



# Pathogenicity and genome-wide sequence analysis reveals relationships between soybean mosaic virus strains

Le Gao<sup>1,2</sup> · Yueying Wu<sup>2</sup> · Jie An<sup>2</sup> · Wenxuan Huang<sup>2</sup> · Xinlei Liu<sup>3</sup> · Yongguo Xue<sup>3</sup> · Xiaoyan Luan<sup>3</sup> · Feng Lin<sup>4</sup> · Lianjun Sun<sup>2</sup>

Received: 12 February 2021 / Accepted: 27 August 2021 / Published online: 13 January 2022  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

## Abstract

Soybean mosaic virus (SMV) is the most prevalent viral pathogen in soybean. In China, the SMV strains SC and N are used simultaneously in SMV resistance assessments of soybean cultivars, but the pathogenic relationship between them is unclear. In this study, SMV strains N1 and N3 were found to be the most closely related to SC18. Moreover, N3 was found to be more virulent than N1. A global pathotype classification revealed the highest level of genetic diversity in China. The N3 type was the most frequent and widespread worldwide, implying that SMV possibly originated in China and spread across continents through the dissemination of infected soybean. It also suggests that the enhanced virulence of N3 facilitated its spread and adaptability in diverse geographical and ecological regions worldwide. Phylogenetic analysis revealed prominent geographical associations among SMV strains/isolates, and genomic nucleotide diversity analysis and neutrality tests demonstrated that the whole SMV genome is under negative selection, with the P1 gene being under the greatest selection pressure. The results of this study will facilitate the nationwide use of SMV-resistant soybean germplasm and could provide useful insights into the molecular variability, geographical distribution, phylogenetic relationships, and evolutionary history of SMV around the world.

## Introduction

Recent years have been quite unusual, with the Coronavirus Disease of 2019 pandemic sweeping the globe and endangering human life and health. In the meantime, however, plant epidemics continue to spread silently, resulting

in severe human food insecurity, especially in developing countries [30, 51]. Plant pathogens are a major threat to agricultural production [14], and viruses comprise nearly half of plant-disease-causing pathogens and are responsible for 30 billion USD in annual economic losses globally [49].

Soybean mosaic virus (SMV), a member of the genus *Potyvirus* within the family *Potyviridae*, is the most prevalent and devastating viral pathogen in all soybean (*Glycine max* (L.) Merr.) production regions worldwide, leading to substantial yield losses and deterioration of seed quality [26, 44, 61, 62]. SMV is seed-borne and aphid-transmitted and can also be transmitted via mechanical inoculation. The diseased soybean seedlings that originate from SMV-infected seeds are the primary inoculum sources. SMV is subsequently transmitted via at least 32 aphid species in a non-persistent manner, eventually resulting in secondary spread within and among soybean fields [26, 31]. SMV infection induces mosaic, chlorosis, rugosity, curl, and necrosis on soybean leaves and causes plant stunting and seed discoloration (seed coat mottling) [22]. Yield reductions are usually reported to range from 8% to 35% under natural field conditions [31]; however, losses of more than 50% and even total crop failure have occurred during severe outbreaks [18].

---

Handling Editor: Stephen John Wylie.

---

Le Gao and Yueying Wu have contributed equally to this work.

---

✉ Le Gao  
gaole@bvca.edu.cn

✉ Lianjun Sun  
sunlj@cau.edu.cn

<sup>1</sup> Department of Horticulture, Beijing Vocational College of Agriculture, Beijing 102442, China

<sup>2</sup> State Key Laboratory of Agrobiotechnology, Beijing Key Laboratory of Crop Genetic Improvement, China Agricultural University, Beijing 100193, China

<sup>3</sup> Institute of Soybean Research, Heilongjiang Provincial Academy of Agricultural Sciences, Harbin 150086, China

<sup>4</sup> Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI 48824, USA

The SMV genome consists of a monopartite, single-stranded, positive-sense RNA molecule of ~9.6 kb in length, possessing a viral genome-linked protein (VPg) covalently bound to the 5' terminus and a poly(A) tail at the 3' end [26, 34]. The viral genome contains a single open reading frame (ORF) encoding a large precursor polyprotein, which is ultimately processed to yield at least 10 mature multifunctional proteins including protein 1 (P1), helper component-proteinase (HC-Pro), protein 3 (P3), 6-kilodalton protein 1 (6K1), cylindrical protein (CI), 6-kilodalton protein 2 (6K2), VPg, nuclear inclusion a-proteinase (NIa-Pro), nuclear inclusion b (NIb), and coat protein (CP) [26, 34]. Additionally, through a frameshift, the SMV genome also encodes a small ORF yielding an 11th protein, termed "pretty interesting *Potyviridae* ORF" (P3N-PIPO), which is produced as a consequence of transcriptional slippage in the P3 cistron [12, 31].

SMV has undergone changes in its pathogenicity during the long-term process of co-evolution with its hosts and the environment; hence, numerous SMV isolates with different levels of pathogenicity exist in nature [23, 66]. Based on their disease reactions on soybean differentials, a large number of SMV isolates have been categorized into seven strains (G1–G7) in the United States and South Korea [6–8, 37, 38, 52, 53, 55] and five strains (A–E) in Japan [60]. In China, SMV has been grouped into two types, namely 22 SC strains (SC1–SC22), which are found nationwide, and three N strains (N1–N3), which are prevalent in Northeast China, on different sets of soybean differentials [41, 46]. In the National Soybean Regional (Uniform) Tests (NSRUT) in China, new soybean cultivars are required to pass SMV resistance assessments before official approval of release; however, SMV SC strains are used for evaluating soybean cultivars from the Yellow-Huai-Hai River Valleys and Southern China, while SMV strains N1 and N3 are used for evaluating the cultivars from Northeast China [48]. The pathogenic relationship between SMV SC and N strains is unclear, and little genomic information about N strains is available, which seriously impedes the nationwide application of the available SMV-resistant soybean germplasm resources.

Since the first determination of full-length nucleotide sequences of SMV [34], complete and partial genome sequencing of SMV isolates has greatly helped us to understand its genomic structure, to identify determinants of virulence, resistance-breaking, and host range, and to study mutation, recombination, phylogenetic relationship, and evolutionary processes [4, 5, 9–11, 18, 19, 24, 27–29, 35, 50, 52–55, 63–68]. In the present study, the pathogenic relationships between SMV N1/N3 and SC strains were clarified via virus resistance assessments of N strains using the uniform soybean differential system, which was used previously for identifying SC strains [41]. Moreover, the virulence of SMV strains N1 and N3 was compared through monitoring

of the pathogenic phenotype and viral accumulation on different susceptible soybean cultivars, and through analysis of the available data from SMV resistance assessments. Finally, the full-length genomic sequences of SMV strains N1 and N3 were determined and compared with other available complete SMV sequences. The results from this study will facilitate the nationwide use of SMV-resistant soybean germplasm, accelerate the progress of soybean resistance breeding in China, and provide useful insights into the molecular variability, geographical distribution, phylogenetic relationships, and evolutionary history of SMV around the world.

## Materials and methods

### Soybean materials and SMV strains

Ten soybean differentials, including Nannong 1138-2, Youbian 30, 8101, Tiefeng 25, Davis, Buffalo, Zaoshu 18, Kwanggyo, Qihuang No. 1, and Kefeng No. 1, which have been used for identifying SMV SC strains [41], were used for resistance assessments of N strains in this study. The soybean seeds were provided by the National Center for Soybean Improvement, Nanjing Agricultural University.

SMV strains N1 and N3, previously identified in Northeast China [46], were obtained from the Jilin Academy of Agricultural Sciences (Northeast Agricultural Research Center of China). Both strains were maintained separately on soybean cv. Nannong 1138-2 (a highly susceptible host) for further biological and molecular analysis.

