Molecular characterization of two isolates of sweet potato leaf curl virus infecting *Ipomoea indica* in China

Awais Zulfiqar, He Ren, Jie Zhang, Jing Li, Xiaofeng Cui and Yan Xie*

Key Laboratory of Rice Biology, Institute of Biotechnology, Zhejiang University, Hangzhou 310029, People’s Republic of China.

Accepted 1 July, 2011

Morning glory (*Ipomoea indica*) plants showing yellow mosaic and vein yellowing symptoms were collected from Jiangsu and Yunnan province of China. The amplified DNA-A components of two isolates named JS1 and Y338 were comprised of 2827 and 2801 nucleotides, respectively, in length with genome organization similar to that of begomoviruses. Complete nucleotide sequence revealed the presence of two ORFs (AV1 and AV2) in the virion sense and four ORFs (AC1, AC2, AC3 and AC4) in the complementary sense. Comparison analysis showed that DNA-A sequence of JS1 isolate was closely related to that of sweet potato leaf curl virus (SPLCV) from United States with nucleotide sequence identity of 97.0% and DNA-A of Y338 showed highest sequence identity at 97.8% with an isolate of SPLCV from China. Phylogenetic analysis showed that JS1 clustered together with SPLCV-US and Y338 showed close relationship with SPLCV-CN. Sequence analysis indicated that isolates JS1 and Y338 were strains of SPLCV.

**Key words:** Sweet potato leaf curl virus, begomoviruses, *Ipomoea indica*, sequence comparison.

**INTRODUCTION**

Members in the family Geminiviridae, named Geminiviruses are important plant pathogens with circular single stranded DNA genome packed into twin shaped icosahedral particles (Moffat et al., 1999). The Geminiviruses fall into four genera, namely Mastrevirus, Curtovirus, Topocuvirus and Begomovirus, on the basis of virus vector species, host range and genome organization. The Begomovirus is the largest genus of this family and comprises of whitefly transmitted geminiviruses which infect dicotyledonous plants (Fauquet et al., 2003, 2008). The genomes of most begomoviruses consist of two components (DNA-A and DNA-B) known as bipartite begomoviruses, although there are some monopartite species that lack DNA-B component. DNA-A encodes the replication-associated protein (Rep); the replication enhancer protein (REN) required for viral DNA replication; the transcriptional activator protein (TrAP) concerned with gene expression control; and the coat protein (CP) which is essential for viral transmission by *Bemisia tabaci* (Hanley-Bowdin et al., 1999). DNA-B encodes products required for inter and intra-cellular movement of the virus in the tissues of the plant hosts (Sharma et al., 2005; Saunders et al., 2000).

Sweet potato (*Ipomoea batatas*) is ranked as the seventh most important food crop worldwide and is an economically important crop in China. Several *Ipomoea* species including *Ipomoea indica*, are grown ornamentally all over the world. The symptoms shown by *Ipomoea*-infecting geminiviruses are host dependent and usually consist of leaf curling and yellow vein. Mild symptoms or even symptomless infections may cause severe damage to sweet potato production (Clark and Hoy, 2006; Valverde et al., 2007). To date, about 20 virus species of distinct families have been described in sweet potato (Valverde et al., 2004).

A number of begomovirus species including sweet potato leaf curl virus (SPLCV), Ipomoea yellow vein virus (IYVV) and sweet potato leaf curl Georgia virus (SPLCGV) have been described as infecting *Ipomoea* species (Lotrakul and Valverde, 1999; Banks et al., 1999; Lotrakul et al., 2003). SPLCV and other related begomoviruses infecting sweet potato have been reported...
Figure 1. Infected *I. indica* showing typical yellow mosaic and yellow vein symptoms.

from all continents of the world like Americas (Fuentes and Salazer, 2003), Asia (Luan et al., 2006), Africa (Miano et al., 2006) and infecting *I. indica* in Europe (Briddon et al., 2006). Many sweet potato growing regions are likely to be infected by begomoviruses, but their incidence and distribution is still unknown. In this work, molecular characterization of two isolates of SPLCV infecting *I. indica* and their phylogenetic relationships with other begomoviruses are described.

