The Complete Genome Sequence of Southern rice black-streaked dwarf virus Isolated from Vietnam

Thi-Sau Dinh1,2, Cuiji Zhou1, Xiuling Cao1, Chenggui Han1, Jialin Yu1, Dawei Li1 and Yongliang Zhang1*
1State Key Laboratory of Agro-Biotechnology, China Agricultural University, Beijing 100193, China
2Faculty of Agriculture Forestry Fisheries, Vinh University, Vinh city, Nghe An province 42000, Vietnam
(Received on March 21, 2012; Revised on July 3, 2012; Accepted on July 11, 2012)

We determined the complete genome sequence of a Vietnamese isolate of Southern rice black-streaked dwarf virus (SRBSDV). Whole genome comparisons and phylogenetic analysis showed that the genome of the Vietnamese isolate shared high nucleotide sequence identities of over 97.5% with those of the reported Chinese isolates, confirming a common origin of them. Moreover, the greatest divergence between different SRBSDV isolates was found in the segments S1, S3, S4 and S6, which differs from the sequence alignment results between SRBSDV and Rice black streaked dwarf virus (RBSDV), implying that SRBSDV evolved in a unique way independent of RBSDV. This is the first report of a complete nucleotide sequence of SRBSDV from Vietnam and our data provides new clues for further understanding of molecular variation and epidemiology of SRBSDV in Southeast Asia.

Keywords: phylogenetic analysis, sequence, Southern rice black-streaked dwarf virus, vietnamese isolate

Southern rice black-streaked dwarf virus (SRBSDV) is proposed to be a novel member of the Fijivirus group 2 in the family Reoviridae (Attoui, 2011; Zhang et al., 2008; Zhou et al., 2008). The double-stranded (dsRNA) of SRBSDV contains 10 segments (S1 to S10), which were named in order of decreasing molecular weight (Zhou et al., 2008). SRBSDV was first observed in Yangxi county, Guangdong province in China in 2001. From then on, this disease became one of the most serious viral diseases of rice in southern and central China, causing severe losses in rice production (Wang et al., 2010; Zhou et al., 2010). In 2008, SRBSDV was also detected in maize plants in northern China (Yin et al., 2011). In 2010, the occurrence of SRBSDV disease was firstly reported in rice fields in the Kumamoto Prefecture of Japan (Choi, 2010). In 2009, serious viral rice disease occurred in many northern provinces of Vietnam, expanding to a total of 5506 hectares. Of this area, the yields of approximately 3510 hectares were totally lost. The causal agent of the rice dwarf disease in the north of Vietnam has also been identified as SRBSDV (Cuong et al., 2009; Hoang et al., 2011). Further investigation suggested that the outbreak of SRBSDV in Vietnam was associated with rising populations of the white-backed plant hopper (Sogatella furcifera), which is the major vector of SRBSDV (Cuong et al., 2009; Zhou et al., 2008). As planthoppers are constantly displaced by wind currents, the spread of the virus diseases they carry will be inevitable.

Symptoms of SRBSDV infection vary depending on the crop age when infected. Characteristic symptoms include dark-green and wrinkled leaves, incomplete tassel, tumor-like protrusions ending in small enations, tiller formation on the upper parts, up-growing rootlets, and in particular the presence of white to black waxy galls along the major veins of the leaves and culms (Wang et al., 2012; Zhou et al., 2008; Zhou et al., 2010).

Up to now, the complete genome sequence of SRBSDV had been obtained only for three isolates, which originated from Guangdong (GD), Hainan (HN) and Hubei (HB) provinces of China, respectively (Wang et al., 2010). As only partial sequences of segments 4 and 10 from several Vietnamese isolates were investigated and the S10 segment was very conserved among different isolates (Cuong et al., 2009; Hoang et al., 2011), more information about the genome sequence of an SRBSDV isolate from Vietnam will further characterize the geographic and molecular variation of this virus, thus facilitating control and prevention of this disease in Southeast Asia.