### Pathogenicity test

Seeds of 10 soybean differentials were individually sown in plastic pots containing moistened nutrition soil mixed with vermiculite and grown at 23–25°C with a photoperiod of 16 h in an insect-proof greenhouse. Seedlings were thinned to 15–20 healthy and uniform plants per pot and mechanically inoculated with SMV strains N1 and N3 as described previously [21]. The inocula were prepared from the symptomatic leaves collected from the corresponding infected Nannong 1138-2 plants, which were homogenized using a sterilized mortar and pestle in 0.01 M phosphate-buffered saline (a mixture of sodium hydrogen phosphate and monopotassium phosphate, pH 7.4) supplemented with a moderate amount of carborundum powder (600-mesh) as an abrasive. Inoculation was performed by gently rubbing the fully expanded unifoliate leaves with the viral suspension using a paintbrush. Leaves were rinsed with tap water shortly after the inoculation, and plants were sprayed regularly with pesticides to prevent cross-infection via aphids. Disease symptoms (i.e., symptomless, mosaic, and necrosis) were monitored starting

7 days post-inoculation (dpi) and recorded at 1-week intervals until the R1 stage (beginning of flowering) [17].

### Virulence comparison

Virus accumulation was detected by quantitative real-time polymerase chain reaction (qRT-PCR) analysis in soybean cvs. Nannong 1138-2 and 8101 that had been challenged with SMV strains N1 and N3. Gene-specific forward (5'-CAGATGGGCGTGGTTATGA-3') and reverse primers (5'-ACAATGGGTTTCAGCGGATA-3') were designed targeting the conserved region of SMV CP using Primer Premier 5.0. *GmTubulin* (accession no. AY907703), amplified with the forward (5'-GGAGTTCACAGAGGCAGAG-3') and reverse primers (5'-CACTTACGCATCACATAGCA-3'), was used as the internal reference control. Samples were collected independently from the corresponding infected young leaves at 7, 14, 21, and 28 dpi, and total RNA extractions followed by first-strand cDNA syntheses were conducted using an RNA Simple Total RNA Kit (Tiangen, China) and PrimeScript<sup>®</sup> RT Master Mix (Takara, Japan), respectively. Subsequently, qRT-PCR was carried out in a reaction mixture with a 20- $\mu$ L final volume, containing 2  $\mu$ L of template cDNA (approximately 50 ng), 0.4  $\mu$ L of each primer (10  $\mu$ M), 10  $\mu$ L of 2 $\times$  SYBR<sup>®</sup> Premix Ex Taq<sup>™</sup> (Takara, Japan), and 7.2  $\mu$ L of sterilized ddH<sub>2</sub>O. Thermal conditions were set to 95 °C for 30 s, followed by 40 cycles at 95 °C for 5 s, 55 °C for 30 s, and 72 °C for 30 s. Samples were analyzed in triplicate on a LightCycler<sup>®</sup> 480 II instrument (Roche, Germany) according to the manufacturer's manual. Transcript levels were quantified using the relative quantification ( $2^{-\Delta\Delta C_t}$ ) method, and data were compared with the internal controls.

Statistical data from SMV resistance assessments deposited in the NSRUT (<https://www.natesc.org.cn/>) over a four-year period (2015-2018) were analyzed. The disease index (DI) of each evaluated soybean cultivar challenged with SMV strains N1 and N3 was calculated as described previously [23]. On the basis of the DI values, the response types of soybean cultivars in terms of virus resistance were classified as HR (highly resistant, DI = 0), R (resistant, 0 < DI  $\leq$  20), MR (moderately resistant, 20 < DI  $\leq$  35), MS (moderately susceptible, 35 < DI  $\leq$  50), S (susceptible, 50 < DI  $\leq$  70), and HS (highly susceptible, 70 < DI  $\leq$  100) [23]. The six classifications were further grouped as resistant (HR, R, and MR) or susceptible (MS, S, and HS).

### Whole-genome sequencing

RNA was isolated from soybean leaves infected with SMV strain N1 or N3 and then converted into cDNA, using the methodology described above. Three gene fragments of each strain, overlapping by at least 250 nt in the adjacent regions,

were amplified by reverse transcription PCR using high-fidelity polymerase KOD FX (Toyobo, Japan) as described previously [68]. The 5' fragment (~3.3 kb) was amplified using the forward primer 5'-AAATTAATAACTMSTYATAAGA-3' and the reverse primer 5'-CCYTGCARYACACTA GTCATTTG-3', the middle fragment (~3.6 kb) was amplified using the forward primer 5'-CTCCACATACGGARA AATG-3' and the reverse primer 5'-CCAACCATRCAA ACMCGTTC-3', and the 3' fragment (~3.2 kb) was amplified using the forward primer 5'-ATGTTTGGGGTYGGC TATGG-3' and the reverse primer 5'-AGGACAACAAAC ATTGCCGYACCT-3' [68]. The amplicons were separated by electrophoresis in a 0.8% agarose gel and visualized using a gel imaging system (Bio-Rad, USA). The bands with the expected sizes were excised and purified using an AxyPrep DNA Gel Extraction Kit (Axygen, USA) and cloned into the pMD19-T vector (Takara, Japan). To ensure accuracy, at least three clones of each fragment were sequenced bidirectionally by TSINGKE Biological Technology Co. Ltd., Beijing, China. The resulting contigs were trimmed and assembled using BioXM 2.6, and the overall pairwise sequence identity between N1 and N3 was calculated using DNAMAN 9.0 at both the nucleotide and amino acid levels.

### Geographical distribution

A total of 104 complete SMV sequences retrieved from the National Center for Biotechnology Information database (<http://www.ncbi.nlm.nih.gov/>) were analyzed. Based on the nucleotide and amino acid sequence differences between N1 and N3, the analyzed SMV strains/isolates were divided into six pathotypes, including the N1 type, intermediate type I, intermediate type II, intermediate type III, intermediate type IV, and the N3 type. The locations where these pathotypes were isolated were marked on a world map to visualize their worldwide geographical distribution.

### Genome-wide analysis

The newly obtained N1 and N3 sequences were compared with other available complete SMV genome sequences. Multiple sequences of SMV nucleotide and deduced amino acids were aligned using BioEdit 7.0 and used for phylogenetic analysis. A phylogenetic tree was built by the neighbor-joining (NJ) method with 1000 bootstrap replicates in MEGA 5.0. To estimate the variations in evolutionary constraints on different regions of the genome, genomic diversity was calculated using DnaSP 5.0 with a sliding window of 100 bp and a step size of 25 bp. To explore the demographic history of SMV populations, Tajima's D, Fu and Li's D\*, and F\* neutrality tests were applied to each SMV-encoded gene in DnaSP 5.0.

## Results

### SMV strains N1 and N3 are most closely related to strain SC18

To compare the pathogenic relationship between SMV N1/N3 and SC strains, 10 soybean differentials previously used for identifying SC strains were inoculated with N strains. Surprisingly, the disease reactions of N1 and N3 were the same, namely cvs. Nannong 1138-2 and 8101 both showed mosaic symptoms and were susceptible, whereas the other eight cultivars were all symptomless and resistant to these two strains (Table 1 and Supplementary Fig. S1). Furthermore, strains N1 and N3 were found to exhibit symptoms and pathogenicity identical to those of strain SC18 on the tested soybean cultivars (Table 1). Therefore, we concluded that strains N1 and N3 were most closely related to SC18 based on their performance on the 10 soybean differentials.

### N3 is more virulent than N1

Although N1 and N3 could systemically infect Nannong 1138-2 and 8101 (Table 1 and Supplementary Fig. S1), the virulence of N1 and N3 differed on these two soybean cultivars. At first, similar mosaic symptoms appeared on Nannong 1138-2 and 8101 inoculated with N1 and N3 at 7 and 14 dpi (Fig. 1); however, symptoms induced by N3 (severe curl) became more prominent and severe than those induced by N1 (moderate crinkling), both on Nannong 1138-2 and 8101 at 21 and 28 dpi (Fig. 1). Subsequently, virus accumulation in Nannong 1138-2 and 8101 infected with N1 and N3 was measured by qRT-PCR at different time points. For Nannong 1138-2, the amount of virus present at 7, 14, and 21 dpi was similar for N1 and N3, whereas at 28 dpi, the amount of N3 was considerably higher than that of N1 (Fig. 1). For 8101, the amount of N3 present was clearly greater than that of N1 at most of the time points, particularly at 28 dpi (Fig. 1). Hence, the differences in viral titers supported the phenotypic observations for the relative virulence of N1 and N3.

