Phyton-
International Journal of
Experimental Botany

Tech Science Press

Doi:10.32604/phyton.2025.066148

ARTICLE

Complete Genomic Sequence Analysis of Sweet Potato Virus 2 Isolates from the Shandong and Jiangsu Provinces in China

Zichen Li[1,#], Jukui Ma[2,#], Minjun Liu[3], Guowei Geng[1,*] and Hongxia Zhang[1,*]

[1]College of Horticulture, Ludong University, Yantai, 264025, China
[2]Key Laboratory of Biology and Genetic Improvement of Sweet Potato, Xuzhou Institute of Agricultural Sciences in Jiangsu Xuhuai Area, Xuzhou, 221131, China
[3]Muping Agricultural Technology Service and Extention Center, Yantai, 264100, China

*Corresponding Authors: Guowei Geng. Email: guowgeng@163.com; Hongxia Zhang. Email: hxzhang@sibs.ac.cn

[#]These authors contributed equally to this work

Received: 31 March 2025; Accepted: 19 May 2025; Published: 27 June 2025

ABSTRACT: Sweet potatoes are significant cash crops, however, their yield and quality are greatly compromised by viral diseases. In this study, the complete genomic sequences of two Sweet Potato Virus 2 (SPV2) isolates from infected sweet potato leaves in the Shandong (designated as SPV2-SDYT, GenBank No. PQ855660.1) and Jiangsu (designated as SPV2-JSXZ, GenBank No. PQ855661.1) provinces in China were obtained using 5′ RACE and RT-PCR amplification. Consistency, phylogeny, codon usage bias, recombination, and selection pressure analyses were conducted using the SPV2-SDYT and SPV2-JSXZ genome sequences. The complete genome sequences of SPV2-SDYT and SPV2-JSXZ were 10561 nucleotides (nt) in length, with respective nucleotide and amino acid identities of 99.25% and 99.12%, respectively. Both isolates were closely related to the SPV2 isolate from China (SPV2-LN). In both SPV2-SDYT and SPV2-JSXZ, the identity of the P1 protein was the highest, whereas that of the P3 protein was the lowest. There were 26 codons with relatively synonymous codon usage (RSCU) values greater than 1 in SPV2-SDYT and 27 codons with RSCU values greater than 1 in SPV2-JSXZ. High-frequency codons in their genomes were predominantly found to end with A/U. Recombination analysis revealed no major recombination sites in either SPV2-SDYT or SPV2-JSXZ. Further selection pressure analysis showed that the non-synonymous substitution rate/synonymous substitution rate (dN/dS) value of all 10 SPV2 proteins was less than 1. This is the first report on the evolutionary relationships of the 17 known SPV2 isolates. Our findings lay the molecular groundwork for preventing and controlling SPV2 infection in root-tuber crops. These findings also contribute to our understanding of the spread and evolution of SPV2, its pathogenic mechanisms, and the development of antiviral strategies against it.

KEYWORDS: Sweet potato virus 2; consistency analysis; phylogenetic analysis; codon usage bias analysis; recombination analysis; selection pressure analysis

1 Introduction

The sweet potato, a twining vine crop plant, holds significant value across various domains, including food, nutrition, economy, and medicine, and plays a crucial role in human livelihood and daily life [1,2]. China boasts the largest sweet potato cultivation area and leads the world in both yield and exports [3]. Sweet potato viral diseases significantly affect the yield and variety of sweet potatoes [4,5]. Currently, over 30 sweet potato viruses have been identified globally, with more than 20 reported in China [6,7]. These viruses are classified into nine families: Potyviridae, Geminiviridae, Caulimoviridae, Bunyaviridae, Bromoviridae,

Comoviridae, Closteroviridae, Luteoviridae, and Flexiviridae. The genus *Potyvirus* in the family Potyviridae comprises sweet potato feathery mottle virus (SPFMV), sweet potato latent virus (SPLV), sweet potato mild speckling virus (SPMSV), sweet potato virus G (SPVG), sweet potato virus C (SPVC), and sweet potato virus 2 (SPV2) [8–11]. SPFMV, SPVG, and SPVC show co-infection in the field. Infection with SPFMV or sweet potato chlorotic stunt virus (SPCSV), both of which belong to the genus *Crinivirus*, typically has a minimal impact on yield. However, co-infection with two or more viruses can result in yield losses of up to 50% [12–14].

SPV2 was first reported in Taiwan, China and Nigeria in 1988. Since then, it has been reported in mainland China, the United States, Australia, Spain, Greece, Korea, East Timor, and South Africa [15–18]. The virus particles of SPV2 are linear, measuring 850 nm in length, and form pinwheel- or scroll-like inclusions within the cytoplasm [19]. SPV2 is a positive single-stranded RNA virus characterized by a poly(A) tail at the 3′ end. It encodes a polyprotein that is cleaved by proteases into 11 mature proteins, including P1, PISPO, HC-Pro, P3, 6K1, CI, 6K2, VPg, NIa-Pro, NIb, and CP. Furthermore, the P3N-PIPO protein is produced via a translational frame-shifting strategy [20,21]. Studies have shown that not all sweet potato viruses of the genus *Potyvirus* possess PISPO and P3N-PIPO proteins. Specifically, SPLV and sweet potato mild mottle virus (SPMMV) do not encode PISPO proteins [22]. Studies have also shown that SPV2 is prevalent in sweet potatoes in China and has been detected in Shandong, Jiangsu, Henan, Liaoning, Shaanxi, Anhui, and other provinces. However, only one SPV2 Chinese sweet potato isolate from Guangxi Province (SPV2-LN, GenBank No. OR842902.1) has been reported to date. In the present study, we obtained the complete genomic sequences of SPV2-SDYT and SPV2-JSXZ and analyzed their consistency, phylogeny, codon usage bias, recombination, and selection pressure. This study aimed to enhance the understanding of the genomic characteristics and evolution of SPV2 isolates in China, thereby offering a theoretical foundation for the prevention and control of SPV2.

2 Materials and Methods

2.1 Viral Sources

Symptomatic sweet potato leaves with chlorotic spots harboring sweet potato virus 2 were respectively collected in Yantai city, Shandong province, China and Xuzhou city, Jiangsu province, China. Using the currently known sequence of SPV2 as primers, RT-PCR amplification was performed and identified by sequencing and comparison as SPV2.

2.2 Total RNA Extraction and RT-PCR Amplification

Total plant RNA was extracted using the M5 SuperPure Total RNA Extraction Reagent (Super TRIgent) (Mei5 Biotech, Beijing, China) according to the manufacturer's instructions. The first strand of cDNA was prepared using M-MLV (H-) Reverse Transcriptase (Vazyme Biotech, Nanjing, China). The cDNA was stored in the refrigerator at −20°C. PCR amplification was performed using it as a template. By comparing the complete genome sequences of 15 other SPV2 sequences from the NCBI database, the primers were designed to amplify 5 overlapping segments of the complete genome (Table 1).