MATERIALS AND METHODS

Virus source and nucleic acid extraction

Naturally infected *I. indica* plants showing yellow mosaic and yellow vein (begomovirus-like symptoms) were collected from two provinces of China (Figure 1). Infected leaf tissues of two samples named JS1 collected from Jiangsu province in 2010 and Y338 from Yunnan province in 2008, were used for nucleic acid extraction according to the method developed by Dellaporta et al. (1983).

PCR amplification, rolling circle amplification (RCA) analysis and cloning

JS1 isolate infecting *I. indica* was determined by RCA diagnosis using Φ29 DNA polymerase from the Templiphi™ DNA Amplification Kit (GE Healthcare, Piscataway, NJ, USA) as described by Guo et al. (2009). A PCR method with RCA product as template and degenerate primer pair PA/PB that amplified the genome of *Ipomoea*-infecting geminivirus covering part of intergenic region (IR) and AV2 gene from total nucleic acid was developed (Xie et al., 2002). Consequently, 500 base pair fragment of JS1 was obtained. The amplicon was cloned into pGEM-T easy vector as described previously (Zhou et al., 2003) and sequenced. On the basis of determined sequence, specific abutting primer pair JS1-F: (5′-CCCTAAGGTTCCTGGCTCGTATTTC-3′)/JS1-R: (5′-CTATCGTGCCCTACACTGGAATGC-3′) was designed for the amplification of full length DNA-A genome which was then cloned and sequenced. PCR did not yield any fragment of Y338 with PA/PB primer pair, then another primer pair BegoAFor1/BegoARev1 (Ha et al., 2006) specific for DNA-A was used to amplify 1.2 kb fragment. On the basis of the determined sequence, specific primer pair Y338F (5′-TGCTGTTGCCCAATTCTTGAG-3′)/Y338R (5′-GGAACGTCCATCTTGAACTCAT-3′) was designed to get 1.6 kb fragment, which was cloned and sequenced.

DNA sequence comparison and phylogenetic analysis

Sequences were assembled and analyzed using the DNAStar software version 6.0 (DNASTar Inc., Madison, WI, USA). Sequence similarity matrixes were performed using BLAST algorithm (http://www.ncbi.nlm.nih.gov/). The open reading frames (ORFs) were found by the online ORF Finder (http://www.ncbi.nlm.nih.gov/orf/orf.html). Selected sequences were subjected to CLUSTAL V method of MegAlign in DNAStar to get multiple sequence alignments. Phylogenetic tree was drawn by neighbor-joining and maximum parsimony methods included in MEGA4.0 (http://www.megasoftware.net/features.html) and was displayed, manipulated and printed using tree view. The abbreviations
Table 1. Geminiviruses sequences used for comparison with two isolates from this study, their genome sequence length, accession number, assigned abbreviation and host species.

<table>
<thead>
<tr>
<th>Virus species</th>
<th>Abbreviation</th>
<th>Length</th>
<th>Host species</th>
<th>Accession code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweet potato leaf curl virus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweet potato leaf curl Canary virus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweet potato leaf curl Lanzarote virus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweet potato leaf curl Spain virus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweet potato leaf curl Georgia virus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States: Georgia: 16</td>
<td>SPLCGV-[US:Geo:16]</td>
<td>2773</td>
<td>Ipomoea batatas</td>
<td>AF326775</td>
</tr>
<tr>
<td>Ipomoea yellow vein virus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other begomoviruses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ageratum yellow vein China virus [Hn 2-19]</td>
<td>AYVCNV[Hn 2-19]</td>
<td>2748</td>
<td>Stachytarpheta</td>
<td>AJ564744</td>
</tr>
<tr>
<td>Stachytarpheta leaf curl virus</td>
<td>StLCuV [Hn 6-1]</td>
<td>2751</td>
<td>Stachytarpheta</td>
<td>AJ564742</td>
</tr>
<tr>
<td>Tobacco leaf curl Yunan virus</td>
<td>TbLCYnV [Y161]</td>
<td>2747</td>
<td>Tomato</td>
<td>AJ566744</td>
</tr>
<tr>
<td>Bean golden mosaic virus [Brazil]</td>
<td>BGMV [Brazil]</td>
<td>2617</td>
<td>Bean</td>
<td>M88686</td>
</tr>
<tr>
<td>Bean dwarf mosaic virus</td>
<td>BDMV</td>
<td>2615</td>
<td>Bean</td>
<td>M88179</td>
</tr>
<tr>
<td>Cucurbit leaf crumple virus [Arizona]</td>
<td>CuLCrV [US]</td>
<td>2632</td>
<td>Cucurbit</td>
<td>AF256200</td>
</tr>
</tbody>
</table>

