Received: 29 September 2023 | Revised: 27 February 2024 | Accepted: 18 April 2024

DOI: 10.1002/pld3.589

RESEARCH ARTICLE

A survey of the *Sli* gene in wild and cultivated potato

Mercedes Ames[1] | Andy Hamernik[1] | William Behling[2] | David S. Douches[2] | Dennis A. Halterman[1] | Paul C. Bethke[1]

[1]US Department of Agriculture, Agricultural Research Service, Vegetable Crops Research Unit, Department of Horticulture, University of Wisconsin, Madison, Wisconsin, USA

[2]Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, Michigan, USA

Correspondence
Paul C. Bethke, US Department of Agriculture, Agricultural Research Service, Vegetable Crops Research Unit, Department of Horticulture, University of Wisconsin, 1575 Linden Drive, Madison, WI 53706, USA.
Email: paul.bethke@ars.usda.gov

Funding information
USDA Agricultural Research Service, Grant/Award Numbers: 0500-00093-001-00-D, 5090-21220-005-000-D; USDA NIFA Specialty Crop Research Initiative, Grant/Award Number: 2019-51181-30021

Abstract

Inbred-hybrid breeding of diploid potatoes necessitates breeding lines that are self-compatible. One way of incorporating self-compatibility into incompatible cultivated potato (*Solanum tuberosum*) germplasm is to introduce the *S-locus* inhibitor gene (*Sli*), which functions as a dominant inhibitor of gametophytic self-incompatibility. To learn more about *Sli* diversity and function in wild species relatives of cultivated potato, we obtained *Sli* gene sequences that extended from the 5′UTR to the 3′UTR from 133 individuals from 22 wild species relatives of potato and eight diverse cultivated potato clones. DNA sequence alignment and phylogenetic trees based on genomic and protein sequences show that there are two highly conserved groups of *Sli* sequences. DNA sequences in one group contain the 533 bp insertion upstream of the start codon identified previously in self-compatible potato. The second group lacks the insertion. Three diploid and four polyploid individuals of wild species collected from geographically disjointed localities contained *Sli* with the 533 bp insertion. For most of the wild species clones examined, however, *Sli* did not have the insertion. Phylogenetic analysis indicated that *Sli* sequences with the insertion, in wild species and in cultivated clones, trace back to a single origin. Some diploid wild potatoes that have *Sli* with the insertion were self-incompatible and some wild potatoes that lack the insertion were self-compatible. Although there is evidence of positive selection for some codon positions in *Sli*, there is no evidence of diversifying selection at the gene level. In silico analysis of Sli protein structure did not support the hypothesis that amino acid changes from wild-type (no insertion) to insertion-type account for changes in protein function. Our study demonstrated that genetic factors besides the *Sli* gene must be important for conditioning a switch in the mating system from self-incompatible to self-compatible in wild potatoes.

KEYWORDS
diploid potato breeding, self-compatibility locus, *S-locus* inhibitor protein modeling, wild species germplasm

1 | INTRODUCTION

Potato (*Solanum tuberosum*) is an important food crop that is consumed by over a billion people each day (Devaux et al., 2021). Potato is an autopolyploid (2n = 4x = 48) that has a high degree of heterozygosity and numerous deleterious alleles in its genome (Spooner et al., 2014; Wu et al., 2023; Zhang et al., 2019). New potato varieties introduced over the past century have been identified using phenotypic selection of F1 populations. This approach limits the detrimental effects of deleterious alleles by maintaining a high degree of heterozygosity. An alternative method for potato breeding based on diploid potatoes has been proposed (Jansky et al., 2016; Lindhout et al., 2011), and diploid potato breeding efforts are underway in the United States, Asia, and Europe (Bradshaw, 2022; Stokstad, 2019; Zhang et al., 2021). This new method envisions potato varieties produced by hybridization of inbred parental breeding lines and distributed in the first generation as botanical seed. Inbred-hybrid breeding of diploid potatoes is predicted to have a higher rate of genetic gain than the phenotypic selection method currently used for breeding tetraploid potatoes (Bethke et al., 2022; Jansky et al., 2016). Maintaining self-compatibility through multiple rounds of inbreeding and trait introgression is a requirement for generating the highly inbred lines used in this breeding scheme.

The tuber-bearing Solanum (section *Petota*) contains over 100 species. These wild potato relatives grow in highly diverse habitats and are a rich resource for disease, pest, and abiotic stress resistance traits (Bashir et al., 2021). Ploidy levels in wild potato relatives range from diploid to hexaploid. Most potato wild relatives are diploid or tetraploid and can be used readily for germplasm enhancement (Spooner et al., 2014). A phylogeny based on nuclear gene data (Spooner et al., 2014) organized the tuber-bearing Solanum into three groups, one comprised by species from North and Central America (Clades 1 and 2 in Spooner & Castillo, 1997), a second with species from northern Peru and Ecuador (Clade 3), and the third with species from the rest of South America and *Solanum verrucosum* from Mexico (clade 4).

Most diploid potatoes, like other members of the Solanaceae, are self-incompatible due to a gametophytic self-incompatibility system controlled by the multi allelic S-locus. The S-locus encodes S-locus RNAse proteins (S-RNAses) that are produced by the female and S-locus F-box proteins (SLFs) produced by the male (Dzidzienyo et al., 2016., McClure et al., 2011). In self-incompatible interactions, S-RNAse activity limits the growth of pollen tubes and prevents fertilization. In compatible interactions, maternal S-RNAses are targeted for degradation by E3 ubiquitin ligase complexes in which paternal SLFs are the F-box subunits (Kubo et al., 2010).

Self-incompatibility is the rule for diploid potatoes, but exceptions occur, and self-compatible individuals have been identified in species that are largely self-incompatible (Cipar et al., 1964 and references therein; Eijlander et al., 2000; Hanneman, 1985; Hosaka & Hanneman, 1998; Ortiz & Mihovilovich, 2020). Examples include *Solanum chacoense and Solanum kurtzianum*. *S. verrucosum* is considered a self-compatible species as most of its populations are self-compatible (Spooner et al., 2014).

Hosaka and Hanneman (1998) determined that a single dominant gene, the S-locus inhibitor (*Sli*) gene, conferred self-compatibility to an individual of the predominantly self-incompatible species *S. chacoense*. *Sli* was mapped to the distal end of chromosome 12 (Clot et al., 2020; Eggers et al., 2021; Hosaka & Hanneman, 1998; Ma et al., 2021). Ma et al. (2021) determined that *Sli* is a non-S-locus F-box gene whose product is capable of interacting with S-locus RNAses and functions as an S-RNAse inhibitor. Eggers et al. (2021) identified a 533 bp insertion in the promoter region upstream of the *Sli* start codon and defined it as a gain of function mutation present exclusively in self-compatible lines. *Sli* sequences with the insertion were expressed in pollen tubes and promoted self-compatibility. Clot et al. (2020) identified the self-compatible haplotype corresponding to the *Sli* gene in several tetraploid potato varieties and concluded it was widespread in cultivated potatoes.

In the present study, we sequenced *Sli* in wild species relatives and cultivated potatoes from the 5′UTR to the 3′UTR using long PCR amplification and long-read, high fidelity amplicon sequencing to determine if the *Sli* gene carrying the 533 bp insertion is present in wild potato germplasm besides *S. chacoense* genotypes M6 (Jansky et al., 2014) and 525-3 (Hosaka & Hanneman, 1998). An additional goal was to characterize *Sli* sequence diversity in a wide range of wild potato species. We used DNA sequence information and phylogenetic analysis to evaluate the origin of the *Sli* haplotype that carries the 533 bp insertion. We also used sequence information to model changes in Sli protein structure and make predictions regarding protein function of *Sli* alleles that contain or lack the 533 bp insertion. Finally, we compared *Sli* DNA sequence data with data on self-compatibility for 116 individuals from 10 wild species and cultivated potato to determine if *Sli* sequences are useful predictors of self-compatibility in wild potato species.

2 | MATERIALS AND METHODS

2.1 | Plant material

Wild potato (2n = 2x = 24) plants were grown from sexual seeds. Cultivated dihaploid potatoes (2n = 2x = 24) from the diploid potato breeding program were grown from tubers. IVP 101, M6, and Atlantic were grown from stem cuttings. All plants were grown in a greenhouse at the University of Wisconsin–Madison. Greenhouse conditions were 25°C during the day, 18°C at night with a 16-h photoperiod provided by natural illumination supplemented by high pressure sodium vapor lights as needed to maintain 500 μmol m^{-2} s^{-1} photosynthetically active radiation. All accessions evaluated in the present study are shown in Table 1.

2.2 | Fertility observations

Fertility observations were conducted after three pollination events with at least 15 flowers pollinated by hand each time. Pollen was collected in gel capsules by buzzing open flowers of an individual plant

TABLE 1 Plant material evaluated for the *Sli* gene sequence.

Species	PI/ID	Number of genotypes evaluated	Ploidy	Collection data
S. andreanum Baker	247360	1	6x	Colombia, Pasto, Nariño. Altitude: 2200 m. Coordinates: 1.6000, −77.1500
	320345	1	2x	Colombia, Cauca. Altitude: 2100 m. Coordinates: 1.2, −76.91666667
	561658	1	2x	Ecuador, Bolivar. Altitude: 3000 m. Coordinates: −1.7, −79.08333333
S. berthaultii Hawkes	458365	1	2x	Salta, Argentina. Altitude: 2250 m. Coordinates: −22.2500, −64.8833
S. brevicaule Bitter	*246536*	*2*	*2x*	Peru, Cuzco. Altitude: 2900 m. Coordinates: −13.3, −72.11666667
	265882	1	4x	Bolivia, Oropesa, Chuquisaca. Altitude: 2750 m. Coordinates: −19.0333, −65.2833
	310933	1	2x	Bolivia, Cercado, Cochabamba. Altitude: 2700 m. Coordinates: −17.4000, −66.1500
	472988	1	2x	Argentina, Tucuman. Altitude: 2700 m. Coordinates: −26.66666667, −65.81666667
	472991	1	2x	Argentina, Jujuy. Altitude: 3300 m. Coordinates: −23.2, −65.45
	473065	2	2x	Argentina, Salta. Altitude: 3900 m. Coordinates: −24.5500, −66.2000
	473368	1	6X	Bolivia, Nor Chicas, Potosi. Altitude: 3100 m. Coordinates: −21.1333, −65.7000
	498115	3	2x	Bolivia, Quillacollo, Cochabamba. Altitude: 3650 m. Coordinates: −17.3333, −66.350
	537026	1	2x	Bolivia, Potosi. Altitude: 3280 m. Coordinates: −19.6, −65.26666667
	545865	1	2x	Bolivia, Potosi. Altitude: 3460 m. Coordinates: −19.56694444, −65.38277778
	545879	1	2x	Bolivia, Campero, Cochabamba. Altitude: 2260 m. Coordinates: −17.98333333, −65.11666667
	545970	4	2x	Bolivia, La Paz. Altitude: 3900 m. Coordinates: −15.6, −69.01666667
	545978	1	2X	Bolivia, Chuquisaca. Altitude: 3000 m. Coordinates: −18.95, −65.31666667
	545997	1	2x	Bolivia, Ibanes, Potosi. Altitude: 3700 m. Coordinates: −18.0167, −66.3833
S. candolleanum Berthault	210046	2	2x	Peru, Ayacucho, Cangallo. Altitude: 2000 m. Coordinates: −13.4000, −73.6833
	266385	1	2x	Peru, Junín. Altitude: 3840 m. Coordinates: −11.6, −75.36666667
	275272	1	2x	Peru, Cuzco. Altitude: 2500 m. Coordinates: −13.33333333, −71.96666667
	365349	1	2x	Peru, Yauyos. Matichacra, Lima. Altitude: 3200 m. Coordinates: −12.3631, −75.8511
	365353	1	2x	Peru, Cuzco. Altitude: 3550 m. Coordinates: −13.33333333, −72.11666667
	365362	3	2x	Peru, Ancash. Altitude: 3500 m. Coordinates: −8.26666667, −77.85
	442696	1	2x	Peru, Puno. Altitude: 3900 m. Coordinates: −15.83333333, −70.03333333
	442697	1	2x	Peru, Apurimac. Altitude: 2600 m. Coordinates: −13.11666667, −74.21666667
	458380	1	2x	Peru, Calca, Cuzco. Altitude: 3350 m. Coordinates: −13.4667, −71.9167
	458381	1	2x	Peru, Cangallo, Ayacucho. Altitude: 2400 m. Coordinates: −13.38333333, −73.25
	458403	1	2x	Peru, Cunyari, Apurimac. Altitude: 2964 m. Coordinates: −13.48, −73.39
	473451	1	2x	Peru, Arequipa. Altitude: 3306 m. Coordinates: −15.5500, −73.6333
	498227	1	2x	Bolivia, La Paz. Altitude: 4000 m. Coordinates: −15.78333333, −68.66666667
S. chacoense Bitter	133619	1	2x	Argentina
	320285	1	2x	Argentina, Córdoba. Altitude: 650m. Coordinates: −32.2167, −65.1167
	320293	1	2x	Argentina, Salta. Altitude: 2400 m. Coordinates: −25.1667, −65.8333
	414143	1	2x	Argentina, Virrey Toledo, Salta. Altitude: 1600 m. Coordinates: −24.9000, −65.6500