Table 1: Primers sequences used in this study

Primer	Sequence (5′-3′)	Purpose
SPV2-F	GTCCATCAGATAAGG	Used for detecting the presence of
SPV2-R	GTCCACATCTGTTGC	SPV2-SDYT and SPV2-JSXZ

(Continued)

Table 1 (continued)

Primer	Sequence (5′-3′)	Purpose
SPV2-9462-F	ATCAATGGTGATTGGGTRATG ATG	Used for 3′ RACE of SPV2-SDYT and SPV2-JSXZ
QT	CCAGTGAGCAGAGTGACGAG GACTCGAGCTCAAGC$(T)_{17}$	
Q_o	CCAGTGAGCAGAGTGACG	
SPV2-441-R	GTGCTCATCAATAGCACAAT CGC	Used for 5′ RACE of SPV2-SDYT and SPV2-JSXZ
SPV2-320-R	GCTACATAGTAGCTATGCC ATTCAG	
Anchor primer	GGCCACGCGTCGACTAG TAC	
Oligo (dG)-anchor primer	GGCCACGCGTCGACTA GTAC$(G)_{16}$	
Oligo (dT)-anchor primer	GGCCACGCGTCGACTA GTAC$(T)_{16}$	
SPV2-5-F	AAATTAACAAAACTCA ATACAACAACAGAAGAATC	Used for amplifying the sequence of SPV2-SDYT and SPV2-JSXZ 1 nt to 3028 nt
SPV2-3028-R	GCYTCAGATTCCCTCA CATTAACC	
SPV2-2739-F	GAGCTGATGGAGGGCA TACCWG	Used for amplifying the sequence of SPV2-SDYT and SPV2-JSXZ 2739 nt to 4999 nt
SPV2-4999-R	GTAGGGCATARCCACT TGTCATTATCG	
SPV2-4589-F	TGAACTTGACTCAGAC ACACAAAGC	Used for amplifying the sequence of SPV2-SDYT and SPV2-JSXZ 4589 nt to 6659 nt
SPV2-6659-R	GAATTTCAGCTTTTGT CTCTGTCGC	
SPV2-6544-F	TGGGTGGAGTCTTTGT TGGAGG	Used for amplifying the sequence of SPV2-SDYT and SPV2-JSXZ 6544 nt to 8777 nt
SPV2-8777-R	CAGCATTTGCTCTCCT ATGTCCC	
SPV2-8587-F	CRTGGAGTGTTGGAAT GACGAAATTC	Used for amplifying the sequence of SPV2-SDYT and SPV2-JSXZ 8587 nt to 10561 nt
SPV2-3-R	GTCTCTGGTATAAGAC AAAAAGTTTATGACTGG	

2.3 5′ RACE and Complete Genome Amplification

The 5′ UTR sequence of SPV2-SDYT and SPV2-JSXZ were amplified using 5′ RACE. For 5′ RACE, the first strand of cDNA was prepared using M-MLV (H-) Reverse Transcriptase (Vazyme Biotech). The specific RNAs in cDNA were removed by RNase H and RNase A (TaKara Biotech, San Jose, CA, USA). After purification by M5 Gel Extraction kit (Mei5 Biotech), add complementary homopolymeric tails to

cDNA with Terminal Deoxynucleotidyl Transferase (TaKara Biotech). Two rounds of PCR amplification were performed, with SPV2-441-R and Oliga-dT/dG as primers for the first round and SPV2-320-R and Anchor-primer as primers for the second round (Table 1). PCR products were purified and ligated into the pMD18-T vector (TaKara Biotech). The complete genome sequences of SPV2-SDYT and SPV2-JSXZ was obtained via sequence assembly after sequencing.

2.4 Sequence Analysis

2.4.1 Sequence and Consistency Analysis

To obtain the complete genome, sequence fragments of SPV2-SDYT and SPV2-JSXZ were assembled using ApE-A plasmid editor v2.0.45 and SnapGene 6.0.2 software. The complete genome sequences of 15 other SPV2 sequences obtained from the NCBI database were aligned together with SPV2-SDYT and SPV2-JSXZ complete genome sequences at the nucleotide and amino acid levels using SnapGene 6.0.2 software. The 5′ UTR, 3′ UTR and ORF sequences of SPV2-SDYT and SPV2-JSXZ and 15 other SPV2 sequences were figured out using ApE-A plasmid editor v2.0.45 software and used for one-to-one alignment with SnapGene 6.0.2 software. The P1, HC-Pro, P3, 6K1, CI, 6K2, VPg, NIa-Pro, NIb, and CP sequences of SPV2-SDYT and SPV2-JSXZ were determined by comparison with other SPV2 sequences known protein sequences and used for one-to-one alignment with SnapGene 6.0.2 software.

2.4.2 Phylogenetic Analysis

Phylogenetic tree was generated using the Neighbor-Joining method, with the No. of Bootstrap Replications set to 1000 and everything else left. The genetic distance was calculated with the pairwise method in MEGA11.

2.4.3 Codon Usage Bias Analysis

GC content of SPV2-SDYT and SPV2-JSXZ was calculated using online EMBOSS software. RSCU values were calculated with MEGA11 software. The RSCU value was higher than 1 if the codon was more preferred and utilised more frequently during gene expression. The RSCU value was 1 if the codon was used without preference. The RSCU value was less than 1 if the codon was low preferred and utilised less frequently during gene expression.

2.4.4 Recombination Analysis

The complete genome sequence alignment used for the recombination analysis was conducted with MEGA11 software. Recombination analysis was performed using seven algorithms, including RDP, CENECONV, BOOTSCAN, MAXCHI, CHIMAERA, SISCAN and 3SEQ in the RDP4 software and everything else left. Recombination existed on the basis of support for more than 6 kinds of algorithms, otherwise, no recombination existed on the basis of support for less than 6 kinds of algorithms [23,24].

2.4.5 Selection Pressure Analysis

For selection pressure analysis, the codon selection pressure of 10 proteins in 17 SPV2 isolates were calculated with SLAC in the online Datamonkey software (http://www.datamonkey.org). The dN/dS values were used to analyze the selection pressure. dN/dS value greater than 1 was indicated positive selection, dN/dS value equal to 1 was indicated neutral selection, and dN/dS value less than 1 was indicated negative selection.

3 Results

3.1 5′ RACE of SPV2-SDYT and SPV2-JSXZ

The 5′ end sequences of SPV2-SDYT and SPV2-JSXZ were obtained via 5′ RACE, with a second round of PCR amplification using the anchor primer and SPV2-320-R (Fig. 1A,B). SPV2-SDYT (5′-AAATTAACAAAACTCAATACAACAACAGAAGAATCAAGCAAACAAAAACAAACTCACGTTTTCAAAGTCTTTGAAGTTTACAATTCTCA-3′) and SPV2-JSXZ (5′-AAATTAACAAAACTCAATACAACAACAGAAGAATCAAGCAAACAAAAACAAACTCACGTTTTCAAAGTCTTTGAAGTTTACAATTCTCA-3′) are shown in Fig. 1A,B, respectively. Identical 5′ UTR sequences were observed in both isolates (Fig. 1A,B).