and GenBank accession numbers of begomoviruses used for comparison and phylogenetic analysis are listed in Table 1.

RESULTS

Genome organization and sequence comparison

Complete nucleotide sequences of DNA-A of two isolates were determined. The DNA-A of JS1 and Y338 contained 2827 nucleotides (nts) and 2801 nts, respectively (GenBank Accession Nos JF768740 and FN806776). JS1 and Y338 had two ORFs in the viral sense strand and four ORFs in the complementary strand separated by an intergenic region (IR). The intergenic region (IR) contains a GC-rich inverted repeat that has the potential to form a stem loop structure, including the conserved nonanucleotide TAATATTAC sequence that contains the nicking site for initiation of virion sense DNA replication (Laufs et al., 1995). The IR that consisted of 277 nts for JS1 and 251 nts for Y338 was similar to those of other geminiviruses.

The DNA-A sequences of JS1 and Y338 shared the
Table 2. Percentage nucleotide and amino acid sequence identities of SPLC-[JS1] and SPLC-[Y338] DNA-A when compared with other previously described *Ipomoea*-infecting begomoviruses.

<table>
<thead>
<tr>
<th>Virus</th>
<th>DNA-A (%)</th>
<th>IR (%)</th>
<th>AV2 (%)</th>
<th>AV1 (%)</th>
<th>AC1 (%)</th>
<th>AC2 (%)</th>
<th>AC3 (%)</th>
<th>AC4 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPLC-[Y338]</td>
<td>92.9/100</td>
<td>83.3/100</td>
<td>99.1/100</td>
<td>96.5/100</td>
<td>97.0/100</td>
<td>90.5/100</td>
<td>91.7/100</td>
<td>85.9/100</td>
</tr>
<tr>
<td>SPLC-[US:Lou:94]</td>
<td>97.0/92.9</td>
<td>92.8/82.5</td>
<td>97.4/96.5</td>
<td>98.8/96.9</td>
<td>98.9/97.0</td>
<td>93.2/93.2</td>
<td>95.8/92.4</td>
<td>95.3/83.5</td>
</tr>
<tr>
<td>SPLC-[KR:J-508:03]</td>
<td>96.4/92.7</td>
<td>89.5/87.3</td>
<td>96.5/96.5</td>
<td>98.8/96.9</td>
<td>99.2/96.7</td>
<td>90.5/90.5</td>
<td>95.1/91.7</td>
<td>96.8/84.7</td>
</tr>
<tr>
<td>SPLC-[CN:F-p3:07]</td>
<td>92.1/90.5</td>
<td>87.2/78.4</td>
<td>94.8/96.5</td>
<td>96.9/97.2</td>
<td>94.8/95.1</td>
<td>85.8/83.1</td>
<td>88.9/84.0</td>
