Abdallah et al. *Horticulture Research* (2020)7:170
https://doi.org/10.1038/s41438-020-00392-z

Horticulture Research
www.nature.com/hortres

ARTICLE Open Access

Self-compatibility in peach [*Prunus persica* (L.) Batsch]: patterns of diversity surrounding the *S*-locus and analysis of SFB alleles

Donia Abdallah[1], Ghada Baraket[1], Veronica Perez[2], Amel Salhi Hannachi[1] and Jose I. Hormaza[3]

Abstract

Self-incompatibility (SI) to self-compatibility (SC) transition is one of the most frequent and prevalent evolutionary shifts in flowering plants. *Prunus* L. (Rosaceae) is a genus of over 200 species most of which exhibit a Gametophytic SI system. Peach [*Prunus persica* (L.) Batsch; $2n = 16$] is one of the few exceptions in the genus known to be a fully self-compatible species. However, the evolutionary process of the complete and irreversible loss of SI in peach is not well understood and, in order to fill that gap, in this study 24 peach accessions were analyzed. Pollen tube growth was controlled in self-pollinated flowers to verify their self-compatible phenotypes. The linkage disequilibrium association between alleles at the *S*-locus and linked markers at the end of the sixth linkage group was not significant ($P > 0.05$), except with the closest markers suggesting the absence of a signature of negative frequency dependent selection at the *S*-locus. Analysis of SFB1 and SFB2 protein sequences allowed identifying the absence of some variable and hypervariable domains and the presence of additional α-helices at the C-termini. Molecular and evolutionary analysis of SFB nucleotide sequences showed a signature of purifying selection in SFB2, while the SFB1 seemed to evolve neutrally. Thus, our results show that the SFB2 allele diversified after *P. persica* and *P. dulcis* (almond) divergence, a period which is characterized by an important bottleneck, while SFB1 diversified at a transition time between the bottleneck and population expansion.

Introduction

A common observation in plants is the adoption, during evolution, of several strategies to prevent selfing, and thus, promote outcrossing limiting the deleterious effects of inbreeding. In fact, the presence of physical barriers between the female and male parts to avoid self-fertilization, such as dichogamy, monoecy, dioecy, or floral heteromorphy, has evolved in both gymnosperms and angiosperms. In addition, the most common way to avoid self-fertilization in angiosperms is self-incompatibility (SI), a system present in more than half of flowering plant species (more than 100 families). This system has been known from at least 1876, when Charles Darwin observed that some plant species were sterile to their own pollen, but fertile when pollinated from pollen of other individuals of the same species. Firstly, those mechanisms were gathered under the term "self-sterility", then the term "self-incompatibility" (SI) was proposed and defined as "*the inability of a plant producing functional gametes to reproduce when it is self-pollinated*"[1].

Maintenance of such ancient system requires strong evolutionary benefits to counteract the advantages of selfing, such as a higher reproductive success. Yet, the loss of SI represents perhaps the most frequent shift in Angiosperm evolution[2]. However, once lost, any system of homomorphic SI is extremely difficult to regain[3–5]. The homomorphic SI system is controlled by a single Mendelian locus (the *S*-locus), which is comprised of tightly

Correspondence: Donia Abdallah (abdallahdonia@yahoo.fr)
[1]Faculté des Sciences de Tunis, Département Biologie, Université de Tunis El Manar, 2092 Tunis, Tunisie
[2]Laboratorio de Agrobiología Juan José Bravo Rodríguez (Cabildo Insular de La Palma), Unidad Técnica del Instituto de Productos Naturales y Agrobiología (IPNA-CSIC), 38700 S/C La Palma, Canary Islands, Spain
Full list of author information is available at the end of the article

© The Author(s) 2020
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

linked genes determining self-recognition specificities and many accessory genes which are also necessary for the proper function of SI[1,6].

Several studies have addressed SI to self-compatibility (SC) transitions. In fact, many phylogenetic analyses using macro- and microevolutionary models have been performed to understand the causes and consequences of these evolutionary shifts. The majority of these studies concluded that SC species have emerged from SI species[7] and a variety of causes of these reversions have been hypothesized. Mutations affecting genes both linked and unlinked to the *S*-locus seem to be the major causes[4]. On the other hand, the loss of SI in natural populations is often associated with the reduction of population size, which leads to the reduction of sexual compatible partners and the number of SI alleles[8,9]. This would tend to reduce gene flow between populations with different mating systems, and may eventually lead to reproductive isolation and speciation[10]. Hence, the relationship between the loss of SI and speciation is of particular interest[11]. Genetic evidence for the relation between loss of SI and speciation has been accumulated in Brassicaceae. For instance, in *Capsella rubella* and *Leavenworthia alabamica* race a4, the loss of SI was associated with both the split from the closely related outcrossers (*Capsella grandiflora* and *Leavenworthia alabamica* race a1, respectively) and a strong genome-wide genetic bottleneck[11–13]. Bottleneck has also played a role in SI breakdown in small founding populations by depletion of genetic diversity[14–16]. Moreover, range expansion, has been postulated as a driving force behind SI loss, as it may favor self-fertilization[16,17].

