Korean Journal of Microbiology (2022) Vol. 58, No. 2, pp. 91-95
DOI https://doi.org/10.7845/kjm.2022.2020

pISSN 0440-2413
eISSN 2383-9902

Complete genome sequence of soybean dwarf virus infecting soybean (*Glycine max* L.)

Yeonhwa Jo[1†], Hoseong Choi[2†], Bong Choon Lee[3], Jin-Sung Hong[4], and Won Kyong Cho[1*]

[1]*College of Biotechnology and Bioengineering, Sungkyunkwan University, Suwon 16419, Republic of Korea*
[2]*Plant Genomics and Breeding Institute, Seoul National University, Seoul 08826, Republic of Korea*
[3]*Crop Foundation Division, National Institute of Crop Science, Rural Development Administration, Wanju 55365, Republic of Korea*
[4]*Department of Applied Biology, Kangwon National University, Chuncheon 24341, Republic of Korea*

콩에서 동정한 soybean dwarf virus의 전체 유전체 서열

조연화[1†] · 최호성[2†] · 이봉춘[3] · 홍진성[4] · 조원경[1*]

[1]성균관대학교 생명공학대학, [2]서울대학교 식물유전체육종연구소, [3]국립식량과학원 작물기초기반과, [4]강원대학교 응용생물학전공

(Received March 21, 2022; Revised May 16, 2022; Accepted June 11, 2022)

We found a soybean plant displaying severe viral disease symptoms, such as stunting, rugosity, and yellowing. To reveal the viral agent, we extracted total RNA from the leaf sample and prepared a ribosome-deleted library followed by RNA-sequencing. Raw data were *de novo* assembled, and the assembled contigs were used in a BLASTX search against a viral protein database. We identified four contigs associated with *Soybean dwarf virus* (SbDV) belonging to the genus *Luteovirus* in the family *Luteoviridae*. We obtained the complete genome of SbDV isolate HS with 5,686 nucleotides (nt), encoding five open reading frames (ORFs). The phylogenetic analysis showed that SbDV isolate HS belongs to Group A containing isolates from the USA and the Czech Republic. Here, we report the complete genome sequence of SbDV isolate HS infecting soybean in Korea obtained by RNA-sequencing.

Keywords: genome, soybean dwarf virus, soybean, virus,

Soybean dwarf virus (SbDV), a pathogenic plant RNA virus, is a member of the genus *Luteovirus* in the family *Luteoviridae* (Terauchi *et al.*, 2001). SbDV was first identified as a severe viral pathogen infecting soybean plants in northern Japan in 1969 (Tamada *et al.*, 1969). Since then, SbDV has been identified in many countries in the world. The main hosts for SbDV are members of the family Fabaceae, including diverse beans, but SbDV also infects plants belonging to the families Chenopodiaceae and Polemoniaceae (Damsteegt *et al.*, 1990). The disease symptoms in the soybean leaves infected by SbDV include severe chlorosis, rugosity, and thickening of leaves (Tian *et al.*, 2017). SbDV is usually transmitted by diverse aphids, such as *Aulacorthum solani* Kaltenbach, *Acyrthosiphon pisum* Harris, and *Nearctaphis bakeri* Cowen (Tian *et al.*, 2017).

Recently, viruses infecting soybean in the Republic of Korea have been intensively examined by RNA-sequencing (Jo *et al.*, 2020). The three major viruses infecting soybean have been found to be *Soybean mosaic virus* (SMV), a member of the genus *Potyvirus* in the family *Potyviridae*; *Soybean yellow mottle mosaic virus* (SYMMV), a member of the genus *Gammacarmovirus* in the family *Tombusviridae*; and *Soybean yellow common mosaic virus* (SYCMV), a member of the genus *Gammacarmovirus* in the family *Tombusviridae*. In addition, several RNA viruses have been identified infecting soybean in Korea, such as peanut stunt virus, peanut mottle virus, tomato spotted wilt virus, bean common mosaic virus, and bean common mosaic necrosis virus. Most soybean plants

[†]These authors contributed equally to this work.
***For correspondence.** E-mail: wonkyong@gmail.com;
Tel.: +82-31-290-7860; Fax: +82-31-290-7870

showing viral disease symptoms were co-infected by different viruses (Jo *et al.*, 2020). Of viruses infecting soybean, SbDV was first identified from soybean in 2003 in Korea (Kim *et al.*, 2006); however, SbDV was not frequently detected as compared to other known viruses infecting soybean. For example, only one study characterizing the genome of SbDV infecting *Vigna angularis* has been reported to date. Fortunately, it is likely that SbDV is not yet widespread in Korea. Thus, it is very important to prevent the spread of SbDV in advance.