(Continues)

TABLE 1 (Continued)

Species	PI/ID	Number of genotypes evaluated	Ploidy	Collection data
	472830	1	2x	Argentina, La Rioja. Altitude: 1140 m. Coordinates: −29.23333333, −66.85
	614703	1	2x	Bolivia, Nor Yungas, La Paz. Altitude: 1243 m. Coordinates: −16.4000, −67.6333
	M6	1	2x	United States
S. chomatophilum Bitter	365339	1	2x	Peru, Cerro Huishca, Junin. Altitude: 3550 m. Coordinates: −10.88333333, −75.95
S. demissum Lindl	275206	1	6x	Mexico, Balleza, Chihuahua. Altitude: 2850 m. Coordinates: 25.4667, −105.2500
	275211	1	6x	Guatemala, Sierra de los Cuchumatanes, Huehuetenango. Altitude: 3500 m. Coordinates: 15.6667, −91.4667
	498012	1	6x	Mexico, Durango. Altitude: 2800 m. Coordinates: 24.2500, −104.3833
	545763	1	6x	Mexico, Oaxaca. Altitude: 2870 m. Coordinates: 17.4667, −96.5000
S. hjertingii Hawkes	251067	1	4x	Mexico, Cerro Potosi, Nuevo Leon. Altitude: 2500 m. Coordinates: 24.8667, −100.2167
	498050	1	4x	Mexico, Sierra de Catorce, San Luis Potosi. Altitude: 2740 m. Coordinates: 23.6900, −100.8900
	570625	2	4x	Mexico, Sierra de Catorce, San Luis Potosi. Altitude: 2600 m. Coordinates: 23.6900, −100.8900
S. hougasii Correll	239423	1	6x	Mexico, Michoacan de Ocampo. Altitude: 2296 m. Coordinates: 19.4667, −102.2500
S. infundibuliforme Phil	442675	1	2x	Argentina, Huichaira, Jujuy. Altitude: 2850 m. Coordinates: −23.5667, −65.4333
	472894	1	2x	Argentina, Serrania de Aparzo, Jujuy. Altitude: 3800 m. Coordinates: −23.1000, −65.1500
	545893	3	2x	Bolivia, Aviles, Tarija. Altitude: 3700 m. Coordinates: −21.8000, −65.1333
	558088	1	2x	Argentina, La Poma, Salta. Altitude: 3700 m. Coordinates: −23.9333, −66.3500
	566766	1	2x	Argentina, Tumbaya, Jujuy. Altitude: 3100 m. Coordinates: −23.6167, −65.5333
	597699	4	2x	Bolivia, Linares, Potosi. Altitude: 3903 m. Coordinates: −19.8333, −65.7000
	597701	1	2x	Bolivia, Nor Chicas, Potosi. Altitude: 3180 m. Coordinates: −20.4000, −65.5667
S. iopetalum (Bitter) Hawkes	275182	1	6x	Mexico, Puebla. Altitude: 1950 m. Coordinates: 20.2500, −98.2167
S. kurtzianum Bitter & Wittm	320271	1	2x	Argentina, Alijilan. Casa del Dique, Catamarca. Altitude: 553 m. Coordinates: −28.1667, −65.5000
	442680	1	2x	Argentina, Catamarca. Altitude: 1600 m. Coordinates: −27.4333, −66.9500
	472923	1	2x	Argentina, Arroyo Arrequetin, San Juan. Altitude: 2600 m. Coordinates: −30.3000, −69.7333
	472924	1	2x	Argentina, Chanural de los Caranchos, Mendoza. Altitude: 1100 m. Coordinates: −33.98333333, −69.23333333
	472928	1	2x	Argentina, Catamarca. Altitude: 1600 m. Coordinates: −27.4000, −66.9333
	472952	1	2x	Argentina, Catamarca. Altitude: 1840 m. Coordinates: −27.9, −66.36666667
	497992	2	2x	N.A.
	558205	1	2x	Argentina, Chilecito, La Rioja. Altitude: 1880 m. Coordinates: −29.2000, −67.6500
	566771	1	2x	Argentina, Chilecito, La Rioja. Altitude: 1400 m. Coordinates: −29.1333, −67.5500
	175435	1	2x	Argentina

TABLE 1 (Continued)

Species	PI/ID	Number of genotypes evaluated	Ploidy	Collection data
S. maglia Schltdl	558316	1	3x	Chile, Petorca, Valparaiso. Altitude: 10 m. Coordinates: −32.5500, −71.4333
S. malmeanum Bitter	458318	1	2x	Argentina, Corrientes. Altitude: 56 m. Coordinates: −29.5667, −57.5333
S. microdontum Bitter	218225	1	2x	Bolivia, Potosi. Altitude: 3888 m. Coordinates: −19.572281, −65.755006
	473170	1	2x	Argentina, Quebrada del Toro, Virrey Toledo, Salta. Altitude: 1600 m. Coordinates: −24.9000, −65.6500
	473363	1	2x	Bolivia, Carrasco, Cochabamba. Altitude: 3025 m. Coordinates: −17.6383, −65.2043
	498128	2	2x	Bolivia, Caballero, Santa Cruz. Altitude: 2550 m. Coordinates: −17.8833, −64.7000
	545901	3	2x	Bolivia, Mendez, Tarija. Altitude: 3000 m. Coordinates: −21.4833, −64.9000
S. pinnatisectum Dunal	347766	1	2x	Mexico, Guanajuato. Altitude: 1770 m. Coordinates: 21.1167, −101.6667
S. polyadenium Greenm	175444	1	2x	Mexico
S. raphanifolium Cárdenas & Hawkes	265862	1	2x	Peru, Sacsayhuaman, Cuzco. Altitude: 3650 m. Coordinates: −13.5167, −71.9833
	296126	1	2x	Peru, Sacsayhuaman, Cuzco. Altitude: 3520 m. Coordinates: −13.5167, −71.9833
	310953	1	2x	Peru, Poroy, Cuzco. Altitude: 3640 m. Coordinates: −13.5000, −72.0500
	310998	1	2x	Peru, Sacsayhuaman, Chincana, Cuzco. Altitude: 3550 m. Coordinates: −13.504664648, −71.975996096
	458383	1	2x	Peru, Anta. Chequerce, Cuzco. Altitude: 3500 m. Coordinates: −13.3833, −72.1333
	473369	1	2x	Peru, Rumicolca, Quispicanchi, Cuzco. Altitude: 3200 m. Coordinates: −13.6000, −71.4500
	473371	2	2x	Peru, Abancay, Apurimac. Altitude: 4050 m. Coordinates: −14.0500, −72.4667
	473467	2	2x	Peru, Pomacandri, Cuzco. Altitude: 3700 m. Coordinates: −14.0333, −71.5667
	473502	1	2x	Peru, Quispicanchis, Keucko, Gpo. Arqueologico del Lago. Altitude: 3700 m. Coordinates: −13, −71
	473526	1	2x	Bolivia, Cerro Potosi. Altitude: 4100 m. Coordinates: −19.616632890192733, −65.75010857818242
	607883	2	2x	Peru, Canchis, Cuzco. Altitude: 3700 m. Coordinates: −14.4931, −71.1642
	607888	1	2x	Peru, Urubamba, Cuzco. Altitude: 3700 m. Coordinates: −13.2272, −72.5053
S. schenckii Juz. & Bukasov	545733	1	6x	Mexico, Puebla. Altitude: 2600 m. Coordinates: 18.9500, −97.6500
	558456	1	6x	Mexico, Oaxaca. Altitude: 2690 m. Coordinates: 17.1500, −96.6000
S. stoloniferum Schltdl	283101	1	4x	Mexico, Majalca, Chihuahua. Altitude: 2150 m. Coordinates: 29.1333, −106.0833
	497994	1	4x	Mexico, Chihuahua. Altitude: 1800 m. Coordinates: 28.9667, −106.4500
	545740	1	4x	Mexico, Durango. Altitude: 2610 m. Coordinates: 22.6669, −104.3500
	545787	1	4x	Mexico, Zacatecas. Altitude: 2400 m. Coordinates: 21.6500, −103.1333
	558466	1	4x	Mexico, Michoacán de Ocampo. Altitude: 2080 m. Coordinates: 19.6833, −101.1167
S. vernei Bitter & Wittm	320330	3	2x	Argentina, Tucuman. Altitude: 2850 m. Coordinates: −26.7667, −65.7500
	320332	2	2x	Argentina, Catamarca. Altitude: 2450 m. Coordinates: −27.35, −66.03333333
	320333	1	2x	Argentina, Jujuy. Altitude: 2960 m. Coordinates: −23.91666667, −65.41666667
	473309	1	2x	Argentina, Salta. Altitude: 3140 m. Coordinates: −23.11666667, −64.53333333

(Continues)

TABLE 1 (Continued)

Species	PI/ID	Number of genotypes evaluated	Ploidy	Collection data
	473311	1	2x	Argentina, Salta. Altitude: 3300 m. Coordinates: −22.0333, −65.0333
	500066	1	2x	Argentina, Corral Grande, Jujuy. Altitude: 3100 m. Coordinates: −24.0000, −65.6167
	500070	1	2x	Argentina, El Remate, Salta. Altitude: 3100 m. Coordinates: −23.2, −64.91666667
S. verrucosum Schltdl	161173	1	2x	Mexico, Michoacan de Ocampo. Altitude: 2296 m. Coordinates: 19.4000, −101.6000
	275258	2	2x	Mexico, Nevado de Colima, Jalisco. Altitude: 3250 m. Coordinates: 19.5500, −103.6333
	275260	2	2x	Mexico, Pachuca, Hidalgo. Altitude: 2700 m. Coordinates: 20.11666667, −98.73333333
	570643	2	2x	Mexico, Cerro Potosi, Nuevo Leon. Altitude: 3200 m. Coordinates: 24.8700, −100.2600
	607844	1	2x	Mexico, Nuevo Leon. Altitude: 2750 m. Coordinates: 23.9833, −99.7333
	607848	1	2x	Mexico, El Chico, Hidalgo. Altitude: 2830 m. Coordinates: 20.1678, −98.6931
S. tuberosum L.	Atlantic	1	4x	United States
	Castle Russet-DH31	1	2x	United States
	Early Rose-DH1	1	2x	United States
	IVP 101	1	2x	United States
	LabelleRusset-DH47	1	2x	United States
	US-W4	1	2x	United States
S. tuberosum L. subsp. *Andigenum*	195158	1	2x	N.A.
	230513	1	2x	Peru, Cohechan, Amazonas. Altitude: 2800 m. Coordinates: −6.21666667, −77.96666667

Note: Plant introduction (PI) numbers are those used by the US Potato Genebank, Sturgeon Bay, Wisconsin. Ploidy estimates and collection data were retrieved from the USDA-ARS Germplasm Resource Information Network (https://www.ars-grin.gov).
Abbreviation: N.A., not available.

using an electric toothbrush. Open flowers on the same plant were pollinated by gently depositing pollen on the stigmas. A plant was considered self-compatible if seeds containing embryos were extracted from one or more berries. If no berries were observed or no seeds were present in the berries, the plant was considered self-incompatible. A subset of self-incompatible genotypes was evaluated for male and female fertility to estimate the extent that fertility barriers other than the S-locus-dependent gametophytic self-incompatibility system contribute to self-incompatibility. Male fertility of 42 genotypes and female fertility of 32 genotypes was evaluated by crossing them with US-W4 and MSFF744-01-1 as males and by crossing them to *S. chacoense* M6 as females. A plant was considered male or female fertile if seeds containing embryos were extracted from one or more berries produced by the cross. For 48 genotypes, only self-pollination experiments were conducted, and male and female fertility was not evaluated. For 35 genotypes, we only had DNA extracted from plants that were no longer extant. Those genotypes could not be evaluated for fertility. Fertility of the Solyntus and DM 1–3 clones were not evaluated in this study but have been reported previously (Peterson et al., 2016; van Lieshout et al., 2020).