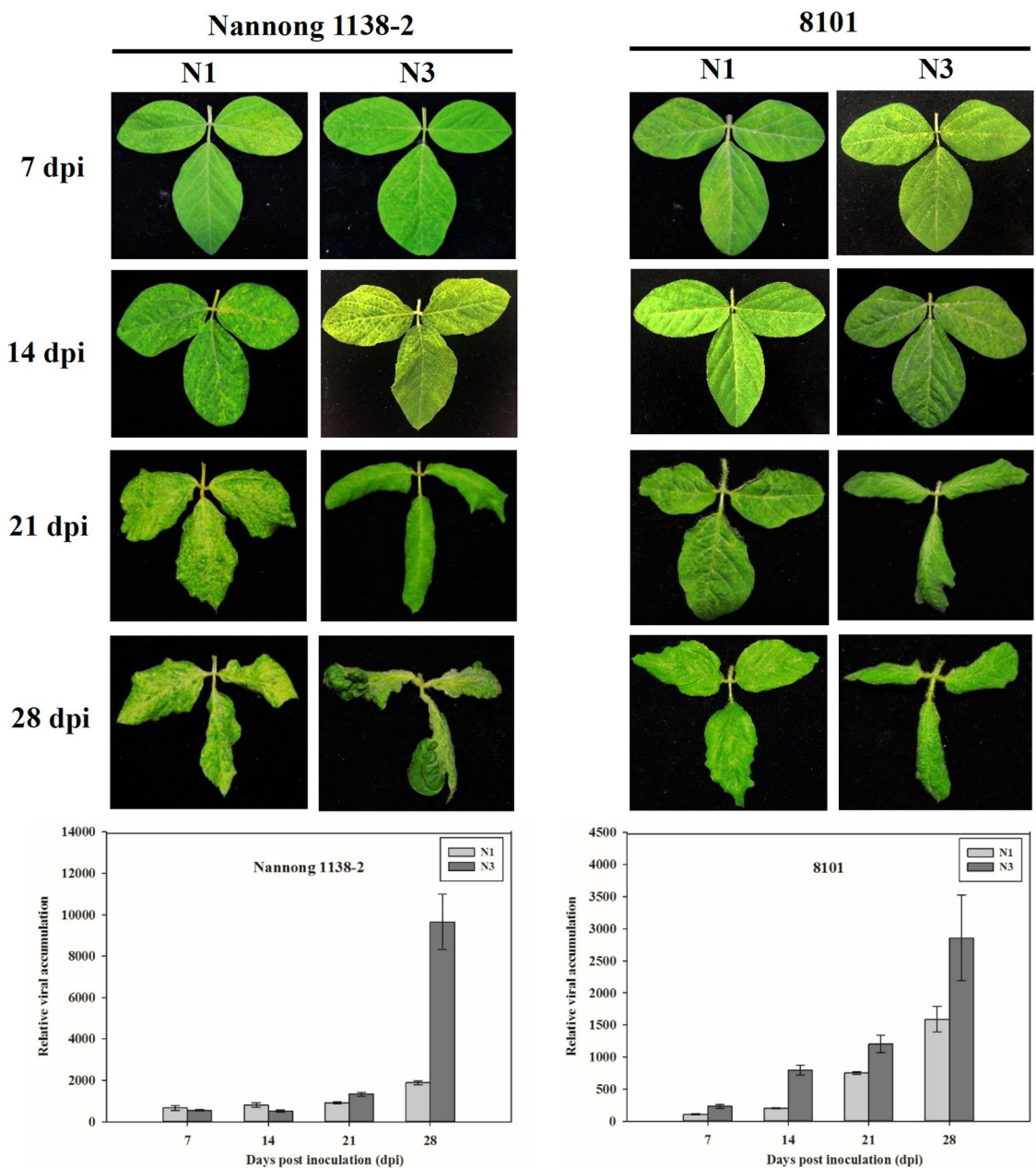
The available statistical data from resistance assessments of SMV strains N1 and N3 in the years 2015–2018 (Supplementary Table S1) were analyzed and are summarized in Table 2. A total of 352 soybean cultivars were assessed for virus resistance to N1 and N3, and no HR and HS types were found (Table 2). Among the 352 cultivars evaluated, 209 (59.4%) were identified as resistant to N1, including R (77, 21.9%) and MR types (132, 37.5%), whereas only 105 (29.8%) were identified as resistant to N3, including R (29, 8.2%) and MR types (76, 21.6%)

**Table 1** Resistance assessments of SMV strains N1 and N3 on 10 soybean differentials

Strain	Disease reaction	Nannong 1138-2	Youbian 30	8101	Tiefeng 25	Davis	Buffalo	Zaoshu 18	Kwanggyo	Qihuang No. 1	Kefeng No. 1
N1	Symptom	—/M	—/—	—/M	—/—	—/—	—/—	—/—	—/—	—/—	—/—
	Pathogenicity	S	R	S	R	R	R	R	R	R	R
N3	Symptom	—/M	—/—	—/M	—/—	—/—	—/—	—/—	—/—	—/—	—/—
	Pathogenicity	S	R	S	R	R	R	R	R	R	R
SC18 <sup>a</sup>	Symptom	—/M	—/—	—/M	—/—	—/—	—/—	—/—	—/—	—/—	—/—
	Pathogenicity	S	R	S	R	R	R	R	R	R	R

Symptoms on inoculated unifoliate leaves/symptoms on non-inoculated upper leaves: —, symptomless; M, mosaic symptom; R, resistant; S, susceptible

<sup>a</sup>The disease reaction of SMV strain SC18 was obtained from Li et al. [41]



**Fig. 1** Symptom appearance and qRT-PCR detection of SMV strains N1 and N3 on soybean cvs. Nannong 1138-2 and 8101 at different time points. Data are expressed as the mean of three biological rep-

licates with error bars indicating the standard deviation (SD). SMV, soybean mosaic virus; dpi, days post inoculation

(Table 2). On the other hand, 143 (40.6%) were identified as susceptible to N1, including MS (131, 37.2%) and S types (12, 3.4%), whereas 247 (70.2%) were identified as susceptible to N3, including MS (134, 38.1%) and S types (113, 32.1%) (Table 2). The average disease index (ADI) of N1 was 31.94 (ranging from 28.04 to 36.82), which was lower than that of N3 (ADI = 42.64, ranging from 40.62 to 45.18) (Table 2). In summary, the 352 soybean cultivars that were evaluated displayed an obviously higher overall level of resistance to N1 than to N3.

Taken together, SMV strain N3 was found to be more virulent than N1 based on the pathogenic phenotype, viral accumulation on soybean cvs. Nannong 1138-2 and 8101 (Fig. 1), and the data of SMV resistance assessments (Table 2).