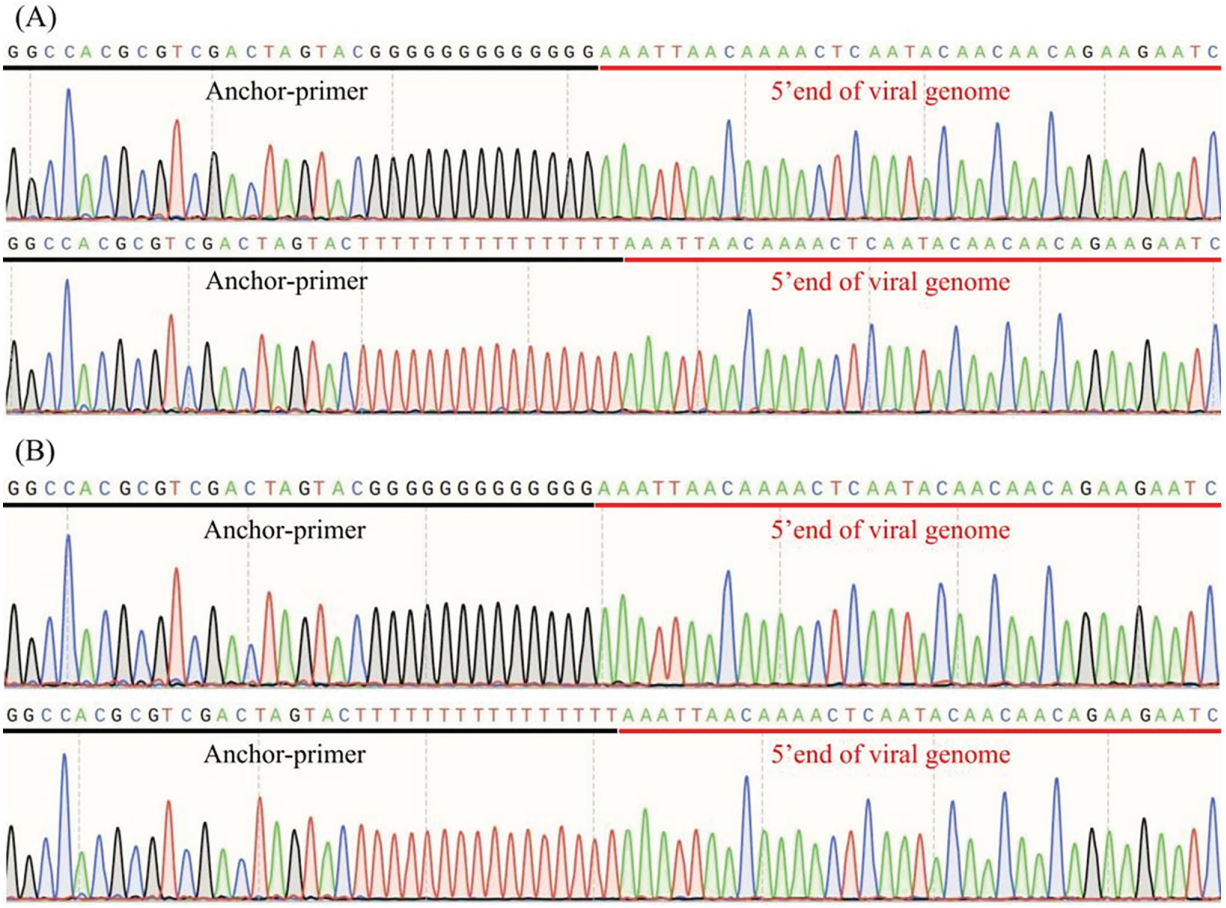

Figure 1: 5′ RACE analysis of SPV2-SDYT and SPV2-JSXZ. **(A)** Anchor-primer and 5′ end of viral genome sequences of SPV2-SDYT; **(B)** Anchor-primer and 5′ end of viral genome sequences of SPV2-JSXZ

3.2 Consistency Analysis of SPV2-SDYT and SPV2-JSXZ with 15 SPV2 Isolates

The sequence identities of SPV2-SDYT and SPV2-JSXZ with 15 other SPV2 isolates were analyzed at both the nucleotide and amino acid levels. The nucleotide and amino acid sequence identities of the ORFs of SPV2-SDYT and SPV2-JSXZ were 99.25% and 99.12%, respectively (Table 2). The complete genome sequence identities of SPV2-SDYT and SPV2-JSXZ with the other 15 SPV2 isolates ranged from 65.92% to 92.39% and 65.89% to 92.35%, respectively. The nucleotide sequence identities of the ORFs of SPV2-SDYT and

SPV2-JSXZ with these 15 isolates were between 66.33% and 92.61% and 66.30% and 92.57%, respectively. For the amino acid sequence identities of the ORFs, SPV2-SDYT and SPV2-JSXZ showed similarities of 62.97%–96.82% and 63.04%–96.55%, respectively, with the other 15 isolates. The 5′ UTR sequence identities of SPV2-SDYT and SPV2-JSXZ with the other 15 isolates were both in the range of 65.25%–69.23%. The 3′ UTR sequence identities for SPV2-SDYT and SPV2-JSXZ showed similarities of 48.74%–99.48% and 48.74%–98.97%, respectively, with the other 15 isolates. Except for the 5′ UTR sequence, both SPV2-SDYT and SPV2-JSXZ showed the highest sequence identity with SPV2-LN (GenBank No. OR842902.1) and the lowest sequence identity with SSBles-74 ZA (GenBank No.MH023310.1). The maximum variation in the above results was observed in the 3′ UTR, indicating the presence of mutation hotspots in this region.

Table 2: Sequence identities (%) between SPV2-SDYT and SPV2-JSXZ with other SPV2 isolates at the nucleotide/amino acid level

SPV2 isolate	Accession number	SDYT				JSXZ			
		Full Length	5′ UTR	Polyprotein ORF	3′ UTR	Full length	5′ UTR	Polyprotein ORF	3′ UTR
China JSXZ	PQ855661.1	99.26	100	99.25/99.12	99.48				
China SDYT	PQ855660.1					99.26	100	99.25/99.12	99.48
USA	NC017970.1	67.40	68.38	67.69/64.51	55.71	67.33	68.38	67.62/64.54	55.71
Australia	KX017448.1	67.11	67.52	67.39/64.43	55.12	67.01	67.52	67.28/64.45	55.12
East Timor	KX017447.1	67.33	65.25	67.65/64.60	55.25	67.17	65.25	67.49/64.59	55.25
USA	JN613807.1	67.40	68.38	67.69/64.50	55.71	67.34	68.38	67.62/64.54	55.71
China	OR842902.1	92.39	65.67	92.61/96.82	99.48	92.35	65.67	92.57/96.55	98.97
Greece	ON055393.1	67.61	67.80	67.91/64.85	55.71	67.45	67.80	67.73/64.92	55.71
	OK181779.1	67.38	66.95	67.69/64.64	55.25	67.24	66.95	67.54/64.66	55.25
South Africa	MN231254.1	67.71	65.25	68.04/64.61	55.71	67.70	65.25	68.02/64.61	55.71
South Africa	MH023310.1	65.92	68.38	66.33/62.97	48.74	65.89	68.38	66.30/63.04	48.74
Spain	KU511270.1	67.39	69.23	67.66/64.20	55.12	67.29	69.23	67.55/64.27	55.12
South Korea	KP115619.1	67.42	66.95	67.74/64.71	54.71	67.15	66.95	67.46/64.79	54.71
South Korea	KP115618.1	67.37	68.38	67.66/64.35	55.25	67.28	68.38	67.56/64.34	55.25
South Korea	KP115617.1	67.29	65.25	67.61/64.80	55.25	67.18	65.25	67.50/64.86	55.25
South Korea	KP115616.1	67.11	65.83	67.41/64.45	55.71	66.95	65.83	67.24/64.47	55.71
South Korea	KP115615.1	67.27	67.80	67.61/64.75	53.33	67.31	67.80	67.64/64.74	53.33

Ten proteins of SPV2-SDYT and SPV2-JSXZ along with 15 other SPV2 isolates were analyzed at both the nucleotide and amino acid levels. The nucleotide and amino acid identities of SPV2-SDYT with the 16 other SPV2 isolates were as follows (Table 3): 6K1: 58.18%–98.79% and 59.62%–98.08%, respectively; 6K2: 62.05%–99.37% and 50.98–100.00%, respectively; CI: 64.28%–99.59% and 64.85%–99.69%, respectively; P1: 88.92%–99.08% and 88.35%–98.71%, respectively; P3: 50.29%–98.30% and 28.10%–97.44%, respectively; Vpg: 62.87%–98.96% and 59.28%–99.48%, respectively; NIa: 61.43%–99.59% and 61.32%–100.00%, respectively; NIb: 63.85%–99.36% and 66.48%–99.23%, respectively; HC-Pro: 61.89%–99.20% and 56.99%–99.13%, respectively; CP: 60.48%–99.89% and 59.79%–100.00%, respectively.