2.3 | *Sli* amplicon generation

2.3.1 | Primer design

Full sequences for the potato locus Identified as the *Sli* gene (PGSC0003DMG400016861) were downloaded from the Sol Genomics Network (Sol Genomics Network) (Fernandez-Pozo et al., 2015), along with exon/intron information and 5-kb upstream and downstream of the *Sli* coding sequence. Sequences from tomato were also obtained from Sol Genomics after running the orthologs tool using the PGSC0003DMG400016861 potato locus including 5-kb upstream and downstream regions of the genome. Full genome assembly for Solyntus V1.1 was downloaded from: Solyntus genome sequence consortium (wur.nl) (van Lieshout et al., 2020). The 533 bp insertion in the *Sli* locus

described by Eggers et al. (2021) was used to localize and extract the region homologous to PGSC0003DMG400016861 from the Solyntus genome. A multiple sequence alignment for the sequences of the Solyntus *Sli* locus, DMv3 *Sli* locus and tomato orthologous *Sli* locus was performed using the MUSCLE plugin in Geneious Prime 2022.0.2 build 2022-01-26 14:24 Java Version 11.0.12+7 (64 bit). Forward primers were designed based on a conserved region of the alignment approximately 1500 bp upstream of the start codon. Reverse primers were designed based on the first exon of an adjacent gene downstream of the *Sli* gene, approximately 400 bp after the *Sli* stop codon. All primers were designed using Primer3Plus V. 3.2.6 (Untergasser et al., 2012) using default settings. Three primer combinations were used for amplification of the *Sli* gene from the five pairs designed. Primers sequences used are reported in the Supporting Information.

2.3.2 | Amplicon generation

Sli locus amplification was done using DNA purified from leaves using the DNeasy Plant Mini kit (Qiagen, Germantown, MD, USA). PrimeSTAR® GXL DNA Polymerase (Takara Bio USA, Inc. San Jose, CA, USA) or LA Taq HS Polymerase (Takara Bio USA, Inc. San Jose, CA, USA) were used. A subset of five samples were amplified using both enzymes (PrimeSTAR® GXL and LA Taq HS Polymerase) in separate reactions to evaluate any differences in amplicon generation. As no differences were found, both enzyme mixes were used depending on their availability. PCR reactions with PrimeSTAR® GXL consisted of 1X Prime Star GXL buffer, 200 µM of dNTPs, 1.25 units of enzyme and 0.3 µM of each primer in a 50 µL total reaction volume. Amplification conditions when using PrimeSTAR® GXL were as follows: 30 cycles at 98°C for 10 s, 60°C for 15 s and 68°C for 4 min. PCR with LA Taq HS polymerase consisted of 1X LA buffer, 250 µM of dNTPs, five units of enzyme and 0.5 µM of each primer in a 50 µL total reaction volume. Amplification conditions when using LA *Taq* HS Polymerase were as follows: One cycle at 94°C for 2 min, 30 cycles at 98°C for 10 s, 64°C for 15 s and 68°C for 4 min followed by a final extension at 72°C for 10 min. Ten-microliter samples of amplified product were separated on a 0.8% agarose gel in TAE 1X buffer along with a 10-kb plus ladder (Thermo Fisher Scientific, Eugene, OR, USA) and stained with 0.5 µg mL^{-1} ethidium bromide. Samples with an amplified product approximately 3 kb in length were purified using AMpure XP Bead-Based Reagent (Beckman Coulter Life Sciences, Indianapolis, IN, USA) following manufacturer's instructions. DNA content of purified amplicons was quantified with the Qubit dsDNA HS reagent (Thermo Fisher Scientific, Eugene, OR, USA). Samples with an amplicon concentration greater than 12.5 ng μL^{-1} were used for sequencing.

2.3.3 | Amplicon sequencing and data analysis

Amplicons were sequenced with a PacBio Sequel II instrument at the UW-Madison Biotechnology Next Gen DNA Sequencing Core facility. HiFi data were generated for indexed and multiplexed samples. Demultiplexing of reads and *.bam file generation for each amplicon were conducted by the sequencing facility. For each amplicon, *.bam files were converted to *.fastq files using the bam2fq command in SamTools (Danecek et al., 2021). Indexing of *.fastq files was performed with the command fqidx in SamTools. Fastq files along with indexed files were analyzed using the PacBio Amplicon Analysis pbAA V1.0.3 (GitHub - PacificBiosciences/pbAA) using the cluster tool with default settings. The pbAA application clusters HiFi reads and generates consensus sequences for each allele. Previously published genomic sequences for the *Sli* gene from Solyntus (van Lieshout et al., 2020), and DM (Potato Genome Sequencing Consortium et al., 2011) were used as guide alleles for the analysis. The two allele guides in fasta file format, listed in the Supplementary material, were indexed using SamTools faidx command and used for clustering in pbAA. All applications were setup through the Conda environment V4.13.0. On average 17,000 circular consensus reads were obtained for each amplicon but only 500 randomly selected sequences that passed pbAA QC and error correction were clustered per each guide by pbAA. Clusters with less than 40X depth were excluded from subsequent analysis.

2.4 | Sequence alignment and gene phylogeny reconstruction

Sequences were aligned using the MAFFT v7.490 (Katoh et al., 2002; Katoh & Standley, 2013) within the alignment function in Geneious Prime 2023.1.2 using default parameters. We reconstructed a phylogenetic tree for the *Sli* gene alignment of sequences 1473 bp upstream (5′UTR) of the start codon to 250 bp downstream of the stop codon (3′UTR); hence, this alignment included the 533 bp insertion located in the 5′UTR. Also, a tree for only the coding sequence of *Sli* was reconstructed. A maximum likelihood tree search under a general time reversible model was calculated using RAxML version 8.2.11 (Stamatakis, 2014) for each dataset. The insertion region, which was absent in most sequences, was considered a gap in the alignments and treated as missing data. Bootstrap analysis with 1000 replicates was also conducted. Bootstrap support is indicated at nodes of the tree.

2.5 | Selection analysis

ω, the ratio (ω = dN/dS) of non-synonymous substitutions per nonsynonymous site (dN) to synonymous substitutions per synonymous site (dS) (Yang & Bielawski, 2000), was calculated for the coding sequence to evaluate evidence of neutral, purifying, or diversifying selection. Selection analyses were conducted using the Datamonkey Server (Weaver et al., 2018). All calculations were done considering a universal genetic code. The coding sequence alignment for selection analysis included additional outgroups belonging to the Solanaceae family downloaded from NCBI. An analysis for detecting recombination using GARD, A Genetic Algorithm for Recombination Detection (Kosakovsky Pond et al., 2006), was conducted prior to selection inference analysis.

BUSTED (branch-site unrestricted statistical test of episodic diversification) (Murrell et al., 2015) was used to perform analysis of episodic diversification in a branch of the phylogenetic tree containing coding sequences from *Sli* genes with a 5′UTR carrying the 533 bp insertion. A gene wide analysis selecting all branches of the maximum likelihood phylogeny was also conducted using BUSTED.

FUBAR (Fast, Unconstrained Bayesian AppRoximation) analysis (Murrell et al., 2013) was used to infer nonsynonymous (dN) and synonymous (dS) substitution rates at each position of the coding sequence alignment as an indicator of either negative or positive pervasive selection. FUBAR analysis used nucleotide general time reversible (GTR) modes with an AIC-c = 13498.23 and logL = −6423.32 for a total of 325 parameters. MEME (mixed effects model of evolution) analysis (Murrell et al., 2012) was used to assess diversifying or positive selection on a per-site basis.

2.6 | Protein diversity, structure, and function prediction analysis

An alignment of the protein coding sequence was performed using MAFFT (Katoh & Standley, 2013., Katoh et al., 2002). Analysis of the secondary structure of the Sli protein was performed using Phyre2 (Protein Homology/AnalogY Recognition Engine) V 2.0. (Kelley et al., 2015). Prediction analysis was conducted in normal mode. Nine different protein sequences were modeled, including the protein encoded by *S. chacoense* M6 that carries the 533 bp insertion. *S. kurtzianum* accession PI 472952 (D19-11) was deliberatively chosen for protein sequence modeling because it is homozygous for *Sli* gene sequences that lack the 533 bp insertion and the plant is self-compatible. Seven additional sequences from self-compatible genotypes were modeled and are listed in Table 6. Phyre Investigator, an application within Phyre^2 was run to analyze sequence variations that could potentially impact protein function using the SuSPect tool (Yates et al., 2014). The fpocket2 tool was used to identify functional parts of the protein, such as the pocket and the amino acid positions predicted to be part of the pocket (Le Guilloux et al., 2009).

2.7 | Accession numbers

Sequence data reported in this article can be found in NCBI under submission number:2594803833 .

3 | RESULTS

3.1 | The diversity and evolutionary history of *Sli* DNA sequences

We obtained 244 genomic sequences of *Sli* from a total of 141 genotypes, 133 accessions of 22 wild species relatives of potato and eight diverse diploid and tetraploid clones of cultivated potato (Table 1). The number of *Sli* alleles identified for each genotype in the study matched the ploidy level or was below it in almost all cases. Exceptions, where the allele number exceeded the presumed ploidy level, were observed for seven out of 20 genotypes of *S. brevicaule* and one out of three genotypes of *S. andreanum*, indicating that these individuals were not diploid (Table S1) or that they have multiple copies of the gene. Polyploid individuals have been reported previously for *S. brevicaule* (Bryan et al., 2017) and *S. andreanum* (Fumia et al., 2022; Spooner et al., 2014).

Across all genotypes, the *Sli* DNA sequence is highly conserved in the 3 exons but harbors numerous nucleotide substitutions, insertions, and deletions in the 5′UTR, the two introns, and the 3′UTR. One prominent insertion upstream of the start codon is the 533 bp insertion characterized by Eggers et al. (2021) and Ma et al. (2021). We identified the 533 bp insertion in seven individuals from five wild species and three clones of cultivated potato (Figure 1 and Table 2). *Sli* with the 533 bp insertion was present in the heterozygous state in six individuals from wild species and in the homozygous state in M6, an inbred line of *S. chacoense* that was used as a positive control. US-W4 and the published Solyntus sequence were also homozygous for *Sli* with the insertion. *Sli* with the 533 bp insertion was present in the heterozygous state in Early Rose DH1, a male sterile dihaploid derived from the female fertile tetraploid Early Rose. The insertion was also present in the heterozygous state in the female fertile tetraploid variety Atlantic (Table 2).

A maximum likelihood phylogenetic tree reconstruction was used to evaluate the likely origin of the introgression of the *Sli* allele with the 533 bp insertion into cultivated potato. The analysis included all the *Sli* sequences from the 10 genotypes that carry alleles with the insertion along with *Sli* sequences from Solyntus (control with insertion), DM (control for no insertion), and *S. lycopersicum* as an outgroup. The phylogenetic tree with the highest log likelihood for *Sli* (−10863.65) (Figure 2) shows that the alleles with the insertion for all the cultivated clones and wild potato genotypes are monophyletic; they cluster together with a bootstrap support of 100%. Remarkably, there are only four unique alleles carrying the insertion and three of these different alleles are present in *S. brevicaule* (PI 473368). A single allele is present in *S. tuberosum* (Atlantic, Solyntus, US-W4 and Early Rose DH1), the inbred *S. chacoense* line M6, *S. brevicaule* (PI 246536 and PI 265882), *S. hougasii* (PI 239423), *S. hjertingii* (PI 251067), and *S. vernei* (PI 320332) (Figure 2).

