Complete Genome Sequencing and Infectious cDNA Clone Construction of Soybean Mosaic Virus Isolated from Shanxi

Defu Wang1†, Liyan Cui2†, Li Zhang2, Zhennan Ma1, and Yanbing Niu1*  
1College of Life Sciences, Shanxi Agricultural University, Taigu 030801, China  
2College of Grassland Science, Shanxi Agricultural University, Taigu 030801, China  
(Received on November 10, 2020; Revised on February 28, 2021; Accepted on March 1, 2021)

Soybean mosaic virus (SMV) is the predominant viral pathogen that affects the yield and quality of soybean. The natural host range for SMV is very narrow, and generally limited to Leguminosae. However, we found that SMV can naturally infect Pinellia ternata and Atractylodes macrocephala. In order to clarify the molecular mechanisms underlying the cross-family infection of SMV, we used double-stranded RNA extraction, rapid amplification of cDNA ends polymerase chain reaction and Gibson assembly techniques to carry out SMV full-length genome amplification from susceptible soybeans and constructed an infectious cDNA clone for SMV. The genome of the SMV Shanxi isolate (SMV-SX) consists of 9,587 nt and encodes a polyprotein consisting of 3,067 aa. SMV-SX and SMV-XFQ008 had the highest nucleotide and amino acid sequence identities of 97.03% and 98.50%, respectively. A phylogenetic tree indicated that SMV-SX and SMV-XFQ018 were clustered together, sharing the closest relationship. We then constructed a pSMV-SX infectious cDNA clone by Gibson assembly technology and used this clone to inoculate soybean and Ailanthus altissima; the symptoms of these hosts were similar to those caused by the virus isolated from natural infected plant tissue. This method of construction not only makes up for the time-consuming and laborious defect of traditional methods used to construct infectious cDNA clones, but also avoids the toxicity of the Potyvirus special sequence to Escherichia coli, thus providing a useful cloning strategy for the construction of infectious cDNA clones for other viruses and laying down a foundation for the further investigation of SMV cross-family infection mechanisms.

Keywords: genome sequencing, Gibson assembly-infectious clone, soybean mosaic virus

Handling Editor: Ho-Jong Ju

Soybean (Glycine max (L.) Merr) is an annual herbaceous plant of the Leguminosae soybean family and represents the world’s main oil crop. Soybean is also rich in soluble fibers, saponins, proteins, fats, and other nutrients that are beneficial to human growth and development. However, viral diseases have become an increasingly serious threat to soybean, causing enormous losses in both yield and quality. According to previous researches, more than 110 plant viruses in the United States are known to infect soybeans; in China over 50 viruses have been reported (Cui et al., 2018; Zhi, 2005). Of those, SMV is the predominant virus known to infect soybean. The range of natural hosts for SMV is particularly narrow and is generally limited to Leguminosae (Zheng et al., 2000; Zhou and Pu, 1990). However, over recent years, researchers have found that SMV can naturally infect Pinellia ternata (Rohozková and Navrátil, 2011; Shentu, 2006; Shi, 2007; Sun et al., 2008). In a previous report, we described the isolation of SMV from susceptible P. ternata (SMV-SXBX), and phylogenetic tree showed that SMV-SXBX and SMV Pinellia isolates were clustered together, but

†These authors contributed equally to this work.
*Corresponding author.  
Phone, FAX) +86-0354-6287205  
E-mail) niuyanbingbest@163.com  
ORCID  
Yanbing Niu  
https://orcid.org/0000-0001-9618-1886  
© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Articles can be freely viewed online at www.ppjonline.org.
far from SMV soybean isolates (Cui et al., 2018). In addition, we have also found that SMV can naturally infect Atractyloides macrocephala (Compositae) in China (data not published). However, the molecular mechanism of SMV cross-family infection is still unknown. As a technology, infectious cloning has been widely used in the study of pathogenicity of plant viruses. At present, we have successfully constructed the infectious cDNA clone of SMV-SXBX (Zhang et al., 2020), but it is not enough to clarify the molecular mechanisms underlying the cross-family infection of SMV. Therefore, this study is dedicated to construction of the infectious cDNA clone of SMV soybean isolate (SMV-SX).