<td>90.6/92.9</td>
</tr>
<tr>
<td>SPLC-[JP:Kum:98]</td>
<td>93.1/89.9</td>
<td>95.1/79.1</td>
<td>96.5/96.5</td>
<td>98.0/96.9</td>
<td>95.0/94.8</td>
<td>69.6/66.9</td>
<td>80.0/77.8</td>
<td>90.6/95.3</td>
</tr>
<tr>
<td>SPLC-[CN:RL31:07]</td>
<td>92.8/97.8</td>
<td>78.4/92.4</td>
<td>99.1/100</td>
<td>96.5/100</td>
<td>97.0/97.8</td>
<td>91.9/98.6</td>
<td>92.4/99.3</td>
<td>85.9/97.6</td>
</tr>
<tr>
<td>SPLC-[CN:RL7:07]</td>
<td>92.9/97.7</td>
<td>84.7/92.4</td>
<td>100/99.1</td>
<td>98.4/100</td>
<td>81.6/97.8</td>
<td>87.8/96.6</td>
<td>91.7/95.8</td>
<td>52.9/95.3</td>
</tr>
<tr>
<td>SPLC-[PR-MN4]</td>
<td>91.8/90.5</td>
<td>84.8/84.1</td>
<td>95.6/94.7</td>
<td>95.3/95.3</td>
<td>96.4/96.7</td>
<td>87.8/87.2</td>
<td>91.7/88.2</td>
<td>87.1/89.4</td>
</tr>
<tr>
<td>IYVV-[ES:Mal:IG3:06]</td>
<td>85.3/86.7</td>
<td>81.5/86.2</td>
<td>97.4/97.4</td>
<td>96.1/95.7</td>
<td>90.9/92.0</td>
<td>65.1/61.5</td>
<td>68.3/67.4</td>
<td>83.5/92.9</td>
</tr>
<tr>
<td>IYVV-[ES:Mal:IG5:06]</td>
<td>83.3/84.2</td>
<td>79.5/85.6</td>
<td>95.7/95.6</td>
<td>92.1/92.1</td>
<td>91.8/91.5</td>
<td>63.1/60.1</td>
<td>68.3/67.4</td>
<td>82.4/89.4</td>
</tr>
<tr>
<td>SPLCCaV-[ES:CI:BG21:02]</td>
<td>86.1/88.1</td>
<td>66.0/72.4</td>
<td>98.3/98.2</td>
<td>97.2/97.6</td>
<td>81.3/82.1</td>
<td>89.9/87.8</td>
<td>94.4/91.7</td>
<td>48.2/55.3</td>
</tr>
<tr>
<td>SPLCESV-[ES:CI:BG5:02]</td>
<td>76.1/75.2</td>
<td>42.3/41.9</td>
<td>85.0/84.1</td>
<td>91.3/90.6</td>
<td>84.9/84.9</td>
<td>61.7/60.1</td>
<td>67.6/66.7</td>
<td>55.3/56.5</td>
</tr>
<tr>
<td>SPLCGV-[US:Geo:16]</td>
<td>76.8/76.6</td>
<td>38.1/48.1</td>
<td>100/99.1</td>
<td>96.1/95.7</td>
<td>80.5/79.7</td>
<td>67.8/65.5</td>
<td>66.2/66.0</td>
<td>45.9/47.1</td>
</tr>
</tbody>
</table>

*Nucleotide sequence identity; †amino acid sequence identity; ‡sequence identities of SPLC-[JS1] when compared with other previously described *Ipomoea*-infecting begomoviruses; §sequence identities of SPLC-[Y338] when compared with other previously described *Ipomoea*-infecting begomoviruses.

