Available online at www.sciencedirect.com
SCIENCE DIRECT®
Virus Research 92 (2003) 113–121
Virus Research
www.elsevier.com/locate/virusres

Sequence and phylogenetic analysis of genome segments S1, S2, S3 and S6 of *Mal de Río Cuarto virus*, a newly accepted *Fijivirus* species

Ana J. Distéfano [a,1], Luis R. Conci [b,2], Marianne Muñoz Hidalgo [a,1], Fabiana A. Guzmán [b,2], Horacio E. Hopp [a,c,1], Mariana del Vas [a,1,*]

[a] *Instituto de Biotecnología, CICVyA, CNIA, INTA, Castelar, Buenos Aires 25 B1712WAA, Argentina*
[b] *Instituto de Fitopatología y Fisiología Vegetal, CICVyA, CNIA, INTA, Cordoba, Argentina*
[c] *Unidad Integrada CNIA-INTA Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina*

Received 3 September 2002; received in revised form 12 November 2002; accepted 25 November 2002

Abstract

Mal de Río Cuarto virus (MRCV) is a newly described species of the genus *Fijivirus*, family *Reoviridae*. The nucleotide sequence of four MRCV genome segments was determined. MRCV S1, S2, S3 and S6 were predicted to encode proteins of 168.4, 134.4, 141.7 and 90 kDa, respectively. MRCV S1 encodes a basic protein that contains conserved RNA-dependent RNA polymerase motifs, and is homologous to *Rice black streaked dwarf virus* (RBSDV), *Fiji disease virus* (FDV) and *Nilaparvata lugens reovirus* (NLRV) polymerases as well as to corresponding proteins of members of other genera of the *Reoviridae*. MRCV S2 codes for a protein with intermediate homology to the ones coded by RBSDV S4 and FDV S3 'B' spike, which is presumably the B-spike protein. MRCV S3 most probably encodes the major core protein and is highly homologous to corresponding proteins of RBSDV S2 and FDV S3. MRCV S6-encoded protein has low homology to the proteins of unknown function coded by RBSDV S6 and FDV S6. The identity levels between all analyzed MRCV coded proteins and their RBSDV counterparts varied between 84.5 and 44.8%. The analysis of the reported sequences allowed a phylogenetic comparison of MRCV with other reovirus and supported its taxonomic status within the genus.

Keywords: MRCV; *Fijivirus*; *Reoviridae*; RNA-dependent RNA polymerase; Core protein

1. Introduction

Mal de Río Cuarto virus (MRCV) causes the most important maize disease in Argentina, a very important corn exporter country. It is a member of the *Fijivirus* genus, *Reoviridae* family, that was originally described as a variant of *Maize rough dwarf virus* (MRDV) (Nome et al., 1981; Milne et al., 1983) and recently accepted as a new virus species (Mertens et al., 2000; Distéfano et al., 2002). MRCV is able to replicate in phloem cells of plants such as maize, wheat and other grasses (Nome et al., 1981; Rodriguez Pardina et al., 1998; Arneodo et al., 2002) as well as in cells of *Delphacodes kuschelli*, the insect that transmits the disease (Remes Lenicov et al., 1985). Its genome consists in 10 linear segments of double-stranded (ds) RNA called S1–S10 ranging in size between 4.5 and 1.8 kbp (Conci, L.R., unpublished results). Fijiviruses form icosahedral double-shelled particles with short surface spikes (A-spikes) on each of the 12 vertices of the icosahedron. Virus particles easily lose the outer layer to give spiked core particles (with 12 B-spikes) that contain one copy of each genome segment, and RNA-dependent RNA polymerase (RdRp) (Mertens et al., 2000).

According to their relatedness, five groups of fijivirus species are recognized. Group 1 has a unique species:

* Corresponding author. Fax: +54-11-4481-2975.
E-mail addresses: adistefa@cicv.inta.gov.ar (A.J. Distéfano), lconci@correo.inta.gov.ar (L.R. Conci), mmunoz@cicv.inta.gov.ar (M. Muñoz Hidalgo), fguzman@correo.inta.gov.ar (F.A. Guzmán), ehopp@correo.inta.gov.ar (H.E. Hopp), mdelvas@cnia.inta.gov.ar (M. del Vas).
[1] C.C. 25 B1712WAA. Castelar, Buenos Aires, Argentina. Fax: +54-11-4481-2975.
[2] Camino 60 cuadras Km 5 ½, 5119 Cordoba, Argentina. Fax: +54-35-1497-4330.

0168-1702/03/$ - see front matter
doi:10.1016/S0168-1702(02)00325-8

Fiji disease virus (FDV), which is the type species of the genus. Group 2 includes *Rice black streaked dwarf virus* (RBSDV), MRDV (which is closely related to RBSDV), *Pangola stunt virus* (PaSV) and MRCV. Groups 3, 4 and 5 have one member each: *Oat sterile dwarf virus* (OSDV), *Garlic dwarf virus* (GDV) and *Nilaparvata lugens reovirus* (NLRV), respectively (Mertens et al., 2000). The nucleotide sequences of some fijivirus genome segments have been obtained: from group 1, FDV genome segments S3, S4 and S9 (S3: Acc.-No.AF049704, S4: Acc.-No.AF049705 GenBank comments, Soo et al., 1998; McMahon et al., 1999) and recently S1, S3 'B' spike, S5 and S6 (Acc.-No.AY029520, Acc.-No.AF359556, Acc.-No.AY029521 and Acc.-No.AF356083 GenBank comments); from group 2, S6–S8 and S10 of MRDV (Marzachi et al., 1991, 1996), S7–S10 counterparts of a Japanese strain of RBSDV (Uyeda et al., 1990; Azuhata et al., 1993; Isogai et al., 1998a) and recently the complete sequence of a Chinese strain of RBSDV (Zhang et al., 2001a,b) and MRCV genome segments S4 and S8 (Distéfano et al., 2002) were obtained; from group 4, S7–S10 of OSDV (Isogai et al., 1998b); from group 5, the whole NLRV genome was reported (Noda et al., 1994; Nakashima et al., 1996).

