Isolation and molecular analysis of SRBSDV isolates infecting rice in Vietnam

Thanh Tran Ngoc[1], Hoang Anh Ta[2], Vinh Do Nang[3] and Neeti Sanan Mishra[1]*

[1]International Center for Genetic Engineering and Biotechnology,
Aruna Asaf Ali Marg, New Delhi, India
And Agricultural Genetic Institute, Hanoi, Vietnam

[2] Plant Protection Research Institute,
Hanoi, Vietnam

[3] Agricultural Genetic Institute,
Hanoi, Vietnam

[4]International Center for Genetic Engineering and Biotechnology,
Aruna Asaf Ali Marg, New Delhi, India

ABSTRACT— ***Southern rice black-streaked dwarf virus (SRBSDV) is a newly described member of the sub-group 2 in the genus Fijivirus, family Reoviridae. SRBSDV genome has 10 segments encoding 13 opened reading frames (ORFs). In this study we cloned and sequenced 6 ORFs from SRBSDV isolates collected from 3 provinces in Vietnam. The sequence analysis suggested that SRBSDV populations in North Vietnam are closely related to those in South China. The viral isolates from Nghe-An and Quang-Tri provinces of Vietnam were closer to each other as compared to the Quang-Ngai province suggesting that the virus seems to be evolving while spreading in a southern direction. Moreover the analysis revealed the presence of DNA binding domains in gene 5.2 indicating that it may be playing an important role in the viral life cycle.***

Keywords— Fijivirus, Vietnam, Rice, Sequence analysis, Evolution

1. INTRODUCTION

Viruses are one of the most destructive pathogens of rice as they are capable of causing substantial yield reduction or complete damage to crop. In Asia alone, rice (*Oryza sativa* L.) cultivation covers 90% of agricultural land [1]. Approximately 15 viruses have been known to affect Asian rice varieties, of which 9 viruses cause severe diseases. Some of the viruses multiply in the insect hosts and are transmitted transovarially. Main vectors bearing these viruses are leaf-hoppers and plant-hoppers such as *Nilaparvata lugens*, *Nephotettix cincticeps*, *Laodelphax striatellus*, and *Sogatella furcifera* [2]. The viruses tend to survive from one season to another through over-wintering plant hoppers which makes their control becomes more difficult [3]. These insect vectors are distributed widely in Asian countries and they migrate long distances, even across the ocean [4 and 5].

The Fijivirus is a genus of the plant Reoviruses along with the Oryzavirus and Phytoreovirus genera. The Reoviridae family is characterized by icosahedral virion particles with diameter ranging from 60-80 nm. Depending on genus, all virions contain a full genome consisting of 10 to 12 linear double-stranded RNA segments [6 and 7]. The Fijiviruses comprise of 5 groups, amongst these, each group contains 1 virus each, except group-2 which has 4 viruses. All four members of group-2 viz: *Mal de Rio Cuarto Virus* (MRCV), *Pangola Stunt Virus* (PaSV), *Maize Rough Dwarf Virus* (MRDV) and *Rice Black-Streaked Dwarf Virus* (RBSDV) are closely related [8]. *Southern Rice Black-Streaked Dwarf Virus* (SRBSDV) is a new species in the group-2 and is closely related to RBSDV in nomenclature, symptoms and sequence.

SRBSDV was first reported in 2001 in Yangxi county of Guangdong province in southern China [9]. It has since spread rapidly through southern China to a number of countries in Southeastern Asia, including Vietnam and Japan [10 and 11]. Severity of its expansion can be estimated from the fact that from the summer 2009 to spring 2010, more than 20000 ha. of land covering 30 provinces of North and Central Vietnam got infected [12]. Recently a fresh alert for SRBSDV infection spread has been recently generated in North and Central Vietnam, with supposed threat to more than 447 ha land in 8 new provinces.