Infectious clones for plant viruses refer to DNA clones that have been constructed artificially using molecular biology techniques and have the same infectious activity as plant viruses under natural conditions. The essence of these infectious clones is that they represent a DNA gene bank that contains the complete genome information of the virus (Boyer and Haenni, 1994). Normally, full-length cDNA clones for plant viruses are constructed using reverse genetics technology and subsequently used to explore RNA viruses at the DNA level, thus overcoming operational difficulties at the RNA level. The first successfully constructed infectious clone of an RNA virus was brome mosaic virus (BMV) (Ahlquist et al., 1984). Since then an increasing number of RNA viruses have been used to construct infectious clones, including cucumber mosaic virus (CMV) (Rizzo and Palukaitis, 1990), tomato aspermy cucumovirus (TAV) (Shi et al., 1997), beet mild yellowing virus (BMVY) (Stephan and Maiss, 2006), tomato yellow leaf curl sardinia virus (TYLCSV) (Kondo and Fujita, 2012), ribgrass mosaic virus (RMV) (Zhu et al., 2006), potato virus M (PVM) (Flatken et al., 2008), and pepper mottle virus (PMV) (Lee et al., 2011). However, the construction of infectious clones for Potyviruses is particularly challenging due to the presence of special viral sequences to Escherichia coli. We expect our findings to provide a reference for the construction of infectious clones for other viruses, and lay a foundation for further research relating to mechanisms of virus infection.

Materials and Methods

Materials. Fresh leaves from soybean plants that showed mosaic symptoms were collected from Jinzhong in Shanxi province in China in July 2018 and stored in liquid nitrogen for subsequent testing.

Extraction of double-stranded RNA and amplification of the SMV-SX genome. Total double-stranded RNA (dsRNA) of symptomatic soybean was extracted according to the method described previously by Krajacić et al. (2007), Li et al. (2007), and Tzanetakis and Martin (2008). Next, we generated first-strand cDNAs from dsRNA via reverse transcription polymerase chain reaction (RT-PCR) with moloney murine leukemia virus reverse transcriptase (Promega, Madison, WI, USA) and an oligo(dT)₁₈ primer at 42°C for 1 h and 70°C for 15 min. The complete viral genome was obtained through several RT-PCR amplifications and covering all gaps with virus-specific primers (Supplementary Table 1). The PCR conditions were set as described previously (Niu et al., 2018). We then performed rapid amplification of cDNA ends (RACE) PCR using a RACE kit (TaKaRa, Dalian, China) in accordance with the manufacturer’s recommendations. The PCR products were purified using an agarose gel DNA Gel Extraction kit (OMEGA, Guangzhou, China) in accordance with the manufacturer’s recommendations. The purified fragment was cloned directly into the pEASY-Blunt vector (TransGen, Beijing, China) and transformed into E. coli strain DH5α for Sanger sequencing.
Bioinformatics analysis of SMV-SX. The amplified sequences were spliced using SeqMan tool from the DNASTAR package (DNASTAR, Inc., Madison, WI, USA). Multiple sequence alignment was carried out for nucleotide and amino acid sequence using the optimal alignment tool in the DNAMAN software package (Lynnon, Biosoft, Canada). Phylogenetic tree performed by a maximum likelihood method with MEGA 7.1 software and ORF finder was used to predict the amino acid sequence encoded by nucleotides.

Construction of pSMV-SX via Gibson assembly. The full-length SMV-SX cDNA was divided into two overlapping amplified fragments using specific primers and ensuring that there was a 25-base-pair overlap (Table 1). A 35 S promoter and a CaMV polyA terminator were inserted into the multiple cloning sites of the binary pGreenII 0000 expression vector to generate pGreen-35S (Hellens et al., 2000). The pGreen-35S was amplified by primers with 25-33 bases that were homologous to the viral genome fragments. The PCR amplification reactions were performed with PrimeSTAR GXL DNA Polymerase (Takara). Amplicons of the expected sizes were purified with a DNA Gel Extraction kit (OMEGA). The viral genome fragments were then cloned into the linearized pGreen-35S vector by Gibson assembly (Tuo et al., 2017). The reaction was performed in a total volume of 10 µl, containing 150 ng of each purified PCR fragment and 5 µl of 2× Gibson Mix (NEB). The reaction mix was incubated at 50°C for 1 h, and then placed on ice prior to transformation into the Rhizobium radiobacter (synonym Agrobacterium tumefaciens).

Transformation of Agrobacterium tumefaciens and transfection. A reaction mix (5 µl) was transformed into 100 µl of Agrobacterium tumefaciens strain GV3101 (Weidi Biotech, Shanghai, China) competent cells (containing the pSoup helper plasmid) by electroporation according to a standard protocol. The transformants were then screened on Luria-Bertani (LB) medium containing 50 mg/l of kanamycin and 50 mg/l of rifampicin for 3 days at 28°C, and then screened by colony PCR using the primers shown in Supplementary Table 1. The plasmids (final 30 µl) were extracted from 80 ml culture of the positive colonies for sequencing.