In July 2020, we found a soybean plant displaying severe viral disease symptoms, such as stunting, rugosity, and yellowing. To reveal the viral agent, we extracted total RNA from the leaf sample showing viral disease symptoms using a RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions. The ribosomal RNA was deleted from the total RNA using TruSeq Stranded Total RNA with a Ribo-Zero Plant Kit (Illumina). After that, we prepared a library for RNA-sequencing using the TruSeq Stranded Total RNA LT Sample Prep Kit (Illumina). The generated library was paired-end (2 × 101 bp) sequenced using the NovaSeq 6000 system (Macrogen). Raw data were *de novo* assembled using the Trinity program with default parameters (Haas *et al.*, 2013), and the assembled contigs were used for a BLASTX search with E-value 1e-10 as a cutoff against the viral protein database. We conducted open reading frame (ORF) prediction for the identified virus-associated contigs using the ORFfinder program (https://www.ncbi.nlm.nih.gov/orffinder/). For phylogenetic tree construction, available complete viral genome sequences were aligned by MAFFT version 7 (Katoh and Standley, 2013) and then trimmed using the trimAl program (Capella-Gutiérrez *et al.*, 2009). Aligned sequences were subjected to the MEGA 7 program for the construction of the phylogenetic tree using the maximum likelihood method with 500 bootstrap replicates (Kumar *et al.*, 2016).

Using the RNA-sequencing result, we identified four viral contigs (1,809 reads) ranging from 1,168 bp to 5,686 bp associated with SbDV. We obtained the complete genome of the SbDV isolate, referred to as HS. The genome of SbDV consisted of positive-sense single-stranded RNA, and SbDV isolate HS was 5,686 nucleotides (nt) in length (Fig. 1A). SbDV isolate HS encoded five ORFs and three untranslated regions (Table 1). ORF1 (position 119 to 1,204) encoded a replicase protein. ORF2 (position 119 to 2,802) was translated from ORF1 by frameshifting (position 1,197) and encoded a replicase protein. ORF3 (position 3,010 to 3,612) encoded a major coat protein, while ORF4 (position 3,026 to 3,595) encoded a movement protein. ORF5 (position 3,010 to 5,082) was a readthrough protein translated by the in-frame translational readthrough from the ORF3 stop codon (position 3,610 to 3,612) and encoded a minor coat protein.

The BLASTN result, using the complete genome sequence of SbDV isolate HS, showed that SbDV isolate HS was closely related to the SbDV isolate SDV-HZ3 (MG600300) identified from red clover in the Czech Republic with 99% sequence coverage and 98.40% nucleotide identity. We generated a phylogenetic tree using 34 available complete genome sequences of SbDV, including HS isolate from this study (Fig. 1B). The phylogenetic tree showed two distinct groups of 34 SbDV isolates. Group A contained 17 SbDV isolates, including the HS isolate. The SbDV isolates in Group A were derived from the Czech Republic (SDV-HZ3 and SDV-HZ1), the USA (Wisc3,

Table 1. Summary of soybean dwarf virus (SbDV) genome organization

Features	Location	Size	Function
5' UTR	1–118	118 nt	UTR
ORF1	119–1,204	361 aa	Replicase protein
ORF2	119–2,802 (Frameshift 1,197)	894 aa	RNA-dependent RNA polymerase P1-P2 fusion protein
UTR	2,803–3,009	207 nt	UTR
ORF3	3,010–3,612	200 aa	Coat protein
ORF4	3,026–3,595	189 aa	Movement protein
ORF5	3,010–5,082 (Translation exception 3,610–3,612)	689 aa	Minor coat protein
3' UTR	5,083–5,682	600 nt	UTR

Open reading frames (ORFs) and untranslated regions (UTRs) are indicated with respective location and size. nt, nucleotides; aa, amino acids.