Major vector of SRBSDV transmission is the white back plant-hopper, *Sogatella furcifera* Horvath, while in some cases the transmission was affected by the small brown plant-hopper, *Laodelphax striatellus* Fallen [13]. The SRBSDV propagation is also enhanced as it is capable of infecting some weeds and other cereal crops in the grass family like barley and maize. The SRBSDV infected plants typically exhibit darkening and deformation of leaves, black-streaked swellings along the veins on the underside of the leaf blades, sheaths and culms, stunting of growth and reduction or complete loss of grain [7 and 10]. The disease manifestation may vary according to the infection of different growth stages; however, symptomatically they look very similar to the RBSDV caused infection. The recent origin and spread of SRBSDV has thus generated significant scientific curiosity to identify and characterize the virus.

SRBSDV like other Fijiviruses has 10 segments of double-stranded RNA that have been named S1-10 according to their lengths [10]. In this study, we describe the molecular comparison of 5 genomic regions of SRBSDV encoding 6 ORFs viz. 5.2, 6.0, 7.1, 7.2, 9.2 and 10.1. These were amplified using RT-PCR from total RNA isolated from rice leaves collected from SRBSDV infected fields in the Nghe-An, Quang-Ngai, Quang-Tri provinces in North midland of Vietnam bordering the Laos region during the summer season. The sequences were analyzed in detail and compared with the available data [12] to present molecular evidence for the divergence in viral isolates indicating that point mutation has probably featured prominently during the southwards spread and divergence of this virus in Vietnam. This information will improve our understanding of virus spread and evolution and therefore assist in the long-term development of durable and safe resistance to the disease.

2. MATERIAL AND METHODS

2.1 *Source of plant material*

Rice leaf samples were collected from SRBSDV infected fields in the Nghe-An, Quang-Ngai, Quang-Tri provinces in North midland of Vietnam bordering the Laos region during the summer crop season. The field infected rice plants appeared stunted and dark green with twisted leaf tips and/or leaf blades, split leaf margins, and white waxy enations along veinlets on the underside of leaf blades and/or on the leaf sheath.

2.2 *RT-PCR*

Total RNA was extracted from infected leaves of rice leaves using an RNeasy kit (Qiagen) according to manufacturer's protocol. cDNAs were synthesized and amplified using Invitrogen SuperScript First-Strand kit. In brief, 1 µg of total RNA were used as a template for first-strand cDNA synthesis using gene specific primers (Table 2). After incubation at 42°C for 60 min, the reaction was stopped by heating at 70°C for 15 min. The resulting cDNA was purified using the PCR Purification Kit (Qiagen) and used as templates for PCR. The amplifications were done using specific primers designed from published genomic sequences (Table 2). The cycling program included a denaturation step at 94°C for 3 min, 30 cycles at 94°C for 30 s, 60°C for 30 s, and 72°C for 2 min, and a final extension step at 72°C for 10 min.

2.3 *Cloning and sequencing*

Amplified DNA fragments were cloned into pGEMT-easy vector (Promega), and transformed into competent Escherichia coli DH5α cells for cloning. Plasmids were isolated using the QIAprep spin mini prep kit (Qiagen), and the inserts were sequenced entirely on both strands with universal primers T7 and SP6. Three replicate clones from each PCR product were sequenced to minimize any error. Sequence data of all the components obtained were compared with other reported isolates of SRBSDV from the Sino-Vietnam region.

2.4 *Sequence analysis*

BLAST algorithms were used to search for sequence similarity in the GenBank (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and EBI databases (http://www.ebi.ac.uk/Tools/blast/). Nucleotide/amino acid sequence similarities were carried out with isolates reported for each component from various SBRSDV isolates. A phylogenetic tree was constructed to determine the relationships between the SRBSDV isolates using CLUSTALW [14]. Nucleotide sequence analysis was performed using DNAman version 6.0 (Lynnon BioSoft, Quebec, Canada). Molecular diversity among the isolates was estimated by calculating the p-distance values for nucleotide and amino-acid comparisons in MEGA5 using the Kimura 2-parameter [15].

All the open reading frames (ORFs) that could potentially express proteins were used in protein-protein BLAST (BLASTp) searches to identify potential homologues of these in the NCBI non-redundant protein sequences database and a nested search strategy to restrict the search to SBRSDV protein sequences. The protein sequences were analyzed using Predict protein for predicting features on the primary sequence [16]. The hydropathy plots were tabulated using Kyte-Doolittle hydrophobicity index [17] that allows for the visualization of hydrophobicity over the length of a peptide sequence.