Table 3: The ten proteins identities (%) of SPV2-SDYT with other SPV2 isolates at the nucleotide/amino acid level

SPV2 isolate	Accession number	SDYT									
		6K1	6K2	CI	P1	P3	Vpg	NIa	NIb	HC-Pro	CP
China JSXZ	PQ855661.1	98.79/96.15	99.37/100.00	99.59/99.69	99.08/98.71	98.30/97.44	98.96/99.48	99.59/100.00	99.36/99.23	99.20/99.13	99.89/100.00
USA	NC017970.1	62.26/61.54	62.72/50.98	64.86/65.47	97.14/96.60	51.59/29.78	64.29/60.00	62.26/61.32	67.83/67.44	62.46/57.42	60.73/59.79
Australia	KX017448.1	58.18/59.62	62.82/54.90	64.80/65.32	96.76/95.15	50.91/30.14	64.43/59.28	61.98/61.73	67.44/67.83	62.42/56.99	61.16/61.51
East Timor	KX017447.1	62.73/61.54	62.65/54.90	64.97/65.47	96.87/95.79	51.72/29.81	64.26/59.79	62.17/62.14	67.48/67.44	62.61/57.64	62.24/61.28
USA	JN613807.1	62.26/61.54	62.72/50.98	64.86/65.47	97.14/96.60	51.59/29.78	64.29/60.00	62.26/61.32	67.83/67.44	62.46/57.42	60.73/59.79
China	OR842902.1	93.59/98.08	88.68/90.57	91.91/98.76	95.95/94.01	87.89/94.32	93.62/97.93	95.20/98.77	94.14/97.30	87.69/97.83	95.56/97.61
Greece	ON055393.1	60.87/61.54	62.65/54.90	65.13/65.47	97.41/96.44	51.91/29.81	64.58/60.51	61.70/61.73	67.22/67.44	62.44/57.42	61.97/63.23
	OK181779.1	60.25/61.54	62.18/54.90	64.86/65.63	97.20/96.44	50.69/29.49	64.85/60.51	61.43/62.14	67.44/67.25	63.21/56.99	61.93/61.17
South Africa	MN231254.1	61.58/59.62	63.87/56.86	64.28/65.01	97.41/96.93	53.39/28.49	66.06/60.51	61.95/62.14	67.42/67.25	62.16/57.21	62.04/61.51
South Africa	MH023310.1	66.01/60.78	64.29/56.86	65.88/64.85	88.92/88.35	52.07/28.10	62.87/60.31	66.84/65.84	63.85/66.48	61.91/56.99	60.48/60.82
Spain	KU511270.1	61.64/61.54	63.80/54.90	64.99/65.32	96.66/95.79	51.59/29.81	64.32/60.51	61.47/61.73	67.64/67.64	62.53/57.64	61.72/61.17
South Korea	KP115619.1	60.38/61.54	62.65/54.90	64.88/65.32	96.82/96.44	52.38/29.81	64.65/60.00	61.86/62.14	67.67/67.64	61.89/57.64	62.59/61.51
South Korea	KP115618.1	61.64/61.54	63.25/54.90	64.69/65.32	97.52/93.85	51.24/30.14	64.20/60.51	61.86/62.14	67.73/67.64	62.21/57.42	62.11/61.17
South Korea	KP115617.1	62.26/61.54	63.03/54.90	65.09/65.47	97.41/96.93	50.29/29.44	64.11/60.51	61.95/62.14	67.58/67.64	62.10/57.42	62.25/61.51
South Korea	KP115616.1	61.64/61.54	62.05/54.90	64.71/65.16	96.55/95.79	51.42/29.20	63.73/60.00	61.90/62.14	67.71/67.54	62.33/57.64	61.11/61.17
South Korea	KP115615.1	62.89/61.54	62.65/54.90	64.93/65.47	96.82/96.76	50.78/29.17	64.81/60.51	61.73/62.14	67.75/67.83	62.17/57.21	61.30/61.51

Similarly, SPV2-JSXZ showed the following ranges of identity with the same isolates (Table 4): 6K1: 57.14%–98.79% and 61.54%–96.15%, respectively; 6K2: 59.26%–99.37% and 50.98%–100.00%, respectively; CI: 64.31%–99.59% and 64.70%–99.69%, respectively; P1: 88.92%–99.08% and 88.67%–98.71%, respectively; P3: 49.96%–98.30% and 28.37%–97.44%, respectively; VPg: 63.34%–98.96% and 59.28%–99.48%, respectively; NIa: 61.31%–99.59% and 61.32%–100.00%, respectively; NIb: 63.81%–99.36% and 66.41%–99.23%, respectively; HC-Pro: 61.95%–99.20% and 56.55%–99.13%, respectively; CP: 60.48%–99.89% and 59.79%–100.00%, respectively.

Table 4: The ten proteins identities (%) of SPV2-JSXZ with other SPV2 isolates at the nucleotide/amino acid level

SPV2 isolate	Accession number	JSXZ									
		6K1	6K2	CI	P1	P3	Vpg	NIa	NIb	HC-Pro	CP
China SDYT	PQ855660.1	98.79/96.15	99.37/100.00	99.59/99.69	99.08/98.71	98.30/97.44	98.96/99.48	99.59/100.00	99.36/99.23	99.20/99.13	99.89/100.00
USA	NC017970.1	62.89/63.46	62.13/50.98	65.03/65.32	97.09/96.93	50.87/30.00	64.34/60.00	62.14/61.32	67.62/67.64	62.73/56.99	60.73/59.79
Australia	KX017448.1	58.24/61.54	63.46/54.90	64.75/65.16	96.71/95.47	50.25/30.29	64.43/59.28	61.86/61.73	67.42/68.02	62.46/56.55	61.16/61.51
East Timor	KX017447.1	63.35/63.46	62.82/54.90	65.06/65.32	96.82/96.12	50.16/30.19	63.83/59.79	62.04/62.14	67.24/67.64	62.82/57.21	62.24/61.28
USA	JN613807.1	62.89/63.46	62.13/50.98	65.03/65.32	97.09/96.93	50.87/30.00	64.34/60.00	62.14/61.32	67.62/67.64	62.66/56.99	60.73/59.79
China	OR842902.1	92.99/94.23	89.31/90.57	91.91/98.44	96.01/94.17	87.13/92.90	93.62/98.45	95.34/98.77	94.40/97.30	87.47/97.39	95.68/97.61
Greece	ON055393.1	57.74/63.46	62.82/54.90	65.29/65.32	97.36/96.76	50.74/30.00	64.53/60.51	61.58/61.73	67.18/67.64	62.73/56.99	61.97/63.23
	OK181779.1	57.14/63.46	62.82/54.90	64.86/65.47	97.03/96.76	51.25/29.43	64.91/60.51	61.31/62.14	67.19/67.44	63.07/56.55	61.93/61.17
South Africa	MN231254.1	62.71/61.54	63.98/56.86	64.31/64.70	97.36/97.25	53.35/28.53	65.62/60.51	61.95/62.14	67.01/67.44	62.17/56.77	62.04/61.51
South Africa	MH023310.1	66.67/62.75	63.87/56.86	65.98/64.70	88.92/88.67	52.19/28.37	63.34/61.54	65.45/65.84	63.81/66.41	61.95/56.55	60.48/60.82
Spain	KU511270.1	62.26/63.46	63.76/54.90	65.16/65.16	96.60/96.12	50.78/30.00	64.38/60.51	61.35/61.73	67.40/67.83	62.58/57.21	61.72/61.17
South Korea	KP115619.1	61.01/63.46	59.26/54.90	64.88/65.16	96.76/96.76	50.82/30.00	64.38/60.00	61.73/62.14	67.07/67.83	62.05/57.21	62.59/61.51
South Korea	KP115618.1	62.26/63.46	59.88/54.90	64.70/65.16	97.47/94.01	50.57/30.29	64.10/60.51	61.73/62.14	67.48/67.83	62.35/56.99	62.11/61.17
South Korea	KP115617.1	62.89/63.46	62.82/54.90	65.14/65.32	97.36/97.25	49.96/30.47	64.43/60.51	61.83/62.14	67.54/67.83	62.24/56.99	62.25/61.51
South Korea	KP115616.1	62.26/63.46	62.18/54.90	64.72/65.01	96.39/95.79	50.24/29.55	64.32/60.00	61.78/62.14	67.46/67.77	62.47/57.21	61.11/61.17
South Korea	KP115615.1	63.52/63.46	63.76/54.90	65.09/65.32	96.55/96.76	50.99/30.19	64.87/60.51	61.61/62.14	67.56/68.02	62.31/56.77	61.30/61.51

In both SPV2-SDYT and SPV2-JSXZ, the P1 protein showed the highest levels of identity, while the P3 protein exhibited the lowest.

