

Complete genome sequence of soybean vein necrosis virus (SVNV) from Alabama

Abdelaal H. A. Shehata,^{1,2} Rachel Livingston,¹ Neha Potnis,¹ Edward J. Sikora,¹ K. M. Martin¹

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT The complete genome of soybean vein necrosis virus (SVNV) isolate SVNV17_Auburn_AL was obtained from soybean using RNA sequencing and rapid amplification of cDNA ends (RACE). The tripartite genome comprises 16,563 nucleotides, representing the Alabama (AL) SVNV genome, and shows high similarity to Iowa, Illinois, and Tennessee isolates, enabling evolutionary analysis.

KEYWORDS soybean vein necrosis virus, full genome, Alabama isolate, RNA sequencing, SNPs, indels, RACE

Soybean vein necrosis virus (*Orthotospovirus glycininecrovenae*, SVNV) is an ambisense ssRNA virus in the genus *Orthotospovirus* first identified in Tennessee in 2008 (1). SVNV consists of three segments: S, M, and L. These encode a nucleocapsid protein (N), nonstructural proteins (NSs and NSm), glycoproteins (G_N and G_C), and an RNA-dependent RNA polymerase (RdRp) (2). The complete sequence of the SVNV17_Auburn_AL isolate was obtained using RNA-Seq and RACE.

In 2023, soybean samples exhibiting symptoms of SVNV were collected. Total RNA was extracted from symptomatic leaves using the previous methodology (3), followed by ribosomal RNA depletion using the Illumina Ribo-Zero Plus rRNA Depletion Kit (Illumina, Cat: 20037135). Libraries were prepared with the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina and sequenced on an Illumina NovaSeq 6000 (150 bp PE, ~47 million reads). Quality control was conducted using FastQC (4), and adapter sequences were removed using BBDuk (<https://sourceforge.net/projects/bbmap/>). Processed reads were mapped to the SVNV-TN genome (GCA_004789395.1) using Bowtie (5). Variants (depth >80, Phred > 100) were called using BCFtools (6) and FreeBayes tool (7), and the average depth was calculated using SAMtools (6). Identified variants were visualized using IGV v2.3.57 (8). The consensus assembly was generated with BCFtools (6). For all tools, the default parameters were used. Missing terminal nucleotides were filled using RACE: 5' ends with the Invitrogen 5' RACE System (ThermoFisher, Cat: 18374058) using segment-specific primers (Table 1), and 3' ends using *E. coli* poly(A) polymerase (NEB) followed by SuperScript III (Invitrogen) synthesis. PCR amplification used Phusion (ThermoFisher, F530S), cloned using the CloneJET PCR Cloning Kit, and 12 colonies were Sanger sequenced using an Applied Biosystems 3730xl sequencer.

The genome of the SVNV17_Auburn_AL comprises 16,563 bases (2,602 bp [S], 4,948 bp [M], and 9,013 bp [L]) with a GC content of 35% and an average depth of 1,669x. The leaders are 58, 57, and 185 bases, while the trailers are 70, 91, and 30 bases for the S, M, and L segments, respectively. The first six bases (AGAGCA) at the 5' ends are identical across all three segments and are complementary to the 3' ends, forming a panhandle similar to other orthotospoviruses (9). Genome comparison between the SVNV17_Auburn_AL isolate and the TN strain revealed 43, 97, and 138 SNPs/indels in the S, M, and L segments (Fig. 1). To determine the impact on the protein level, the ORFs' sequence was translated with Expasy (10) and aligned to the TN strain with Clustal

Editor Jelle Matthijnsens, Katholieke Universiteit Leuven, Leuven, Belgium

Address correspondence to K. M. Martin, kmm0173@auburn.edu.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 28 May 2025

Accepted 14 August 2025

Published 5 September 2025

[This article was published on 5 September 2025 with an error in the title. The title was corrected in the current version, posted on 15 September 2025.]