Prunus L. (Rosaceae), a genus of over 200 species of deciduous and evergreen trees and shrubs with economically important fruit and nut crops[18], exhibits an RNase-based Gametophytic SI system (GSI). In this system, pollen rejection occurs in the style if the *S*-allele of the haploid pollen matches one of the *S*-alleles present in the diploid pistil. The gene that controls the female function encodes a stylar ribonuclease (*S*-RNase), whereas a pollen specific F-Box gene (SFB) has been identified as the responsible of the pollen function[19]. SI is generally the rule in this genus and most species are partially or fully self-incompatible. However, in several mostly self-incompatible species, SC genotypes are observed and the transition from SI to SC has been attributed to different causes in the different species. This reversion is mostly related to mutations in pistil and pollen *S*-locus determinants[20,21]. For instance, an insertion upstream of the *S*6m-RNase in sour cherry (*P. cerasus*)[22] and a similar mutation in Japanese plum (*P. salicina*), *Se*-RNase[23], reduces the *S*-RNase expression levels leading to an insufficient accumulation of *S*-RNase in the pistil which breaks the recognition function[21]. In sweet cherry (*P. avium*)[24] and Japanese apricot (*P. mume*)[25], the SC phenotype was associated with indels in the SFB codifying region causing a frame-shift in translation that produces a nonfunctional truncated protein[20]. In apricot (*P. armeniaca*), two different mutations conferring SC, an insertion in the SFBc allele that produces an SFBc truncated protein and a mutation in *S*-locus unlinked factors, also called modifier genes (m), both independently have been shown to cause the loss of pollen-*S* function[21,26]. In almond (*P. amygdalus*), SC has been attributed to an inactive *Sf*-RNase protein as a result of an *Sf* allele. Nevertheless, a similar *Sf*a (active) allele encodes an active *Sf* protein. This apparent paradox was resolved by the discovery that *Sf*i and *Sf*a are epialleles differing by the methylation of a single nucleotide upstream of the coding sequence[27,28].

An exception to the widespread SI in the genus is peach [*Prunus persica* (L.) Batsch; $2n = 16$], a fully self-compatible diploid species with no recent whole-genome duplication. SI to SC reversion in peach remains a puzzling issue and few works have addressed this topic. The wild ancestor of cultivated peach remains unknown and it is probably extinct, although closely related species such as *P. davidiana*, *P. kansuensis* and *P. mira* are cultivated in some regions in China[29]. A first question regarding the evolution of SC in peach is if domestication of this crop that took place about 7000 years ago in China[30] could have been involved in the SC transition. In fact, several thousand years of domestication have produced more than 1000 cultivars of *P. persica* worldwide, with significant phenotypic differences in fruit size, flavor, and flower type[31]. However, Tao et al.[32] proposed that if human selection pressure for SC had been the main reason behind the reversion to SC, we would also expect many SC selections in other species of the genus with a large history of cultivation, such as almond, cherry and plum, and this is not the case. At the molecular level, the loss of SI in peach is mainly attributed to a deficiency in pollen *S*-gene expression (SFB) codifying a nonfunctional truncated protein. To date, only four SFB alleles have been identified in peach. Earlier studies reported two SFB alleles, SFB1 and SFB2[21,32,33]. Later, two additional alleles, SFB3 and SFB4, have been identified[34–36]. Tao et al.[32] have suggested a possible pressure of weak selection for SC at the beginning of peach speciation.

The main objective of this study was to understand the mechanism of the SI to SC transition in *P. persica*. For that, twenty-four peach accessions from different origins were used. The first step was to verify the self-compatible phenotype of the used samples through observation of pollen tube growth in self-pollinated flowers. The second step was to elucidate the pattern of genetic diversity based on nine SSR loci surrounding the *S*-locus at the sixth

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

linkage group. The final step was to sequence and analyze the nucleotide and peptide sequences of the SFB alleles of the studied peach genotypes.

Results

Pollination tests

Pollination tests for each genotype were carried out on flowers collected at the balloon stage (Fig. 1a). Each flower was self-pollinated and no intercrosses were made. Pollen germination on the stigma was successful in all cultivars tested (Fig. 1b). After germination, the pollen tubes were arrested in the style of self-incompatible plum (Cidre) used as reference (Fig. 1c), while the pollen tubes reached the bases of the styles in all self-pollinated peach flowers as well as in flowers of "Bedri", a self-compatible plum confirming the SC phenotype of the peach genotypes studied in this work (Fig. 1d).

Variability patterns around the *S*-locus

Genetic diversity

Since in peach, as in other *Prunus* species, the genes directly involved in SI system are clustered at one end of the sixth linkage group[36], we explored the pattern of diversity along this region using the gametophytic self-incompatibility locus (*S*-locus) and nine SSR markers flanking the *S*-locus mapped in the *Prunus*-TE-F2 linkage map (http://www.rosaceae.org/peach/genome) (Fig. 2a, b).

A total of 39 alleles (*Ao*) in the 10 loci studied were amplified in the 24 peach accessions (Table 1). The observed heterozygote fraction ($Ho = 0.4$) was lower than the expected ($He = 0.44$). The PaCons1F/R2 primer, specific to the second intron of the *S*-RNAse gene, amplified the lowest number of alleles ($Ao = 2$). The heterozygote level was very low ($Ho = 0.19$) giving a low power of discrimination ($Pd = 0.19$). The CPPCT023 locus, which is ~38 cM upstream the *S*-locus, showed the highest number of alleles ($Ao = 6$) and power of discrimination ($Pd = 0.72$) (Table 1).