Sli alleles that lack the insertion are grouped into different clusters in Figure 2. One cluster contains alleles exclusively from cultivated potato, with 100% bootstrap support. There is another cluster with 98% bootstrap support containing alleles from wild potato accessions. A sister of the cluster of *Sli* alleles carrying the insertion has low <70% bootstrap support (Figure 2).

Clot et al. (2020) suggested that one or more alleles with the insertion were introgressed into cultivated potatoes by the mid-1800s, well before the archiving of germplasm in genebanks. Alternatively, alleles with the insertion could have been acquired by wild species through inadvertent genetic transfer after the material was collected and before the material was transferred to the US

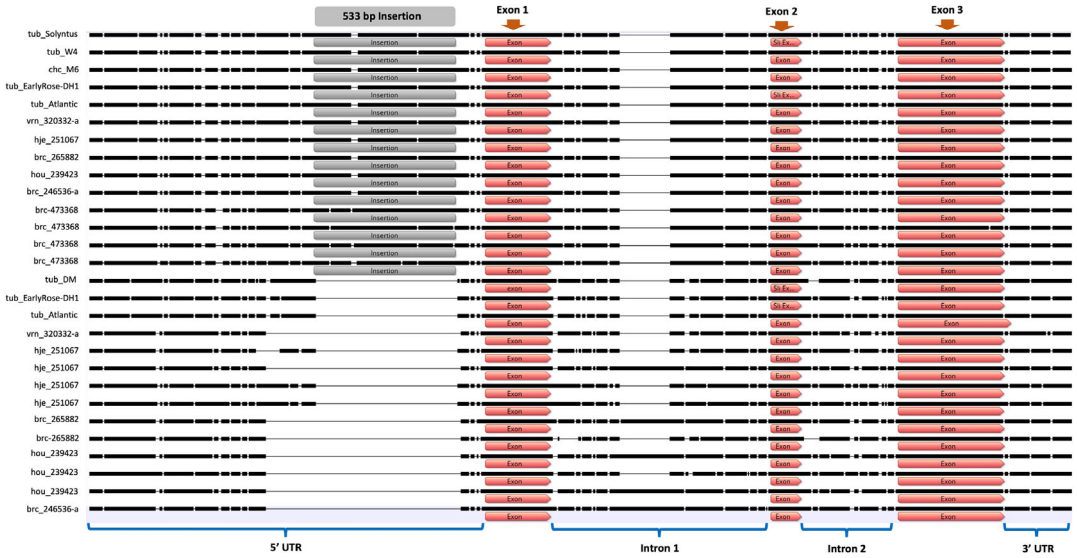

FIGURE 1 *Sli* gene alignment diagram for all accessions carrying one or more *Sli* alleles with the 533 bp insertion (gray box) in the 5′UTR region. Alleles without the insertion for the same accessions are represented in the bottom half of the alignment. Also diagramed is the gene model for the *Sli* gene, which consists of three exons, two introns, and the 3′UTR and 5′UTR regions.

potato genebank. A list of the accessions where at least one *Sli* allele carrying the insertion was identified, along with their fertility status, ploidy, zygosity, original collection data and genebank acquisitions times are reported in Table 2. The material was collected over a wide geographic range (Figure 3) and was donated to the US Potato Genebank at different times and from different sources. Thus, it seems unlikely that *Sli* alleles with the insertion were inadvertently introgressed into these accessions of wild species before they were deposited into the US potato genebank.

A broad maximum likelihood phylogenetic tree reconstruction that includes the 244 alleles discovered in the 141 genotypes along with *Sli* sequence data from Solyntus, DM and *S. lycopersicum* as an outgroup, was used to establish likely evolutionary relationships among *Sli* genomic sequences. The phylogenetic tree with the highest log likelihood for *Sli* (−60281.36) is shown in Figure S1. As above, the *Sli* tree reconstruction shows a single origin for *Sli* alleles carrying the 533 bp insertion with 100% bootstrap analysis support. Outside of the branch containing *Sli* alleles with the 533 bp insertion, the arrangement of species in this maximum likelihood phylogeny based only on the *Sli* gene is similar to the arrangement of species in the phylogeny for tuber-bearing Solanum based on morphological, molecular, crossing and field observation data (Spooner et al., 2014). For instance, a clade with 100% bootstrap support is composed of accessions from northern Peru and Ecuador such as *S. andreanum* and species known to be of hybrid origin from clades 3 and 4 such as *S. iopetalum* (Figure S1). This grouping of species was identified previously using 12 ortholog DNA markers (Spooner et al., 2018). Similarly, with a bootstrap support of 80%, the analysis supports a single origin for *Sli* alleles in accessions from a subset of species belonging to Clade 4. These species were also clustered together in Spooner et al. (2018). The tree reconstruction based on *Sli* also shows a level of interrelatedness and introgression among diploid species of Clade 4 similar to that documented previously in multiple studies (Huang et al., 2019; Spooner & van den Berg, 1992).

3.2 | Identification of DNA sequence under selection

The analysis with GARD (Kosakovsky Pond et al., 2006) did not find evidence of recombination (Δ c-AIC vs the null model = 0.00), hence we proceed with the gene-wide selection analysis (Murrell et al., 2015) to detect episodic diversification within *Sli* in which the expected non-synonymous substitution rate is greater than the synonymous rate. The analysis with BUSTED included site-to-site synonymous rate variation and was conducted on the DNA coding sequences. The analysis started by building a phylogeny that was tested with no prior hypothesis of any branch being under positive selection (entire tree). Subsequently it tested a pre-specified branch of the tree containing *Sli* alleles with the 533 bp insertion (insertion branch). Of 240 *Sli* sequences, 160 unique coding sequences ranging from 798 to 825 bp long were used for the analysis. There was no support for episodic diversifying selection for the entire *Sli* gene based on the likelihood ratio test for the entire tree ($p = 0.4799$) and the insertion branch ($p = 0.3683$). Results for rates distribution for branch-site combination analysis are shown in Table 3.

Site selection analysis of cDNA sequences was conducted to identify codon positions in *Sli* that are under positive or negative selection. Analysis with FUBAR found evidence of pervasive positive diversifying selection at five codon sites with a posterior probability equal or higher than 0.9 (Table 4). Analysis with MEME found evidence of positive diversifying selection at 11 sites with a *p*-value

TABLE 2 Wild and cultivated potatoes carrying the *Sli* allele with the 533 bp insertion.

Species	PI/ID	Ploidy	Fertility	*Sli* alleles state	Collection place	Collection date	Received by NPGS	Donor
Wild species accessions								
Solanum brevicaule	246536-a	2x	SI (MF, FF)	Heterozygote	Urubamba, Cusco, Peru	01 Jan–31 Dec 1958	11 Mar 1958	D. S. Correll, Maryland, USA, USDA
	265882	4x	Unknown	Heterozygote	Oropesa, Chuquisaca, Bolivia	01 Jan–31 Dec 1960	27 May 1960	Max Plank Institute Germany
	473368	6x	SC	Heterozygote	Tupiza, nor Chicas, Potosi, Bolivia	16 Mar 1961	Mar 1982	J. G. Hawkes, University of Birmingham, UK
Solanum chacoense	M6	2x	SC	Homozygote	Inbred line	N.A.	N.A.	N.A.
Solanum hjertingii	251067	4x	Unknown	Heterozygote	Nuevo Leon, Mexico	25 Aug 1958	10 Sep 1958	J. G. Hawkes, University of Birmingham, UK
Solanum hougasii	239423	6x	Unknown	Heterozygote	Michoacan de Ocampo, Mexico	01 Jan–31 Dec 1957	07 May 1957	Rockefeller Foundation
Solanum vernei	320332	2x	SI (MF, FF)	Heterozygote	Catamarca, Argentina	20 Feb 1966	Sep 1966	J. G. Hawkes, University of Birmingham, UK
Cultivated clones								
Solanum tuberosum L	Atlantic	4x	FF	Heterozygote	Cultivated variety	N.A.	N.A.	N.A.
	Early Rose DH1	2x	SI (MS, FF)	Heterozygote	Dihaploid extracted from 4X variety	N.A.	N.A.	N.A.
	US-W4	2x	SC	Homozygote	Dihaploid extracted from 4X breeding line	N.A.	06 July 1939	N.A.
	Solyntus[a]	2x	SC	Homozygote	Inbred line	N.A.	N.A.	N.A.

Note: Plant introduction (PI) numbers are those used by the US Potato Genebank, Sturgeon Bay, Wisconsin (NPGS). Ploidy estimates were retrieved from the USDA-ARS Germplasm Resource Information Network (https://www.ars-grin.gov). Entries with unknown fertility were only available as frozen DNA samples.

Abbreviations: FF, female fertile; FS, female sterile; MF, male fertile; MS, male sterile; N.A., not applicable; SC, self-compatible; SI, self-incompatible.

[a]Solyntus *Sli* gene sequence was retrieved from NCBI and was not generated in this study.

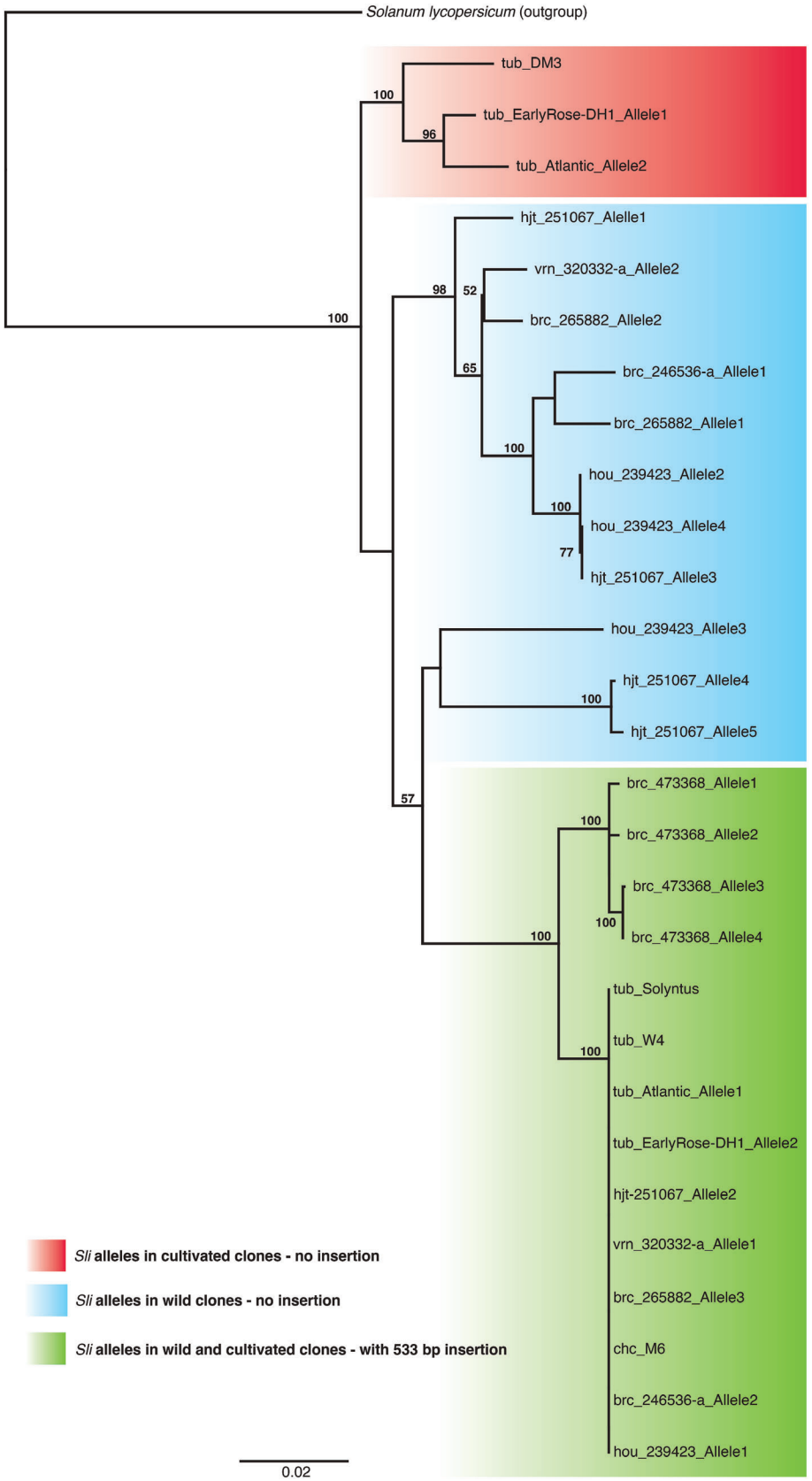

FIGURE 2 Maximum likelihood tree reconstruction of the *Sli* gene alleles for wild and cultivated potato accessions that harbor one or more alleles with the 533 bp insertion in the 5′UTR region. The sequence alignment of the entire gene, from 5′UTR to 3′UTR, has been used in this reconstruction. The numbers at tree nodes represent bootstrap support values. The number at the bottom of the tree represents the rate of nucleotide substitutions in the tree. Abbreviations in the tree: tub = *S. tuberosum*, hjt = *S. hjertingii*, vrn = *S. vernei*, brc = *S. brevicaule*, hou = *S. hougasii*, chc = *S. chacoense*.