**Sequence variations between N1 and N3**

The genomes of SMV strains N1 and N3 were completely cloned and sequenced (Supplementary Fig. S2), and both

**Table 2** Resistance of soybean cultivars to SMV strains N1 and N3 in the National Soybean Regional (Uniform) Tests

Year	n	N1						N3							
		HR	R	MR	MS	S	HS	ADJ <sup>a</sup>	HR	R	MR	MS	S	HS	ADJ <sup>a</sup>
2015	60	0	18	31	11	0	0	28.04	0	5	18	23	14	0	40.62
2016	79	0	17	37	25	0	0	31.07	0	5	20	36	18	0	42.06
2017	97	0	27	35	30	5	0	31.82	0	14	16	37	30	0	42.71
2018	116	0	15	29	65	7	0	36.82	0	5	22	38	51	0	45.18
Total	352	0 (0%)	77 (21.9%)	132 (37.5%)	131 (37.2%)	12 (3.4%)	0 (0%)	31.94 ± 3.64	0 (0%)	29 (8.2%)	76 (21.6%)	134 (38.1%)	113 (32.1%)	0 (0%)	42.64 ± 1.90

n, number of soybean cultivars evaluated; HR, highly resistant; R, resistant; MR, moderately resistant; MS, moderately susceptible; S, susceptible; HS, highly susceptible; ADJ, average disease index

<sup>a</sup>ADJ was calculated by averaging the corresponding disease indices of the evaluated soybean cultivars in each year (Supplementary Table S1), and total ADI with standard deviation was calculated by averaging the ADIs of these four years

**Table 3** Nucleotide and amino acid sequence differences between SMV strains N1 and N3

Region	Position of nucleotide (amino acid)	N1	N3
HC-Pro	1439 (436)	<u>A</u> AT ( <b>Asn</b> )	AGT ( <b>Ser</b> )
	1914 (594)	GG <u>G</u> (Gly)	GG <u>A</u> (Gly)
CI	4377 (1415)	AAT ( <b>Asn</b> )	AAC ( <b>Asn</b> )

Underlining indicates the positions of nucleotide differences, and the amino acid variations are highlighted in bold

**Table 4** Pathotype classification of SMV strains/isolates based on the nucleotide and amino acid differences between N1 and N3

Pathotype	HC-Pro <sub>1439</sub>	HC-Pro <sub>1914</sub>	CI <sub>4377</sub>	HC-Pro <sub>436</sub>
N1 type	A	G	T	N
Intermediate type I	G	G	T	S
Intermediate type II	G	G	C	S
Intermediate type III	G	A	T	S
Intermediate type IV	G	G	C	R
N3 type	G	A	C	S

N, asparagine; S, serine; R, arginine

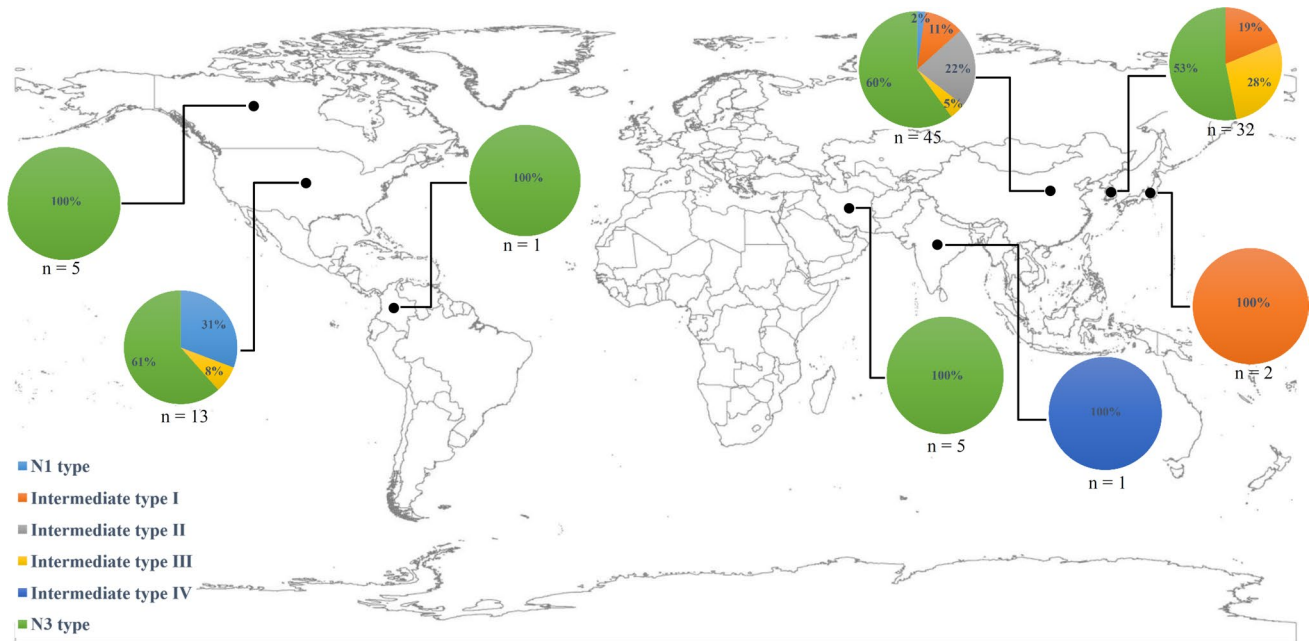
had a genome of 9589 nucleotides encoding a 3067-amino-acid polyprotein (Supplementary Texts S1 and S2). The complete sequences of N1 and N3 were deposited in the GenBank database with accession numbers MN623289 and MN623290, respectively. Sequence alignments showed that N1 and N3 were almost identical, sharing 99.97% sequence identity at both the nucleotide and amino acid levels, with only three nucleotide differences and one amino acid difference (Table 3, Supplementary Figs. S3 and S4). Two nucleotide differences in HC-Pro (A or G at position 1439 and G or A at position 1914) and one nucleotide difference in CI (T or C at position 4377) resulted in a single amino acid difference in HC-Pro, namely Asn (N) or Ser (S) at position 436 (Table 3, Supplementary Figs. S3 and S4).

### The N3 type is the most frequent and widespread type worldwide

A total of 104 SMV strains/isolates were analyzed in this study (Supplementary Table S2). Based on the nucleotide and amino acid sequence differences between N1 and N3 (Tables 3 and 4), SMV strains/isolates were classified as six pathotypes consisting of N1 type (5), intermediate type I (13), intermediate type II (10), intermediate type III (12), intermediate type IV (1), and N3 type (63) (Table 5 and Supplementary Table S3), which were found in China (45), Korea (32), Japan (2), Iran (5), India (1), Canada (5), the USA (13), and Colombia (1) (Table 5, Supplementary Table S3, and Fig. 2). Among the SMV pathotypes, the

**Table 5** Distribution of SMV pathotypes worldwide

Pathotype	China	Korea	Japan	Iran	India	Canada	USA	Colombia	Total
N1 type	1	0	0	0	0	0	4	0	5
Intermediate type I	5	6	2	0	0	0	0	0	13
Intermediate type II	10	0	0	0	0	0	0	0	10
Intermediate type III	2	9	0	0	0	0	1	0	12
Intermediate type IV	0	0	0	0	1	0	0	0	1
N3 type	27	17	0	5	0	5	8	1	63
Total	45	32	2	5	1	5	13	1	104



**Fig. 2** Worldwide geographical distribution of SMV pathotypes. The number (n) of the analyzed SMV strains/isolates and the percentage of the pathotypes are indicated for each country, including China (n

= 45), Korea (n = 32), Japan (n = 2), Iran (n = 5), India (n = 1), Canada (n = 5), USA (n = 13), and Colombia (n = 1). SMV, soybean mosaic virus