Linkage disequilibrium analysis

A total of 29.5% of intra-chromosome pair comparisons showed significant LD in the genotypes studied. LD association was calculated by considering the significance of LD blocks between each of the alleles at the first and second loci. A total of 65% of the blocks showed significant LD over the tested loci in the analyzed accessions.

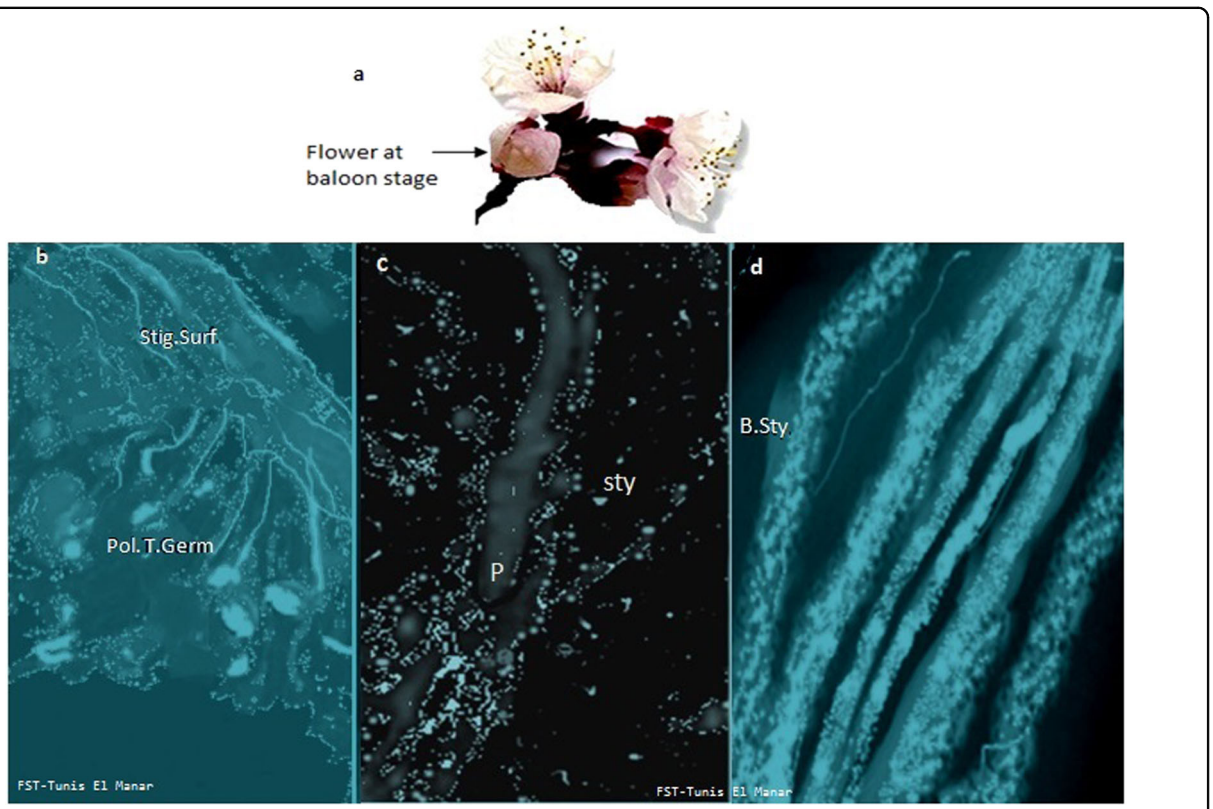

Fig. 1 Pollen grain germination and pollen tube growth. a The flowers of each genotype were collected at the balloon stage, **b** the pollen grain germination was successful at the surface of stigma in self-pollinated flowers, **c** pollen tube growth was arrested at the style of self-incompatible plum "Cidre" used as reference, and **d** the pollen tubes reached the base of the style in peach and self-compatible plum "Bedri" used as reference. Stig Surf surface of stigma. Pol T Germ pollen tube germination. Sty style. P pollen tube. B Sty base of style. Scale bars = 50 µm