Phylogenetic relationships between fijiviruses remain unclear, mainly because of the lack of sequence information of whole genomes of some species (GDV, PaSV) and of several segments of other species (FDV, MRCV, MRDV, OSDV). This paper reports the sequence analysis of the three largest MRCV genome segments (S1, S2, S3) and that of segment S6. Comparative sequence studies with other members of the family *Reoviridae* were performed. The subsequent phylogenetic analysis supported the new MRCV taxonomic status as an independent species of the *Fijivirus* genus.

2. Materials and methods

2.1. Virus source, MRCV library construction and DNA sequencing

Lyophilized roots from naturally infected field-grown maize plants of the Argentinean province of Córdoba were used as a starting material. dsRNA purification was done by differential adsorption to CF11 cellulose as described by Dodds (1993). A random-primed cDNA library was prepared using a commercial cDNA synthesis kit (Librarian II, Invitrogen, USA). Briefly, 3 μg of total dsRNA was used as template, heat denatured and reverse transcribed using AMV RT. After second strand cDNA synthesis, non-palindromic BstXI adapters were ligated to cDNA ends and the product was cloned into cut pcDNA II vector (Invitrogen). Probes were obtained by first strand RT synthesis in the presence of [α-^{32}P] dCTP, using low melting point agarose (BioRad, USA)-purified MRCV S1–S3 and S6 as templates. Once several clones of the random library were identified as belonging to MRCV, S1–S3 and S6 inserts were sequenced by the dideoxynucleotide chain-termination method (Sanger et al., 1977) using fmol DNA Sequencing System (Promega Biotech, USA) following provider's specifications and protocols. Several clones were produced and sequenced across the central region of S1 and S6 and completely spanning segments S2 and S3. Each base pair position was sequenced from at least two independent clones. After sequence assembly, it was evident that approximately 1.7 kbp from the 5′ end and 0.8 kb from the 3′ end from genome segment S1 were absent from the random MRCV library. Thus, a second, this time specific, cDNA library of the missing regions of S1 was constructed by using two specific oligonucleotides as primers (5′-ATAGCTCAGACATTTCGTTCAG-3′ nt 1754–1733 and 5′-CTTTGATTACGCAGTTGAGC-3′ nt 3769–3788, complementary and identical to the S1 coding strand, respectively). Similarly, a specific MRCV S6 library was constructed using two specific oligonucleotides as primers (5′-TCACACTTCTCCGATTCCAA-3′, nt 1758–1739, complementary to S6 coding strand, and 5′-CAAGAGCAATTGTCGATTGAAG-3′, nt 1856–1877, identical to the coding strand). For the specific libraries construction, we used a cDNA synthesis kit (Universal Ribo Clone, Promega Biotech, USA) and cloned into dephosphorylated EcoRV-digested pBluescript II (KS$^+$) (Stratagene, USA). As before, each base pair position was sequenced from at least two independent clones.

2.2. Amplification and cloning of 5′ and 3′ terminal sequences

To amplify and analyze MRCV ends, 5′ and 3′ rapid amplification of cDNA ends (RACE) was performed (Frohman et al., 1988) using 1 μg of total denatured MRCV dsRNA genome as a template. For 5′ RACE, specific MRCV S1, S2, S3 and S6 internal antisense primers were: S1: 5′-CAAGATTTTTCAATACCACC-3′, nt 479–459; S2: 5′-GGGAGTCATCAACTAAGGAA-3′ nt 541–522; S3: 5′-TCTAAGTAATGCTGGGTGAT-3′ nt 610–591; S6: 5′-ACCATTTGTTCTCATCTTTA-3′ nt 464–445. Nested specific primers were: S1: 5′-TAATCTCTTCGGCATTCGCA-3′, nt 340–321; S2: 5′-TGCCTGGGTCAAGTATTCAA-3′ nt 159–139; S3: 5′-TCTTGGCTTGTATCGTGCTT-3′ nt 281–262; S6: 5′-ACAGTTTGAGTTTGAATAGC-3′ nt 402–383. For 3′ RACE, specific MRCV primers were: S1: 5′-TTTAGTTGATGCTGCTGGTG-3′, nt 4023–4042; S2: 5′-TCGTAACTTTGAACCTGATG-3′ nt 3465–3484; S3: 5′-TTGAGAACCGTGATTTGAAC-3′ nt 3357–3376; S6: 5′-CAAGAGCAATTGTCGATTGAAG-3′ nt

1856–1877. Nested specific primers were: S1: 5′-TTCCCATCTAAATCAACTCT-3′ nt 4380–4399; S3: 5′-ATTTCGTAGTGATGGGCGGT-3′ nt 3487–3505; S6: 5′-TTGAAGCAATTCGTACTTTG-3′ nt 2106–2125. The amplified products were purified (prep-A-gene, BioRad, USA), cloned into pGEMt easy vector (Life Technologies, USA) and sequenced as detailed before. At least four independent clones of each end were sequenced and analyzed.

2.3. Computer analysis of the sequenced data

Sequence data were assembled and analyzed by EDIT SEQ and SEQMAN programs (Lasergene, DNAstar). The nucleotide sequences reported in this paper were submitted to GenBank under accession numbers AF499925 (MRCV S1), AF499926 (MRCV S2), AF499928 (MRCV S3) and AF499927 (MRCV S6). Protein sequence analysis was performed by PROTEAN program (Lasergene). Conserved motifs were identified using PROSITE (Hofmann et al., 1999).

Databases were searched using FASTA (Pearson, 1990) and BLAST (Altschul et al., 1990) programs. Secondary RNA structures where predicted with the MFOLD program, version 3.1 (Zuker et al., 1999; Mathews et al., 1999) available at http://bioinfo.math.rpi.edu/~fold/rna/form1.cgi. Multiple sequence analysis was carried out with CLUSTALW version 1.8 (Thompson et al., 1994). Phylogenetic trees were constructed with the unweighted pair-group method with arithmetical average algorithm (UPGMA) (Zukerlandl and Pauling, 1965) and neighbor-joining method (Saitou and Nei, 1987) and bootstrapped ($n = 100$), using PHYLIP software package version 3.5c (Felsenstein, 1993).