Viral genomic dsRNAs were extracted directly from naturally infected field-grown rice plants from Thua Thien Hue province in central Vietnam using the method described previously (Dodds et al., 1984). Primers for cloning segments S1 to S10 of Vietnamese isolate were designed to obtain the full length sequence of each genome segment based on the previously reported SRBSDV sequences available in GenBank of NCBI (Wang et al., 2010). Two step RT-PCR was employed to clone cDNA fragments of

*Corresponding author.
Phone) +86-10-62733190, FAX) +86-10-62732012
E-mail) cauzhangyl@cau.edu.cn
SRBSDV using the primers listed in Supplementary Table 1. The first strand cDNA was synthesized with MLV reverse transcriptase (Promega, Madison, WI, USA) and subjected to thermocycling amplification using LA Taq DNA polymerase (TaKaRa, Dalian, China). The 3’ and 5’ terminal sequences of SRBSDV Vietnamese isolate were obtained using the single primer amplification technique (SPAT) as described by Zhou et al. (2008). Briefly, an oligodeoxyribonucleotide primer, P-linker (5’-PO4-ttccttagcagtgtaacctgtGTCAGTCGACAGCA-NH2-3’), was ligated to the 3’ ends of both strands of viral dsRNA using T4 RNA ligase (TaKaRa). Oligonucleotides that failed to ligate were removed by spin column chromatography on cellulose nitrate CF-11. Up-stream specific primers for each genome segment were based on the sequence previously determined elsewhere and the downstream general primers were either 5’-tgcgatcagccagttca-3’ (complementary to the 3’ end of the P-linker sequence) or 5’-ggaactgacacagagtgatcagctgcat-3’ (complementary to the 5’ end of the P-linker sequence). The PCR products were recovered with QIAquick Gel Extraction Kit (QIAGEN, Hilden, Germany) and cloned into pGEM-T Easy vector (Promega, Madison, WI, USA) according to the manufacturer’s instructions, followed by transformation into Escherichia coli Trans-T1 (TransGen Biotech, Beijing, China). In order to obtain the consensus sequences and avoid inconsistencies, three or more clones from each of the two separate PCR reactions were sequenced.

The entire DNA sequence of each segment of SRBSDV was then assembled by matching of overlapping sequences. Complete nucleotide sequences of each genomic RNA segment were determined and submitted to GenBank (accession numbers JQ692572-JQ692581 for segments S1 to S10, respectively). The sequences were then compared to equivalent sequences from other SRBSDV isolates present in GenBank. Identities at nucleotide (nt) and amino acid (aa) levels were calculated using the DNAANAL software (vers. 5.2.2, Lynnon BioSoft, Quebec, Canada). Phylogenetic analysis were conducted with Bioedit (vers. 7.0.4.1) (Hall, 1999) and Clustal W method (Thompson et al., 1994), and visualized with MEGA (vers. 4.1) (Kumar et al., 2008). The phylogeny reconstruction was generated by the neighbor-joining method and displayed using the MEGA 4.1 program, with bootstrap value of 500 replicates.

The viruses included for comparison in this study are Fiji disease virus (FDV) (Harding et al., 2006; McQualter et al., 2004; Soo et al., 1998), Oat sterile dwarf virus (OSDV) (Isogai et al., 1998b), Nilaparvata lugens reovirus (NLRV) (Nakashima and Noda, 1994; Nakashima et al., 1996; Noda et al., 1994), Mal de Rio Cuarto virus (MRCV) (Distefano et al., 2002; Distefano et al., 2005; Guzmán et al., 2007), Maize rough dwarf virus (MRDV) (Marzachi et al., 1991, 1999) and Clustal W method (Thompson et al., 1994), and visualized with MEGA (vers. 4.1) (Kumar et al., 2008). The phylogeny reconstruction was generated by the neighbor-joining method and displayed using the MEGA 4.1 program, with bootstrap value of 500 replicates.