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

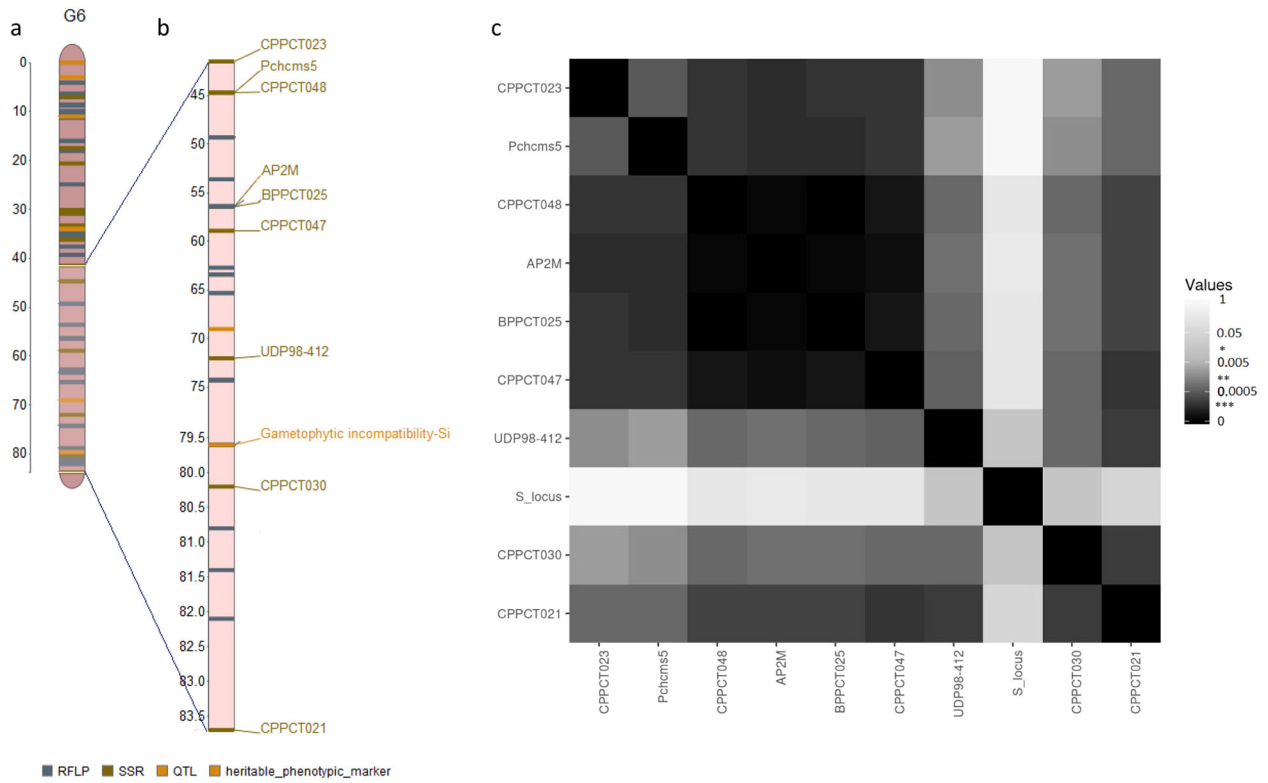

Fig. 2 Linkage disequilibrium extent at the end of the sixth LG. a Schematic representation of the sixth chromosome in *Prunus*, **b** genetic position of the tested loci mapped in the *Prunus*-TE-F2 linkage map (http://www.rosaceae.org/peach/genome), **c** heatmap representing the distribution of *P* values per pairs of the ten tested loci. *: significance of *P* value

Table 1 Characteristics and genetic diversity of the 10 loci amplified in the 24 peach accessions studied in this work

Locus	Loci characteristics		Genetic diversity			
	Genetic position (cM)[a]	References	*Ao*	*Ho*	*He*	PD
CPPCT023	41.50	Dirlewanger et al.[85]	6	0.5	0.53	0.72
Pchcms5	44.70	Sosinski et al.[86]	3	0.42	0.39	0.48
CPPCT048	44.70	Dirlewanger et al.[85]	5	0.53	0.49	0.6
AP2M	56.40	Aranzana et al.[87]	4	0.39	0.35	0.57
BPPCT025	56.40	Dirlewanger et al.[85]	3	0.42	0.48	0.47
CPPCT047	58.90	Dirlewanger et al.[85]	5	0.45	0.49	0.61
UDP98-412	72.00	Testolin et al.[88]	4	0.32	0.37	0.52
PaCons1F/R2	79.6	Sonneveld et al.[89]	2	0.19	0.33	0.19
CPPCT030	80.20	Dirlewanger et al.[85]	4	0.39	0.42	0.52
CPPCT021	83.70	Dirlewanger et al.[85]	3	0.45	0.48	0.47
Average	–	–	3.9	0.4	0.44	0.5

Ao number of observed alleles, *Ho* observed heterozygosity, *He* expected heterozygosity, *PD* power of discrimination
[a]Genetic position of the SSR loci in centimorgan (cM) in the *Prunus*-TE-F2 linkage map (http://www.rosaceae.org/peach/genome)

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026



Fig. 3 Example of SFB allele amplification in 13 peach samples with PsSFB-F1 and PsSFB-R1. L: 1 KB ladder. R1, R2, and R3: Plum samples used as control with previously known *S*-genotypes. R1 Fortune, R2 Santa Rosa, R3 Beauty. Kha Kharfi, Mes Meski, Bar Bargou, Kar Khoukh Arbi, Bou Boutabgaya, Amb Amber, Ess Essifi, NP Negra Palmera, RM Rojo Mollar, VNZ Venezolano, AM Amarillo Melocoton, RR Rubby Rich, SL Spring Lady

The heatmap showed that LD association between pair of alleles at the *S*-locus and those linked was not significant ($P > 0.05$), except with the closest locus UDP98-412 (Fig. 2c).

SFB allele analysis

PCR amplification and sequencing

The SFB gene was amplified in the 24 peach accessions. Two bands were obtained, one of 1150 bp and one of 1270 bp (Fig. 3). The obtained fragments were purified and sequenced. Obtained sequences were compared to the data available in Genbank using BlastX and, thus, aligned with the peach SFB1, SFB2, SFB3, and SFB4 alleles. The band of 1270 bp shared more than 98% similarity with the peach SFB1 allele (AB252414) whereas the band of 1150 bp shared more than 98% similarity with the peach SFB2 allele (AB252416).