3. Results

The complete sequences of MRCV S1, S2, S3 and S6 were obtained and analyzed. Computer analysis revealed that each genome segment contains one major open reading frame (ORF). Flanking sequences of the initiation codons of all four ORFs were seen to deviate from the consensus sequence for translational initiation by eukaryotic ribosomes (Kozak, 1984, 1986). MRCV S1 and S2 ORFs had a G at the +4 position of their AUG codons and lacked an A at the −3 position. Conversely, MRCV S3 and S6 ORFs had an A at the −3 position and lacked a G at the +4 position. Properties of MRCV S1, S2, S3 and S6 and of the previously reported S4 and S8 (Distéfano et al., 2002), as well as different features of their deduced encoded proteins are detailed in Table 1. As occurred for NLRV S2 and S3 (Nakashima et al., 1996), MRCV S2 (3617 bp) and S3 (3826 bp) did not migrate according to their expected Mr in SDS-PAGE. A similar behavior was previously described for a number of segments of various members of *Reoviridae* family (Mertens et al., 2000) and could reflect sequence-based differences in bending or other conformational features of the dsRNA helixes that affect the migration in gels.

As described for MRCV S4 and S8 (Distéfano et al., 2002) and genome segments of other members of *Fijivirus* genus such as MRDV and RBSDV (Marzachi et al., 1991, 1996; Uyeda et al., 1990; Azuhata et al., 1993; Isogai et al., 1998a; Zhang et al., 2001a,b), MRCV segments S1, S2, S3 and S6 have the conserved 8-nucleotide sequence 5′AAGUUUUU3′ at the 5′ ends, and the conserved sequence 5′CAGCUnnnGUC3′ at the 3′ ends. Imperfect inverted repeats were identified at the 5′ and 3′ terminal regions of all MRCV studied segments. The MFOLD program predicted that these repeats are able to form stable secondary structures (displaying a panhandle, a stem loop and a non-base-paired 3′-tail). The stem loop shape was particularly distinctive of each individual segment including the ones reported for MRCV S4 and S8 (Distéfano et al., 2002) (data not shown). These secondary structures presumably act as replication and packaging signals (Patton and Spencer, 2000).

The largest genome segment (MRCV S1) potentially encodes a 168.4 kDa basic protein (pI of 8.22) and contains RdRp distinctive sequence motifs. The GDD motif, well known as the core motif of RdRps (Kamer and Argos, 1984) was found at amino acid (aa) position 866–868. Motifs A, B, C (Poch et al., 1989), motifs 1, 2, and 3 (Bruenn, 1991), and motifs I, IV, V and VI (Koonin, 1992) previously described for RdRps were also identified. Fig. 1 shows the region between positions 657 and 905 of the MRCV S1 coded protein and the equivalent region of the RdRps of other reoviruses that contains all the described conserved motifs. MRCV S1 coded protein had 77.9% identity with RBSDV S1, 64.2% with FDV S1 and 32.3% with NLRV S1 (all of them *Fijivirus*)-coded RNA polymerases. In addition, we detected identity with RNA polymerases coded by viruses belonging to other genera of the *Reoviridae* family: 19% (in 1200 aa) with *Rice ragged stunt virus* S4 (RRSV, *Oryzavirus*), 21.1% (in 1179 aa) and 21% (in 680 aa) with *Lymantria dispar cypovirus* S1 (LdCPV) and *Bombix mori cytoplasmic polyhedrosis virus* S2 (BmCPV) (both of them *Cypovirus*), 25.6% (in 543 aa) and 24.7% (in 522 aa) with *Eyach virus* S1 (EYAV) and *Colorado tick fever virus* S1 (CTFV) (both of them *Coltivirus*). At the nucleotide level, MRCV S1 had 73.6%, 69.3% (in 3821 nt) and 54.7% (in 2161 nt) identity with RBSDV S1, FDV S1 and NLRV S1, respectively. We did not find a second possible ORF on the MRCV S1 complementary strand, as was reported for RBSDV S1 (Zhang et al., 2001b).

MRCV S2 has an extremely low G+C content of 31.41%. It encodes a single 134.39 kDa protein which

Table 1
Features of the MRCV genome segments and their deduced coded proteins

Genome segment	Accession number	Length (bp)	GC content (%)	Non-coding region (5′/3′)	Number of codons	Deduced coded protein (kDa)	pI[a]
S1	AF499925	4501	32.30	37/72	1464	168.40	8.22
S2	AF499926	3617	31.41	33/71	1171	134.39	6.24
S3	AF499928	3826	33.40	46/102	1226	141.66	6.52
S4[b]	AF395873	3566	32.70	15/116	1145	131.67	8.12
S6	AF499927	2638	37.67	79/192	789	90.0	4.66
S8[b]	AF395872	1931	34.90	24/134	591	68.26	6.68

[a] Isoelectric point.
[b] From Distéfano et al. (2002).

has 56.8, 40.3 and 20.4% identity to RBSDV S4, FDV S3 'B' spike and NLRV S2-encoded proteins, respectively. The latter is proposed to be the B-spike protein located on the surface of the core (Nakashima et al., 1996). MRCV S2-deduced protein showed 19% identity (in 263 aa) with the protein coded by the 3.81-segment of the unclassified insect virus *Diadromus puschellus reovirus* (DpRV) (Rabouille et al., 1994) by comparisons based on BLAST algorithm. A leucine zipper pattern [L-x(6)-L-x(6)-L-x(6)-L] was identified between positions

```
(i)                -          A          B          C          D
(ii)               -          1          2          3          -
(iii)              I          IV         V          VI         -

MRCV        (657)  IDRRARVI...DMKGMDAH...SGFFATSAQHTLFL...HSVMGDDVL...LSRIF  (905)
RBSDV       (657)  IDRRARVI...DMKGMDAH...TGFFATSAQHTLFL...HSVMGDDVL...LSRMY  (909)
FDV         (661)  IDRRARVI...DMKGMDAH...SGFFATSAQHTLFL...HSVMGDDVF...LSRLF  (914)
NLRV        (646)  IDRRGRII...DMSGMDAH...SGLFATSGQHTMFL...NYVMGDDIF...YSKYS  (894)
RRSV        (500)  IGRRQRAI...DASVQASV...SGQPFTTVHHTFTL...LTVQGDDTR...GFKVS  (735)
RDV         (643)  AWRPVRPI...DCSSWDQT...SGRLDTFFMNSVQN...FQVAGDDAI...PQKTV  (890)
LdCPV-14    (491)  IDRRQRAI...DISGMDAS...SGRADTSTHHTVLL...MRILGDDVR...YSKNS  (729)
BmCPV-1     (481)  SDRRQRAI...DASYVTTN...SGRADTSTHHTVLL...IKILGDDIM...DTKDI  (798)
CTFV        (584)  VGRRPRVI...DVKGMDSS...SGLLNTADQHTFLG...GSVLGDDQV...ETKFL  (848)
EYAV        (584)  VGRRPRVI...DVKGMDSS...SGLLNTADQHTFLG...GSVLGDDQV...ETKFL  (848)
GCRV        (526)  VQRRARSI...DASITYNY...SGSTATSTEHTANN...YVCQGDDGI...YLKLY  (791)
CSV         (492)  VQRRARSI...DASITYAA...SGSTATSTEHTANN...YVCQGDDGM...YLKLY  (757)
AHSV-9      (516)  PIKSTRIV...DYSEFDTH...SGENSTLIANSLHN...EQYVGDDTL...MSKTL  (807)
BTV-2       (515)  PIKATRTI...DYSEYDTH...SGENSTLIANSMHN...EQYVGDDTL...PSKTM  (804)
SiRV-A      (455)  PGRRTRII...DVSQWDSS...SGEKQTKAANSIAN...IRVDGDDNY...KVKAL  (669)
GBR         (496)  PARQTRGI...DVSQWDAS...SGEKTTKIGNSFAN...LRVDGDDNV...RVKAL  (720)
PoRV-C      (452)  PGRRTRAI...DVSQWDAS...SGEKQTKIMNSIAN...IRVDGDDNY...KVKAL  (662)
MRV-1       (521)  VQRRPRSI...DASITWDF...SGSTATSTEHTANN...YVCQGDDGL...GWKYD  (772)

RdRps consensus    --R--R-I...D-S---d-...SG---T----T--n.....----GDD--...—-K--
                     K    V     K         T          S                     R
```