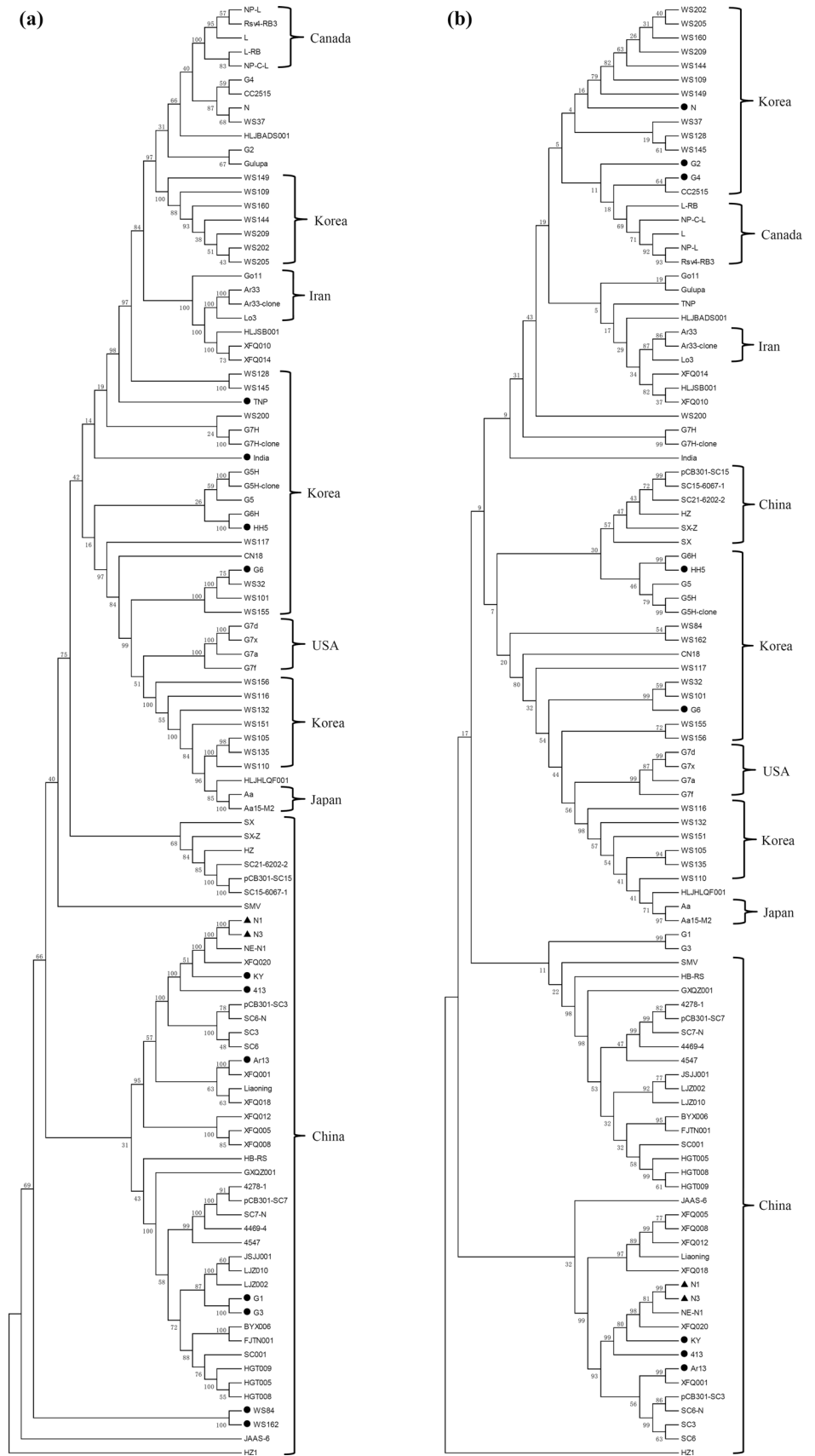
N1 type was found infrequently, while the N3 type was the most frequent by far. The other types (except for intermediate type IV) were found at a similar frequency (Table 5 and Supplementary Table S3). It was observed that nearly all of the pathotypes were present in China, while there were only three pathotypes in Korea and the USA and only one pathotype in the other countries (Table 5, Supplementary Table S3, and Fig. 2). In particular, it is notable that the N3 type was found in most of the countries and is widely distributed globally (Table 5, Supplementary Table S3, and Fig. 2).

**Phylogenetic analysis and the geographical distribution of SMV**

Phylogenetic relationships among the available complete SMV sequences (Supplementary Table S2) were analyzed by aligning the nucleotide and amino acid sequences (Fig. 3)

and constructing phylogenetic trees, which showed that N1 and N3 were the most closely related to each other at both the nucleotide (Fig. 3a) and amino acid levels (Fig. 3b). Moreover, phylogenetic analysis revealed a significant geographical association among SMV strains/isolates. All five SMV sequences from Canada, two sequences from Japan, and most of the sequences from China and Iran clustered together (Fig. 3). The four G7 strains/isolates (G7a, G7d, G7f, and G7x) from the USA were consistently classified as one subgroup (Fig. 3), and the other sequences from the USA did not cluster closely, due to their sequence diversity. SMV strains/isolates from Korea could be roughly divided into three subgroups including seven sequences (WS109, WS144, WS149, WS160, WS202, WS205, and WS209), nine sequences (CN18, G5, G5H, G5H-clone, G6H, WS32, WS101, WS117, and WS155), and six sequences (WS105, WS110, WS116, WS132, WS135, and WS151), at both the

**Fig. 3** Phylogenetic analysis of SMV strains/isolates based on the full-length nucleotide (a) and amino acid (b) sequences. The phylogenetic trees were constructed using MEGA 5.0. The neighbor-joining method with 1000 bootstrap replicates was used to determine phylogenetic relationships. The newly sequenced N1 and N3 are indicated by black triangles, and strains/isolates with black circles indicate that they were not from the indicated country. SMV, soybean mosaic virus





nucleotide and amino acid levels. No pronounced geographical correlation was found among the analyzed SMV strains/isolates when performing the phylogenetic analysis based on the nucleotide and amino acid sequences of each individual gene (Supplementary Figs. S5 and S6), which was probably attributable to the limited amount of genetic information.

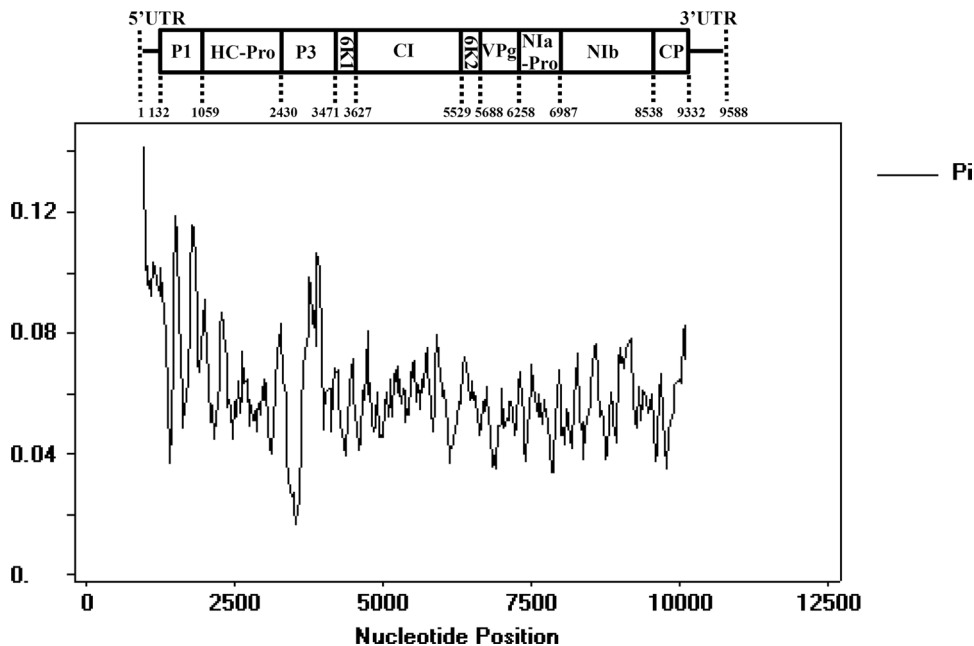
### Selection constraints and virus population demography

Nucleotide variability evaluation and neutrality tests were carried out to investigate the evolution of SMV populations and variation under natural selection pressure. The results showed an overall  $P_i$  value less than 1 (Fig. 4), implying that the whole SMV genome is under negative selection, and some regions of the 5'UTR, P1, and P3 exhibited higher polymorphism than the other regions (Fig. 4), suggesting that these regions are less evolutionarily constrained. Tajima's  $D$ ,  $F_u$  and Li's  $D^*$  and  $F^*$  values for each SMV gene were all negative (Table 6), suggesting that all of the SMV genes are under negative selection pressure to varying degrees and that the SMV population is increasing. P1 may be under the greatest selection pressure, as it was found to have the lowest Tajima's  $D$  value (Table 6). The  $F_u$  and Li's  $D^*$  and  $F^*$  values for P1, NIa-Pro, NIB, and 3'UTR were all significant, and the value for P1 was more highly significant, which is consistent with the result of Tajima's  $D$  test (Table 6). Moreover,  $F_u$  and Li's  $D^*$  and  $F^*$  values were all lower than Tajima's  $D$  values (Table 6), indicating that the mutation rate in SMV may have increased more recently.