The SFB alleles obtained in this work were deposited in Genbank under accession numbers KY629929–KY629934 and MN125684–MN125701 (Supplementary Material 1). The SFB genotypes of different samples were identified and detailed in Table 2.

Structural features of peach SFB alleles

To explore the different domains of the SFB *Prunus* gene, we carried out alignments of the deduced amino acid sequences with plum SFB peptides alleles (Fig. 4). The plum SFB peptide sequences showed five domains: the F-box domain localized at the N-termini, two variable regions (designated V1 and V2) located downstream the F-Box, and two hypervariable regions (designated HVa and HVb) at the C-termini[37]. However, the peach SFB1 allele lacked the HVb region whereas the peach SFB2 allele lacked the V2, HVa, and HVb regions (Fig. 4).

Protein secondary structure analyses showed that, overall, the proportion of the β-strands (E) and coils was higher than α-helices in both proteins (Fig. 5). SFB1 protein structure showed 10 α-helices and 23 β-strands (Fig. 5a), while protein structure prediction generated for SFB2 showed 11 α-helices and 19 β-strands (Fig. 5b). The first 60 amino acid fragment showed no differences in the locations and the number of α-helices and β- strands. Variation was observed in the number, nature and position of secondary structural elements from the 70th amino acid. The main difference is related to the presence of additional α-helices in the SFB2 peptide sequence at the 237–243th (α8-helix) and the 299–312nd (α9-helix) amino acid fragments. The C-termini region of both proteins showed considerable differences in the locations and the number of α-helices and β-strands. However, they shared an α-helix conformation at the 411–422nd (α10-helix) fragment in SFB1 and the 342–350th (α10-helix) fragment in SFB2.

In addition, the obtained structures were compared to peptide sequences of self-incompatible almond (SFBk)

Table 2 Origin and *S*-genotypes of the 24 peach accessions analyzed in this work

Origin	Sample	*S*-genotype
Tunisia	Kharfi	S1S2
Tunisia	Meski	S1S2
Tunisia	Bargou	S1S2
Tunisia	Khoukh Arbi	S1S2
Tunisia	Boutabgaya	S1S2
Tunisia	Amber	S1S2
Tunisia	Essifi	S1S2
Tunisia	Bargo limaoui	S2S2
Tunisia	Khoukh ahmer	S1S1
Tunisia	Platine	S2S2
Spain	Amarillo Melocoton	S1S1
Spain	Blanco Mollar	S1S1
Spain	Mollar	S1S1
Spain	Negra Palmera	S2S2
Spain	Venezolano	S2S2
Spain	Rojo Mollar	S2S2
Spain	Amarillo Merollo	S1S2
USA	Rubby Rich	S1S1
USA	Spring Lady	S1S1
USA	Sun Late	S1S2
USA	Fleur De Star	S1S2
USA	Scup	S1S2
USA	Queen Crest	S1S2
USA	Rich May	S1S2