highest nucleotide sequence identity (97%) with SPLCV from United States: SPLC-[US: Lou: 94] and SPLCV from China: SPLC-[CN: RL31: 07], respectively (Table 2). Highest sequence identity of both isolates with SPLCV suggest that JS1 and Y338 are two strains of SPLCV, and named SPLC-[JS1] and SPLC-[Y338]. SPLC-[JS1] and SPLC-[Y338] shared 92.9% sequence similarity with each other. Further sequence comparisons showed that DNA-A of SPLC-[JS1] and SPLC-[Y338] had more than 89.9% similarity with all strains of SPLCV and shared 75.2 to 88.1% identity with other previously described *Ipomoea*-infecting begomoviruses (Table 2). When individual encoded proteins were compared, AV1, AC2 and AC3 of SPLC-[JS1] shared highest sequence identity with that of SPLCV-[US: Lou: 94] at 98.8, 93.2 and 95.8%, respectively. The amino acid sequences of AC1 and AC4 were closely related to that of SPLCV-[KR: J-508:03] at 99.2 and 98.8%, respectively while AV2 of SPLC-[JS1] shared 100% amino acid sequence identities with that of SPLCV-[CN: RL: 07] and SPLCGV-[US: Geo: 16]. Comparisons of ORFs of SPLCV-[Y338] indicated that it shared the highest amino acid sequence identities with that of SPLCV-[CN: RL31:07] for AV1 (100%), AV2 (100%), AC1 (97.8%), AC2 (98.6%), AC3 (99.3%) and AC4 (97.6%) (Table 2).

**Intergenic region (IR) analysis**

The IRs of SPLCV-[Y338] and SPLCV-[JS1] contain four imperfect copies of the iterative elements, three direct and one inverted and their arrangement is similar to that of Old World begomoviruses (Table 3). Three directly repeated copies of iterated sequences (TGGAGAC) in IRs of SPLCV-[Y338] and SPLCV-[JS1] were upstream from the TATA box which was at 2692 to 2697 nucleotides. An inverted repeat consisting of GTCTCAATA was found as a 3’distal inverted copy placed right ward of three direct repeats. Iterative elements and corresponding iteron-related domains in the N-terminal regions of replication protein (Rep IRD) of *Ipomoea*-infecting begomoviruses is identified in Table 3 and shown to be characteristic for seven groups according to Argüello-Astorga et al. (1994).

**Phylogenetic analysis**

Relationship among SPLCV-[JS1], SPLCV-[Y338] and other begomoviruses were analyzed using full length DNA-A sequences. A phylogenetic tree inferred from the complete genome sequences of *Ipomoea*-infecting begomoviruses and selected other begomoviruses available in GenBank is
Table 3. Iterative elements and corresponding iteron-related domains in the N-terminal regions of replication associated protein (Rep IRD) of *Ipomoea*-infecting begomoviruses.

<table>
<thead>
<tr>
<th>Virus isolate</th>
<th>Iterative element</th>
<th>Rep IRD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPLCV(CN:Y338, F-p3, JS1)</td>
<td>ATTTGGAGAC AATTGGAGAC TGGAGAC TATATA</td>
<td>GTCTCCAAAT MAPP(K)RFKIQ</td>
</tr>
<tr>
<td>SPLCV(JP:Kum, Kyo)</td>
<td>ATTTGGtgac AATTGGTGAC TGGTGAC TATA</td>
<td>GTCACCAAAT MAPP(K)RFRIS</td>
</tr>
<tr>
<td>SPLCV(KR:J-508)</td>
<td>AATGGGTGGA AATTGGTGGA GGTGGA TATA</td>
<td>TCCACCTAAT MPRAGRFNIN</td>
</tr>
<tr>
<td>SPLCV(US:Lou:94)</td>
<td>AATCGGAGG AATTGGAGG ATTTGGAGG TATATATA</td>
<td>CCTCCAAAT MPRKGFRVQ</td>
</tr>
<tr>
<td>SPLCV(PR-MN4) IYVV(ES:IG3, IG5)</td>
<td>ATTTGGTGAC AATTGGTGAC TGGTGAC TATA</td>
<td>GTCACCAAAT MAPP(K)RFRIS</td>
</tr>
<tr>
<td>SPLCESV(BG5, IG2)</td>
<td>AATCGGAGG AATTGGAGG ATTTGGAGG TATATA</td>
<td>CCTCCAAAT MPRKGFRVQ</td>
</tr>
<tr>
<td>SPLCLaV(BG27, BG30)</td>
<td>AATCGGAGG AATTGGAGG ATTTGGAGG TATATA</td>
<td>CCTCCAAAT MPRKGFRVQ</td>
</tr>
<tr>
<td>SPLCCaV(BG21, BG25)</td>
<td>AATCGGAGG AATTGGAGG ATTTGGAGG TATATA</td>
<td>CCTCCAAAT MPRKGFRVQ</td>
</tr>
<tr>
<td>SPLCGV(US:Geo:16)</td>
<td>ATTTGGTGC ATTTGGTGTC TGGTGTC TATA</td>
<td>GACACCAAAT MPRQPGFRVS</td>
</tr>
</tbody>
</table>