The viruses included for comparison in this study are Fiji disease virus (FDV) (Harding et al., 2006; McQualter et al., 2004; Soo et al., 1998), Oat sterile dwarf virus (OSDV) (Isogai et al., 1998b), Nilaparvata lugens reovirus (NLRV) (Nakashima and Noda, 1994; Nakashima et al., 1996; Noda et al., 1994), Mal de Rio Cuarto virus (MRCV) (Distefano et al., 2002; Distefano et al., 2005; Guzmán et al., 2007), Maize rough dwarf virus (MRDV) (Marzachi et al., 1991, 1999) and Clustal W method (Thompson et al., 1994), and visualized with MEGA (vers. 4.1) (Kumar et al., 2008). The phylogeny reconstruction was generated by the neighbor-joining method and displayed using the MEGA 4.1 program, with bootstrap value of 500 replicates.

<table>
<thead>
<tr>
<th>Segment</th>
<th>GD isolate</th>
<th>HN isolate</th>
<th>HB isolate</th>
<th>RBSVD</th>
<th>MRCV</th>
<th>MRDV</th>
<th>FDV</th>
<th>OSDV</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>97.9 (96.4)</td>
<td>98.8 (99.0)</td>
<td>99.3 (99.5)</td>
<td>78.1 (86.0)</td>
<td>73.1 (79.1)</td>
<td>n/a</td>
<td>63.1 (63.4)</td>
<td>n/a</td>
</tr>
<tr>
<td>S2</td>
<td>98.8 (98.1)</td>
<td>99.3 (99.1)</td>
<td>99.7 (99.8)</td>
<td>78.6 (89.6)</td>
<td>74.1 (83.7)</td>
<td>n/a</td>
<td>61.0 (56.6)</td>
<td>n/a</td>
</tr>
<tr>
<td>S3</td>
<td>97.7 (96.7)</td>
<td>97.6 (98.4)</td>
<td>98.4 (98.6)</td>
<td>72.7 (73.6)</td>
<td>63.7 (56.8)</td>
<td>n/a</td>
<td>50.7 (37.6)</td>
<td>n/a</td>
</tr>
<tr>
<td>S4</td>
<td>99.2 (99.4)</td>
<td>97.5 (97.0)</td>
<td>98.7 (99.1)</td>
<td>77.9 (86.3)</td>
<td>71.0 (74.3)</td>
<td>n/a</td>
<td>60.8 (54.7)</td>
<td>n/a</td>
</tr>
<tr>
<td>S5</td>
<td>98.4</td>
<td>98.9</td>
<td>99.2</td>
<td>69.4</td>
<td>64.0</td>
<td>n/a</td>
<td>47.7</td>
<td>n/a</td>
</tr>
<tr>
<td>S6</td>
<td>98.7 (96.7)</td>
<td>98.9 (98.9)</td>
<td>99.2 (98.6)</td>
<td>68.3 (62.7)</td>
<td>55.5 (42.5)</td>
<td>n/a</td>
<td>36.9 (22.9)</td>
<td>n/a</td>
</tr>
<tr>
<td>S7</td>
<td>99.8</td>
<td>99.8</td>
<td>99.3</td>
<td>71.7</td>
<td>62.4</td>
<td>73.0</td>
<td>51.8</td>
<td>35.9</td>
</tr>
<tr>
<td>S8</td>
<td>99.5 (99.8)</td>
<td>99.1 (99.2)</td>
<td>99.2 (99.3)</td>
<td>72.1 (71.9)</td>
<td>60.4 (55.5)</td>
<td>71.5 (71.1)</td>
<td>46.8 (38.5)</td>
<td>39.2 (22.5)</td>
</tr>
<tr>
<td>S9</td>
<td>99.9</td>
<td>99.1</td>
<td>98.9</td>
<td>74.2</td>
<td>67.2</td>
<td>74.8</td>
<td>42.5</td>
<td>34.5</td>
</tr>
<tr>
<td>S10</td>
<td>99.9 (99.1)</td>
<td>97.4 (99.0)</td>
<td>98.0 (98.1)</td>
<td>77.5 (71.8)</td>
<td>64.6 (61.2)</td>
<td>77.8 (73.2)</td>
<td>34.3 (36.7)</td>
<td>22.8 (20.4)</td>
</tr>
</tbody>
</table>