3. RESULTS

3.1 *Identification and molecular characteristics of novel isolates of SRBSDV*

The total RNA was isolated and sequences of 5 genomic regions of SRBSDV encoding 6 ORFs viz. 5.2, 6.0, 7.1, 7.2, 9.2 and 10.1 were amplified using RT-PCR. The amplifications were done using specific primers designed from the published genomic sequences (Table 2). The amplified products were cloned and sequenced. The sequence details and accession numbers of all fifteen genes are summarized in Table 3. Sequence comparisons of the three genomes and their phylogenetic analysis showed that all three isolates were closely related to each other, though gene sequences of viral isolates from Nghe-An and Quang-Tri provinces of Vietnam were closer to each other as compared to those from the Quang-Ngai province (Table S1).

3.2 *Sequence analysis of the genes from new strains of SRBSDV and their comparison with other Fijiviruses*

Figure 1 shows a genomic map of SRBSDV for each component, while Table 3 summarizes the nucleotide and protein information obtained for the 5 genes of all three isolates. The sequences of each gene from the three SRBSDV isolates were compared to published sequences (Table S1 and Figure 2). To determine the relationships between the SRBSDV isolates, a phylogenetic tree was constructed using CLUSTALW [16]. Molecular diversity among the isolates was estimated by calculating the p-distance values for nucleotide comparisons in MEGA5 using the Kimura 2-parameter [17]. This analysis reveals that genes 6.0 and 7.1 are the most conserved among the three species, while gene 9.2 exhibited high sequence divergence. Genes 5.2 and 7.2 of viral isolates from Nghe-An and Quang-Tri provinces of Vietnam were closer to each other as compared to the Quang-Ngai province. On analyzing the sequences of Genes 6.0 and 7.2 obtained from different isolates of SRBSDV and RBSDV it was observed that despite of high similarity between the sequences the SRBSDV sequences grouped away from the RBSDV sequences (Figure S1).

3.3 *Phylogenetic and molecular analysis of the proteins encoded by the viral ORFs*

The amino acid sequence of protein encoded by each gene, used in this study, was compared to published SRBSDV protein sequences. The protein sequences were analyzed using Predict protein for predicting features on the primary sequence [16]. Analysis of the amino acids revealed that the encoded proteins of the viral isolates under study also shared high sequence identity with each other (Table S2 and Figure 2).

Gene 5.2 is predicted to encode a protein (P5.2) of 205 amino acids. This protein is specific to the SRBSDV group. The amino acid sequence of P5.2 from Quang-Ngai isolate was 97.1% similar to that of Quang-Tri isolate and 96.1% similar to that of Nghe-An isolate (Table S2). It was observed that the sequences of the three isolates show positional differences at 8 amino acid positions. However, phylogenetic analysis revealed that the sequence of Quang-Ngai isolate was closer to the Guangdong isolate which is considered to be the parental strain, while the sequences of Quang-Tri and Nghe-An isolates grouped away from it (Figure 2A). The function of this protein is not known. Primary analysis of the amino acid sequence revealed that it encodes a protein of 23.7 Kda with a theoretical pI of 8.3. This protein shows poor water stability (Figure 3A). It has a mixed structure with a putative Leucine zipper pattern (LQLKQLDLTQSLVKLEQNFNQ) and five potential phosphorylation sites (10:SPR, 184:SER, 48:SRQE, 83:TITE and 167:TMSD). It also has a potential DUF domain (68-127 amino acid) with similarities to COG5179 whose members are annotated as being transcription initiation factor TFIID subunit 1. It is thus likely to be playing an important role in viral life cycle.