Fig. 1. Conserved aa sequences of RdRps of reoviruses. The motifs presented by (i) Poch et al. (1989), (ii) Bruenn (1991) and (iii) Koonin (1992) are shown at the top. Consensus between all reported RdRps are summarized at the bottom. Upper case letters indicate an overall consensus or two aa alternatives in a certain position. Lower case letters indicate the position in which three aa possibilities were found. Numbers in parenthesis on the left and right indicate the starting and finishing position of the aligned sequences. The *Reoviridae* sequences used for comparison were: from genus *Fijivirus*: MRCV (AF499925); RBSDV—China Zhejiang (AJ294757), FDV (AY029520) and NLRV (D49693); from genus *Oryzavirus*: RRSV—Thailand (U66714); from genus *Phytoreovirus*: RDV—China (U73201); from genus *Cyporeovirus*: LdCPV (AF389452) and BmCPV (AF323782); from genus *Coltivirus*: CTFV—Florio N-7180 (AF133428) and EYAV (AF282467); from genus *Aquareovirus*: *Grass carp reovirus* (GCRV) (AF260511S2) and *Chum salmon reovirus* (CSV) (AF418295); from genus *Orbivirus*: *African horse sickness virus* serotype 9 (AHSV-9) (U94887) and *Bluetongue virus* serotype 2 (BTV-2) (L20508); from genus *Rotavirus*: Simian rotavirus SA11 (SiRV-A) (P22679), Group B rotavirus-IDIR (GBR) (P35942) and Porcine rotavirus (PoRV-C) (M74216); from genus *Orthoreovirus*: *Mammalian orthoreovirus* subgroup 1 (MRV-1) (M24734).

475–496 of MRCV S2-deduced protein. This pattern was located at a different position than that of FDV S3 'B' spike coded protein and was not present in the homologous RBSDV S4 and NLRV S2 coded protein. At the nucleotide level, the identities between MRCV S2 and the corresponding RBSDV and FDV segments were 64.2 and 57.1% (in 3027 nt), respectively. No identity was detected with the NLRV genomic counterpart.

MRCV S3 codes for a highly conserved 141.66 kDa protein. It has 84.5% identity with RBSDV S2, 57.6% with FDV S3 and 22.4% with NLRV S3 major core proteins, respectively (Nakashima et al., 1996; Isogai et al., 1998a; Zhang et al., 2001b). In addition, 19% (in 520 aa) and 16% identities (in 1040 aa) were detected by BLAST with the proteins coded by RRSV S3 and the 4.23-segment of DpRV (Bigot et al., 1995), in that order. At the nucleotide level, the identities between MRCV S3 and RBSDV S2 or FDV S3 were 76.5 and 64.5%, respectively, and were undetectable in corresponding segments of NLRV, RRSV and DpRV.

MRCV S6 has a high G+C content (37.67%) and the capacity to code for a 90 kDa protein acid protein (pI of 4.66) with 44.8 and 27.4% identities with the proteins encoded by RBSDV S6 and FDV S6, respectively. There was no detectable identity with any of the NLRV proteins or to other known proteins. The function of these proteins is unknown. After analyzing the secondary structure of the MRCV S6 coded protein, a long uninterrupted alpha-helix region encompassing the complete C-terminal half was predicted. This feature was also predicted at similar positions in the proteins coded by RBSDV S6 and FDV S6. This was noteworthy due to the fact that the C-terminal sequences of these related proteins were not aligned by the FASTA program. The nucleotide sequence identity between MRCV S6 and RBSDV S6 was of 57.9%, and no detectable similarities were found while comparing with FDV or NLRV sequences. Nevertheless, the size of MRCV S6 is more similar to that of NLRV S6 than to NLRV S5 or NLRV S7.

The homologies between MRCV coded proteins and their RBSDV, FDV and NLRV counterparts along with their putative function and localization are displayed on Table 2. MRCV S4 and S8 previously described coded proteins, are also included. When comparing MRCV with RBSDV, the homology between different segments varied markedly: S3, S1 and S4 were highly conserved, S8 and S2 intermediately conserved and S6 very poorly conserved. Accordingly, the same was observed when comparing with FDV equivalent segments.

