, *Laodelphax striatellus* (Nakashima & Noda, 1995). The most distinct difference between NLRV and RBSDV is that RBSDV has the ability to multiply in rice plants while NLRV does not. Among the reported sequences of S7, S8 and S10 in RBSDV, only the deduced amino acid sequence of ORF2 of RBSDV S7 showed no similarity with any sequences of NLRV. The product of RBSDV S7 ORF2 may have a different function from the gene products of NLRV and may be involved in multiplication in rice plants.

As far as viruses are classified based on their comprehensive characteristics, we consider NLRV should be classified in the genus *Fijivirus*. However, if NLRV is excluded from the genus because of its inability to multiply in plants, a new genus composed of non-phytopathogenic hopper-borne reoviruses will be required in the family *Reoviridae*.

References

Azuhata, F., Uyeda, I., Kimura, I. & Shikata, E. (1993). Close similarity between genome structures of rice black-streaked dwarf and maize rough dwarf viruses. *Journal of General Virology* **74**, 1227–1232.

Boccardo, G., Hatta, T., Francki, R. I. B. & Grivell, C. J. (1980). Purification and some properties of reovirus-like particles from leafhoppers and their possible involvement in wallaby ear disease of maize. *Virology* **100**, 300–313.

Bruenn, J. A. (1991). Relationships among the positive strand and double-strand RNA viruses as viewed through their RNA-dependent RNA polymerases. *Nucleic Acids Research* **19**, 217–226.

Falk, B. W., Kim, K. S. & Tsai, J. H. (1988). Electron microscopic and physicochemical analysis of a reo-like virus of the planthopper *Peregrinus maidis*. *Intervirology* **29**, 195–206.

Fickett, J. W. (1982). Recognition of protein coding regions in DNA sequences. *Nucleic Acids Research* **10**, 5303–5318.

Fukumoto, F. (1992). Difference in electrophoretic pattern of the rice dwarf virus genome segments in agarose and polyacrylamide gel electrophoresis. *Annals of the Phytopathological Society of Japan* **58**, 83–86.

Fukusho, A., Yu, Y., Yamaguchi, S. & Roy, P. (1989). Completion of the sequence of bluetongue virus serotype 10 by the characterization of a structural protein, VP6, and a non-structural protein, NS2. *Journal of General Virology* **70**, 1677–1689.

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

Hibino, H. (1989). Insect-borne viruses of rice. *Advances in Disease Vector Research* **6**, 209–241.

Hirano, H. & Watanabe, T. (1990). Microsequencing of proteins electrotransferred onto immobilizing matrices from polyacrylamide gel electrophoresis: application to an insoluble protein. *Electrophoresis* **11**, 573–580.

Kamer, G. & Argos, P. (1984). Primary structural comparison of RNA-dependent polymerases from plant, animal, and bacterial viruses. *Nucleic Acids Research* **12**, 7269–7282.

Koonin, E. V. (1992). Evolution of double-stranded RNA viruses: a case for polyphyletic origin from different groups of positive-stranded RNA viruses. *Seminars in Virology* **3**, 327–339.

Kowalik, T. F. & Li, J. K. K. (1991). Bluetongue virus evolution: sequence analysis of the genomic S1 segments and major core protein VP7. *Virology* **181**, 749–755.

Kozak, M. (1984). Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin *in vivo*. *Nature* **308**, 241–246.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eucaryotic ribosomes. *Cell* **44**, 283–292.

Marzachi, C., Boccardo, G. & Nuss, D. L. (1991). Cloning of the maize rough dwarf virus genome: molecular confirmation of the plant-reovirus classification scheme and identification of two large nonoverlapping coding domains within a single genomic segment. *Virology* **180**, 518–526.

Milne, R. G. & Lovisolo, O. (1977). Maize rough dwarf and related viruses. *Advances in Virus Research* **21**, 267–341.

Mitchell, D. B. & Both, G. W. (1990). Completion of the genomic sequence of the simian rotavirus SA11: nucleotide sequence of segments 1, 2, and 3. *Virology* **177**, 324–331.

Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A. & Summers, M. D. (1995). *Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses. Archives of Virology* Supplementum 10, 208–239.

Nakashima, N. & Noda, H. (1994). Nucleotide sequence of *Nilaparvata lugens* reovirus genome segment S8 coding for the major outer capsid protein. *Journal of General Virology* **75**, 2803–2806.

Nakashima, N. & Noda, H. (1995). Nonpathogenic *Nilaparvata lugens* reovirus is transmitted to the brown planthopper through rice plant. *Virology* **207**, 303–307.