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

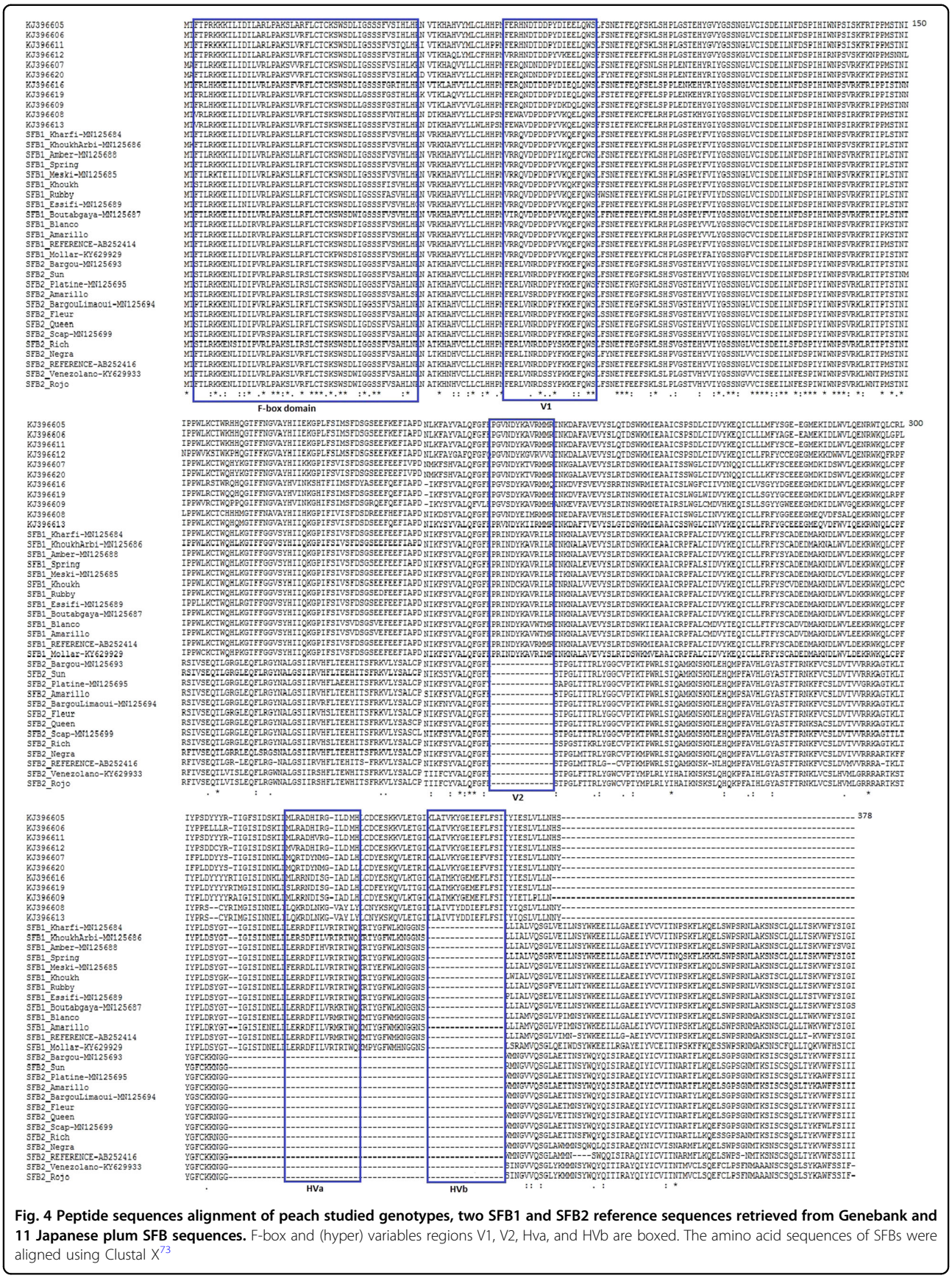

Fig. 4 Peptide sequences alignment of peach studied genotypes, two SFB1 and SFB2 reference sequences retrieved from Genebank and 11 Japanese plum SFB sequences. F-box and (hyper) variables regions V1, V2, Hva, and HVb are boxed. The amino acid sequences of SFBs were aligned using Clustal X[73]

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

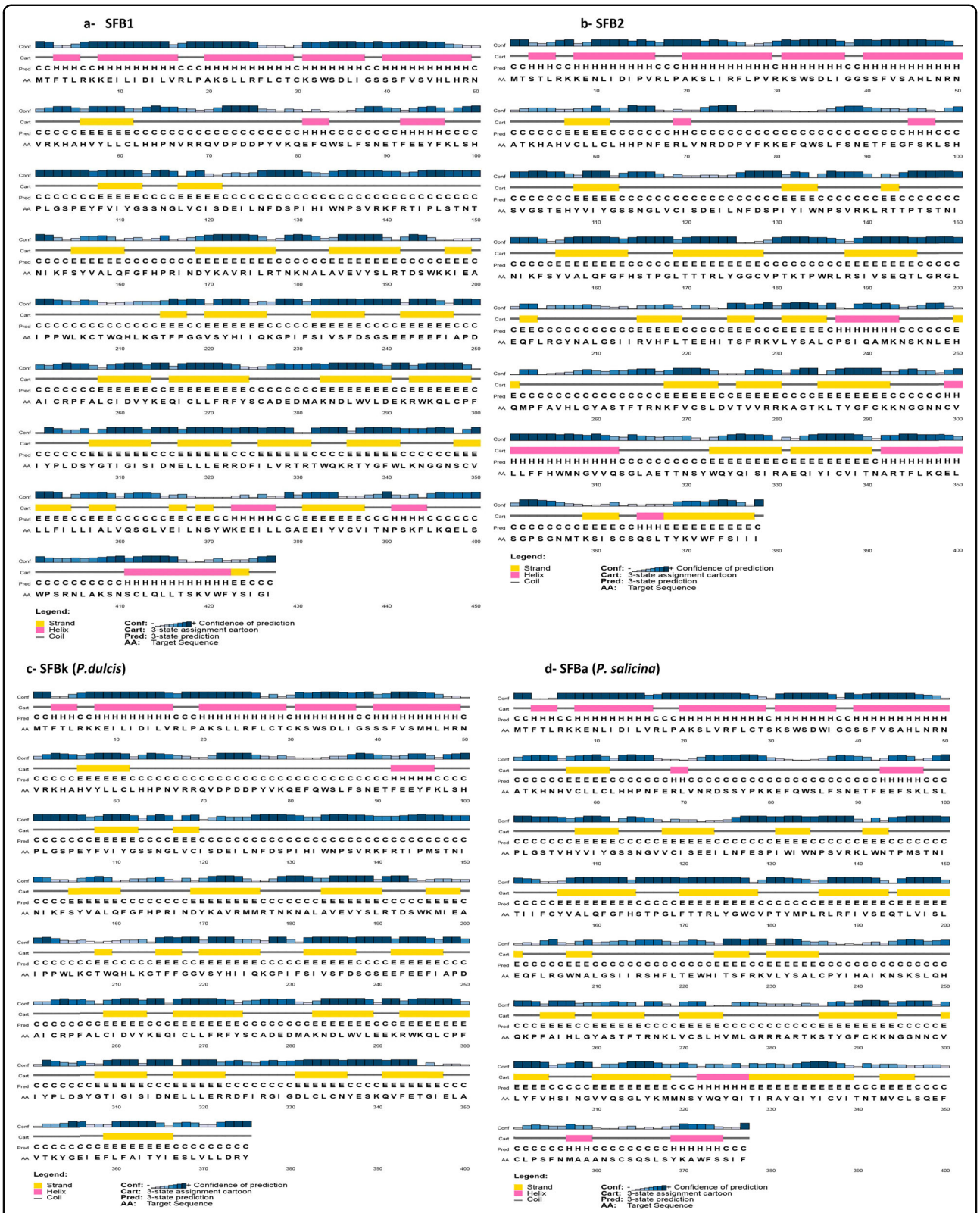

Fig. 5 Secondary protein structures predictions of consensus SFB1 sequence (**a**) SFB2 sequence(**b**), SFBk (**c**), and SFBa (**d**). The figures showed the distribution of the β-strands (E), coils and α-helices along the four proteins. Structures were done on the website: http://bioinf.cs.ucl.ac.uk/psipred