Likewise, **Gene 6.0** contains an ORF of 2382 bp that is predicted to encode a protein (P6.0) of 793 amino acids. The amino acid sequence of P6.0 from Quang-Ngai isolate was 99.5% similar to that of Quang-Tri isolate and 98.6% similar to that of Nghe-An isolate (Table S2). It was observed that the sequences of the three isolates show positional differences at 17 amino acid positions. Phylogenetic analysis revealed that the sequence of Nghe-An isolate was closer to the Guangdong homolog, while the Quang-Tri and Quang-Ngai isolates grouped away from it (Figure 2B). Primary analysis of the protein sequence revealed that it encodes a protein of 90.05 Kda with a theoretical pI of 4.72. This is a putative myosin-like protein having host RNAi suppression activities. This protein can be classified as a water soluble protein (Figure 3B). It also has a mixed structure with a putative Leucine zipper pattern (LQLKQLDLTQSLVKLEQNFNQ). The protein shows a high potential of being phosphorylated due to the presence of two potential CAMP (cAMP- and cGMP-dependent protein kinase) phosphorylation sites (17:RRVT, 542:RRYS), nine potential PKC (Protein Kinase C) phosphorylation sites (53:TKR, 60:SNK, 262:TLK, 480:SQK, 487:TSK, 593:SDR, 599:SLK, 619:TER and 639:TTR), nineteen potential CKII (Caesin Kinase II) phosphorylation sites (45:SSSE, 68:TVAE, 179:TFLD, 207:TDSD, 224:SPSD, 276:SNAE, 337:SKSE, 439:TVND, 456:TSTE, 461:SETD, 499:SWAD, 507:TECD, 27:TYED, 619:TERD, 658:SVVE, 696:SGVD, 721:SLCD, 728:TVSE and 745:STHD) and one TYR (Tyrosine kinase) phosphorylation site (601:KQLDDAY).

Segment 7 encodes two gene fragments. **Gene 7.1** contains an ORF of 1074 bp that is predicted to encode a protein (P7.1) of 357 amino acids. The amino acid sequence of P7.1 from Quang-Ngai isolate was 99.2% similar to both the Quang-Tri and Nghe-An isolates (Table S2). It was observed that the sequences of the three isolates show positional

differences at 3 amino acid positions. Further the phylogenetic analysis showed that these three sequences were close to the Guangdong homolog but Quang-Ngai isolate branched away from the Quang-Tri and Nghe-An isolates (Figure 2C). P7.1 is known to form tubular structures in insect cells. Primary analysis of the amino acid sequence revealed that it encodes a protein of 40.5 Kda with a theoretical pI of 6.1. This protein exhibits poor water solubility (Figure 3C). It has twelve potential sites for phosphorylation by CAMP kinase (262:RKDS and 284:RKPS), PKC (47:SPK, 185:SFK, 260:SVR, 282:SPR and 307:SVK) and CKII (16:TKNE, 102:SNQD, 142:TSSE, 206:STSE and 307:SVKE).

Gene 7.2 contains an ORF of 930 bp that is predicted to encode a protein (P7.2) of 309 amino acids. The amino acid sequence of P7.2 from Quang-Ngai isolate was 99.7% similar to both the Quang-Tri and Nghe-An isolates (Table S2). It was observed that the sequences of the three isolates show positional difference at only 1 amino acid position. Further the phylogenetic analysis showed that like P7.1 all the three P7.2 sequences were close to the Guangdong homolog but Quang-Ngai isolate branched away from the Quang-Tri and Nghe-An isolates (Figure 2D). P7.2 is a non-structural protein of 36.4 Kda with a theoretical pI of 4.5. It exhibits poor water solubility (Figure 3D). The protein has two potential PKC phosphorylation sites (295:SMR and 299:TKK) and four potential CKII phosphorylation sites (16:TSLE, 46:SLSE, 109:SQFE and 220:SIPD).