To analyze MRCV relationships with other reovirus, phylogenetic trees with MRCV S1, S2, S3, and S6 sequences were constructed using the UPGMA method (Fig. 2). Bootstrap analysis ($n = 100$) of the data indicated that major branches of every tree were well supported. The phylogenetic analysis was repeated using the neighbor-joining method and the result was consistent with the trees of Fig. 2 (data not shown). The topology of the trees showed that MRCV is closely related to RBSDV, and to a lesser extent to FDV and NLRV. Particularly conserved sequences such as MRCV S1 (Fig. 2a) and MRCV S3 (Fig. 2c) allowed us to establish the phylogenetic relationship of these fijiviruses with members of other genera such as LdCPV (*Cypovirus*), RRSV (*Oryzavirus*), DpRV (unclassified reovirus) and rice dwarf virus (RDV) (*Phytoreovirus*). In order to study the phylogenetic relationships within the family *Reoviridae*, a tree was constructed using the overall aa sequences of RdRp proteins of all reported viruses of this family by neighbor-joining method (Fig. 3). MRCV closely clustered with other fijiviruses and was more related to members of the *Cypovirus* (BmCPV, LdCPV), *Oryzavirus* (RRSV) and *Coltivirus* (CTFV, EYAV) genera than to any other genus of *Reoviridae*.

4. Discussion

Genomic analysis of the previously reported MRCV segments S4 and S8 led us to propose that although MRCV was closely related to MRDV and RBSDV, it should be considered as a different *Fijivirus* species (Distéfano et al., 2002). In order to further extend the characterization of MRCV genome, the nucleotide and deduced aa sequences of four additional genome segments were obtained and analyzed, including the one coding for the viral RdRp. Because of their strong conservation and the fact that the presence of defined RdRp motifs almost undoubtfully identifies them, the RNA polymerase sequences are widely used for evolutionary studies. In this context, the analysis of MRCV RNA polymerase gene sequence and its comparison with homologous genes within the genus significantly completed the evolutionary picture. The previously described RdRps motifs were identified in the protein sequence coded by MRCV S1 and, as shown in Fig. 1, were conserved among reoviruses. Phylogenetic analysis based on aa sequences of representative reovirus RdRps (Fig. 3) showed the existence of two major evolutionary divisions of polymerase proteins among *Reoviridae* family: one grouped *Orthoreovirus*, *Aquareovirus*, *Orbivirus* and *Rotavirus*, and the second one grouped *Fijivirus*, *Cypovirus*, *Oryzavirus* and *Coltivirus*. Noticeably RDV, a plant reovirus belonging to *Phytoreovirus*, grouped with members of *Orbivirus* and *Rotavirus* genera and was only distantly related to *Oryzavirus* and *Fijivirus* plant reoviruses. These results are in accordance with the observation that RRSV *Oryzavirus* polymerase is more related to NLRV than to RDV (*Phytoreovirus*), BoRV (*Rotavirus*) or *Bluetongue virus* (BTV) (*Orbivirus*) (Upadhyaya et al., 1998).

Table 2
Identity between MRCV and related viruses coded proteins and their possible function and location

	RBSDV[a]	FDV	NLRV	Function	Location
MRCV S1	77.9%	64.2%	32.3%	RNA polymerase	Core
MRCV S2	56.8%	40.3%	20.4%	B-spike	Inner shell
MRCV S3	84.5%	57.6%	22.4%	Major core structural	Core
MRCV S4[b]	73.9%	56.1%	23.0%	Unknown	Unknown
MRCV S6	44.8%	27.4%	NI[c]	Unknown	Unknown
MRCV S8[b]	59.6%	ND[d]	23.8%	NTP-binding	Core

MRCV and related virus genome segment counterparts are: MRCV S1-RBSDV S1-FDV S1-NLRV S1; MRCV S2-RBSDV S4-FDV S3 'B' spike-NLRV S2; MRCV S3-RBSDV S2-FDV S3-NLRV S3; MRCV S4-RBSDV S3-FDV S4-NLRV S4; MRCV S6-RBSDV S6-FDV S6-NLRV S6; MRCV S8-RBSDV S8-NLRV S7.

[a] Chinese strain (Zhejiang).

[b] From Distéfano et al. (2002).

[c] NI, no identity detected.

[d] ND, no data available.