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

Table 3 Patterns of nucleotide diversity and mutational variation in peach SFB nucleotide sequences

	Nucleotide diversity					Mutations								
	N	Poly.S	Sing	Par.Inf	π^a	*M*	Syn.M	N.Syn.M	$\pi_S{}^a$	$\pi_A{}^a$	π_A/π_S	$K_S{}^a$	$K_A{}^a$	K_A/K_S
SFB1	13	161	112	49	0.08	175	39	136	0.025	0.027	1	0.023	0.025	1
SFB2	13	228	95	133	0.05	244	58	186	0.028	0.055	2	0.05	0.025	0.5

N total number of sequences, *Poly.S* polymorphic sites, *Sing* number of singletons, *Par.Inf* parsimony informative sites, π nucleotide diversity, *M* total number of mutations, *Syn.M* number of synonymous mutations, *N.Syn.M* number of non-synonymous mutations, π_S nucleotide diversity at synonymous sites, π_A nucleotide diversity at non-synonymous sites, K_S divergence between sequences at synonymous sites, K_A divergence between sequences at non-synonymous sites
[a]Estimates with Jukes and Cantor correction

Table 4 Summary of neutrality tests

	McDonald–Kreitman test						Tajima			Fu and Li	
	Fixed		Polymorphic		f_A/f_S	P_A/P_S	D_T	D_S	D_A	*D**	*F**
	f_S	f_A	P_S	P_A							
SFBp1	5	18	34	118	3.6	3.4	−1.71	−1.5	−0.9	−1.7	−2
SFBp2	19	31	53	141	1.63	2.6	−1.15	−0.97	−1.5	−0.45	−0.72

f_S fixed synonymous substitutions between sequences, f_A fixed non-synonymous substitutions between sequences, P_S polymorphic synonymous substitutions between sequences, P_A polymorphic non-synonymous substitutions between sequences, D_T Tajima statistic for all sites, D_S Tajima statistic at synonymous sites, D_A Tajima statistic at non-synonymous sites. *D** and *F** Fu and Li statistics

and self-incompatible plum (SFBa) (Fig. 5c, d respectively), both considered as the ancestral copies of SFB1 and SFB2 respectively[31]. The first 60 amino acids conserve the same location and number of α-helices and β- strands. After that, differences in number and position of the β-strands between each protein and its ancestral copy were observed. The main difference between SFB2 and SFBa resides in the additional α-helices found in SFB2 (α8-and α9-helices). Moreover, the comparison of the C-termini of the four protein structures showed the presence of the additional α-helices found at SFB1 (α10-helix) and SFB2 (α10-helix).

Polymorphism and divergence

Polymorphism in SFB alleles In SFB1, 161 polymorphic sites were identified, of which 112 are singletons and 49 are parsimony informative sites. The nucleotide diversity showed a value of 0.08. In SFB2, 228 polymorphic sites were revealed, of which 95 were identified as singletons and 133 were revealed as parsimony informative sites with a nucleotide diversity value of 0.05 (Table 3).

The majority of mutations occurred in non-synonymous positions for both SFB1 and SFB2 (N.Syn.M > Syn.M) (Table 3). In SFB1, nucleotide diversity showed similar values between non-synonymous ($\pi_A = 0.027$) and synonymous ($\pi_S = 0.025$) sites. In contrast, in SFB2, nucleotide diversity was higher at non-synonymous ($\pi_A = 0.55$) than at synonymous ($\pi_S = 0.028$) sites (Table 3).

Divergence at SFB alleles Divergence at non-synonymous and synonymous sites (K_A and K_S, respectively) provides information about the form of sequence evolution in a giving gene. Thus, pairwise sequence divergences were calculated.

As shown in Table 3, synonymous and non-synonymous sites are evolving at equal rates for SFB1 since $K_A/K_S = 1$. In SFB2, divergence at synonymous sites showed a value of $K_S = 0.05$, while the divergence at non-synonymous sites showed a value of $K_A = 0.025$.

Evolutionary implications

Neutrality tests: SFB alleles are selectively non-equivalent Result of the McDonald and Kreitman (MK) test in SFB1 sequences supported the null hypothesis[38] since the ratio of polymorphic sites at non-synonymous and synonymous sites ($P_A/P_S = 3.4$) equaled the number of fixed differences at non-synonymous and synonymous sites ($f_A/f_S = 3.6$). However, in SFB2 sequences, the ratio f_A/f_S was lower than the ratio P_A/P_S with values of 1.63 and 2.6, respectively (Table 4).