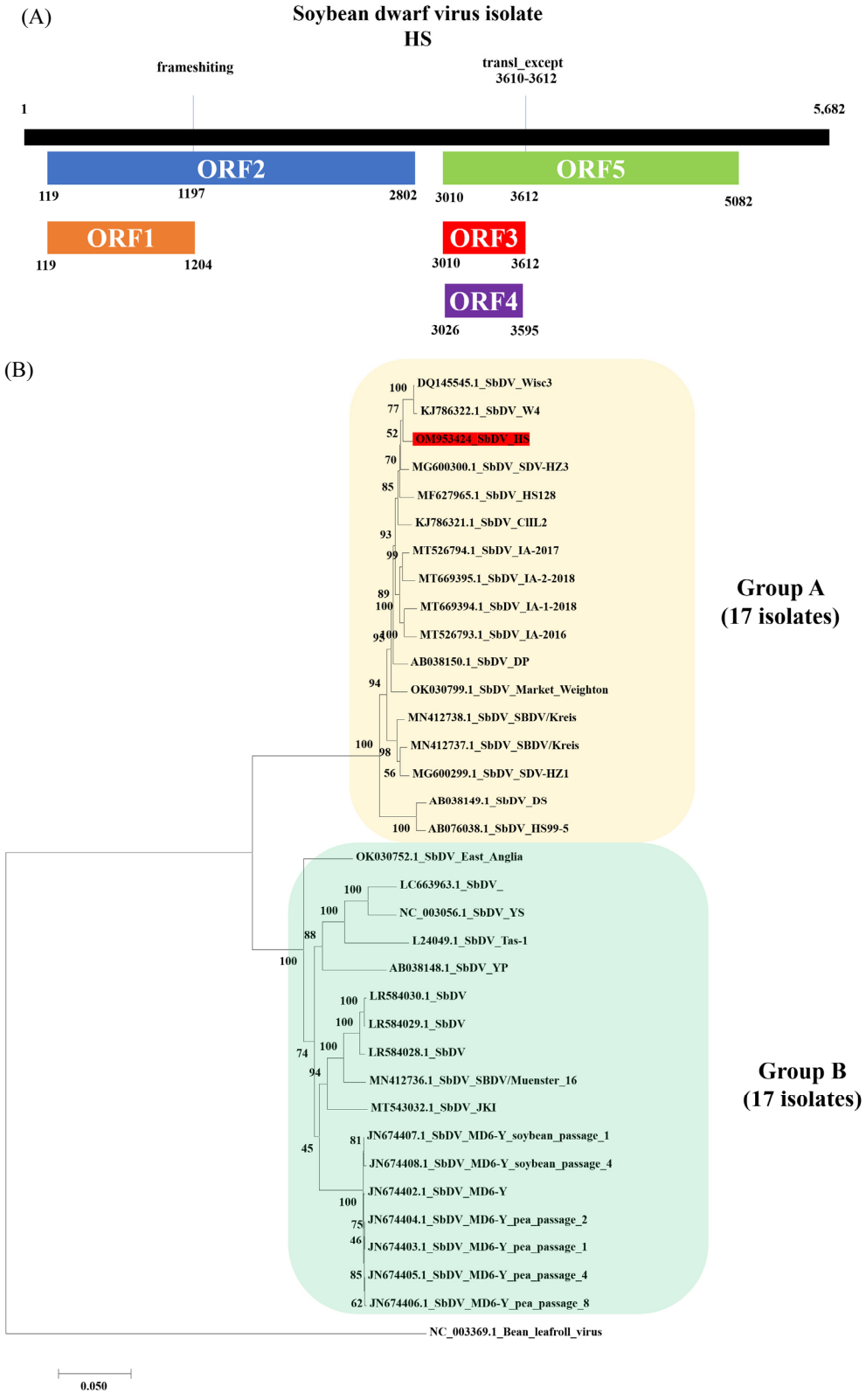

Fig. 1. Genomic organization and phylogenetic relationship of SbDV isolate HS with known SbDV isolates. (A) Genomic organization of SbDV isolate HS. The numbers indicate the nucleotide positions of individual ORFs in the SbDV genome. (B) Phylogenetic relationship of 34 SbDV isolates based on complete genome sequences. The phylogenetic tree was constructed using the MEGA 7 program with the maximum likelihood method and 500 bootstrap replicates.