Nault, L. R. & Ammar, E. D. (1989). Leafhopper and planthopper transmission of plant viruses. *Annual Review of Entomology* **34**, 503–529.

Noda, H., Ishikawa, K., Hibino, H. & Omura, T. (1991*a*). A reovirus in the planthopper, *Nilaparvata lugens*. *Journal of General Virology* **72**, 2425–2430.

Noda, H., Ishikawa, K., Hibino, H., Kato, H. & Omura, T. (1991*b*). Nucleotide sequences of genome segments S8, encoding a capsid protein, and S10, encoding a 36K protein, of rice gall dwarf virus. *Journal of General Virology* **72**, 2837–2842.

Noda, H., Nakashima, N. & Omura, T. (1994). Cloning of the *Nilaparvata lugens* reovirus genome: conserved terminal nucleotide sequences and nucleotide sequence of genome segment S10. *Journal of General Virology* **75**, 221–225.

Nuss, D. L. & Dall, D. J. (1990). Structural and functional properties of plant reovirus genomes. *Advances in Virus Research* **38**, 249–306.

Ofori, F. A. & Francki, R. I. B. (1985). Transmission of leafhopper A virus, vertically through eggs and horizontally through maize in which it does not multiply. *Virology* **144**, 152–157.

Pearson, W. R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods in Enzymology* **183**, 63–98.

Poch, O., Sauvaget, I., Delarue, M. & Tordo, N. (1989). Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO Journal* **8**, 3867–3874.

Roy, P., Fukusho, A., Ritter, G. D. & Lyon, D. (1988). Evidence for genetic relationship between RNA and DNA viruses from the sequence homology of a putative polymerase gene of bluetongue virus with that of vaccinia virus: conservation of RNA polymerase genes from diverse species. *Nucleic Acids Research* **16**, 11759–11767.

Roy, P. Marshall, J. J. A. & French, T. J. (1990). Structure of the bluetongue virus genome and its encoded proteins. *Current Topics in Microbiology and Immunology* **162**, 43–87.

Shikata, E. (1989). Plant reoviruses. In *Plant Viruses*, vol. 1, pp. 207–234. Edited by C. L. Mandahar. Boca Raton: CRC Press.

Sōgawa, K. (1982). The rice brown planthopper: feeding physiology and host plant interactions. *Annual Review of Entomology* **27**, 49–73.

Suzuki, N., Watanabe, Y., Kusano, T. & Kitagawa, Y. (1990). Sequence analysis of the rice dwarf phytoreovirus segment S3 transcript encoding for the major structural core protein of 114 kDa. *Virology* **179**, 455–459.

Suzuki, N., Tanimura, M., Watanabe, Y., Kusano, T., Kitagawa, Y., Suda, N., Kudo, H., Uyeda, I. & Shikata, E. (1992). Molecular analysis of rice dwarf phytoreovirus segment S1: interviral homology of the putative RNA-dependent RNA polymerase between plant- and animal-infecting reoviruses. *Virology* **190**, 240–247.

Takahashi, Y., Tomiyama, M., Hibino, H. & Omura, T. (1994). Conserved primary structures in core capsid proteins and reassembly of core particles and outer capsids between rice gall dwarf and rice dwarf phytoreoviruses. *Journal of General Virology* **75**, 269–275.

Uyeda, I., Azuhata, F. & Shikata, E. (1990). Nucleotide sequence of rice black-streaked dwarf virus genome segment 10. *Proceedings of the Japan Academy Series B* **66**, 37–40.

Uyeda, I., Kimura, I. & Shikata, E. (1995). Characterization of genome structure and establishment of vector cell lines for plant reoviruses. *Advances in Virus Research* **45**, 249–279.

Van Seuningen, I. & Davril, M. (1990). Electrotransfer of basic proteins from nondenaturing polyacrylamide acid gels to nitrocellulose: detection of enzymatic and inhibitory activities and retention of protein antigenicity. *Analytical Biochemistry* **186**, 306–311.

Weiner, J. R. & Joklik, W. K. (1989). The sequences of serotype 1, 2, and 3 L1 genome segments and analysis of the mode of divergence of the reovirus serotypes. *Virology* **169**, 194–203.

Weiner, J. R., Bartlett, J. A. & Joklik, W. K. (1989). The sequences of reovirus serotype 3 genome segments M1 and M3 encoding the minor protein μ2 and the major nonstructural protein μNS, respectively. *Virology* **169**, 293–304.

(*Received 18 May 1995; Accepted 4 September 1995*)