3.3 Phylogenetic and Genetic Distance Analysis

To clarify the phylogenetic relationship between SPV2-SDYT and SPV2-JSXZ and the other 15 reported SPV2 strains, a phylogenetic tree was constructed using the NJ method in MEGA11 (Fig. 2). The results showed that the samples clustered into two groups. The first group included the United States (GenBank No. NC_017970.1, GenBank No. JN613807.1), Australia (GenBank No. KX017448.1), Spain (GenBank No.

KU511270.1), Greece (GenBank No. ON055393.1), Korea (GenBank No. KP115619.1, GenBank No. KP115618.1, GenBank No. KP115617.1, GenBank No. KP115616.1, GenBank No. KP115615.1), Timor-Leste (GenBank No. KX017447.1), and South African isolates (GenBank No. MN231254.1, GenBank No. MH023310.1). Fourteen isolates were included in this study. Of these, the South African isolate (GenBank No. MH023310.1) was far from the others, indicating that it was distantly related to the other isolates in this group. The second group included SPV2-SDYT (GenBank No. PQ855660), SPV2-JSXZ (GenBank No. PQ855661), and SPV2-LN (GenBank No. OR842902.1), indicating that the three Chinese isolates were distantly related to the 14 isolates from other countries and that SPV2 evolution may have been geographically influenced.

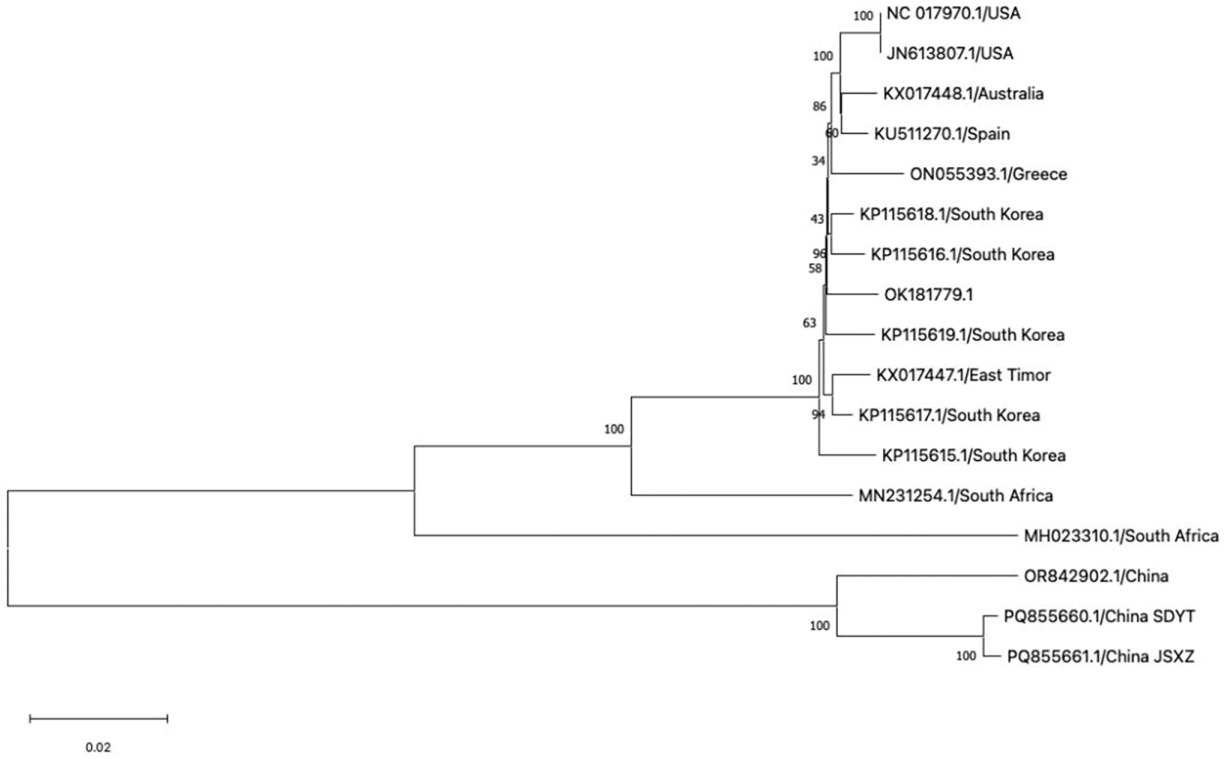

Figure 2: Phylogenetic tree of the 17 SPV2 isolates from different countries. The bootstrap value is 1000. The scale bar stands for a genetic distance of 0.020

To further clarify the degree of genetic variation in the 17 SPV2 isolates that had been reported so far by NCBI, including SPV2-SDYT and SPV2-JSXZ, genetic distance was calculated using the pairwise method in MEGA11 (Table 5). The results showed that the genetic distances of SPV2-SDYT and SPV2-JSXZ from SPV2-LN (GenBank No. OR842902.1) were the lowest, with values of 0.0498 and 0.0502, respectively, whereas their genetic distance from SSBles-74 ZA (GenBank No. MH023310.1) was the highest, with values of 0.2908 and 0.2912, respectively. The Fst values of the 17 SPV2 isolates were greater than 0.25, except for that of SPV2-LN (GenBank No. OR842902.1), indicating a significant degree of genetic differentiation among SPV2 populations. These results are consistent with those of the phylogenetic tree analysis, indicating that geography may be an important factor influencing the genetic evolution of SPV2.

Table 5: Genetic distances of 17 SPV2 isolates

	PQ855660.1	PQ855661.1	NC_017970.1	KX017448.1	KX017447.1	JN613807.1	OR842902.1	ON055393.1	OK181779.1	MN231254.1	MH023310.1	KU511270.1	KP115619.1	KP115618.1	KP115617.1	KP115616.1	KP115615.1
PQ855660.1																	
PQ855661.1	0.0047																
NC_017970.1	0.2696	0.2700															
KX017448.1	0.2708	0.2710	0.0111														
KX017447.1	0.2687	0.2695	0.0154	0.0146													
JN613807.1	0.2696	0.2700	0.0000	0.0111	0.0154												
OR842902.1	0.0498	0.0502	0.2718	0.2722	0.2717	0.2718											
ON055393.1	0.2680	0.2684	0.0176	0.0172	0.0185	0.0176	0.2712										
OK181779.1	0.2697	0.2704	0.0150	0.0147	0.0150	0.0150	0.2723	0.0185									
MN231254.1	0.2679	0.2690	0.0688	0.0681	0.0641	0.0688	0.2714	0.0730	0.0681								
MH023310.1	0.2908	0.2912	0.1522	0.1542	0.1543	0.1522	0.2930	0.1574	0.1549	0.1496							
KU511270.1	0.2694	0.2697	0.0099	0.0091	0.0130	0.0099	0.2726	0.0156	0.0137	0.0660	0.1536						
KP115619.1	0.2695	0.2702	0.0149	0.0142	0.0145	0.0149	0.2742	0.0185	0.0147	0.0677	0.1559	0.0128					
KP115618.1	0.2675	0.2678	0.0113	0.0109	0.0109	0.0113	0.2709	0.0146	0.0111	0.0653	0.1524	0.0093	0.0109				
KP115617.1	0.2677	0.2682	0.0124	0.0118	0.0085	0.0124	0.2704	0.0156	0.0123	0.0620	0.1545	0.0101	0.0118	0.0082			
KP115616.1	0.2704	0.2708	0.0129	0.0120	0.0126	0.0129	0.2735	0.0163	0.0129	0.0662	0.1547	0.0110	0.0125	0.0081	0.0098		
KP115615.1	0.2693	0.2702	0.0171	0.0165	0.0165	0.0171	0.2729	0.0208	0.0170	0.0677	0.1546	0.0148	0.0155	0.0131	0.0141	0.0150	