Agroinfiltration was performed as previously described (Sparkes et al., 2006; Yan et al., 2012). A positive single colony of agrobacterium was picked from each infectious clone in order to inoculate 10 ml of LB medium supplemented with 50 mg/l of kanamycin and 50 mg/l of rifampicin. Bacteria were grown overnight at 28°C until they reached an OD600 of 1.0-1.5. The cultures were then pelleted and washed using infiltration buffer (10 mM MES pH 5.6, 10 mM MgCl2, and 100 µM a cetosyringone). Then, cells were diluted with infiltration buffer to a final OD600 of 0.6-0.8 and incubated for 3 h at room temperature in the dark prior to the agroinfiltration of soybean and A. altissima plants using 1 ml- syringes without a needle.

The plasmid inoculation approach was performed as previously described (Kang et al., 2016; Seo et al., 2009). Plasmid DNA of pSMV-SX was prepared using the Plasmid Maxi Kit (Qiagen, Shanghai, China). Approximately 50 µg plasmid DNA was diluted in 50 mM potassium phosphate, pH 7.5, to a total volume of 100 µl and inoculated onto two leaves per plant by direct rub-

Table 1. Primer pairs for creating SMV-SX infectious cDNA clones

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence (5′-3′)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGr35S-smvF</td>
<td>CCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</td>
<td>For vector of SMV</td>
</tr>
<tr>
<td>pGr35S-R</td>
<td>CTCCTCCAAATGAAATGAACTTCCT</td>
<td>For SMV fragment 1</td>
</tr>
<tr>
<td>SMV-F1</td>
<td>AGGAAGTTCATTTTCATTTGGAGAGG</td>
<td>For SMV fragment 2</td>
</tr>
<tr>
<td>SMV-R3716</td>
<td>AGAGACGCTGAAGACTCCCTTTATTTT</td>
<td>For RT-PCR detection</td>
</tr>
<tr>
<td>SMV-F3697</td>
<td>AGAGACGCTGAAGACTCCCTTTATTTT</td>
<td>For RT-PCR detection</td>
</tr>
<tr>
<td>SMV-R1</td>
<td>AGGGACAACACATCGCCGTA</td>
<td>For RT-PCR detection</td>
</tr>
<tr>
<td>SMV-CP-F</td>
<td>TTACAATCCAGCAAGGAGAAGGAA</td>
<td>For RT-PCR detection</td>
</tr>
<tr>
<td>SMV-CP-R</td>
<td>GGTCAAAACCCACACTAAACAAAAG</td>
<td>For RT-PCR detection</td>
</tr>
</tbody>
</table>

The sequence marked with a horizontal line represents a homologous sequence.

SMV, soybean mosaic virus; SMV-SX, SMV Shanxi isolate; RT-PCR, reverse transcription polymerase chain reaction.
inoculation with carborundum. All the inoculated plants were grown in a controlled environment with a 16 h light cycle at 28°C and then for 8 h in the dark at 25°C. After 45 days inoculation, the symptoms of inoculated plant leaves were photographed and further detected by RT-PCR technique.

**RT-PCR.** Total RNA was extracted from the upper non-inoculated leaves of the inoculated plants for RT-PCR using Tap DNA polymerase (TransGen) and a specific set of primers: SMV-CP-F and SMV-CP-R (Table 1). The PCR products were then transformed into *E. coli* DH5α for sequencing.

**Results**

**Structure characteristics of the SMV-SX genome.** First, dsRNA was extracted from diseased soybean leaves and reverse transcribed into cDNA to act as a template. We then used 11 pairs of specific primers and 2 pairs of RACE primers to amplify the entire SMV-SX genome. A series of fragments were produced, all of the expected sizes 820, 1,121, 776, 993, 1,142, 1,037, 953, 1,027, 1,904, and 1,136 bp (Fig. 1). The SMV-SX genome consists of 9,587 nt and was determined by assembling the 13 amplified fragments using DNASTAR software.