### Discussion

The SMV SC and N strains were classified on different sets of soybean differentials [41, 46], resulting in an inability to compare their pathogenicity and posing severe limitations on exchanging and introducing SMV-resistant soybean cultivars across China. In the present study, SMV strains N1 and N3 were found to be most closely related to strain SC18 (Table 1), which was strongly supported by previous research showing that SC18 accounted for 90.5% of the total isolates collected from Heilongjiang province in northeastern China and was predominant and widespread in this region [40]. Thus, we believe that soybean cultivars resistant to either SC18 or N1/N3 could be used interchangeably, and the mapped  $R$  genes conferring resistance to SC18 [39] could also provide resistance to N1 and N3. This information will help to overcome the problems caused by the pathogenic differentiation and host specialization exhibited by SMV SC and N strains and to facilitate the nationwide application of outstanding soybean materials, ultimately accelerating the soybean-breeding progress for improving SMV resistance in China. Nevertheless, the same problems caused by non-uniform SMV strains still exist in other countries, making it difficult to understand the pathogenic divergence of SMV strains/isolates [23, 66, 68]. Consequently, we look forward to future international cooperation to establish a global standard set of soybean differentials in a system for unifying the SMV classification. This will be extremely helpful in exploring the physiological effects and geographical distribution of SMV strains around the world, as well as

**Fig. 4** Sliding-window analysis of  $P_i$  values for each gene of SMV strains/isolates.  $P_i$  values were calculated using DnaSP 5.0 and are shown in a sliding window of 100 bp with a step size of 25 bp. The SMV genome is shown to scale above the graph. SMV, soybean mosaic virus



**Table 6** Neutrality tests for each SMV gene region of 104 strains and isolates

	5'UTR	P1	HC-Pro	P3	6K1	CI	6K2	VPg	Nla-Pro	Nlb	CP	3'UTR
Tajima's D	-0.94798	-1.60648	-1.00155	-1.04521	-1.19158	-1.07624	-0.91236	-1.03432	-1.18374	-1.25775	-0.99488	-1.04708
Fu and Li's D*	-1.62244	-4.15269**	-2.12148	-1.63957	-1.72368	-2.31390	-1.26851	-2.26244	-2.53108*	-2.62186*	-1.97870	-2.50102*
Fu and Li's F**	-1.60118	-3.62035**	-1.94506	-1.64956	-1.80257	-2.11136	-1.34632	-2.06717	-2.32364*	-2.41413*	-1.85231	-2.25241

The test statistic was calculated using DnaSP 5.0. Statistical significance: \*,  $P < 0.05$ ; \*\*,  $P < 0.02$

conveniently making use of the elite SMV-resistant soybean germplasm resources across borders.

HC-Pro has been shown to be strongly associated with symptomatology in several plant-potyvirus systems. Atreya et al. [3] provided evidence that amino acid substitutions within the N-terminus of HC-Pro of tobacco vein mottling virus (TVMV) affected helper component activity, virus accumulation, and symptom expression in infected tobacco plants, especially the mutation  $K_{307}E$ , which not only completely abolished aphid transmissibility but also noticeably affected TVMV virulence. Replacement of two-thirds of the TVMV HC-Pro with that from another potyvirus attenuated symptoms and reduced virus accumulation on *Nicotiana benthamiana*, suggesting that HC-Pro may have universal importance in regulating potyvirus virulence [2]. Introduction of several mutations into HC-Pro of tobacco etch virus decreased virus accumulation and symptom development in tobacco plants [13]. Fragments were exchanged and site-directed mutagenesis was conducted in HC-Pro of zucchini yellow mosaic virus, and a unique mutation ( $R_{180}I$ ) in the highly conserved motif (from FRNK to FINK) dramatically decreased symptom severity in various cucurbit species, including squash, cucumber, melon, and watermelon, by directly influencing the levels and regulatory functions of microRNA populations [20, 58]. Genome sequence analysis of the progeny virus of SMV strain G7H implied that a single amino acid mutation ( $P_{359}S$ ) in HC-Pro led to changes in symptoms in soybean cv. Jinpumkong-2 [56]. Similar results were obtained in the current study, and based on the symptom severity, viral accumulation, and the available data in NSRUT, SMV strain N3 was shown to be more virulent than strain N1 (Fig. 1, Table 2, and Supplementary Table S1). Whole-genome sequencing showed that only three nucleotides (two in HC-Pro and one in CI) differed between N1 and N3, with only a single amino acid difference ( $N_{436}S$ ) in HC-Pro between N1 and N3. The other two nucleotide differences were synonymous substitutions (Table 3, Supplementary Figs. S3 and S4), indicating that the difference in virulence between N1 and N3 was very likely determined by the variation in HC-Pro. In combination, these results convince us that HC-Pro acts as a determinant of symptoms and virulence in the SMV-soybean pathosystem.

Soybean, which has a 5000-year history of cultivation, originated and was domesticated in ancient China and was disseminated early to North, East, and South Asia, and afterwards from Northeast China to Europe and the Americas after the year 1700 [32, 33, 45, 59]. The origin of SMV has been supposed to correspond to that of soybean, namely in South and East Asia, particularly in China [25, 26, 68]. In this study, China was found to have several different SMV pathotypes, displaying the highest level of genetic diversity, while the number of pathotypes was limited in the other countries (Table 5, Supplementary Table S3, and

Fig. 2), implying that SMV might have originated in China. Moreover, we presume that the dissemination of soybean has facilitated the spread of SMV across continents via seed-transmission, as soybean is the natural host of SMV. The present study showed that the N3 type is the most common and widespread worldwide (Table 5, Supplementary Table S3, and Fig. 2), suggesting that the enhanced virulence of N3 (Fig. 1 and Table 2) derived from the genetic variation in HC-Pro (Table 3, Supplementary Figs. S3 and S4) facilitated its spread and increased its adaptability in diverse geographical and ecological regions worldwide.

As no uniform system for SMV strain classification is available worldwide, genome sequencing is the most effective approach for exchanging information about the multitudinous SMV strains/isolates [66, 68]. Therefore, the complete sequences of N1 and N3 obtained in this study will complement the current sequence information of SMV. The results from the sequence analysis could broaden and enrich our knowledge about the molecular variability, geographical distribution, phylogenetic relationships, and evolutionary history of SMV. Besides, directional selection pressure created by the widespread application of SMV-resistant soybean cultivars has given rise to the frequent occurrence of resistance-breaking SMV strains/isolates [9, 18, 50]. Furthermore, interspecific genetic exchanges between SMV and other potyviruses have been continually detected. Chen et al. [4] isolated an SMV-like virus from *Pinellia ternate*, which was shown to result from recombination between SMV and dasheen mosaic virus. An SMV strain that probably originated from the recombination between SMV and bean common mosaic virus (BCMV) or a BCMV-like virus is prevalent in soybean-growing areas of China [5, 64, 65, 67, 68]. Jiang et al. [35] reported that an SMV isolate that was formed by the recombination between SMV and watermelon mosaic virus could cause different diseases in soybean and *N. benthamiana* plants. These observations emphasize the potential risk in soybean productions and the vital role of genome sequencing in enabling us to discover resistance-breaking and recombinant SMV variants, providing an efficient strategy for monitoring and prevention of SMV.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00705-021-05271-z>.