3.4 Codon Preference in SPV2-SDYT and SPV2-JSXZ

Codon usage bias analysis was performed to explore the preferences of SPV2-SDYT and SPV2-JSXZ for different codons during translation. Relative synonymous codon usage (RSCU) values were calculated using the MEGA11 software. The results showed that the RSCU values of SPV2-SDYT and SPV2-JXSZ were similar, with the most and least abundant amino acids being leucine (Leu: 8.8%) and tryptophan (Trp: 1.2%), respectively (Fig. 3A,B). The first, second, and third codon GC contents (GC1, GC2, and GC3, respectively) in SPV2-SDYT were 50.79%, 36.75%, and 42.94%, respectively. The first, second, and third codon GC contents (GC1, GC2, and GC3, respectively) in SPV2-JSXZ were 50.70%, 36.69%, and 42.99%, respectively. The frequency of the third codon GC content in SPV2-SDYT and SPV2-JSXZ was lower than that in A/U, indicating that the codon terminal bases were mainly A/U. It can be inferred that the codon preference of the two isolates was greatly affected by the pressure of natural selection. There were 26 codons with RSCU values greater than 1 and 35 codons with values less than 1 in SPV2-SDYT. There were 27 codons with RSCU values greater than 1 and 34 codons with values less than 1 in SPV2-JSXZ. There were fewer codons with RSCU values greater than 1 than with values less than 1, indicating that the overall codon usage bias was relatively weak for SPV2-SDYT and SPV2-JSXZ.

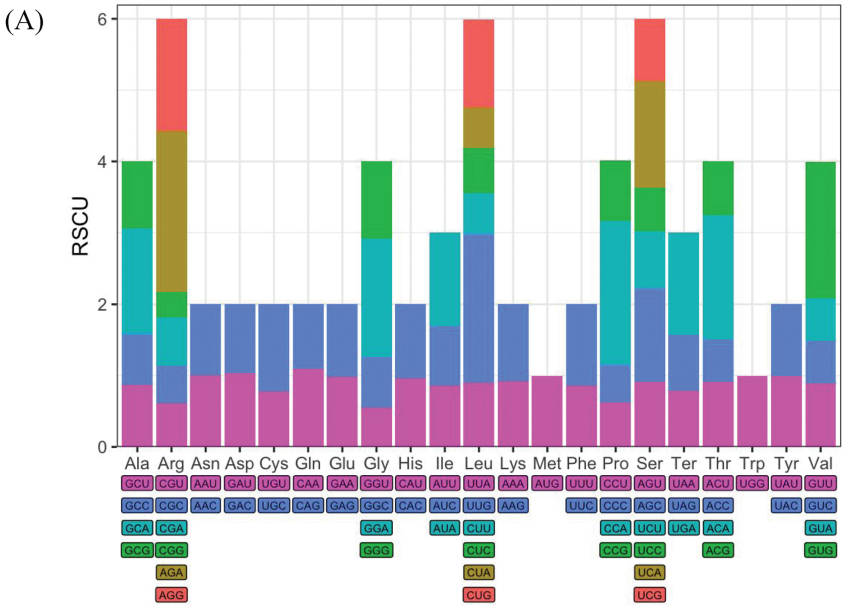

Figure 3: (Continued)

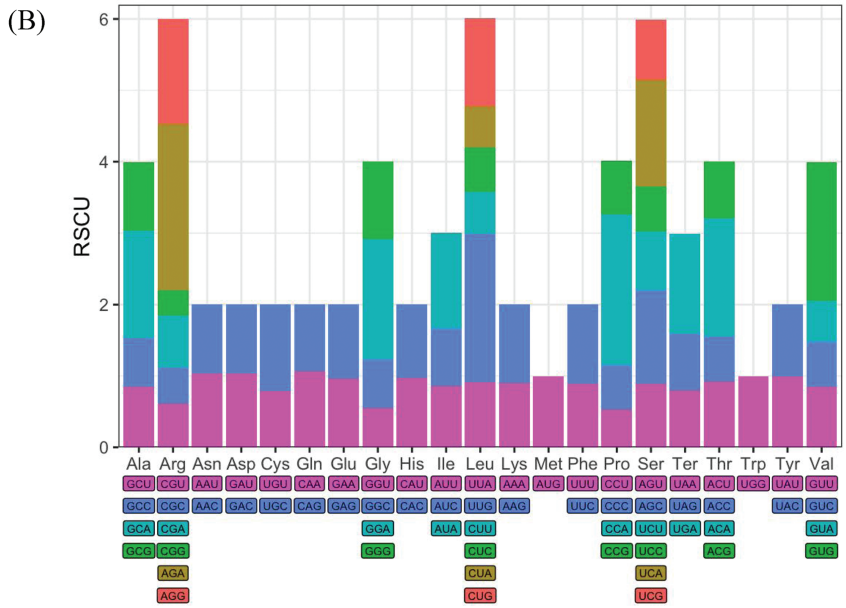

Figure 3: Condon preference analysis. **(A)** The RSCU value of codons in SPV2-SDYT; **(B)** The RSCU value of codons in SPV2-JSXZ. RSCU value greater than 1 means that the codon is relatively frequently used

3.5 Recombination Analysis

We further performed recombination analysis using RDP4 and identified six recombination sites in 17 SPV2 isolates, with three recombination sites found in the South African isolate (GenBank No. MH023310.1), one recombination site found in the South African isolate (GenBank No. MN231254.1), and two recombination sites found in the Greek isolate (GenBank No. ON055393.1) (Table 6). Among the six recombination sites, four were supported by seven types of algorithms and two were supported by six types of algorithms, with SSBles-74 ZA (GenBank No. MH023310.1), which showed the most recombination sites and covered the complete genome, implying an increase in genetic diversity (Fig. 4).

Table 6: Summary of recombination sites occurred in SPV2

Recombinant sequence (s)	Region	Parental sequences minor/Major	*p*-value for the seven detection methods in RDP4						
			RDP	GENECONV	BootScan	MaxChi	Chimaera	SiScan	3Seq
MH023310.1	1406–1976	Unknown (PQ855661.1)/ MN231254.1	2.67E−75	2.71E−93	8.24E−78	8.05E−24	1.17E−11	5.14E−26	3.73E−13
	1405–10688	JN613807.1/ Unknown (KP115619.1)	1.32E−53	4.07E−52	6.87E−54	1.14E−05	3.57E−11	6.42E−15	2.49E−13
	9350–9523	MN231254.1/ Unknown (KX017448.1)	1.44E−12	NS	4.65E−08	1.11E−03	1.70E−05	2.79E−03	5.86E−09
MN231254.1	2526–9587	KP115617.1/ Unknown (ON055393.1)	3.20E−74	1.07E−33	1.44E−21	2.26E−23	3.35E−24	2.44E−16	1.41E−63

(Continued)

Table 6 (continued)

Recombinant sequence (s)	Region	Parental sequences minor/Major	*p*-value for the seven detection methods in RDP4						
			RDP	GENECONV	BootScan	MaxChi	Chimaera	SiScan	3Seq
ON055393.1	10090–10547	Unknown (MH023310.1)/ KP115618.1	1.93E−47	4.09E−21	2.00E−21	1.07E−12	2.35E−11	2.14E−09	2.49E−13
	7411–7788	Unknown (MN231254.1)/ KP115619.1	4.77E−27	4.87E−31	NS	5.18E−03	1.85E−05	3.40E−11	1.24E−13