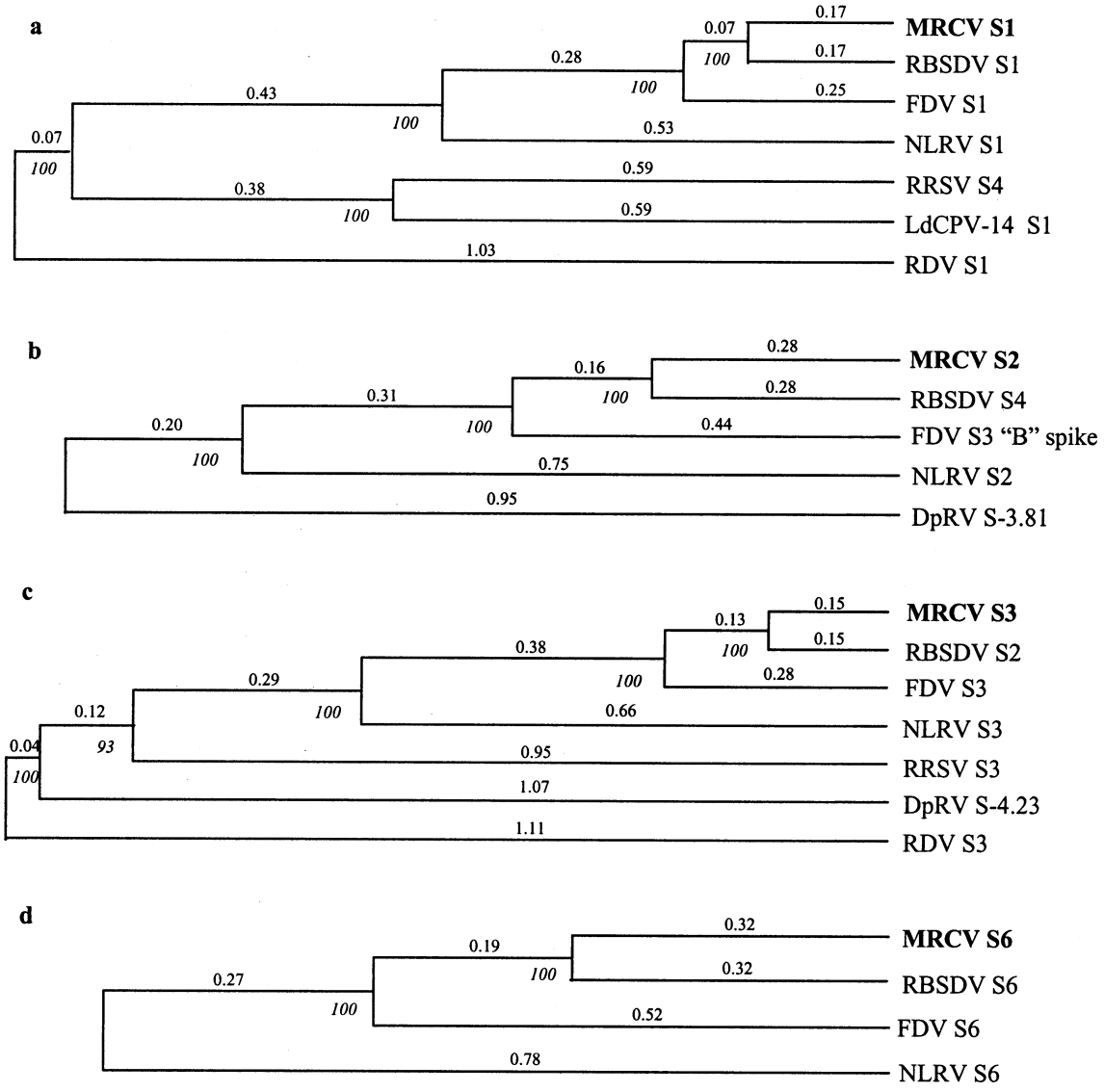

Fig. 2. Phylogenetic trees constructed with MRCV S1 (a), MRCV S2 (b), MRCV S3 (c) and MRCV S6 (d) and related sequences using the UPGMA method. The branch lengths are shown on top of the lines. Figures below the lines (in italics) indicate the frequency of the cluster after bootstrap analysis (100 replicates). DpRV designates *Diadromus puschellus reovirus*; all other virus abbreviations are defined in the legend of Fig. 1.

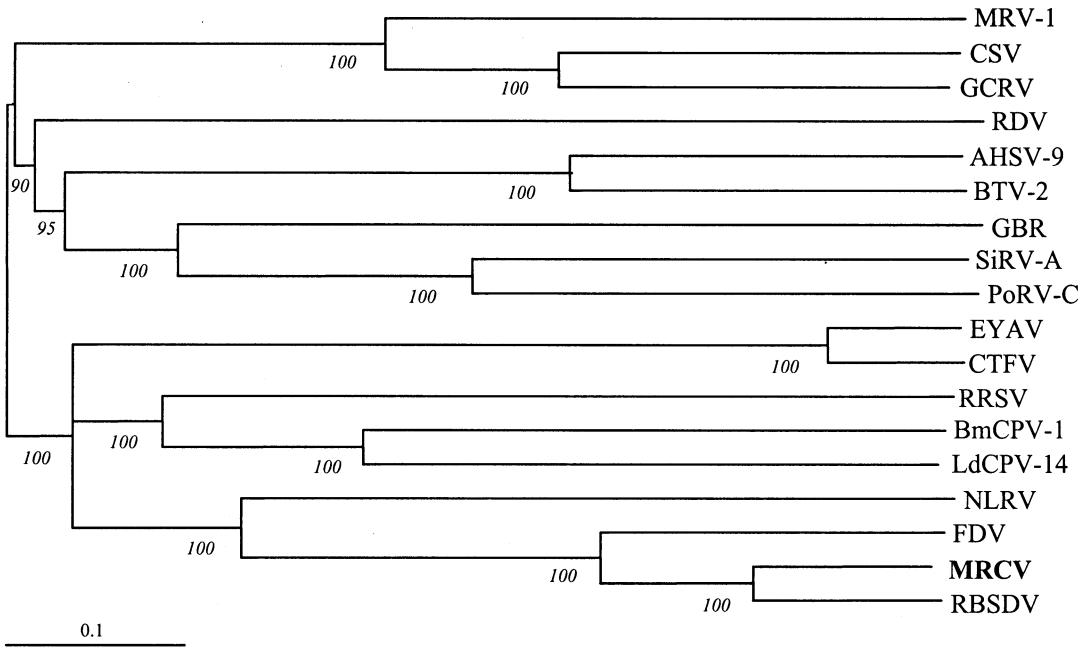

Fig. 3. Neighbor-joining phylogenetic tree based on aa sequences of all representative reovirus RdRps (abbreviations are listed in the legend of Fig. 1). The branch lengths are shown on top of the lines. Figures below the lines (in italics) indicate the frequency of the cluster after bootstrap analysis (100 replicates).

Due to the sequence identity of MRCV S3-coded protein with the major core proteins coded by RBSDV S2 (Isogai et al., 1998a; Zhang et al., 2001b) and NLRV S3 (Nakashima et al., 1996), we assumed that MRCV S3 probably also coded for this important structural protein. The identity between the proteins coded by MRCV S3 and its RBSDV counterpart was extremely high (84.5%). This level of identity was comparatively higher than that observed among the RNA polymerases of both viruses (77.9%), and represented the highest sequence homology so far detected between MRCV and RBSDV. A similar case was reported before in animal reoviruses: when comparing *African horse sickness virus* (AHSV) and BTV (both of them *Orbivirus*), putative viral polymerase conservation is second, in terms of homology, after the inner core protein VP3 (Vreede and Huismans, 1998). For viruses belonging to other genera of the same family, such as *Rotavirus*, the major core protein is also strongly conserved but to a lesser extent than the viral RNA polymerase (Bremont et al., 1992). Most of the work regarding transcription, encapsidation and replication of reoviruses was done on rotaviruses, where it was shown that one copy of each of all genome segments of dsRNA are highly organized inside the core. The major core protein VP2 has an unspecific affinity for RNA, and its assembly into cores is necessary for RdRP activity (Patton et al., 1997). Thus, the requirement of the core protein in dsRNA synthesis results in the coordination of RNA replication and virion morphogenesis in such a way that dsRNA are not produced unless cores are available in to which they can be packaged (Patton and Spencer, 2000). The strong conservation between the putative major core protein coded by MRCV S3 with RBSDV S2 and FDV S3 might be understood when the role of this protein in structure architecture and its possible involvement in virus replication and packaging was considered.