**Acknowledgements** This work was supported by grants from the National Natural Science Foundation of China (32001571), the General Project of Science and Technology Plan of Beijing Municipal Education Commission, the Fundamental Research Funds for the Central Universities (2020TC178), the High Level Vocational Schools with Chinese Characteristics and Major Construction Plan (Horticultural Technology Professional Group), and the Beijing Characteristic High Level College Construction Project (Urban Horticulture Major Group). We thank Dongquan Guo for providing SMV strains N1 and N3 at the Institute of Agricultural Biotechnology, Jilin Academy of Agricultural Sciences.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Ahangaran A, Habibi MK, Mohammadi GHM, Winter S, Garcia-Arenal F (2013) Analysis of soybean mosaic virus genetic diversity in Iran allows the characterization of a new mutation resulting in overcoming *Rsv4*-resistance. *J Gen Virol* 94:2557–2568
- Atreya CD, Pirone TP (1993) Mutational analysis of the helper component-proteinase gene of a potyvirus: effects of amino acid substitutions, deletions, and gene replacement on virulence and aphid transmissibility. *Proc Natl Acad Sci USA* 90:11919–11923
- Atreya CD, Atreya PL, Thornbury DW, Pirone TP (1992) Site-directed mutations in the potyvirus *HC-Pro* gene affect helper component activity, virus accumulation, and symptom expression in infected tobacco plants. *Virology* 191:106–111
- Chen J, Zheng H, Lin L, Adams MJ, Antoniw JF, Zhao M, Shang Y, Chen J (2004) A virus related to soybean mosaic virus from *Pinellia ternata* in China and its comparison with local soybean SMV isolates. *Adv Virol* 149:349–363
- Chen Y, Wu M, Ma F, Chen J, Wang B (2017) Complete nucleotide sequences of seven soybean mosaic viruses (SMV), isolated from wild soybeans (*Glycine soja*) in China. *Adv Virol* 162:901–904
- Cho EK, Chung BJ (1976) Studies on identification of soybean virus diseases in Korea I. Preliminary studies on a soybean virus disease. *Korean J Plant Prot* 15:61–68
- Cho EK, Chung KW (1986) Strains of soybean mosaic virus causing soybean necrotic disease in Korea. *Korean J Breed* 18:150–153
- Cho EK, Goodman RM (1979) Strains of soybean mosaic virus: classification based on virulence in resistant soybean cultivars. *Phytopathology* 69:467–470
- Choi BK, Koo JM, Ahn HJ, Yum HJ, Choi CW, Ryu KH, Chen P, Tolín SA (2005) Emergence of *Rsv*-resistance breaking soybean mosaic virus isolates from Korean soybean cultivars. *Virus Res* 112:42–51
- Chowda-Reddy RV, Sun H, Chen H, Poysa V, Ling H, Gijzen M, Wang A (2011) Mutations in the P3 protein of soybean mosaic virus G2 isolates determine virulence on *Rsv4*-genotype soybean. *Mol Plant Microbe Interact* 24:37–43
- Chowda-Reddy RV, Sun H, Hill JH, Poysa V, Wang A (2011) Simultaneous mutations in multi-viral proteins are required for soybean mosaic virus to gain virulence on soybean genotypes carrying different *R* genes. *PLoS ONE* 6:e28342
- Chung BY, Miller WA, Atkins JF, Firth AE (2008) An overlapping essential gene in the *Potyviridae*. *Proc Natl Acad Sci* 105:5897–5902
- Cronin S, Verchot J, Haldeman CR, Schaad MC, Carrington JC (1995) Long-distance movement factor: A transport function of the potyvirus helper component proteinase. *Plant Cell* 7:549–559
- Dodds PN (2010) Genome evolution in plant pathogens. *Science* 330:1486–1487
- Domier LL, Hobbs HA, McCoppin NK, Bowen CR, Steinlage TA, Chang S, Wang Y, Hartman GL (2011) Multiple loci condition seed transmission of soybean mosaic virus (SMV) and SMV-induced seed coat mottling in soybean. *Phytopathology* 101:750–756
- Eggenberger AL, Stark DM, Beachy RN (1989) The nucleotide sequence of a soybean mosaic virus coat protein-coding region