1 GD isolate, Guangdong isolate (GenBank accession No. FN563983-FN563988 and EU784840-EU784843); HN isolate, Hainan isolate (GenBank accession No. FN563989-FN563996, EU522339 and EU522360) and HB, Hubei isolates of SRBSDV (GenBank accession No. HM585270-HM585279); MRCV, Mal de Rio Cuarto virus (GenBank accession No. AF499925-AF499928, AF395872, AF395873, AY607586, AY607587, AY923115 and DQ023312); MRDV, Maize rough dwarf virus (GenBank accession No. L76560-L76562 and X55701); FDV, Fiji disease virus (GenBank accession No. AY029520, AF049704, AF359556, AF049705, AY029521, AF356083, AY789927, AY297693, AF050086 and AY297694); RBSVD, Rice black-streaked dwarf virus (GenBank accession No. AJ294757, AJ409145-AJ409148, AJ293984, AJ297642, AJ297430, AJ297431 and AJ297433); OSDV, Oat sterile dwarf virus (GenBank accession No. AB011024-AB011027).

2 Percent nucleotide identities are shown for the complete segment comparisons and (in brackets) the percentage amino acid identities between the predicted ORFs. The amino acid sequences deduced from S1-S4, S6, S8 and S10 were named P1-P4, P6, P8 and P10, respectively. Similarly, proteins encoded by S5, S7 and S9 were named P5-1/2, P7-1/2 and P9-1/2, respectively.
1996), Rice black-streaked dwarf virus (RBSDV) (Wang et al., 2003; Zhang et al., 2001) and three isolates of SRBSDV (Wang et al., 2010) (Table 1). All the viruses mentioned above belong to the genus *Fijivirus*, and MRCV, MRDV and RBSDV are classified as *Fijivirus* group 2.

Among the putative structural proteins encoded by SRBSDV, P1, P2 and P4 are putative RNA-dependent RNA polymerase, core protein and outer-shell B-spike protein, respectively; P3 is a putative capping enzyme (Wang et al., 2010; Zhang et al., 2001); P8 and P10 are putative core and major outer capsid proteins, respectively (Isogai et al., 1998a; Wang et al., 2010). Among the putative nonstructural proteins encoded by SRBSDV, P6 is a viral RNA-silencing suppressor (Lu et al., 2011), P7-1 is the major constituent of the tubules and has the intrinsic ability to self-interact to form tubules in non-host insect cells (Liu et al., 2011); while P9-1 of SRBSDV was essential for viroplasm formation and viral replication (Jia et al., 2012) (Table 1). Hoang et al (2011) obtained partial sequences of segments S4 and S10 from different Vietnamese SRBSDV isolates and comparisons showed that both of them, especially S10, had high sequence identity among these isolates, suggesting that these segments were not the best alternative for comparison analysis and that more needs to be known about sequence diversity in other segments of the SRBSDV genome. For this reason, we cloned and obtained the full genome sequence of SRBSDV Vietnamese isolate.