Previously we analyzed the MRCV S4 complete sequence and remarked its high conservation with respect to FDV S4 coded protein and to a RBSDV S3 partial sequence (L36524) (Distéfano et al., 2002). As the full RBSDV S3 sequence is now available (Zhang et al., 2001b), the overall aa homology between MRCV S4 and RBSDV S3 coded proteins could be established. The relatively high homology between them (73.9%) placed this protein sequence among the most conserved sequences, supporting previous conclusions.

In contrast to the high conservation levels between MRCV S1, S3 and S4 protein sequences to their RBSDV and FDV counterparts, the identity values between MRCV S2 coded protein with the RBSDV and FDV homologous proteins was intermediate. Its corresponding NLRV segment codes for the B-spike (Nakashima et al., 1996), and probably the protein coded by MRCV S2 plays the same role. In spite of the fact that the function of spikes is quite well understood in *Rotavirus* (cell attachment and membrane penetration Ludert et al., 1996; Denisova et al., 1999), and *Orthoreovirus* (entry into the cell, route of exit for nascent transcripts Dryden et al., 1993), the role of the two types of *Fijivirus* spikes (A and B) was not characterized yet.

With MRCV and RBSDV both being members of *Fijivirus* group 2, the identity found between both viruses S6 coded proteins (44.8%) was surprisingly low. Also, a low identity to FDV S6 coded protein and no identity to any NLRV coded proteins were detected. In accordance, no identity was detected between the protein coded by RBSDV S6 and any of NLRV coded proteins (Zhang et al., 2001b). This comparatively low conservation of the protein sequence and the fact that MRCV and its related viruses do not share insect vectors species, led us to consider the possibility that MRCV S6 coded protein could be involved in some step of the virus cycle in the insect vector.

The comparative level of homology between different MRCV and RBSDV or FDV encoded proteins varied noticeably (Table 2), adding evidence at an interspecific level, for an independent evolution of the different reovirus genome segments. This was formerly proposed at the intraspecific level by Chappell et al. (1994) and Bonneau et al. (2001) after studying *Mammalian orthoreoviruses* (MRV) and BTV, respectively. The last authors demonstrated that individual BTV genome segments evolve independently of one another by genetic drift in a host-specific manner.

Phylogenetic trees based on nucleotide sequences of MRCV genome segments S1, S2, S3 and S6 (Fig. 2) were consistent with each other and with the ones previously reported for MRCV S4 and S8 (Distéfano et al., 2002).

This work reported two strongly conserved proteins: the RdRp coded by MRCV S1 and the putative major core protein coded by MRCV S3. The possible spike-B protein encoded by MRCV S2 was intermediately conserved, and the protein of unknown function encoded by MRCV S6 was relatively poorly conserved. The analysis of the four MRCV segments presented here contributes to establish its genetic relationships with others viruses of the *Reoviridae* family giving additional support to the previous proposal that MRCV should be considered as a new *Fijivirus* species (Distéfano et al., 2002).

Acknowledgements

This work was supported by PICT 3489 and PICT 3488 from the Agencia Nacional de Promoción Científica y Técnica and by PIP 4460 from the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). AJD holds a doctoral fellowship from CONICET. HEH is a career member of the Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC) and Professor at the Departamento de Fisiología, Biología Molecular y Celular, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires. We are grateful to Drs Ruth Heinz and Guido König for the critical reading of the manuscript.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Arneodo, J.D., Lorenzo, E., Laguna, I.G., Abdala, G., Truol, G.A., 2002. Cytopathological characterization of *Mal de Río Cuarto virus* in corn, wheat and barley. Fitopatol. Brasil 27, 298–302.

Azuhata, F., Uyeda, I., Kimura, I., Shikata, E., 1993. Close similarity between genome structures of rice black-streaked dwarf and maize rough dwarf viruses. J. Gen. Virol. 74, 1227–1232.

Bigot, Y., Drezen, J.M., Sizaret, P.Y., Rabouille, A., Hamelin, M.H., Periquet, G., 1995. The genome segments of DpRV, a commensal reovirus of the wasp *Diadromus pulchellus* (Hymenoptera). Virology 210, 109–119.

Bonneau, K.R., Mullens, B.A., MacLachlan, N.J., 2001. Occurrence of genetic drift and founder effect during quasispecies evolution of the VP2 and NS3/NS3A genes of bluetongue virus upon passage between sheep, cattle, and *Culicoides sonorensis*. J. Virol. 75, 8298–8305.

Bremont, M., Juste-Lesage, P., Chabanne-Vautherot, D., Charpilienne, A., Cohen, J., 1992. Sequences of the four larger proteins of porcine group C rotavirus and comparison with the equivalent group A rotavirus proteins. Virology 186, 684–692.

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

Chappell, J.D., Goral, M.I., Rodgers, S.E., dePamphilis, C.W., Dermody, T.S., 1994. Sequence diversity within the reovirus S2 gene: reovirus genes reassort in nature, and their termini are predicted to form a panhandle motif. J. Virol. 68, 750–756.

Denisova, E., Dowling, W., LaMonica, R., Shaw, R., Scarlata, S., Ruggeri, F., Mackow, E.R., 1999. Rotavirus capsid protein VP5* permeabilizes membranes. J. Virol. 73, 3147–3153.

Distéfano, A.J., Conci, L.R., Muñoz Hidalgo, M., Guzmán, F., Hopp, H.E., del Vas, M., 2002. Sequence analysis of two complete genome segments of *Mal de Río Cuarto virus* (MRCV) supported the definition of a putative new species of the *Fijivirus* genus. Arch. Virol. 147, 1699–1709.

Dodds, J., 1993. dsRNA in diagnosis. In: Matthews, F. (Ed.), Diagnosis of Plant Virus Disease. CRC Press Inc, pp. 273–294.

Dryden, K.A., Wang, G., Yeager, M., Nibert, M.L., Coombs, K.M., Furlong, D.B., Fields, B.N., Baker, T.S., 1993. Early steps in reovirus infection are associated with dramatic changes in supramolecular structure and protein conformation: analysis of virions and subviral particles by cryoelectron microscopy and image reconstruction. J. Cell Biol. 122, 1023–1041.

Felsenstein, J., 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Department of Genetics, University of Washington, Seattle.

Frohman, M.A., Dush, M.K., Martin, G.R., 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA 85, 8998–9002.