- and its expression in *Escherichia coli*, *Agrobacterium tumefaciens* and tobacco callus. *J Gen Virol* 70:1853–1860
17. Fehr WR, Caviness CE (1977) Stages of soybean development. Special Report No. 80. Cooperation Extension Service, Agriculture and Home Economic Experiment Station, Iowa State University, Ames
  18. Gagarinova AG, Babu M, Poysa V, Hill JH, Wang A (2008) Identification and molecular characterization of two naturally occurring soybean mosaic virus isolates that are closely related but differ in their ability to overcome *Rsv4* resistance. *Virus Res* 138:50–56
  19. Gagarinova AG, Babu M, Stromvik MV, Wang A (2008) Recombination analysis of soybean mosaic virus sequences reveals evidence of RNA recombination between distinct pathotypes. *Virology* 375:1–8
  20. Gal-On A (2000) A point mutation in the FRNK motif of the potyvirus helper component-protease gene alters symptom expression in cucurbits and elicits protection against the severe homologous virus. *Phytopathology* 90:467–473
  21. Gao L, Ding X, Li K, Liao W, Zhong Y, Ren R, Liu Z, Adhimoalam K, Zhi H (2015) Characterization of soybean mosaic virus resistance derived from inverted repeat-SMV-*HC-Pro* genes in multiple soybean cultivars. *Theor Appl Genet* 128:1489–1505
  22. Gao L, Luo J, Ding X, Wang T, Hu T, Song P, Zhai R, Zhang H, Zhang K, Li K, Zhi H (2020) Soybean RNA interference lines silenced for eIF4E show broad potyvirus resistance. *Mol Plant Pathol* 21:303–317
  23. Gao L, Sun S, Li K, Wang L, Hou W, Wu C, Zhi H, Han T (2018) Spatio-temporal characterisation of changes in the resistance of widely grown soybean cultivars to soybean mosaic virus across a century of breeding in China. *Crop Pasture Sci* 69:395–405
  24. Gao L, Zhai R, Zhong Y, Karthikeyan A, Ren R, Zhang K, Li K, Zhi H (2015) Screening isolates of soybean mosaic virus for infectivity in a model plant, *Nicotiana benthamiana*. *Plant Dis* 99:442–446
  25. Gibbs AJ, Trueman JWH, Gibbs MJ (2008) The bean common mosaic virus lineage of potyviruses: where did it arise and when? *Adv Virol* 153:2177–2187
  26. Hajimorad MR, Domier LL, Tolin SA, Whitham SA, Saghai Maroof MA (2018) Soybean mosaic virus: a successful potyvirus with a wide distribution but restricted natural host range. *Mol Plant Pathol* 19:1563–1579
  27. Hajimorad MR, Eggenberger AL, Hill JH (2003) Evolution of soybean mosaic virus-G7 molecularly cloned genome in *RsvI*-genotype soybean results in emergence of a mutant capable of evading *RsvI*-mediated recognition. *Virology* 314:497–509
  28. Hajimorad MR, Eggenberger AL, Hill JH (2005) Loss and gain of elicitor function of soybean mosaic virus G7 provoking *RsvI*-mediated lethal systemic hypersensitive response maps to P3. *J Virol* 79:1215–1222
  29. Hajimorad MR, Eggenberger AL, Hill JH (2008) Adaptation of soybean mosaic virus avirulent chimeras containing P3 sequences from virulent strains to *RsvI*-genotype soybeans is mediated by mutations in HC-Pro. *Mol Plant Microbe Interact* 21:937–946
  30. He S, Krainer KMC (2020) Pandemics of people & plants: which is the greater threat to food security? *Mol Plant* 13:933–934
  31. Hill JH, Whitham SA (2014) Control of virus diseases in soybeans. *Adv Virus Res* 90:355–390
  32. Hymowitz T, Harlan JR (1983) Introduction of soybean to North America by Samuel Bowen in 1765. *Econ Bot* 37:371–379
  33. Hymowitz T, Newell CA (1981) Taxonomy of the genus *Glycine*, domestication and uses of soybeans. *Econ Bot* 35:272–288
  34. Jayaram C, Hill JH, Miller WA (1992) Complete nucleotide sequences of two soybean mosaic virus strains differentiated by response of soybean containing the *Rsv* resistance gene. *J Gen Virol* 73:2067–2077
  35. Jiang H, Li K, Dou D, Gai J (2017) Characterization of a soybean mosaic virus variant causing different diseases in *Glycine max* and *Nicotiana benthamiana*. *Adv Virol* 162:549–553
  36. Khatabi B, Fajolu OL, Wen RH, Hajimorad MR (2012) Evaluation of North American isolates of soybean mosaic virus for gain of virulence on *Rsv*-genotype soybeans with special emphasis on resistance-breaking determinants on *Rsv4*. *Mol Plant Pathol* 13:1077–1088
  37. Kim JS, Lee EJ (1991) A new virulent strain of soybean mosaic virus infecting SMV resistant soybean cultivar, Deogyou. *Korean J Plant Pathol* 7:37–41
  38. Kim YH, Kim OS, Lee BC, Moon JK, Lee SC, Lee JY (2003) G7H, a new soybean mosaic virus strain: its virulence and nucleotide sequence of *CI* gene. *Plant Dis* 87:1372–1375
  39. Li K, Ren R, Adhimoalam K, Gao L, Yuan Y, Liu Z, Zhong Y, Zhi H (2015) Genetic analysis and identification of two soybean mosaic virus resistance genes in soybean [*Glycine max* (L.) Merr]. *Plant Breed* 134:684–695
  40. Li K, Xia Y, Wang D, Yang Y, Ren R, Gao L, Zhang K, Zhi H (2014) Analysis of dynamic change of soybean mosaic virus strains in Heilongjiang province of China. *Soybean Sci* 33:880–884
  41. Li K, Yang Q, Zhi H, Gai J (2010) Identification and distribution of soybean mosaic virus strains in Southern China. *Plant Dis* 94:351–357
  42. Lim W, Kim Y, Kim KH (2003) Complete genome sequences of the genomic RNA of soybean mosaic virus strains G7H and G5. *Plant Pathol J* 19:171–176
  43. Lin J, Yang Y, Hou W, Yang C, Xie L, Zhi H, Zhang M (2015) Sequence and pathogenicity of recombined soybean mosaic virus isolate from Hebei province, China. *Acta Agron Sin* 41:1657–1662
  44. Liu J, Fang Y, Pang H (2016) The current status of the soybean-soybean mosaic virus (SMV) pathosystem. *Front Microbiol* 7:1906
  45. Liu X, He J, Wang Y, Xing G, Li Y, Yang S, Zhao T, Gai J (2020) Geographic differentiation and phylogeographic relationships among world soybean populations. *Crop J* 8:260–272
  46. Lyu W, Zhang M, Wei P, Xie S, Guo J, Jiang Y, Geng Y (1985) Classification and distribution of strains of soybean mosaic virus in Northeast China. *Acta Phytopathol Sin* 15:225–229
  47. Mesa HJ, Montoya MM, Gutiérrez PA (2018) Molecular characterization of soybean mosaic virus (SMV) infecting purple passion fruit (*Passiflora edulis* f. *edulis*) in Antioquia, Colombia. *Arch Phytopathol Plant Prot* 51:617–636
  48. Ministry of Agriculture of the People's Republic of China (2007) Regulations for the regional tests of crop varieties—Soybean (NY/T1299-2007). Ministry of Agriculture of the People's Republic of China, Beijing
  49. Nicaise V (2014) Crop immunity against viruses: outcomes and future challenges. *Front Plant Sci* 5:660
  50. Saruta M, Kikuchi A, Okabe A, Sasaya T (2005) Molecular characterization of A2 and D strains of soybean mosaic virus, which caused a recent virus outbreak in soybean cultivar Sachiutaka in Chugoku and Shikoku regions of Japan. *J Gen Plant Pathol* 71:431–435
  51. Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A (2019) The global burden of pathogens and pests on major food crops. *Nat Ecol Evol* 3:430–439
  52. Seo JK, Lee HG, Kim KH (2009) Systemic gene delivery into soybean by simple rub-inoculation with plasmid DNA of a soybean mosaic virus-based vector. *Adv Virol* 154:87–99
  53. Seo JK, Lee HG, Choi HS, Lee SH, Kim KH (2009) Infectious in vivo transcripts from a full-length clone of soybean mosaic virus strain G5H. *Plant Pathol J* 25:54–61

54. Seo JK, Lee SH, Kim KH (2009) Strain-specific cylindrical inclusion protein of soybean mosaic virus elicits extreme resistance and a lethal systemic hypersensitive response in two resistant soybean cultivars. *Mol Plant Microbe Interact* 22:1151–1159
55. Seo JK, Ohshima K, Lee HG, Son M, Choi HS, Lee SH, Sohn SH, Kim KH (2009) Molecular variability and genetic structure of the population of soybean mosaic virus based on the analysis of complete genome sequences. *Virology* 393:91–103
56. Seo JK, Sohn SH, Kim KH (2011) A single amino acid change in HC-Pro of soybean mosaic virus alters symptom expression in a soybean cultivar carrying *Rsv1* and *Rsv3*. *Adv Virol* 156:135–141
57. Shi Y, Hong X, Chen J, Adams MJ, Zheng H, Lin L, Qin B, Chen J (2005) Further molecular characterisation of potyviruses infecting aroid plants for medicinal use in China. *Adv Virol* 150:125–135
58. Shibolet Y, Haronsky E, Leibman D, Arazi T, Wassenegger M, Whitham SA, Gaba V, Gal-On A (2007) The conserved FRNK box in HC-Pro, a plant viral suppressor of gene silencing, is required for small RNA binding and mediates symptom development. *J Virol* 81:13135–13148
59. Singh R, Hymowitz T (1999) Soybean genetic resources and crop improvement. *Genome* 42:605–616
60. Takahashi K, Tanaka T, Iida W, Tsuda Y (1980) Studies on virus diseases and casual viruses of soybean in Japan. *Bull Tohoku Natl Agric Exp Stn* 62:1–130
61. Whitham SA, Qi M, Innes RW, Ma W, Lopes-Caitar V, Hewezi T (2016) Molecular soybean-pathogen interactions. *Annu Rev Phytopathol* 54:443–468
62. Widyasari K, Alazem M, Kim KH (2020) Soybean resistance to soybean mosaic virus. *Plants* 9:219
63. Yang Q, Li K, Zhi H, Gai J (2014) Genetic diversity of Chinese soybean mosaic virus strains and their relationships with other plant potyviruses based on *P3* gene sequences. *J Integr Agric* 13:2184–2195
64. Yang Y, Gong J, Li H, Li C, Wang D, Li K, Zhi H (2011) Identification of a novel soybean mosaic virus isolate in China that contains a unique 5' terminus sharing high sequence homology with bean common mosaic virus. *Virus Res* 157:13–18
65. Yang Y, Lin J, Zheng G, Zhang M, Zhi H (2014) Recombinant soybean mosaic virus is prevalent in Chinese soybean fields. *Adv Virol* 159:1793–1796
66. Zhang H, Cui X, Chen X, Zhi H, Zhang S, Zhao L (2012) Determination of the complete genomic sequence and molecular biological analysis of soybean mosaic virus. *Can J Plant Pathol* 34:288–297
67. Zhang L, Shang J, Jia Q, Li K, Yang H, Liu H, Tang Z, Chang X, Zhang M, Wang W, Yang W (2019) Genetic evolutionary analysis of soybean mosaic virus populations from three geographic locations in China based on the *P1* and *CP* genes. *Adv Virol* 164:1037–1048
68. Zhou G, Shao Z, Ma F, Wu P, Wu X, Xie Z, Yu D, Cheng H, Liu Z, Jiang Z, Chen Q, Wang B, Chen J (2015) The evolution of soybean mosaic virus: an updated analysis by obtaining 18 new genomic sequences of Chinese strains/isolates. *Virus Res* 208:189–198

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.