Sequence comparison between RBSDV and Vietnamese isolates showed the same segment identities trends as reported previously, of which S1, S2 and S10 are the most conserved while S5 and S6 are the least (Wang et al., 2010). In the phylogenetic analysis using the corresponding regions of all other plant-infecting fijiviruses, all the SRBSDV isolates clustered in one distinct clade well-separated from isolates of other species in the genus (Fig. 1). Comparing the different segments with their counterparts among the 4 SRBSDV isolates indicated that the Vietnamese isolate was highly similar to HB isolate in all 10 genome segments with over 98.4% nt and 98.0% aa identities. Sequence alignment between Vietnamese and GD isolates revealed that segments S3 and S6 were most divergent, with 97.7% and 97.8% nt identities and 96.7% aa identity, respectively. The most sequence divergence between Vietnamese and HN isolate was located in segments S3 and S4, with 97.6% and 97.5% nt identities and 98.4% and 97.0% aa identities, respectively. Interestingly, S3 and S4 were the most divergent segments at the nt level whereas S6 and S9 were most divergent at the aa level between the Vietnamese and HB isolates. Moreover, it is worth noting that P9-1, encoded by S9, had a relatively low aa identity between the Vietnamese and HN isolates (97.4%) (Table 1).

A previous report showed that S5 and S6 were the most divergent segments between isolates GD and HN (Wang et al., 2010), whereas sequence comparison in our work revealed that the most divergence among the four SRBSDV isolates was found in segments S1, S3, S4 and S6, which also differs from the sequence alignment results between SRBSDV and RBSDV. These results implied a unique evolution pattern of SRBSDV. Recombination or reassortment occurs in many RNA viruses and can be of major

![Fig. 1. Phylogenetic tree of SRBSDV Vietnamese (VNM) isolate and other fijiviruses constructed with genomic sequences of segments S1, S3, S4 and S6 (A-D), respectively. The phylogeny reconstruction was generated by the neighbor-joining method and displayed using the MEGA4.1 program, with bootstrap value of 500 replicates. Bootstrap values are shown at nodes. Isolates are indicated in the tree by virus name-accession number. The names of SRBSDV as well as some RBSDV isolates were also listed.](image-url)
evolutionary significance (Jiang et al., 2008; Jonson et al., 2011; Miranda et al., 2000; Simon-Loriere and Holmes, 2011). Li et al. (2012) reported that four natural recombinants were detected and the recombinant breakpoints were identified on S10 RNA of 21 RBSDV isolates (Li et al., 2012). Using the software RDP3 (ver. 3.44) (Li et al., 2012; Martin et al., 2010), the sequences of segments S1, S3, S4 and S6 of SRBSDV, respectively, were examined to identify putative recombinants and recombination. Only recombination events that were detected by at least five different methods were accepted. Results showed that recombination events had occurred in all of these four segments of SRBSDV isolates from different geographical regions, indicating that recombination affects several aspects of molecular variation and evolution of SRBSDV. This likely resulted from the frequent migration of the white-backed plant hopper vector harboring SRBSDV from one place to another and transmission of SRBSDV among host plants such as maize and rice. Overall, the regions which varied most among the four SRBSDV isolates were closely related to its pathogenesis, which involved virus-vector-host interactions as well as environmental factors. Phylogenetic analysis indicated that SRBSDV isolates from China and Vietnam present different evolutionary distances in the segments S1, S3, S4 and S6, further suggesting the diversity and complexity in the evolution of SRBSDV.

Collectively, we report the first complete genome sequence of an SRBSDV isolate from Vietnam. Our results enriches the data for study of molecular variation and epidemiology of SRBSDV in Southeast Asia. Further investigation of the biological characters needs to be done in order to better illustrate the extent of variation in all 10 genomic RNA segments of SRBSDV and its relationship to variation in pathogenicity of the virus.

Acknowledgments

This work was partly supported by the National Natural Science Foundation of China (Grant No. 31100115) and the Fundamental Research Funds for the Central Universities of China (Grant No. 2010JS074). We also thank Yongqiang Li for technical assistance in the recombination analysis.

References


Liu, Y., Jia, D., Chen, H., Chen, Q., Xie, L., Wu, Z. and Wei, T. 2011. The P7-1 protein of southern rice black-streaked dwarf...