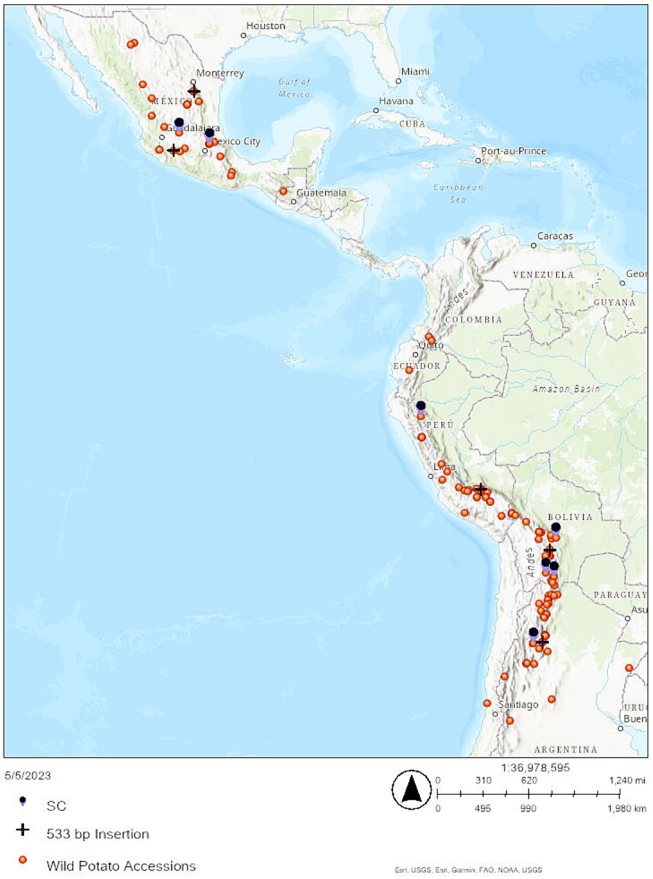

FIGURE 3 Geographical distribution of the wild potato accessions surveyed for the *Sli* gene. Self-compatible genotypes and genotypes carrying the 533 bp insertion *Sli* allele are shown on the map.

equal or lower than 0.1. MEME identified the five sites found by FUBAR as well as six additional sites. Codons in *Sli* identified by MEME and FUBAR as being under positive diversifying selection are presented in Table 4.

3.3 | Diversity of Sli protein sequences

For 242 *Sli* genomic sequences, the predicted Sli protein sequence had 266 amino acids. One allele variant from Atlantic had 275 amino acids and one from *S. vernei* (PI 320333) had 267 amino acids. Twelve genotypes had predicted proteins that were truncated, ranging from 44 amino acids in *S. andreanum* PI (247360) to 243 amino acids in *S. iopetalum* (PI 275182) (Table S1) and four of the 12 accessions were homozygous for a truncated protein *S. polyadenium* (PI 175444), *S. microdontum* (PI 473170), *S. brevicaule* (PI 545879), and *S. kurtzianum* (PI 558205) (Table S1). The Sli amino acid sequence had reduced variability compared to DNA sequence of the gene but in total we identified 133 distinct protein alleles of 266 to 275 amino acids (not including outgroup) that appear to be complete based on their DNA sequence. A Maximum likelihood tree constructed from coding sequences encoding unique Sli protein variants is shown in Figure 4. The tree contains few resolved species relationships due to the high overall similarity of exon regions. However, the three protein variants encoded by the *Sli* alleles with the 533 bp insertion cluster together with a 99% bootstrap support. Two variants are present in *S. brevicaule* (PI 473368) an hexaploid genotype. The remaining 10 protein sequences are identical to each other and they are represented by *S. hougasii* PI 239423 (Figure 4). The high bootstrap support of this cluster of three sequences suggests that the *Sli* allele with the insertion likely had a single origin.

3.4 | The structure of Sli proteins

In silico protein modeling was used to explore potential differences in 3-D protein structure among *Sli* alleles. Domain analysis was conducted using the *Sli* sequences from *S. kurtzianum* D19-11 (PI 472952) and M6. Both genotypes are diploid and homozygous for a single allele of *Sli*. D19-11 contains an allele without the insertion and M6 contains an allele with the insertion. The model for the Sli protein of D19-11 is shown in Figure 5. Two protein domains were predicted for Sli from D19-11. The first domain extended from position one to 91 and the second domain from position 97 to 265. The model with the highest confidence (100%) was predicted by Phyre[2] using 164 residues (62% coverage). Only 10 amino acids differ between the protein translated from the M6 *Sli* allele (with the insertion) and the D19-11 allele (without the insertion) (Table 5). Protein domain analysis of Sli from M6 generated the same structure and with the same confidence level as for that for D19-11. Mutational sensitivity analysis followed the same trend in both proteins and none of the 10 sites that differ between the two proteins had a medium or high mutational sensitivity in this analysis (Table 5). However, two of these 10 sites, amino acids 42 and 56, were among the five identified by both MEME and FUBAR as being under positive selection (Table 4).

Pocket detection analysis predicted a pocket, a typical feature of active sites, in the Sli protein. The predicted pocket in the protein from M6 consisted of 10 amino acids, many fewer than the 19 amino acids in the predicted pocket of D19-11 (Figure 5a,b and Table 6). Only four amino acids (ALA 104, ARG 105, GLU 131, and VAL 132) overlapped between the pockets detected in D19-11 and M6 (Table 6).

Given that the pockets in the D19-11 and M6 Sli proteins were considerably different in size, we expanded the pocket analysis to include additional genotypes without the insertion and the two additional genotypes with the insertion. All predicted Sli protein structures were similar to M6 and D19-11. Two distinctive pockets with different amino acid compositions were found (Table 6). Genotypes brc_246536b, krt_472928, and chc_133619 had predicted pockets like that of D19-11. Eleven additional predicted pockets were composed by the same amino acids present in M6's pocket. There was no association between the pocket type and fertility as self-compatible genotypes had both types of pocket (Table 6).

TABLE 3 Results from analysis using BUSTED algorithm for alignment-wide evidence of episodic diversification in the *Sli* gene for all branches of the maximum likelihood phylogenetic tree and for the branch containing *Sli* alleles with the 533 bp insertion.

All branches (entire tree)					
Unconstrained model fit (dN/dS > 1 allowed)			**Constrained model fit (dN/dS > 1 not allowed)**		
Selection mode	**dN/dS**	**Proportion**	**Selection mode**	**dN/dS**	**Proportion**
Negative selection (ω1)	0.07902	76.430%	Negative selection (ω1)	0.06827	0.15962%
Negative selection (ω2)	0.08802	0.1853%	Negative selection (ω2)	0.07191	73.396%
Diversifying selection (ω3)	1.108	23.384%	Neutral evolution (ω3)	1.000	26.445%
Log (L)	−6052.82		Log (L)	−6052.86	
AIC-c	12,792.9		AIC-c	12,791	
Parameters	341		Parameters	340	
Pre-specified branch of the tree containing *Sli* alleles with the 533 bp insertion (insertion branch)					
Unconstrained model fit (dN/dS > 1 allowed)			**Constrained model fit (dN/dS > 1 not allowed)**		
Selection mode	**dN/dS**	**Proportion**	**Selection mode**	**dN/dS**	**Proportion**
Negative selection (ω1)	0.000	87.498%	Negative selection (ω1)	0.000	24.864%
Negative selection (ω2)	0.000	6.1368%	Negative selection (ω2)	0.000	23.368%
Diversifying selection (ω3)	11.14	6.3655%	Neutral evolution (ω3)	1.000	51.768%
Log (L)	−4989.73		Log (L)	−4990.04	
AIC-c	10,644.7		AIC-c	10,643.3	
Parameters	330		Parameters	329	

Note: Unconstrained model: $\omega 1 \leq \omega 2 \leq 1 \leq \omega 3$, constrained model: $\omega 1 \leq \omega 2 \leq 1 = \omega 3$.
Abbreviations: AID-c, Akaike information criterion corrected for small size; Log (L), log-likelihood of the models tested.

TABLE 4 Results from analysis using MEME and FUBAR algorithms for site selection evidence of positive diversifying selection (positive) in codon sites of the *Sli* gene.

Codon position	**MEME support *p* value**	**Selection type**	**FUBAR support Posterior probability**	**Selection type**
39	0.10	Positive	0.60	None
42	**0.01**	**Positive**	**0.99**	**Positive**
46	0.00	Positive	0.84	None
56	**0.02**	**Positive**	**0.96**	**Positive**
63	0.04	Positive	0.89	None
175	0.08	Positive	0.73	None
210	0.04	Positive	0.88	None
212	0.05	Positive	0.83	None
218	**0.04**	**Positive**	**0.92**	**Positive**
242	**0.02**	**Positive**	**0.99**	**Positive**
251	**0.05**	**Positive**	**0.93**	**Positive**

Note: Positions in bold text had strong support using both methods. None = sites not detected under positive or negative selection by FUBAR.

3.5 | Self-compatibility observations and Sli

We could not evaluate self-compatibility of those genotypes for which we had DNA but not the individual from which the DNA was extracted. Of the 116 extant individuals we evaluated for self-compatibility (Table S1), most were diploid wild potatoes and self-incompatible. Only 17 individuals were self-compatible (Table 7) and only three of them, M6, a hexaploid accession of *S. brevicaule* and US-W4, carried a *Sli* allele with the 533 bp insertion. The remaining 14 self-compatible accessions, from wild species, *S. tuberosum* L. and *S. tubersoum* L. subsp. Andigenum, lacked the 533 bp insertion or encoded a truncated protein 30 amino acids shorter at the carboxy-