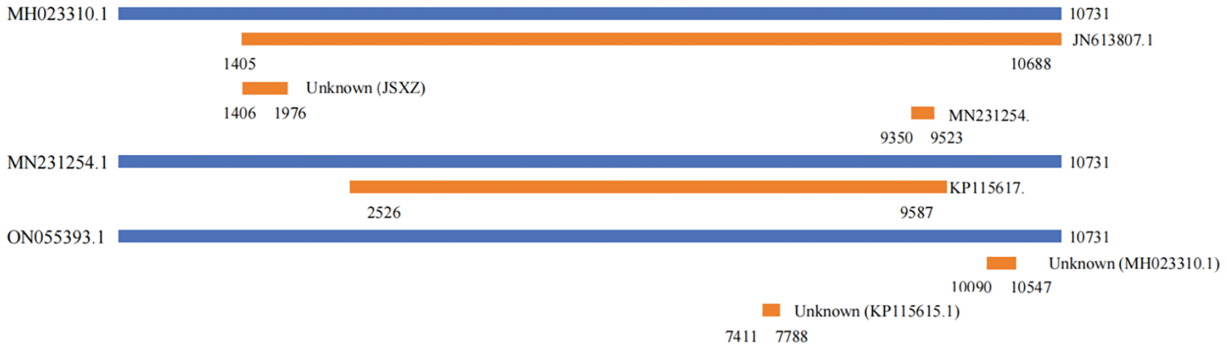

Figure 4: Recombinant analysis of SPV2 complete genomic sequence

3.6 Selection Pressure Analysis

To analyze the selection pressure of SPV2 during the evolutionary process, dN/dS values were calculated for 10 proteins, including P1, HC-Pro, P3, 6K1, CI, 6K2, Vpg, NIa, NIb, and CP, in SPV2 isolates using the online Datamonkey software. The results showed that the dN/dS values of all SPV2 proteins were less than 1, and none had negative selection sites. The highest number of negative selection sites was observed in the CI protein (160 sites), whereas the lowest number of negative selection sites was observed in the CP protein (10 sites) (Table 7). The purifying selection pressure was unevenly distributed across the complete genome, with the strongest purifying selection observed in HC-Pro (dN/dS = 0.0554) and the weakest in P1 (dN/dS = 0.340) (Table 7). These findings indicate that negative selection plays a crucial role in the evolution of SPV2.

Table 7: Selection pressure analysis of SPV2

Gene	d_N/d_S	Sites under positive selection	Sites under negative selection
P1	0.340	0	20
HC-Pro	0.0554	0	36
P3	0.141	0	61
6K1	0.0889	0	14
CI	0.0638	0	160
6K2	0.123	0	12

(Continued)

Table 7 (continued)

Gene	d_N/d_S	Sites under positive selection	Sites under negative selection
Vpg	0.113	0	32
NIa	0.0853	0	43
NIb	0.103	0	83
CP	0.193	0	10

4 Discussion

SPV2 is widespread in China, and mixed infections with other viruses can have a serious impact on the development of the sweet potato industry and cause severe economic losses worldwide. In this study, we obtained the complete genome sequences of SPV2-SDYT and SPV2-JSXZ using 5′ RACE and RT-PCR amplification and analyzed the structural characteristics and phylogenetic relationships of the SPV2 genome by aligning sequence identities, measuring genetic distances, and constructing phylogenetic evolutionary trees (Figs. 1A,B and 2). Genome size may differ within the same species owing to various factors, such as gene duplication and the environment [25]. The complete genome sequences of the 17 SPV2 isolates ranged from 10,561 to 10,732 nt, with SPV2-SDYT and SPV2-JSXZ being 10,561 nt in length. According to the phylogenetic tree, geoenvironmental factors strongly influenced the results. We found that one group contained three Chinese isolates and the other contained isolates from other countries. Among the isolates from other countries, the South African isolate was distantly related, whereas the isolates from the United States, Australia, Spain, Greece, Korea, and East Timor were closely related (Fig. 2). This may be related to the high genetic diversity of the SPV2 strains. The appearance of SPV2-SDYT and SPV2-JSXZ not only enriched the SPV2 Chinese isolates and showed interspecific evolution but also provided a scientific basis for understanding the origin and evolution of viruses.

Codon usage bias can be affected by various factors, including GC content, gene size and location, tRNA abundance and interactions, gene recombination rates, mutations, natural selection, and mRNA folding. As high tRNA abundance is better able to participate in translation and improve translation efficiency, most highly expressed genes tend to use codons with high tRNA abundance [26]. We observed that codon preference was closely related to tRNA abundance and interactions in SPV2-SDYT and SPV2-JSXZ (Fig. 3A,B). GC content is another important factor that influences codon preference. Codons containing G and C were mostly found in genomes with high GC content, whereas those containing A and U were mostly found in genomes with low GC content [27]. The frequency of codon GC content in SPV2-SDYT and SPV2-JSXZ was lower than that in A/U, suggesting that the preferred codons in the genome were mainly A/U. The codon with the largest RSCU in SPV2-SDYT and SPV2-JSXZ was AGA, at 2.25 and 2.33, respectively.

Recombination plays a crucial role in the evolution of viral genomes by enhancing genetic diversity and facilitating the emergence of new viruses. It also serves as a valuable tool for vaccine development and gene therapy [28]. Six recombination events were observed in SSBles-74 ZA, SPV2, and SPV2-SP1, indicating that recombination was not the main driving force for SPV2 evolution (Table 6) (Fig. 4).

Selection pressure is a driving force of biological evolution. The nonsynonymous substitution rate (dN) can result in amino acid changes, whereas the synonymous substitution rate (dS) cannot [29]. The direction and rate of gene evolution, identification of functionally important genes, and functional divergence of genes are affected by dN/dS values [30,31]. We observed that the dN/dS values for all SPV2 proteins were less than 1, indicating that negative selection had occurred (Table 7). Negative selection pressure may reduce variation,

suggesting that SPV2 has evolved to retain only mutations that are favorable for survival. Our observations also indicated that negative selection may be the main driving force of SPV2.

5 Conclusion

In this study, complete genomic sequences of SPV2 isolates from the Shandong and Jiangsu provinces in China were obtained using 5′ RACE and RT-PCR amplification. To the best of our knowledge, this is the first report of the complete genome sequence of sweet potatoes from China. Consistency, phylogeny, codon usage bias, recombination, and selection pressure analyses indicated that they were closely related to the SPV2 isolate from China (SPV2-LN). The complete genome sequence identities of SPV2-SDYT and SPV2-JSXZ with the SPV2 isolate from China (SPV2-LN) were 92.39% and 92.35%, respectively. The three Chinese isolates belonged to the same group in the phylogenetic tree, and geographic factors affected the evolution of SPV2. High-frequency codons in the genomes were predominantly found to end with A/U. Recombination was not the main driver of SPV2 evolution, negative selection may have been the main driver. SPV2 can infest sweet potatoes singly and cause mild symptoms, whereas the synergism of sweet potatoes with other sweet potato viruses can cause severe symptoms and greater economic losses.

Acknowledgement: We are grateful for the critical reading by Prof. Nianwei Qiu at Qufu Normal University.

Funding Statement: This work was funded by the National Natural Science Foundation of China (32100132), Shandong Province Natural Sciences Foundation of China (ZR2021QC008), 'Youth Innovation Team Program' in College of Shandong Province of China (2022KJ119), and supported by Young Talent of Lifting Engineering for Science and Technology in Shandong, China (SDAST2024QT085).

Author Contributions: Zichen Li: Investigation, methodology, resources, writing—original draft. Jukui Ma: Conceptualization, formal analysis, supervision. Minjun Liu: methodology, validation. Guowei Geng: Conceptualization, methodology, validation, formal analysis, investigation, writing—original draft, writing—review & editing, visualization, funding acquisition. Hongxia Zhang: Conceptualization, methodology, resources, writing—original draft, writing—review & editing, supervision, project administration. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The authors confirm that the data supporting the findings of this study are available within the article.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Bovell-Benjamin AC. Sweet potato: a review of its past, present, and future role in human nutrition. Adv Food Nutr Res. 2007;52(9):1–59. doi:10.1016/S1043-4526(06)52001-7.
2. Behera S, Chauhan VBS, Pati K, Bansode V, Nedunchezhiyan M, Verma AK, et al. Biology and biotechnological aspect of sweet potato (*Ipomoea batatas* L.): a commercially important tuber crop. Planta. 2022;256(2):40. doi:10.1007/s00425-022-03938-8.
3. Yang Y, Chen Y, Bo Y, Liu Q, Zhai H. Research progress in the mechanisms of resistance to biotic stress in sweet potato. Genes. 2023;14(11):2106. doi:10.3390/genes14112106.
4. Kokkinos CD, Clark CA, Mcgregor CE. The effect of sweet potato virus disease and its viral components on gene expression levels in sweetpotato. J Am Soc Hortic Sci. 2006;131(5):657–66. doi:10.1007/978-3-540-38918-76602.
5. Jones RAC. Global plant virus disease pandemics and epidemics. Plants. 2021;10(2):233. doi:10.3390/plants10020233.