Hofmann, K., Bucher, P., Falquet, L., Bairoch, A., 1999. The PROSITE database, its status in 1999. Nucleic Acids Res. 27, 215–219.

Isogai, M., Uyeda, I., Lee, B., 1998a. Detection and assignment of proteins encoded by rice black streaked dwarf fijivirus S7, S8, S9 and S10. J. Gen. Virol. 79, 1487–1494.

Isogai, M., Uyeda, I., Lindsten, K., 1998b. Taxonomic characteristics of fijiviruses based on nucleotide sequences of the *oat sterile dwarf virus* genome. J. Gen. Virol. 79, 1479–1485.

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

Koonin, E.V., 1992. Evolution of double-stranded RNA viruses: a case of polyphyletic origin from different groups of positive stranded RNA viruses. Semin. Virol. 3, 327–339.

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.

Ludert, J.E., Feng, N.G., Yu, J.H., Broome, R.L., Hoshino, Y., Greenberg, H.B., 1996. Genetic mapping indicates that VP4 is the rotavirus cell attachment protein in vitro and in vivo. J. Virol. 70, 487–493.

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

Marzachi, C., d'Aquilio, S.A.M., Boccardo, G., 1996. The double-stranded RNA genome of *Maize rough dwarf fijivirus* contains both mono and dicistronic segments. Eur. J. Plant Pathol. 102, 601–605.

Mathews, D.H., Sabina, J., Zuker, M., Turner, D.H., 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. 288, 911–940.

McMahon, J.A., Dale, J.L., Harding, R.M., 1999. Taxonomic implications for Fijiviruses based on the terminal sequences of *Fiji disease fijivirus*. Arch. Virol. 144, 2259–2263.

Mertens, P.P.C., Arella, M., Attoui, H., Belloncik, S., Bergoin, M., Boccardo, G., Booth, T.F., Chiu, W., Diprose, J.M., Duncan, R., Estes, M.K., et al., 2000. Family *Reoviridae*, in: van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E., Estes, M.K., Lemon, S., McGeoch, D., Wickner, R.B., Mayo, M.A., Pringle, C.R., Maniloff, J., (eds.), Virus Taxonomy. Classification and Nomenclature of Viruses. Seventh Report of the International Committee on Taxonomy of Viruses, Academic Press, New York San Diego, pp. 395–435.

Milne, R.G., Boccardo, G., Dal Bó, E., Nome, F., 1983. Association of *Maize rough dwarf virus* with Mal de Río Cuarto in Argentina. Phytopathology 73, 1290–1292.

Nakashima, N., Koizumi, M., Watanabe, H., Noda, H., 1996. Complete nucleotide sequence of the *Nilaparvata lugens* reovirus: a putative member of the genus *Fijivirus*. J. Gen. Virol. 77, 139–146.

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. J. Gen. Virol. 75, 221–225.

Nome, S.F., Lenardón, S.L., Raju, B.C., Laguna, I.G., Lowe, S.K., Docampo, D., 1981. Association of reovirus-like particles with 'Enfermedad de Río IV' of maize in Argentina. Phytopathol. Z. 101, 7–15.

Patton, J.T., Jones, M.T., Kalbach, A.N., He, Y.W., Xiaobo, J., 1997. Rotavirus RNA polymerase requires the core shell protein to synthesize the double-stranded RNA genome. J. Virol. 71, 9618–9626.

Patton, J.T., Spencer, E., 2000. Genome replication and packaging of segmented double-stranded RNA viruses. Virology 277, 217–225.

Pearson, W.R., 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol. 183, 63–68.

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

Rabouille, A., Bigot, Y., Drezen, J.M., Sizaret, P.Y., Hamelin, M.H., Periquet, G., 1994. A member of the *Reoviridae* (DpRV) has a ploidy-specific genomic segment in the wasp *Diadromus pulchellus* (Hymenoptera). Virology 205, 228–237.

Remes Lenicov, A.M., de Teson, A., Dagoberto, E., Huguet, N., 1985. Hallazgo de uno de los vectores del Mal de Río Cuarto en maíz. Gac. Agropec. 25, 251–258.

Rodriguez Pardina, P., Giménez Pecci, M.P., Laguna, I.G., 1998. Wheat: a new natural host for the *Mal de Río Cuarto Virus* in the endemic disease area, Río Cuarto, Córdoba province, Argentina. Plant Dis. 82, 149–152.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic tree. Mol. Biol. Evol. 4, 406–425.

Sanger, F., Nicklen, S., Coulson, A., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

Soo, H.M., Handley, J.A., Maugeri, M.M., Burns, P., Smith, G.R., Dale, J.L., Harding, R.M., 1998. Molecular characterization of Fiji disease fijivirus genome segment 9. J. Gen. Virol. 79, 3155–3161.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Upadhyaya, N.M., Ramm, K., Gellatly, J.A., Li, Z., Kositratana, W., Waterhouse, P.M., 1998. Rice ragged stunt oryzavirus genome segment S4 could encode an RNA dependent RNA polymerase and a second protein of unknown function. Arch. Virol. 140, 1945–1956.

Uyeda, I., Azuhata, F., Shikata, E., 1990. Nucleotide sequence of rice black-streaked dwarf virus genome segment 10. P. Jpn Acad. 66, 37–39.

Vreede, F.T., Huismans, H., 1998. Sequence analysis of the RNA polymerase gene of African horse sickness virus. Arch. Virol. 143, 413–419.

Zhang, H., Chen, J., Lei, J., Adams, M.J., 2001a. Sequence analysis shows that a dwarfing disease on rice, wheat and maize in China is caused by *Rice black-streaked dwarf virus*. Eur. J. Plant Pathol. 107, 563–567.

Zhang, H.M., Chen, J., Adams, M.J., 2001b. Molecular characterization of segments 1–6 of *Rice black-streaked dwarf virus* from China provides the complete genome. Arch. Virol. 146, 2331–2339.

Zuker, M., Mathews, D.H., Turner, D.H., 1999. Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide. In: Barciszewski, J., Clark, B.F.C. (Eds.), RNA Biochemistry and Biotechnology. NATO ASI Series. Kluwer Academic Publishers, pp. 11–43.

Zukerlandl, E., Pauling, L., 1965. Molecules as documents of evolutionary history. J. Theor. Biol. 8, 357–366.