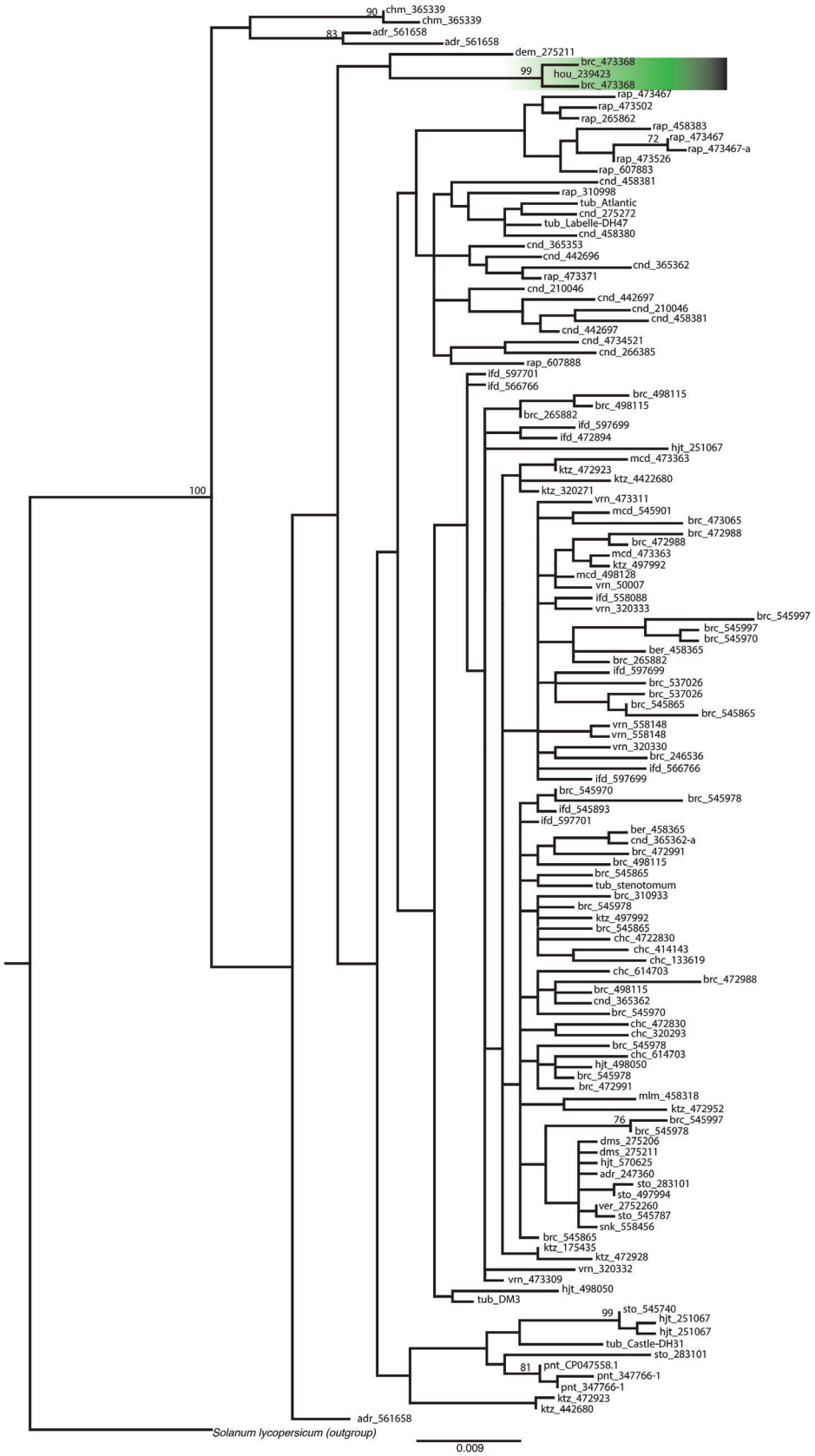

FIGURE 4 A maximum likelihood tree constructed from coding sequences encoding unique Sli protein variants. The three protein variants encoded by the *Sli* alleles with the 533 bp insertion cluster together with a 99% bootstrap support and are shown in the shaded green area. Abbreviations in the tree: adr = *S. andreanum*, *ber* = *S. berthaultii*, brc = *S. brevicaule*, *cnd* = *S. candollenaum*, chc = *S. chacoense*, chm = *S. chomatophilum*, dms = *S. demissum*, hjt = *S. hjertingii*, hou = *S. hougasii*, *ifd* = *S. infundibuliforme*, *ktz* = *S. kurtzianum*, *mlm* = *S. malmeanum*, *mcd* = *S. microdontum*, *pnt* = *S. pinnatissectum*, *rap* = *S. raphanifolium*, *snk* = *S. schenckii*, *sto* = *S. stoloniferum*, vrn = *S. vernei*, tub = *S. tuberosum*.

terminal side than the full-length protein (*S. polyadenium* PI 175444). Of the 42 self-incompatible individuals evaluated for male fertility, and the 32 evaluated for female fertility, 38 were male fertile, 29 were female fertile and 26 were both male and female fertile. For these individuals at least, S-locus-dependent self-incompatibility appears to be the primary cause of self-incompatibility.

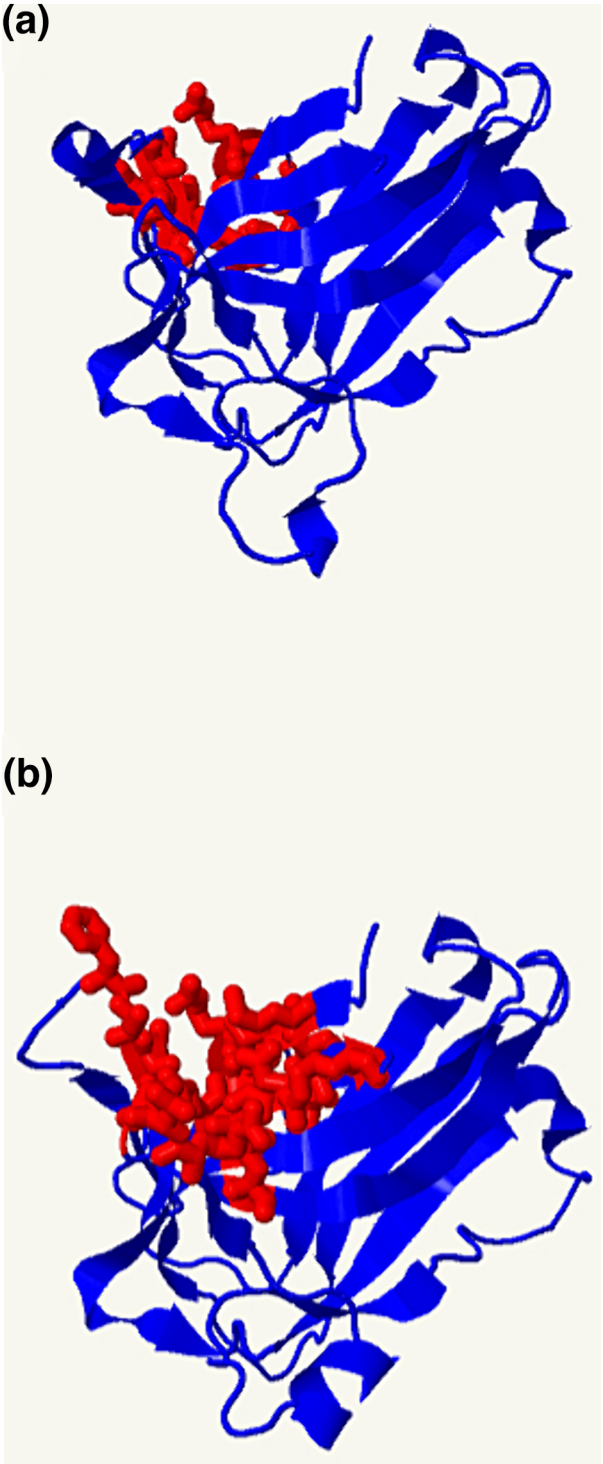

FIGURE 5 Predicted pockets of the Sli protein for *S. kurtzianum* D19-11 showing a large pocket type (a) and *S. chacoense* M6 showing a small pocket type (b).

We found self-compatible genotypes carrying *Sli* alleles with and without the 533 bp insertion. Predicted protein structures from alleles with and without the insertion seem comparable and functional based on in silico modeling. To further explore *Sli* function, we looked at transcription of *Sli* in a self-compatible individual homozygous for alleles that lack the 533 bp insertion (D19-11) and an individual homozygous for alleles with the insertion (M6). We detected cDNA of *Sli* in stylar tissue before pollination (T0) and 6 h post-pollination (T6) for both D19-11 (no insertion) and M6 (with insertion) (Figure S3) using non-quantitative PCR. The DNA sequences of the cDNA products matched the cDNA sequences predicted from the corresponding *Sli* gene.

4 | DISCUSSION

The main objective of this research was to determine if the *Sli* gene carrying the 533 bp insertion is present in wild potato germplasm other than the inbred *S. chacoense* genotypes M6 (Jansky et al., 2014) and 525-3 (Hosaka & Hanneman, 1998). We also wanted to characterize *Sli* sequence diversity in a wide range of wild potato species. Finally, we wanted to know if *Sli* sequences are useful predictors of self-compatibility in wild potato species.

Our results showed clearly that there are two types of *Sli* alleles, those with the 533 bp insertion and those without the insertion. The frequency of the gene with the insertion was low in the wild species we evaluated. Of 136 wild species genotypes examined, an allele with the insertion was found in nine. Those nine, however, were collected over an expansive geographical range that included Northeast and Southwest Mexico, Southern Peru, Southern Bolivia, and Northern Argentina (Figure 3).

The data presented here support the hypothesis that *Sli* gene diversification occurred during speciation within section *Petota* (Figure S1). The *Sli* phylogenetic tree has a branching pattern that is similar to the phylogenetic tree for tuber-bearing potatoes based on molecular and morphological data (Section *Petota*) (Spooner et al., 2014). Similarly, the level of reticulation observed in Clade 4 (Spooner et al., 2016) explained by interspecific hybridization events is observed in the *Sli* gene tree with identical alleles shared by accessions in different species (Figure S1). There is strong evidence to support the finding that *Sli* sequences with the 533 bp insertion trace back to a single origin (Figures 2 and 4). Yet within this group, minor sequence diversification is apparent (Figures 1 and 2). The likely monophyletic origin of the alleles with the insertion is interesting given that it goes beyond the limits of species and because studies trying to reconstruct the evolutionary history of potatoes with single or a few genes have shown very low support for monophyletic clades, especially among South American diploid potatoes (Ames & Spooner, 2010; Fajardo & Spooner, 2011; Spooner et al., 2018).

It has been argued that the S-RNAses that *Sli* interacts with are very old and diversified before speciation in potatoes (Dzidzienyo et al., 2016). Diversification of S-RNAses would increase outcrossing and diversification of taxa early during the evolution of potatoes (Dzidzienyo et al., 2016). In this context, the *Sli* gene may have evolved, or its expression may have been modified during speciation as a means to restore self-compatibility when species were expanding geographically, when pollinators were scarce, or when it was advantageous to preserve species identity. Additional research is needed to

TABLE 5 Amino acid positions that differ between a Sli protein sequence encoded by the *Sli* allele with no insertion and the *Sli* allele with the insertion and their mutational sensitivity.

	S. kurtzianum D19-11 (no insertion)		*S. chacoense* M6 (insertion)	
Position in protein	Amino acid	Mutational sensitivity (0 = low, 9 = high)	Amino acid	Mutational sensitivity (0 = low, 9 = high)
10	Gly	1	Asp	2
42[a]	Phe	1	Val	1
54	Ala	0	Asp	1
56[a]	Ile	3	Asn	1
110	Ser	2	Thr	3
167	Ala	3	Ser	3
169	Asp	1	Asn	1
202	Glu	0	Gly	0
214	Arg	3	Cys	3
249	Arg	1	Gln	0

[a]Sites detected as being under positive selection (see Table 4).

evaluate if alterations in gene expression of *Sli* and S-locus S-RNAses can allow for switching between incompatible and compatible genotype.

Sli with the insertion was present at low frequency in the wild species examined. The presence of an allele carrying the insertion was genotype-specific and not all the genotypes evaluated from the same accession carried this allele. It is unknown if the allele with the insertion is present at a higher frequency in in-situ wild potato populations, where self-compatibility could provide an advantage for reproduction, for instance where species are expanding their range of geographical distribution, or in areas experiencing a decline in pollinators.

While the *Sli* allele with the insertion confers self-compatibility in cultivated potato, the presence of this allele was neither necessary nor sufficient for self-compatibility in wild species relatives of cultivated potato. Of the 12 self-compatible individuals of wild species and two individuals from *S. tuberosum* subsp. Andigenum identified in this study, only one contained the allele with the insertion. Thus, other genetic factors must be important for conditioning a switch in the mating system from self-incompatible to self-compatible, such as changes in S-RNAses gene expression as shown in tomato (Broz et al., 2017) and in physalis (Pretz & Smith, 2021).

Several studies have demonstrated that the rate of protein evolution is inversely proportional to its level of expression, with highly expressed genes having lower estimated rates of ω, regardless of their functional importance (Pál et al., 2001; Zhang & Yang, 2015). This phenomenon has been shown in distant and recent evolutionary events (Shibai et al., 2022). Considering that the split between the tomato and potato clades is relatively recent, only eight million years ago (Särkinen et al., 2013), it is possible that the rule could be applied to the *Sli* gene and high levels of tissue specific expression could be associated with low rates of ω. Quantitative *Sli* gene expression experiments along with protein activity assays are needed to confirm this hypothesis.