shown in Figure 2. The phylogenetic tree showed that the whitefly transmitted geminiviruses were clustered into two groups. One group consisted of *Ipomoea*-infecting begomoviruses and the second group contained other Old World and New World begomoviruses from different hosts. Further analysis showed that SPLCV-[JS1] and SPLCV-[Y338] clustered together with strains of SPLCV to form a subgroup in the phylogenetic tree (Figure 2). Phylogenetic analyses showed that *Ipomoea*-infecting begomoviruses are more conserved than other Old World and New World begomoviruses. To further understand the sequence variation among different strains of *Ipomoea*-infecting begomoviruses, the full length sequences of 21 *Ipomoea*-infecting begomoviruses were analyzed by DnaSP version 4.10.3. Figure 3 shows that the variation in whole genome sequence was not uniform. The highest variation peak was located among AC4, N terminal of AC1 and IR, and the second variation peak was located among C-terminal of AV1, AC2 and AC3, while the position overlapped by AV2 and AV1 was more conserved (Figure 3).

**DISCUSSION**

To date, begomoviruses are classified on the basis of genome sequences; generally, isolates showing less than 89% sequence identity are considered to be distinct species, while those having more than 89% identity are considered to be strains of the same species (Fauquet et al., 2008). Comparison analysis of DNA-A of SPLCV-[JS1] and SPLCV-[Y338] with other reported begomoviruses showed that they shared high sequence identity (89.9 to 97.8%) with SPLCV, suggesting that SPLCV-[JS1] and SPLCV-[Y338] are strains of SPLCV. The IR is the part of the genome that shows the greatest variation among different begomoviruses. Apart from conserved nonanucleotide sequence and the TATA boxes, IRs of different viruses have iteron elements which vary in length, sequence number and orientation. The IR of SPLCV-[JS1] and SPLCV-[Y338] shared identical iterative sequences and identical modular organization of the iterons with SPLCV-[US: Lou: 94], SPLCV-[JP:Kum, Kyo] and SPLCV-[KR:J-508], which shows that they may have common ancestor.

Begomoviruses infect *Ipomoea* species worldwide and have been reported in Asia, America, Africa and Europe. Naturally, sweet potato begomoviruses have only been isolated from *Convolvulaceae* family plants which have not been reported as host of other geminiviruses. Therefore, it remains to uncover, where and how sweet potato geminiviruses had a chance to exchange their sequences with other geminiviruses. It seems that it happened in unknown plant species. Global occurrence of SPLCV indicates that sweet potato geminiviruses have already been spread by vegetative propagation and export of their roots all over the world. Further characterization of more isolates from different geographical regions may lead to a better understanding of the origin, variability and relationships among different sweet potato
Figure 2. Phylogenetic tree obtained using the neighbor-joining method with 1000 bootstrap replications available in the MEGA4.0 based on alignments of the complete nucleotide sequences of the DNA-A of SPLCV-[JS1], SPLCV-[Y338] and other reported begomoviruses.

Figure 3. Sliding window plot showing the distribution of genetic variation estimated by nucleotide diversity (Pi) for Ipomoea-infecting begomoviruses. A window size of 100 and a step of 25 nucleotides were used. The relative positions of the ORFs of viral DNA genome are illustrated above the plot in linear DNA format.
geminiviruses.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant No. 31070128), the National Excellent Doctoral Dissertation of China (No. 200768) and the Fundamental Research Funds for the Central Universities. Awais Zulfiqar acknowledges the Higher Education Commission (HEC) of Pakistan for awarding him with the PhD fellowship.

REFERENCES