The genomic organization of SMV-SX is typical for members of the *Potyvirus* genus and is consistent with that of other SMV sequences, with a large open reading frame (ORF) encoding a putative polyprotein of 3,067 amino acids. The ORF begins at nt position 133 and ends with a stop codon at position 9332-9335. The 5′ and 3′ untranslated regions consist of 132 and 252 nt, respectively. The polyprotein has nine putative cleavage sites and can be cleaved into ten mature proteins (P1, HC-Pro, P3, 6K1, CI, 6K2, VPg, NIa, NIb, and CP). The cleavage site in the P1/
HC-Pro dipeptide is predicted to be cleaved by the P1 serine proteinase, whereas the cleavage site in the HC-Pro/P3 is cleaved by the HC-Pro cysteine proteinase. In addition, the cleavage sites of the seven other dipeptides are cleaved by the NIa-Pro cysteine proteinase. The predicted cleavage sites are as follows: IQHY309/S (P1/HC-Pro), YRVG766/G (HC-Pro/P3), VSAQ1113/A (P3/6K1), VKVQ1165/S (6K1/CI), VQLQ1799/S (CI/6K2), VSTQ1852/G (6K2/CI), and MSRR1856/S (CI/6K3).
SMV Infectious cDNA Clone Construction

NIa-VPg), VEME2042/S (NIa-VPg/NIaPro), VTVOQ 2285/G (NIa-Pro/NIb), and VSLQ2802/S (NIb/CP) (Fig. 2).

SMV-SX sequence similarity. The homology and similarity of SMV-SX with other SMV isolates were analyzed using DNAMAN software and showed that the nucleotide and amino acid identities were 75.19-97.03% and 81.70-98.50%, respectively (Table 2). SMV-SX had the highest identity to SMV-XFQ008 from China. The identities of the nucleotide and amino acid sequences were 97.03% and 98.50%, respectively. Phylogenetic analysis also showed that the SMV-SX isolate clustered with SMV-XFQ008 (Fig. 3).

Construction of a pSMV-SX plasmid by Gibson assembly. Using the 3′-RACE reverse transcription product of the soybean dsRNA as a template, we used two large pairs of specific primers (SMV-F1/SMV-R3716 and SMV-F3697/SMV-R1) (Table 1) to amplify fragments of 3,738 bp and 5,894 bp, respectively. Using the pGreen-35S plasmid as a template, and using specific primers (pGr35S-smvF/pGr35S-R), we then obtained the target band of 3,233 bp by reverse PCR (Fig. 4A-C). Gibson assembly technology was then used to ligate the two SMV fragments into the

Fig. 3. Phylogenetic tree based on the nucleotide sequences of SMV-SX and other SMV isolates. Phylogenetic tree was constructed by the maximum likelihood method with 1,000 bootstrap replicates. SMV, soybean mosaic virus; SMV-SX, SMV Shanxi isolate.
linearized pGreen-35S vector. The recombination vector was subsequently electroporated into Agrobacterium GV3101 (pSoup) competent cells. Positive clones were then identified by colony PCR using specific primers SMV-F6 and SMV-R6 (Fig. 4D, Supplementary Table 1) and plasmid extraction (Fig. 4E).

**Activity analysis of the infectious pSMV-SX clone.** Agrobacterium infiltration and plasmid inoculation were used to inoculate soybean varieties (Kefeng No. 1 and 1138-2 at the 3-4 leaf stage) and *A. altissima*. After 45 days, the inoculated plants all showed symptoms of susceptibility (Fig. 5D-F). We then extracted RNA from the symptomatic leaves of all plants and used this as a template for RT-PCR using SMV-CP-F and SMV-CP-R primers, the electrophoresis and sequencing results showed that the constructed SMV-SX infectious clone had infectious activity (Fig. 5G-I).

**Discussion**

Plant viruses have caused a range of serious diseases in a wide variety of plants, thus causing global economic losses of up to 30 billion dollars annually (Sastry and Zitter, 2014). More than 90% of viral nucleic acids in plants are single-stranded RNA, and dsRNA is produced when replicated (Zhou and Li, 1995). The principle that dsRNA can be specifically adsorbed by cellulose CF-11 at a certain concentration of alcohol can be used to identify plant viruses (Tian and Wang, 1995). dsRNA technology overcomes the limitations associated with the use of single-stranded RNA in traditional RT-PCR, is stable during extraction, and is not affected by polyphenols and polysaccharides. This technology is not only simple and convenient to operate, but can also detect known, unknown, and low concentration viruses. In addition, the combination of dsRNA and sequence-independent amplification technology can achieve the purpose of detecting multiple viruses simultaneously (Kou et al., 2017). Niu et al. (2003), Wang et al. (2016), and Yang et al. (2017), all used this technology to detect pear vein yellow virus (PVYV), youcai mosaic virus (YoMV), and malva vein clearing virus (MVCV).