Copyright © 2025 Shehata et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

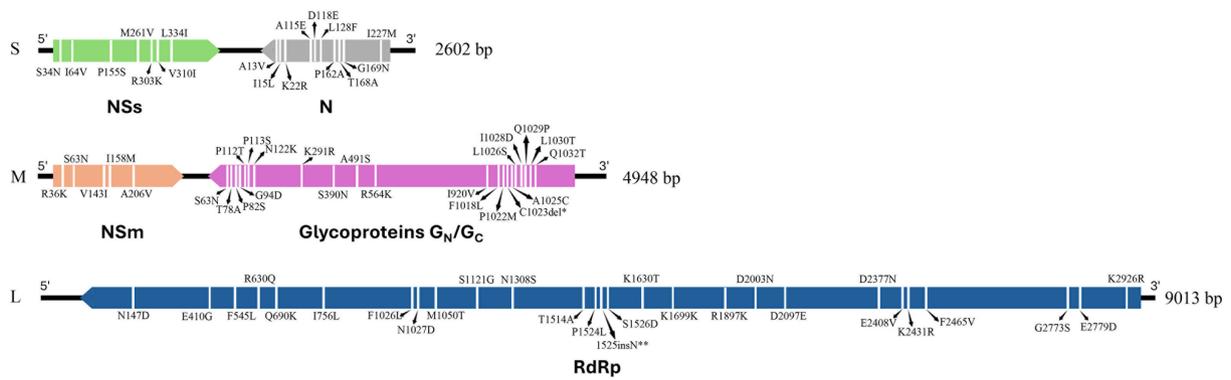


FIG 1 Amino acid mutations detected in the genome of the Alabama isolate SVNV17_Auburn_AL based on RNA-Seq data. Green indicates the nonstructural silencing suppressor protein (NSs), gray the nucleocapsid protein (N), orange the nonstructural movement protein (NSm), magenta the glycoproteins (G_N and G_C), and blue the RNA-dependent RNA polymerase (RdRp). Asterisks (*) represent deletions; double asterisks (**) represent insertions. Tapered ends denote the direction of translation for each open reading frame (ORF).

TABLE 1 Soybean vein necrosis virus segment-specific primers and PolyT-R, used for rapid amplification of 5' and 3' RACE protocols

No	Primer ID	Sequence	Amplified contig Usage
1	SVNV-S-NSs-300-R	CCTGGGAAAATTCATCGCTCGACAT	300 bp 5' - cDNA ^a
2	SVNV-S-NSs-274-R	GTCGACAATAGTATCTGGATCCTCAA	274 bp 5' - PCR ^a
3	SVNV-S-N-2329-F	CAATAGCCTTTCTGTTTTGAGAAAG	274 bp 3' - PCR
4	SVNV-M-NSm-300-R	TGTACCAGACTCATATACCTAATT	300 bp 5' - cDNA ^a
5	SVNV-M-NSm-274-R	CAGATGATCCTGTAGTCTCTCACAA	274 bp 5' - PCR ^a
6	SVNV-M-Gn-4682-R	ACCAAGCTTTTGCCAAATTTTCCTTG	274 bp 3' - PCR
7	SVNV-L-RdRp-300-R	GAGTAGTTCAGTTGGAACTAGTTCA	300 bp 5' - cDNA ^a
8	SVNV-L-RdRp-274-R	AGTGGTTCGAAAAGCTAAATGGTCAAA	274 bp 5' - PCR ^a
9	SVNV-L-RdRp-8737-F	CTTCAACTTAACATGGTTGTAGATT	274 bp 3' - PCR
10	PolyT-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTT	NA 3' - PCR

^aPrimers marked with an asterisk were used for the 5' end following the RACE protocol instructions: the first primer (labeled "cDNA") was used for first-strand cDNA synthesis, and the second primer located slightly downstream of the first primer (labeled "PCR") was used for subsequent PCR amplification.

Omega (11), identifying 17, 26, and 27 amino acid changes in the S, M, and L segment ORFs (3) (Fig. 1). Based on BLASTn (12), compared with the previous SVNV isolates from AL (AL-2 and AL-3) (S-MT669382.1, M-MT548010.1, and L-MT548011.1), our S, M, and L segments share 96%, 97%, and 98% nucleotide identity, indicating that this isolate is genetically distinct. Notably, our segments are more closely related to strains from Tennessee (NC_055178.1, 97.93%), Illinois (MT293139.1, 98.28%), and Iowa (MT536771.1, 98.78%), underscoring the genetic variability of SVNV across regions and highlighting the importance of continued genomic surveillance.