Gene 9.2 contains an ORF of 630 bp that is predicted to encode a protein (P9.2) of 209 amino acids. This is a highly conserved protein as amino acid sequence of P9.2 from Quang-Ngai isolate was 99.5% similar to that of Quang-Tri isolate and 100% similar to that of Nghe-An isolate (Table S2). It was observed that the sequences of the three isolates show positional difference at 1 amino acid only. Phylogenetic analysis revealed that the Quang-Ngai sequence was closer to the Guangdong homolog, while the Quang-Tri and Nghe-An grouped away from it (Figure 2E). Primary analysis of the amino acid sequence revealed that it encodes a protein of 24.2 Kda with a theoretical pI of 4.9. The function of this protein is not known. It is predicted to be poorly soluble in water (Figure 3E) and has a transmembrane helix. The protein has three potential CAMP phosphorylation sites (27:SAK, 96:TLK and 162:STK) and four potential CKII phosphorylation sites (36:SDFD, 47:SDSE, 150:SLTE and 163:TKIE).

Gene 10.1 contains an ORF of 1674 bp that is predicted to encode a protein (P10.1) of 557 amino acids. The amino acid sequence of P10.1 from Quang-Ngai isolate was 99.4% similar to that of Quang-Tri isolate and 99.8% similar to that of Nghe-An isolate (Table S2). It was observed that the sequences of the three isolates show positional difference at 3 amino acids only. Phylogenetic analysis revealed that the Quang-Ngai and Quang-Tri sequences were closer to the Guangdong homolog, while the Nghe-An sequence grouped away from it (Figure 2F). Primary analysis of the amino acid sequence revealed that it encodes a protein of 62.6 Kda with a theoretical pI of 7.3. It functions as an outer capsid protein and is predicted to be water soluble (Figure 3F). The protein has potential phosphorylation sites for CAMP kinase (530: KKAT), PKC (132:TSK, 154:TLK, 211:SGK, 297:TWK, 354:SLR, 400:TNK and 409:SYR), CK II (133:SKLD, 147:TLTE, 179:SGLE, 227:STTD, 275:SALD, 282:SLLE, 293:SFIE, 358:TEIE, 456:TIDD and 542:SNSE) and TYR (213:KLRIDGGY, 305:KDEDQDEYand 436:KLGVEKAY).

4. DISCUSSION

The SRBSDV is closely related to RBSDV in nomenclature, symptoms and sequence. The virus was transmitted to between rice seedlings and from rice to maize seedlings by the white backed plant-hopper, *Sogatella furcimera*. It is now recognized as a new species in the group 2. Its recent origin and spread has generated significant scientific curiosity to identify and characterize the virus. The 10 SRBSDV segments encode 13 ORFs (Figure 1). The smallest fragment (S10) is 1.798 bp whereas the largest fragment (S1) is 4500 bp [9]. The genome encodes for six putative structural proteins (P1, P2, P3, P4, P8 and P10) and five putative non-structural proteins (P6, P7-1, P7-2, P9-1 and P9-2). The functions of two proteins (P5.1 and P5.2) are not known [9, 13 and 18] (Table 1), although a recent study has shown that MRCV P5-2, could function as regulator of host gene expression [19].

In an earlier work, a one-step reverse transcription-polymerase chain reaction (RT-PCR) protocol was used to confirm the presence of SRBSDV from 29 provinces among 5 agro-ecological regions in North and Central Vietnam. Moreover partial sequences of RNA segments 4 and 10 from several isolates showed very low genetic divergences between isolates from Vietnam and China, suggesting a common origin [11].

In another study, the molecular diversity of segment S10 was analysed because it encodes the outer capsid protein known to be involved in vector specificity and because interactions with the vector play an important role in fijivirus evolution [12]. Moreover partial sequences of RNA segments 4 and 10 from several isolates showed very low genetic divergences between isolates from Vietnam and China, suggesting a common origin. Analysis of the Chinese isolates indicated that segments S5 and S6 have the maximum molecular diversity [20]. In this study, the nucleotide and protein sequences of six genes 5.2, 6.0, 7.1, 7.2, 9.2 and 10.1, from the three Vietnam isolates were isolated and sequenced.