The tests for selection using BUSTED on the *Sli* coding sequence alignment determined that the entire *Sli* gene is not under positive selection. Analysis with MEME and FUBAR, however, indicated that positive selection acted on specific nucleotide positions. FUBAR detected five amino acid sites under positive diversifying selection, and all of them overlapped with the ones detected by MEME. MEME also identified additional sites under positive selection, likely because the MEME algorithm is able to detect both episodic and pervasive positive selection (Murrell et al., 2012).

Two of the sites detected by MEME and FUBAR as being under positive selection are part of the ten amino acids that differ between *Sli* in D19-11 and M6 (Tables 4 and 5, sites 42 and 56). Three amino acids (218, 242, and 251) in the second protein domain are also under positive selection based on our results. This domain is thought to determine the specificity of interaction with S-RNAses (Vieira et al., 2019).

Our study identified two pocket types, but these had no correlation between *Sli* with the insertion and *Sli* without the insertion. The two predicted pocket types overlap only in a core set of four amino acids. Perhaps the differences in the amino acids present in the pockets detected would explain the specificities in Sli–S-RNAse recognition, but further investigation is needed to show this. Also, further investigation is needed to determine if differences in their ability to recognize and inhibit S-RNAses are quantitative or temporal. Tissue-specific *Sli* expression in pollen tubes (Eggers et al., 2021) or modifications to expression levels in combination with a reduction in S-locus diversity (Igic et al., 2008) may have a synergistic role in conferring self-compatibility. Overall, a better understanding of the interactions between Sli and S-RNAses at the molecular level is needed before we can make well-supported conclusions regarding their evolutionary models.

5 | SUMMARY

This study demonstrated that the *Sli* gene carrying the 533 bp insertion is present in wild potato germplasm at a low frequency. The low

TABLE 6 Amino acid composition of large and small functional pockets detected in Sli protein variants.

Variant	Insertion	Fertility	M 101	I 102	S 103	A 104	R 105	L 107	I 109	T 111	V 113	T 121	H 124	F 129	S 130
ktz_472952 (D19-11)	A	SC	L	L	L	L	L							L	L
brc_246536b	A	SC	L	L	L	L	L							L	L
ktz_472928	A	SC	L	L	L	L	L							L	L
chc_133619	A	SC	L	L	L	L	L						L	L	L
pnt_347766-1	A	SC				S	S	S	S	S	S	S			
pnt_347766-2	A					S	S	S	S	S	S	S			
ktz_472923-1	A	SC				S	S	S	S	S	S	S			
ktz_472923-2	A					S	S	S	S	S	S	S			
chm_365339_1	A	Unknown				S	S	S	S	S	S	S			
chm_365339_2	A					S	S	S	S	S	S	S			
ifd_558088	A	SI				S	S	S	S	S	S	S			
ber_458365-1	A	SI				S	S	S	S	S	S	S			
ber_458365-2	A					S	S	S	S	S	S	S			
chc_M6	P	SC				S	S	S	S	S	S	S			
brc_473368-1	P	SC				S	S	S	S	S	S	S			
brc_473368-2	P					S	S	S	S	S	S	S			

Note: Insertion refers to the protein variant encoded by *Sli* alleles with or without the 533 bp insertion in the 5′UTR: A = absent, P = present. Shaded amino acid residues are shared between the large and small protein pocket types.
Abbreviations: L, amino acid residue is present in the large type of protein pocket; S, amino acid residue is present in the small type of protein pocket; SC, source genotype is self-compatible; SI, source genotype is self-incompatible.

TABLE 6 (Continued)

Variant	E 131	V 132	A 133	V 163	R 218	V 219	D 220	W 222	E 224	Q 259	G 260	I 261	E 262	R 264
ktz_472952 (D19-11)	L	L		L	L	L	L		L	L	L	L	L	L
brc_246536b	L			L	L	L	L	L	L	L	L	L	L	L
ktz_472928	L	L		L	L	L	L	L	L	L	L	L	L	L
chc_133619	L	L		L	L	L	L		L	L	L	L	L	L
pnt_347766-1	S	S	S											
pnt_347766-2	S	S	S											
ktz_472923-1	S	S	S											
ktz_472923-2	S	S	S											
chm_365339_1	S	S	S											
chm_365339_2	S	S	S											
ifd_558088	S	S	S											
ber_458365-1	S	S	S											
ber_458365-2	S	S	S											
chc_M6	S	S	S											
brc_473368-1	S	S	S											
brc_473368-2	S	S	S											

Note: Insertion refers to the protein variant encoded by *Sli* alleles with or without the 533 bp insertion in the 5′UTR: A = absent, P = present. Shaded amino acid residues are shared between the large and small protein pocket types.
Abbreviations: L, amino acid residue is present in the large type of protein pocket; S, amino acid residue is present in the small type of protein pocket; SC, source genotype is self-compatible; SI, source genotype is self-incompatible.

variability of sequences carrying this insertion and their well supported phylogenetic clustering indicates that they have a single origin. Although the *Sli* allele with the insertion confers self-compatibility in cultivated potato germplasm, the presence of this allele is not a guarantee of self-compatibility in wild potato species and a lack of *Sli* alleles with the insertion is not a guarantee of self-incompatibility. In

TABLE 7 Self-compatible genotypes.

Species	PI/ID	Ploidy	Insertion	Sli protein
S. brevicaule	473368	6X	P	Complete
	246536-b	2x	A	Complete
S. chacoense	133619	2x	A	Complete
	M6	2x	P	Complete
S. kurtzianum	175435	2x	A	Complete
	472923	2x	A	Complete
	472928	2x	A	Complete
	472952	2x	A	Complete
S. pinnatisectum	347766	2x	A	Complete
S. polyadenium	175444	2x	A	Truncated
S. tuberosum	US-W4	2x	P	Complete
	IVP 101	2x	A	Complete
S. tuberosum L. subsp. Andigenum	195158	2x	A	Complete
	230513	2x	A	Complete
S. verrucosum	161173	2x	A	Complete
	275260-a	2x	A	Complete
	607848	2x	A	Complete

Note: Plant introduction (PI) indicates the accession in the US Potato genebank that was the source of the individual clone phenotyped. Genotypes maintained clonally by the genebank are indicated by name of the clone (ID). a and b designate specific genotypes of an accession with multiple genotypes in the study. Not all genotypes in the accession were self-compatible. Insertion: A = Absent, P = present.

silico protein analysis indicated that multiple Sli protein variants are predicted to be functional.

AUTHOR CONTRIBUTIONS

Mercedes Ames generated *Sli* sequence data, performed DNA sequence analysis, selection analysis and protein modeling, wrote and edited the manuscript, and prepared all figures and tables. Andy Hamernik conducted fertility assays and reviewed the manuscript. William Behling and David S. Douches contributed germplasm, provided guidance on experimental approaches, and edited the manuscript. Dennis A. Halterman and Paul C. Bethke developed the research project, secured funding, and wrote or edited multiple versions of the text and figures.

ACKNOWLEDGMENTS

This research was funded by ARS project number 5090-21220-005-000-D and by USDA NIFA Specialty Crop Research Initiative grant (2019-51181-30021). This research used resources provided by the SCINet project of the USDA Agricultural Research Service, ARS project number 0500-00093-001-00-D.

CONFLICT OF INTEREST STATEMENT

The authors declare they have no conflict of interest.

PEER REVIEW

The peer review history for this article is available in the Supporting for this article.

DATA AVAILABILITY STATEMENT

Sequence data are available from NCBI Popset: 2594803833. Fertility data are available in the Supporting Information and raw data can be shared upon request to authors.

ORCID

Mercedes Ames https://orcid.org/0000-0002-5376-5117
Andy Hamernik https://orcid.org/0009-0006-7006-2778
William Behling https://orcid.org/0000-0003-1757-9699
David S. Douches https://orcid.org/0000-0002-0636-5356
Dennis A. Halterman https://orcid.org/0000-0002-2428-9393
Paul C. Bethke https://orcid.org/0000-0001-7507-9962

REFERENCES

Ames, M., & Spooner, D. M. (2010). Phylogeny of *Solanum* series *Piurana* and related species in *Solanum* section *Petota* based on five conserved ortholog sequences. *Taxon*, *59*(4), 1091–1101. https://doi.org/10.1002/tax.594009

Bashir, I., Nicolao, R., & Heiden, G. (2021). *Wild potatoes: A genetic reservoir for potato breeding* (pp. 215–240). Elsevier eBooks. Published online January 1. https://doi.org/10.1016/b978-0-12-822137-2.00012-6

Bethke, P. C., Halterman, D., Francis, D. M., et al. (2022). Diploid potatoes as a catalyst for change in the potato industry. *American Journal of Potato Research.*, *99*(5–6), 337–357. https://doi.org/10.1007/s12230-022-09888-x

Bradshaw, J. E. (2022). Breeding diploid F1 hybrid potatoes for propagation from botanical seed (TPS): Comparisons with theory and other crops. *Plants.*, *11*(9), 1121. https://doi.org/10.3390/plants11091121

Broz, A. K., Randle, A. M., Sianta, S. A., Tovar-Méndez, A., McClure, B., & Bedinger, P. A. (2017). Mating system transitions in *Solanum habrochaites* impact interactions between populations and species. *The New Phytologist*, *213*(1), 440–454. https://doi.org/10.1111/nph.14130

Bryan, G. J., McLean, K., Waugh, R., & Spooner, D. M. (2017). Levels of intra-specific AFLP diversity in tuber-bearing potato species with different breeding systems and ploidy levels. *Frontiers in Genetics.*, *8*, 8. https://doi.org/10.3389/fgene.2017.00119

Cipar, M. S., Peloquin, S. J., & Hougas, R. W. (1964). Inheritance of incompatibility in hybrids between *Solanum tuberosum* haploids and diploid species. *Euphytica*, *13*(2), 163–172. https://doi.org/10.1007/bf00033305

Clot, C. R., Polzer, C., Prodhomme, C., Schuit, C., Engelen, C. J. M., Hutten, R. C. B., & van Eck, H. J. (2020). The origin and widespread occurrence of Sli-based self-compatibility in potato. *Theoretical and Applied Genetics.*, *133*(9), 2713–2728. https://doi.org/10.1007/s00122-020-03627-8

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2) giab008. https://doi.org/10.1093/gigascience/giab008

Devaux, A., Goffart, J.-P., Kromann, P., Andrade-Piedra, J., Polar, V., & Hareau, G. (2021). The Potato of the Future: Opportunities and Challenges in Sustainable Agri-food Systems. *Potato Research*, *64*(4), 681–720. https://doi.org/10.1007/s11540-021-09501-4

Dzidzienyo, D. K., Bryan, G. J., Wilde, G., & Robbins, T. P. (2016). Allelic diversity of S-RNase alleles in diploid potato species. *Theoretical and Applied Genetics.*, *129*(10), 1985–2001. https://doi.org/10.1007/s00122-016-2754-7

Eggers, E. J., van der Burgt, A., van Heusden, S. A. W., de Vries, M. E., Visser, R. G. F., Bachem, C. W. B., & Lindhout, P. (2021). Neofunctionalisation of the *Sli* gene leads to self-compatibility and facilitates precision breeding in potato. *Nature Communications*, *12*(1), 4141. https://doi.org/10.1038/s41467-021-24267-6

Eijlander, R., ter Laak, W., Hermsen, J. G. T., Ramanna, M. S., & Jacobsen, E. (2000). Occurrence of self-compatibility, self-incompatibility and unilateral incompatibility after crossing diploid *S. tuberosum* (SI) with *S. verrucosum* (SC): I. Expression and inheritance of self-compatibility. *Euphytica*, *115*(2), 127–139. https://doi.org/10.1023/a:1003902907599

Fajardo, D., & Spooner, D. M. (2011). Phylogenetic relationships of *Solanum* series *Conicibaccata* and related species in *Solanum* section *Petota* inferred from five conserved ortholog sequences. *Systematic Botany*, *36*, 163–170. https://doi.org/10.1600/036364411X553252

Fernandez-Pozo, N., Menda, N., Edwards, J. D., Saha, S., Tecle, I. Y., Strickler, S. R., Bombarely, A., Fisher-York, T., Pujar, A., Foerster, H., Yan, A., & Mueller, L. A. (2015). The sol genomics network (SGN)—From genotype to phenotype to breeding. *Nucleic Acids Research*, *43*-(Database issue), D1036–D1041. https://doi.org/10.1093/nar/gku1195

Fumia, N., Pironon, S., Rubinoff, D., Khoury, C. K., Gore, M. A., & Kantar, M. B. (2022). Wild relatives of potato may bolster its adaptation to new niches under future climate scenarios. *Food and Energy Security*, *11*(e360). https://doi.org/10.1002/fes3.360

Hanneman, R. E. (1985). Self-fertility in *Solanum chacoense*. *American Journal of Potato Research*, *62*, 428–429.