ACKNOWLEDGMENTS

This work was funded by the Alabama Soybean Board.

This study did not involve humans or animals and was therefore exempt from institutional ethics review.

AUTHOR AFFILIATIONS

¹Department of Entomology and Plant Pathology, Auburn University, Auburn, Alabama, USA

²Department of Plant Pathology, Faculty of Agriculture, Assiut University, Assiut, Egypt

AUTHOR ORCID*s*

Abdelaal H. A. Shehata  <http://orcid.org/0000-0002-5498-2009>

K. M. Martin  <http://orcid.org/0000-0002-6654-2668>

FUNDING

Funder	Grant(s)	Author(s)
Alabama Soybean Board		K. M. Martin

AUTHOR CONTRIBUTIONS

Abdelaal H. A. Shehata, Methodology, Writing – original draft | Rachel Livingston, Methodology | Neha Potnis, Data curation, Formal analysis | Edward J. Sikora, Resources, Writing – review and editing | K. M. Martin, Resources, Conceptualization

DATA AVAILABILITY

The complete sequences of SVNV-AL in this paper are available at the NCBI under GenBank accession numbers: [PV592793](https://www.ncbi.nlm.nih.gov/nuclseq/PV592793), [PV592794](https://www.ncbi.nlm.nih.gov/nuclseq/PV592794), and [PV592795](https://www.ncbi.nlm.nih.gov/nuclseq/PV592795). The SVNV-infected soybean RNA-Seq library sequence is under BioProject: [PRJNA1246688](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1246688), BioSample: [SAMN47792314](https://www.ncbi.nlm.nih.gov/biosample/SAMN47792314), and SRA: [SRR32991661](https://www.ncbi.nlm.nih.gov/sra/SRR32991661).

REFERENCES

1. Tzanetakis I, We R, Newman M, Hajimorad R. 2009 Soybean vein necrosis virus: a new threat to soybean production in Southeastern United States, p 5131–5131. In *American Phytopathological Society (APS Press)*. St. Paul, Minnesota, USA.
2. Zhou J, Kantartzi SK, Wen R-H, Newman M, Hajimorad MR, Rupe JC, Tzanetakis IE. 2011. Molecular characterization of a new tospovirus infecting soybean. *Virus Genes* 43:289–295. <https://doi.org/10.1007/s11262-011-0621-9>
3. Shehata AHA, Mayfield MA, Sikora EJ, Martin KM. 2025. Characterization of soybean vein necrosis virus (SVNV) proteins: sequence analysis of field strains and comparison of localization patterns in differing cell types. *Virology (Auckl)* 604:110450. <https://doi.org/10.1016/j.virol.2025.110450>
4. Simons A. 2010. A quality control tool for high throughput sequence data. a quality control tool for high throughput sequence data 10:f1000research. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
5. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:1–10. <https://doi.org/10.1186/gb-2009-10-3-r25>
6. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* 10:giab008. <https://doi.org/10.1093/gigascience/giab008>
7. Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. arXiv. <https://doi.org/10.48550/arXiv.1207.3907>
8. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* 29:24–26. <https://doi.org/10.1038/nbt.1754>
9. de Haan P, Kormelink R, de Oliveira Resende R, van Poelwijk F, Peters D, Goldbach R. 1991. Tomato spotted wilt virus L RNA encodes a putative RNA polymerase. *J Gen Virol* 72 (Pt 9):2207–2216. <https://doi.org/10.1099/0022-1317-72-9-2207>
10. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31:3784–3788. <https://doi.org/10.1093/nar/gkg563>
11. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using clustal Omega. *Mol Syst Biol* 7:539. <https://doi.org/10.1038/msb.2011.75>
12. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/0022-2836\(05\)80360-2](https://doi.org/10.1016/0022-2836(05)80360-2)