Analysis of specific gene sequences of the isolates used in this study suggests that these SBRSDV isolate are very similar to the other known SRBSDV isolates. It was observed that genes 5.2 and 7.2 of viral isolates from Nghe-An and

Quang-Tri provinces of Vietnam were closer to each other as compared to the Quang-Ngai province. At the nucleotide level genes 6.0 and 7.2 were the most conserved among the three species while gene 9.2 exhibited high sequence divergence. The difference at the nucleotide level points to position specific single nucleotide changes. However this was not reflected at the protein level as proteins encoded by gene 7.1, 7.2 and 9.2, were highly conserved within the viral isolates. The analysis revealed that gene 5.2 has DNA binding domains and may be playing an important role in the viral life cycle. The study also demonstrates the presence of an uneven sequence variability amongst the SBRSDV isolates from Vietnam. This indicates that the nucleotide variations are a result of recent introduction in the viral genome and the virus is in the evolving phase as it is spreading to different regions.

Thus the sequence analysis suggests that SRBSDV populations in North Vietnam are closely related to those in South China. The viral isolates from Nghe-An and Quang-Tri provinces of Vietnam were closer to each other as compared to the Quang-Ngai province suggesting that the virus seems to be evolving while spreading in a southern direction. It has been proposed that the movement of WBPH has been the main factor in the spread of SRBSDV between provinces in North Vietnam and South China as it is a typical long-distance migration pest [11]. Earlier studies have suggested that japonica rice landraces carry ovicidal and sucking-inhibited resistance genes for the white backed plant-hopper though indica rice or hybrid rice have little resistance to the pest [21]. The wide cultivation of susceptible hybrid rice varieties and increasing temperatures favour the possibility for SBRSDV infection and transmission, which is a potential threat to rice production in South Asia. Thus there is an urgent need for more information in order to develop successful and sustainable strategies to manage and restrict these plant viruses.

5. ACKNOWLEDGEMENT

This work was partly funded by Department of Biotechnology (DBT), Government of India. TTN received fellowship from ICGEB for carrying out this work. We thank the staff of PPRI for extending the RNA isolated from SBRSDV infected field samples for use in this study. We thank Prof. Sudhir K. Sopory and Dr. Sunil K. Mukherjee for their help with the study.

6. REFERENCES

Type your main text in 10-point Times New Roman, single-spaced with single line spacing. Do not use double-spacing. The spacing before and after paragraphs is 6 pt. All paragraphs should be indented by 0.2". Be sure your text is fully justified—that is, flush left and flush right. Please do not place any additional blank lines between paragraphs.

Figure and table captions should be 10-point Times New Roman non-boldface. Initially capitalize only the first word of each figure caption and table title. Figures and tables must be numbered separately. For example: "**Figure 1:** UML Sequence Diagram for Heath Information Service", "Table 1: Input data". Figure captions are to be below the figures. Table titles are to be centered above the tables.

7. REFERENCES

[1] Zeigler R, Barclay A. The Relevance of Rice. Rice2008; 1: 3-10.
[2] Hibino H. Biology and epidemiology of rice viruses. Annu. Rev. Phytopathol.1996; 34: 249-274.
[3] Shinkai A. Studies on insect transmission of rice virus diseases in Japan. Bulletin Natl. Institute Agri. Sci. Tokyo.1962; C14: 1–12.
[4] Kisimoto R. Long distance migration of plant hopper Sogatella furcifera and Nilaparvata lugens. Tropical Agri. Res. Service.1971; 5: 201–216.
[5] Otuka A, Matsumura M, Watanabe T, Dinh VT. A migration analysis for rice planthoppers, Sogatella furcifera (Horváth) and Nilaparvata lugens (Stål) (Homoptera: Delphacidae), emigrating from northern Vietnam from April to May. Appl. Entomol. Zool.2008; 43: 527-534.
[6] Renault S, Stasiak K, Federici B, Bigot Y. Commensal and mutualistic relationships of reoviruses with their parasitoid wasp hosts. J. Insect Physiol.2005; 51: 137-148.
[7] Cuong HV, Hai NV, Man VT, Matsumoto M. Rice dwarf disease in North Vietnam in 2009 is caused by southern rice black-streaked dwarf virus (SRBSDV). Bulletin Institute Tropical Agri., Kyushu Univ.2009; 32: 85-92.
[8] Boccardo G, Milne RG. Plant reovirus group. Commonwealth Mycol. Institute/Association Applied Biol.1984; 294
[9] Wang Q, Yang J, Zhou G-H, Zhang H-M, Chen J-P, Adams M.J. The Complete Genome Sequence of Two Isolates of Southern rice black-streaked dwarf virus, a New Member of the Genus Fijivirus. J. Phytopathol.2010; 158: 733-737.
[10] Zhang HM, Yang J, Chen JP, Adams MJ. A black-streaked dwarf disease on rice in China is caused by a novel fijivirus. Arch. Virol.2008; 153: 1893-1898.