W4, ClIL2, IA-2017, IA-2-2018, IA-1-2018, and IA-2016), Japan (DP, DS, and HS99-5), the UK (Market Weighton), Germany (SBDV/Kreis Stormarn_18 and SBDV/Kreis Stormarn_16), and Korea (HS128 and HS). Most isolates in Group A were identified from soybean, and some isolates were derived from *Trifolium pratense* L. (SDV-HZ3 and SDV-HZ1), *Pisum sativum* (Market Weighton, SBDV/Kreis Stormarn_18, and SBDV/Kreis Stormarn_16), and *Vigna angularis* (HS128). Most isolates in Group A were identified from *Pisum sativum*. In contrast, Group B contained SbDV isolates from Japan (YS and YP), the UK (East Anglia SbD), Germany (JKI ID 23556 and SBDV/Muenster_16), Australia (LR584028, LR584029, and LR584030), and the USA (MD6-Y).

Taken together, we report the complete genome of SbDV isolate HS identified from soybean in Korea by RNA-sequencing.

Nucleotide sequence accession number

The complete genome sequence of soybean dwarf virus isolate HS has been deposited in GenBank under the accession number OM953424.

적 요

왜화, 융기, 황색 등 바이러스성 질환 증상이 심한 콩 식물을 발견했다. 바이러스병을 일으키는 병원체를 밝히기 위해, 감염된 잎 샘플에서 총 RNA를 추출하고, 리보솜 삭제 라이브러리를 준비한 후 RNA 염기서열 분석을 실시하였다. Raw data를 *de novo* 조립하였으며, 조립된 contig들을 이용하여 바이러스 단백질 데이터베이스에 매치하여 BLASTX 검색을 수행하였다. *Luteoviridae*과의 *Luteovirus*속의 콩위축바이러스(*Soybean dwarf virus*, SbDV)와 관련된 4개의 contig들을 동정하였다. 5개의 ORF (Open Reading Frame)를 코드하는 5,686개의 뉴클레오티드를 가진 SbDV 분리주 HS의 완전한 게놈을 얻었다. 계통학 분석 결과, SbDV 분리 HS는 미국과 체코로부터 분리되는 A 그룹에 속하는 것으로 나타났다. 본 연구에서는 RNA-sequencing을 이용해 한국의 콩에 감염되어 있는 SbDV 분리 HS의 전체 유전체 서열을 보고한다.

Acknowledgments

This work was carried out with the support of the "Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ01498301)" conducted by the Rural Development Administration, Republic of Korea and Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through Crop Viruses and Pests Response Industry Technology Development Program, funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA) (320038-3).

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Capella-Gutiérrez S, Silla-Martínez JM, and Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973.

Damsteegt V, Hewings A, and Sindermann A. 1990. Soybean dwarf virus: experimental host range, soybean germ plasm reactions, and assessment of potential threat to U.S. soybean production. *Plant Dis.* **74**, 992–995.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, *et al.* 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512.

Jo Y, Yoon YN, Jang YW, Choi H, Lee YH, Kim SM, Choi SY, Lee BC, and Cho WK. 2020. Soybean viromes in the Republic of Korea revealed by RT-PCR and next-generation sequencing. *Microorganisms* **8**, 1777.

Katoh K and Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780.

Kim SM, Lee JB, Lee YH, Choi SH, Choi HS, Park JW, Lee JS, Lee GS, Moon JK, Moon JS, *et al.* 2006. First report of soybean dwarf virus on soybean (*Glycine max*) in Korea. *Res. Plant Dis.* **12**, 213–220.

Kumar S, Stecher G, and Tamura K. 2016. MEGA 7: molecular evolutionary genetics analysis version 7.0 for bigger datasets.

Mol. Biol. Evol. **33**, 1870–1874.

Tamada T, Goto T, Chiba I, and Suwa T. 1969. Soybean dwarf, a new virus disease. *Ann. Phytopath. Soc.* **35**, 282–285.

Terauchi H, Kanematsu S, Honda K, Mikoshiba Y, Ishiguro K, and Hidaka S. 2001. Comparison of complete nucleotide sequences of genomic rnas of four soybean dwarf virus strains that differ in their vector specificity and symptom production. *Arch. Virol.* **146**, 1885–1898.

Tian B, Gildow FE, Stone AL, Sherman DJ, Damsteegt VD, and Schneider WL. 2017. Host adaptation of soybean dwarf virus following serial passages on pea (*Pisum sativum*) and soybean (*Glycine max*). *Viruses* **9**, 155.