Techniques to clone infectious plant viruses have been applied to many aspects of plant virus research. First, those techniques have been used to investigate the molecular biology of viruses; Galiakparov et al. (2003) used infectious cloning techniques to investigate the proteins encoded by the five ORFs of grapevine virus A (GVA). Second, those techniques have been used to investigate viral pathogenic factors; Chiang et al. (2007) used infectious cloning technology to show that three amino acids in the P1 and HC-Pro regions of PRSV are closely related to the pathogenicity of the virus. Third, this technology has been used to detect the five ORFs of grapevine virus A (GVA). Second, those techniques have been used to investigate viral pathogenic factors; Chiang et al. (2007) used infectious cloning technology to show that three amino acids in the P1 and HC-Pro regions of PRSV are closely related to the pathogenicity of the virus. Third, this technology has been used to construct plant expression vectors (Chapman et al., 1992) and gene silencing vectors (Liu and Lomonossoff, 2002). Fourth, these techniques have been used to develop attenuated strains. By using reverse genetics to make site-
directed mutations of infectious clones, it was possible to obtain attenuated strains, thus achieving cross-protection (Chiang et al., 2007).

However, studies have shown that there are special sequences in the Potyvirus that can cause varying degrees of toxicity to E. coli (Chikh Ali et al., 2011; Johansen and Lund, 2008). However, the specific mechanism underlying these effects has yet to be identified. This makes the construction of infectious clones of SMV more difficult. In the current study, Gibson assembly technology was used to rapidly construct an infectious SMV clone shuttle plasmid. This construction overcomes the time-consuming and laborious nature of the traditional method used to construct infectious clones, and also avoided the toxicity of SMV infectious clones to E. coli. To our knowledge, this study represents the first construction of a full-length infectious cDNA of SMV that is capable of infecting soybean and Altissima by using Gibson assembly technology. The agroinfectious clone described herein can therefore be useful for further studies of plant-virus interaction and the characterization of viral genes. Our findings also lay down the foundation for studying the mechanisms associated with cross-family SMV infection.

SMV only has a very narrow range of hosts, and generally only infects Leguminosae. Over recent years, our laboratory has found that SMV can infect P. ternata

**Fig. 5.** Plants inoculated with the infectious SMV-SX clone all showed symptoms and reverse transcription polymerase chain reaction detection. (A) Normal Ailanthus altissima plant. (B) Normal soybean 1138-2 plant. (C) Normal soybean Kefeng No. 1 plant. (D) Susceptible A. altissima plant showing mosaic symptoms. (E) Susceptible soybean 1138-2 plant showing mosaic symptoms. (F) Susceptible soybean Kefeng No. 1 showing mosaic symptoms. (G) Amplification results from the SMV-CP primer in A. altissima. (H) Amplification results from the SMV-CP primer in soybean 1138-2. (I) Amplification results from SMV-CP primer in soybean of Kefeng No. 1. SMV, soybean mosaic virus; SMV-SX, SMV Shanxi isolate. M, DL 2000 DNA marker; lanes 1-3, inoculated plants; CK, non-inoculated plant.
(Araceae) and A. atractylodes (Compositae), although the mechanism underlying cross-family infection remains unclear. Sequence comparison revealed that there was a significant difference between the P1 protein of a SMV–P. ternata isolate and a SMV soybean isolate (data not published). Other researches have suggested that this non-conservative protein plays a key role in the successful adaptation of Potyviruses into a wide range of host species (Rohožkoá and Navrátil, 2011; Valli et al., 2007). At present, our laboratory is conducting further research on P1 protein to investigate its impact on cross-family SMV infection. Based on our previously constructed infectious clone of SMV soybean isolate, the molecular mechanism of SMV cross-family infection was to clarify by exchanging the P1 protein and other differential proteins.

Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

This study was supported by National Natural Science Foundation of China (No. 31772130), China Agriculture Research System (Grant No. CARS-21) and Modern Agro-industry Technology Research System (No. 2020-05). Authors also thank to Dr Tuo Decai and Dr Niu Jingping for providing the vector pGreen-35S and soybean seeds. The authors would like to express their gratitude to EditSprings (https://www.edit springs.com/) for the expert linguistic services provided.

Electronic Supplementary Material

Supplementary materials are available at The Plant Pathology Journal website (http://www.cppjonline.org/).

References


88:1016-1028.