Hosaka, K., & Hanneman, R. E. (1998). Genetics of self-compatibility in a self-incompatible wild diploid potato species *Solanum chacoense*. 1. Detection of an S locus inhibitor (Sli) gene. *Euphytica*, *99*, 191–197. https://doi.org/10.1023/A:1018353613431

Huang, B., Ruess, H., Liang, Q., Colleoni, C., & Spooner, D. M. (2019). Analyses of 202 plastid genomes elucidate the phylogeny of Solanum section *Petota*. *Scientific Reports*, *9*(1), 4454. https://doi.org/10.1038/s41598-019-40790-5

Igic, B., Lande, R., & Kohn, J. R. (2008). Loss of self-incompatibility and its evolutionary consequences. *International Journal of Plant Sciences.*, *169*(1), 93–104. https://doi.org/10.1086/523362

Jansky, S., Charkowski, A. O., Douches, D. S., et al. (2016). Reinventing potato as a diploid inbred line-based crop. *Crop Science.*, *56*(4), 1412–1422. https://doi.org/10.2135/cropsci2015.12.0740

Jansky, S. H., Chung, Y. S., & Kittipadukal, P. (2014). M6: A diploid potato inbred line for use in breeding and genetics research. *Journal of Plant Registrations.*, *8*(2), 195–199. https://doi.org/10.3198/jpr2013.05.0024crg

Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*(14), 3059–3066. https://doi.org/10.1093/nar/gkf436

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, *10*(6), 845–858. https://doi.org/10.1038/nprot.2015.053

Kosakovsky Pond, S. L., Posada, D., Gravenor, M. B., Woelk, C. H., & Frost, S. D. W. (2006). Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Biology and Evolution*, *23*(10), 1891–1901. https://doi.org/10.1093/molbev/msl051

Kubo, K., Entani, T., Takara, A., Wang, N., Fields, A. M., Hua, Z., Toyoda, M., Kawashima, S. I., Ando, T., Isogai, A., Kao, T. H., & Takayama, S. (2010). Collaborative non-self recognition system in S-RNase–based self-incompatibility. *Science*, *330*(6005), 796–799. https://doi.org/10.1126/science.1195243

Le Guilloux, V., Schmidtke, P., & Tuffery, P. (2009). Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics*, *10*, 168. Published 2009 Jun 2. https://doi.org/10.1186/1471-2105-10-168

Lindhout, P., Meijer, D., Schotte, T., Hutten, R. C. B., Visser, R. G. F., & van Eck, H. J. (2011). Towards F1 hybrid seed potato breeding. *Potato Research.*, *54*(4), 301–312. https://doi.org/10.1007/s11540-011-9196-z

Ma, L., Zhang, C., Zhang, B., Tang, F., Li, F., Liao, Q., Tang, D., Peng, Z., Jia, Y., Gao, M., Guo, H., Zhang, J., Luo, X., Yang, H., Gao, D., Lucas, W. J., Li, C., Huang, S., & Shang, Y. (2021). A nonS-locus F-box gene breaks self-incompatibility in diploid potatoes. *Nature Communications*, *12*(1), 4142. https://doi.org/10.1038/s41467-021-24266-7

McClure, B., Cruz-García, F., & Romero, C. (2011). Compatibility and incompatibility in S-RNase-based systems. *Annals of Botany.*, *108*(4), 647–658. https://doi.org/10.1093/aob/mcr179

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S. L., & Scheffler, K. (2013). FUBAR: A fast, unconstrained Bayesian AppRoximation for inferring selection. *Molecular Biology and Evolution*, *30*(5), 1196–1205. https://doi.org/10.1093/molbev/mst030

Murrell, B., Weaver, S., Smith, M. D., Wertheim, J. O., Murrell, S., Aylward, A., Eren, K., Pollner, T., Martin, D. P., Smith, D. M., Scheffler, K., & Kosakovsky Pond, S. L. (2015). Gene-wide identification of episodic selection. *Molecular Biology and Evolution*, *32*(5), 1365–1371. https://doi.org/10.1093/molbev/msv035

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., & Kosakovsky Pond, S. L. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics*, *8*(7), e1002764. https://doi.org/10.1371/journal.pgen.1002764

Ortiz, R., & Mihovilovich, E. (2020). Genetics and cytogenetics of the potato. In H. Campos & O. Ortiz (Eds.), *The potato crop. Its agricultural, nutritional and social contribution to humankind* (pp. 219–247). Springer. ISBN: 978-3-030-28683-5.

Pál, C., Papp, B., & Hurst, L. D. (2001). Highly expressed genes in yeast evolve slowly. *Genetics*, *158*(2), 927–931. https://doi.org/10.1093/genetics/158.2.927

Peterson, B. A., Holt, S. H., Parker, F., Laimbeer, E., Doulis, A. G., Coombs, J., Douches, D. S., Hardigan, M. A., Buell, C. R., & Veilleux, E. (2016). Self-fertility in a cultivated diploid potato populations examined with the Infinium 8303 potato single-nucleotide polymorphism array. *The Plant Genome*, *9*(3). https://doi.org/10.3835/plantgenome2016.01.0003

Potato Genome Sequencing Consortium, Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., Zhang, G., Yang, S., Li, R., Wang, S., & Orjeda, G. (2011). Genome sequence and analysis of the tuber crop potato. *Nature*, *475*(7355), 189–195. https://doi.org/10.1038/nature10158

Pretz, C., & Smith, S. D. (2021). Intraspecific breakdown of self-incompatibility in *Physalis acutifolia* (Solanaceae). *AoB Plants*, *14*(1), plab080. https://doi.org/10.1093/aobpla/plab080

Särkinen, T., Bohs, L., Olmstead, R. G., & Knapp, S. (2013). A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): A dated 1000-tip tree. *BMC Evolutionary Biology.*, *13*(1), 214. https://doi.org/10.1186/1471-2148-13-214

Shibai, A., Kotani, H., Sakata, N., Furusawa, C., & Tsuru, S. (2022). Purifying selection enduringly acts on the sequence evolution of highly expressed proteins in *Escherichia coli*. *G3 Genes|Genomes|Genetics*, *12*(11). https://doi.org/10.1093/g3journal/jkac235

Spooner, D. M., & Castillo, R. C. (1997). Reexamination of series relationships of South American wild potatoes (Solanaceae: *Solanum* sect. *Petota*): Evidence from chloroplast DNA restriction site variation. *American Journal of Botany.*, *84*(5), 671–685. https://doi.org/10.2307/2445904

Spooner, D. M., Ghislain, M., Simon, R., Jansky, S. H., & Gavrilenko, T. (2014). Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. *The Botanical Review.*, *80*(4), 283–383. https://doi.org/10.1007/s12229-014-9146-y

Spooner, D. M., Ruess, H., Arbizu, C. I., Rodríguez, F., & Solís-Lemus, C. (2018). Greatly reduced phylogenetic structure in the cultivated potato clade (*Solanum* section *Petota* pro parte). *American Journal of Botany*, *105*(1), 60–70. https://doi.org/10.1002/ajb2.1008

Spooner, D. M., & van den Berg, R. G. (1992). An analysis of recent taxonomic concepts in wild potatoes (*Solarium* sect. *Petota*). *Genetic Resources and Crop Evolution*, *39*(1), 23–37. https://doi.org/10.1007/bf00052651

Spooner, D. M., Alvarez, N., Peralta, I. E., & Clausen, A. M. (2016). Taxonomy of wild potatoes and their relatives in southern South America (*Solanum* sects. *Petota* and *Etuberosum*). *Systematic Botany Monographs*, *100*, 1–240. http://www.jstor.org/stable/26155422

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stokstad, E. (2019). The new potato. *Science*, *363*(6427), 574–577. https://doi.org/10.1126/science.363.6427.574

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—New capabilities and interfaces. *Nucleic Acids Research*, *40*(15), e115. https://doi.org/10.1093/nar/gks596

van Lieshout, N., van der Burgt, A., de Vries, M. E., ter Maat, M., Eickholt, D., Esselink, D., van Kaauwen, M. P. W., Kodde, L. P., Visser, R. G. F., Lindhout, P., & Finkers, R. (2020). Solyntus, the new highly contiguous reference genome for potato (*Solanum tuberosum*). *G3: Genes, Genomes, Genetics*, *10*(10), 3489–3495. https://doi.org/10.1534/g3.120.401550

Vieira, J., Rocha, S., Vázquez, N., López-Fernández, H., Fdez-Riverola, F., Reboiro-Jato, M., & Vieira, C. P. (2019). Predicting specificities under the non-self gametophytic self-incompatibility recognition model. *Frontiers in Plant Science*, *10*, 10. https://doi.org/10.3389/fpls.2019.00879

Weaver, S., Shank, S. D., Spielman, S. J., Li, M., Muse, S. V., & Kosakovsky Pond, S. L. (2018). Datamonkey 2.0: A modern web application for characterizing selective and other evolutionary processes. *Molecular Biology and Evolution*, *35*(3), 773–777. https://doi.org/10.1093/molbev/msx335

Wu, Y., Li, D., Hu, Y., Li, H., Ramstein, G. P., Zhou, S., Zhang, X., Bao, Z., Zhang, Y., Song, B., Zhou, Y., Zhou, Y., Gagnon, E., Särkinen, T., Knapp, S., Zhang, C., Städler, T., Buckler, E. S., & Huang, S. (2023). Phylogenomic discovery of deleterious mutations facilitates hybrid potato breeding. *Cell*, *186*(11), 2313–2328.e15. https://doi.org/10.1016/j.cell.2023.04.008

Yang, Z., & Bielawski, J. P. (2000). Statistical methods for detecting molecular adaptation. *Trends in Ecology & Evolution.*, *15*(12), 496–503. https://doi.org/10.1016/s0169-5347(00)01994-7

Yates, C. M., Filippis, I., Kelley, L. A., & Sternberg, M. J. E. (2014). SuSPect: Enhanced prediction of single amino acid variant (SAV) phenotype using network features. *Journal of Molecular Biology.*, *426*(14), 2692–2701. https://doi.org/10.1016/j.jmb.2014.04.026

Zhang, C., Wang, P., Tang, D., Yang, Z., Lu, F., Qi, J., Tawari, N. R., Shang, Y., Li, C., & Huang, S. (2019). The genetic basis of inbreeding depression in potato. *Nature Genetics.*, *51*(3), 374–378. https://doi.org/10.1038/s41588-018-0319-1

Zhang, C., Yang, Z., Tang, D., Zhu, Y., Wang, P., Li, D., Zhu, G., Xiong, X., Shang, Y., Li, C., & Huang, S. (2021). Genome design of hybrid potato. *Cell*, *184*(15), 3873–3883.e12. https://doi.org/10.1016/j.cell.2021.06.006

Zhang, J., & Yang, J. R. (2015). Determinants of the rate of protein sequence evolution. *Nature Reviews Genetics.*, *16*(7), 409–420. https://doi.org/10.1038/nrg3950

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ames, M., Hamernik, A., Behling, W., Douches, D. S., Halterman, D. A., & Bethke, P. C. (2024). A survey of the *Sli* gene in wild and cultivated potato. *Plant Direct*, *8*(5), e589. https://doi.org/10.1002/pld3.589