[11] Kumamoto K. New Virus Disease Found in Japan. Yomiuri Online, International Rice Research Institute website.2010; http://ricehoppers.net/2010/09/new-virus-disease-found-in-japan/.
[12] Hoang AT, Zhang H, Yang J, Chen J, Hébrard E, Zhou G, Vinh VN, Cheng J. Identification, characterization, and distribution of southern rice black-streaked dwarf virus in Vietnam. Plant Disease.2011; 95: 1063-1069.
[13] Zhou G, Wen J, Cai D, Li P, Xu D, Zhang S. Southern rice black-streaked dwarf virus: A new proposed Fijivirus species in the family Reoviridae. Chinese Science Bulletin.2008; 53: 3677-3685.
[14] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nuc. Acids Res.1994; 22: 4673-4680.
[15] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol. Biol. Evol.2011; 28: 2731-2739.
[16] Field D, Tiwari B, Booth T, Houten S, Swan D, Bertrand N, Thurston M. Open software for biologists: from famine to feast. Nature Biotechnol.2006; 24: 801-804.
[17] Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol.1982; 157: 105-132.
[18] Jia D, Chen H, Zheng A, Chen Q, Liu Q, Xie L, Wu Z, Wei T. Development of an insect vector cell culture and RNA interference system to investigate the functional role of fijivirus replication protein. J. Virol.2012; 86: 5800-5807.
[19] Maroniche GA, Mongelli VC, Llauger G, Alfonso V, Taboga O, del Vas M. In vivo subcellular localization of Mal de Rı´o Cuarto virus (MRCV) non-structural proteins in insect cells reveals their putative functions. Virol.2012; 430: 81–89.
[20] Wang Q, Yang J, Zhou GD, Zhang HM, Chen JP, Adams MJ. The complete genome sequence of two isolates of a new rice-infecting fijivirus from China. J. Phytopathol.2010; 158: 733-737.
[21] Cheng Z, Li S, Gao R, Sun F, Liu W, Zhou G, Wu J, Zhou X, Zhou Y. Distribution and genetic diversity of Southern rice black-streaked dwarf virus in China. Virol. J.2013; 10: 307.

8. LEGEND TO FIGURES

Figure 1: Map of SRBSDV segments showing the complete gene size and ORF positions. The grey bars represent the ORFs while the white bars represent the UTRs. The second ORF coded by gene segments 5, 7 and 9 are represented by black bars.

Figure 2: Phylogenetic analysis of the nucleotide and protein sequences of SRBSDV genes obtained from different isolates. The different represent segments (A) 5.2 (B) 6.0 (C) 7.1, (D) 7.2 (E) 9.2 (F) 10.1

Figure 3: Hydropathic analysis of the protein sequences of SRBSDV genes. The panels represent segments (A) 5.2 (B) 6.0 (C) 7.1, (D) 7.2 (E) 9.2 (F) 10.1

9. ELECTRONIC SUPPLEMENTARY MATERIAL

Table S1. The nucleotide sequence comparison of individual genes between different isolates of the SRBSDV from Vietnam.

Figure S1: Phylogenetic analysis of genes 6.1 and 7.2 of different SRBSDV and RBSDV isolates. The different panels represent (A) 6.1 (B) 7.2

Table S2: The amino acid sequence comparison of individual gene products between different isolates of the SRBSDV from Vietnam. The analysis for each gene is provided separately. **A.** Gene 5.2, **B**. Gene 6.0, **C**. Gene 7.1, **D**. Gene 7.2, **E**. Gene 9.2, **F**. Gene 10.1.