6. Wei KJ, Jiang AM, Jiang S, Huang YJ, Jiang SY, Su XL, et al. New isolate of sweet potato virus 2 from *Ipomoea nil*: molecular characterization, codon usage bias, and phylogenetic analysis based on complete genome. Virol J. 2024;21(1):222. doi:10.1186/s12985-024-02500-0.
7. Liu H, Zhao F, Qiao Q, Zhang D, Wang Y, Wang S, et al. Complete genome sequence of a divergent sweet potato chlorotic stunt virus isolate infecting *Calystegia hederacea* in China. Arch Virol. 2021;166(7):2037–40. doi:10.1007/s00705-021-05076-0.
8. Mukasa SB, Rubaihayo PR, Valkonen JPT. Interactions between a crinivirus, an ipomovirus and a potyvirus in coinfected sweetpotato plants. J Plant Pathol. 2010;55(3):458–67. doi:10.1111/j.1365-3059.2006.01350.x.
9. Untiveros M, Fuentes S, Salazar LF. Synergistic interaction of sweet potato chlorotic stunt virus (Crinivirus) with Carla-, Cucumo-, Ipomo-, and potyviruses infecting sweet potato. Plant Dis. 2007;91(6):669–76. doi:10.1094/PDIS-91-6-0669.
10. Chase O, Javed A, Byrne MJ, Thuenemann EC, Lomonossoff GP, Ranson NA, et al. CryoEM and stability analysis of virus-like particles of potyvirus and ipomovirus infecting a common host. Commun Biol. 2023;6(1):433. doi:10.1038/s42003-023-04799-x.
11. Clark CA, Davis JA, Abad JA, Cuellar WJ, Fuentes S, Kreuze JF, et al. Sweetpotato viruses: 15 years of progress on understanding and managing complex diseases. Plant Dis. 2012;96(2):168–85. doi:10.1094/PDIS-07-11-0550.
12. Bednarek R, David M, Fuentes S, Kreuze J, Fei Z. Transcriptome analysis provides insights into the responses of sweet potato to sweet potato virus disease (SPVD). Virus Res. 2021;295:198293. doi:10.1016/j.virusres.2020.198293.
13. Wang S, Liu ST, Qiao Q, Zhang DS, Qin YH, Zhang ZC. Methodology for identification of disease resistance of sweet potato cultivars to sweet potato virus disease and yield loss estimation. J Plant Prot. 2014;41(2):176–81. doi:10.13802/j.cnki.zwbhxb.2014.02.009.
14. Zhang ZC, Qiao Q, Qin YH, Zhang DS, Tian YT. First evidence for occurrence of sweet potato virus disease (SPVD) caused by dual infection of Sweet potato feathery mottle virus and Sweet potato chlorotic stunt virus in China. Acta Phytopathol Sinica. 2012;42(3):328–33. doi:10.13926/j.cnki.apps.2012.03.014.
15. Trenado HP, Lozano G, Valverde RA, Navas-Castillo J. First report of sweet potato virus G and sweet potato virus 2 infecting sweet potato in Spain. J Plant Disease. 2010;91(12):1687–7. doi:10.1094/PDIS-91-12-1687C.
16. Tairo F, Jones RAC, Valkonen JPT. Potyvirus complexes in sweetpotato: occurrence in Australia, serological and molecular resolution, and analysis of the sweet potato virus 2 (SPV2) component. Plant Dis. 2006;90(9):1120–8. doi:10.1094/PD-90-1120.
17. Kwak HR, Kim J, Kim MK, Seo JK, Jung MN, Kim JS, et al. Molecular characterization of five potyviruses infecting korean sweet potatoes based on analyses of complete genome sequences. Plant Pathol J. 2015;31(4):388–401. doi:10.5423/PPJ.OA.04.2015.0072.
18. Qin Y, Zhang Z, Qiao Q, Zhang D, Tian Y, Wang Y. Molecular variability of sweet potato chlorotic stunt virus (SPCSV) and five potyviruses infecting sweet potato in China. Arch Virol. 2013;158(2):491–5. doi:10.1007/s00705-012-1503-8.
19. Urcuqui-Inchima S, Haenni AL, Bernardi F. Potyvirus proteins: a wealth of functions. Virus Res. 2001;74(1–2):157–75. doi:10.1016/s0168-1702(01)00220-9.
20. Miao R, Ye Z, MacFarlane S, Li Y, Mao Q, Tian Y, et al. The P3N-PIPO protein encoded by wheat yellow mosaic virus is a pathogenicity determinant and promotes its pathogenicity through interaction with NbRLK6 in nicotiana benthamiana. Viruses. 2022;14(10):2171. doi:10.3390/v14102171.
21. Rodamilans B, Valli A, Mingot A, San León D, López-Moya JJ, García JA. An atypical RNA silencing suppression strategy provides a snapshot of the evolution of sweet potato-infecting potyviruses. Sci Rep. 2018;8(1):15937. doi:10.1038/s41598-018-34358-y.
22. Li F, Xu D, Abad J, Li R. Phylogenetic relationships of closely related potyviruses infecting sweet potato determined by genomic characterization of Sweet potato virus G and Sweet potato virus 2. Virus Genes. 2012;45(1):118–25. doi:10.1007/s11262-012-0749-2.
23. Martin DP, Lemey P, Posada D. Analysing recombination in nucleotide sequences. Mol Ecol Resour. 2011;11(6):943–55. doi:10.1111/j.1755-0998.2011.03026.x.

24. Maag JLV. gganatogram: an R package for modular visualisation of anatograms and tissues based on ggplot2. F1000Research. 2018;7:1576. doi:10.12688/f1000research.16409.2.
25. Faizullah L, Morton JA, Hersch-Green EI, Walczyk AM, Leitch AR, Leitch IJ. Exploring environmental selection on genome size in angiosperms. Trends Plant Sci. 2021;26(10):1039–49. doi:10.1016/j.tplants.2021.06.001.
26. Wan XF, Xu D, Kleinhofs A, Zhou J. Quantitative relationship between synonymous codon usage bias and GC composition across unicellular genomes. BMC Evol Biol. 2004;4(1):19. doi:10.1186/1471-2148-4-19.
27. Parvathy ST, Udayasuriyan V, Bhadana V. Codon usage bias. Mol Biol Rep. 2022;49(1):539–65. doi:10.1007/s11033-021-06749-4.
28. Stedman KM. Viral recombination: ecology, evolution, and pathogenesis. Viruses. 2018;10(7):358. doi:10.3390/v10070358.
29. Shen X, Song S, Li C, Zhang J. Synonymous mutations in representative yeast genes are mostly strongly non-neutral. Nature. 2022;606(7915):725–31. doi:10.1038/s41586-022-04823-w.
30. Mugal CF, Wolf JBW, Kaj I. Why time matters: codon evolution and the temporal dynamics of d*N*/d*S*. Mol Biol Evol. 2014;31(1):212–31. doi:10.1093/molbev/mst192.
31. Spielman SJ, Wilke CO. The relationship between dN/dS and scaled selection coefficients. Mol Biol Evol. 2015;32(4):1097–108. doi:10.1093/molbev/msv003.