Based on the frequency spectrum of polymorphism, a second class of neutrality tests was conducted. Tajima[39] *D* and Fu and Li[40,41] *D** and *F** statistics showed values of −1.71, −1.7, and −2.0, respectively, in SFB1 and values of −1.15, −0.45, and −0.72, respectively, in SFB2 (Table 4). Tajima *D* statistic was calculated at synonymous and

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

non-synonymous sites separately. In SFB1, a negative D value of $D_{A1} = -0.9$ was observed at non-synonymous sites and a higher value of $D_{S1} = -1.5$ was revealed at non-synonymous ones. However, in SFB2, a value of $D_{A2} = -1.5$ was found at non-synonymous sites, while the D value at synonymous sites was $D_{S2} = -0.97$ (Table 4).

Evolutionary relationships between SFB1 and SFB2 sequences AUPGMA dendrogram and a Minimum Spanning Network were drawn in order to trace the evolutionary relationships between SFB1 and SFB2 alleles and their ancestral copies. In fact, the SFB1 allele is a pollen part mutant (PPM) related to the almond (*Prunus dulcis*) SFBk allele, while the SFB2 allele is a PPM version related to the Japanese plum (*Prunus salicina*) SFBa allele[31]. Two sequences of SFBk and SFBa were retrieved from GenBank data libraries under accessions numbers AB252408 (SFBk) and AB252410 (SFBa).

The UPGMA tree divided the sequences into two main groups (Fig. 6a). The first gathered the SFB1 and SFBk sequences, while the SFB2 and SFBa sequences fell into the second group. The topology of the dendrogram showed that the divergence between SFBa and SFB2 occurred before the divergence between SFBk and SFB1.

The minimum spanning network confirmed the UPGMA results and showed a higher number of mutational events between the SFB2 and SFBa alleles than between SFB1 and SFBk alleles (Fig. 6b). These mutational events represent the parsimony informative sites described in Table 3.

In order to estimate the approximate age of SFB1 and SFB2 alleles, the rho statistic implemented in the NETWORK software was computed between the ancestral nodes and the descendants' sequences. The mutational rate used in this work was 10^{-3} mutation per year for GSI *S*-alleles[40]. The approximate age of SFBa compared to all SFB2 sequences was 13,1571.429 years, whereas the age of SFBk compared to all SFB1 sequences obtained was 66,642.857 years. The approximate ages of the oldest SFB peach alleles were estimated by calculating the rho statistic between the closest SFB sequences (SFB1_control and SFB2 control) to the ancestral alleles. The approximate age of SFB1_control compared to SFBk was 25,500 years, whereas the approximate age of SFB2_control was 53,000 years.

Discussion

In this study, 24 peach cultivars were used to analyze the loss of the SI system in this species. Pollination tests were carried out to confirm the self-compatible phenotype of the studied genotypes. Genetic diversity around the *S*-locus was tested using nine surrounding SSR loci. In addition, we analyzed the SFB alleles, SFB1 and SFB2, since it has been proposed that SC in peach is mainly due to defective SFB alleles[32].

Genetic diversity of peach cultivars around the *S*-locus

Linkage disequilibrium (LD) analysis at one end of the sixth peach LG showed nonsignificant LD between the *S*-locus and the surrounding SSR loci. In contrast, significant LD blocks on LG6, particularly between the *S*-locus and closely related SSR markers, were observed in other *Prunus* species such as wild and sweet cherry[42] as well as in other species in the Bassicaceae such as *Arabidopsis thaliana*, *Arabidopsis lyrata*[43], *Capsella rubella*[43,44], and *Biscutella neustriaca*[44,45]. This significant LD is due to the influence of negative frequency dependent selection (NFDS) force in the genomic neighborhood of the *S*-locus. NFDS is a form of balancing selection by which rare alleles escape loss by drift. This selective force can extend the high polymorphism at the selected locus to closely linked neutral sites[46]. Thus, the absence of significant LD between the *S*-locus and closely related SSR markers in *P. persica* species suggests the absence of NFDS forces at the *S*-locus in this species.

Analysis of the SFB gene in *P. persica*

In this study, two of the four SFB alleles reported so far in peach were revealed in the 24 accessions studied. The absence of SFB3 and SFB4 alleles may confirm that SFB3 may be specific to ornamental cultivars, while SFB4 is restricted to wild genotypes as revealed by Hanada et al.[34] although other studies such as that of Gu et al.[35] have reported the presence of SFB3 and SFB4 in cultivated varieties.

On the other hand, the results showed a predominance of S1S2 genotypes in Tunisian cultivars which may confirm previous observations on the presence of a geographical feature of the *S*-genotypes in peach[34,35]. In contrast, in self-incompatible *Prunus* species such plum[47] and sweet cherry[48], a lack of correlation between *S*-alleles and the geographic origin of the cultivars was observed indicating that *S*-alleles coalesced before *Prunus* speciation[47]. Thus, *P. persica* SFB alleles seem to have diversified after peach speciation.

A common structure in the SFB peptide sequence in different *Prunus* species is the presence of the F-box and the hypervariable and variable regions[35,49]. The lack of variable and hypervariable regions in SFB peptide sequences confirmed that the SFB peach gene encodes a truncated protein. Similar results were obtained by Gu et al.[35] where the SFB1 and SFB3 alleles lacked the HVb region, whereas the SFB2 and SFB4 alleles lacked the V2, HVa and HVb regions. In fact, the V2, HVa, and HVb regions are located at the 3′ end of the SFB gene and play a crucial role in the allele-specific recognition function implying that the C-terminal region may be exposed on

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

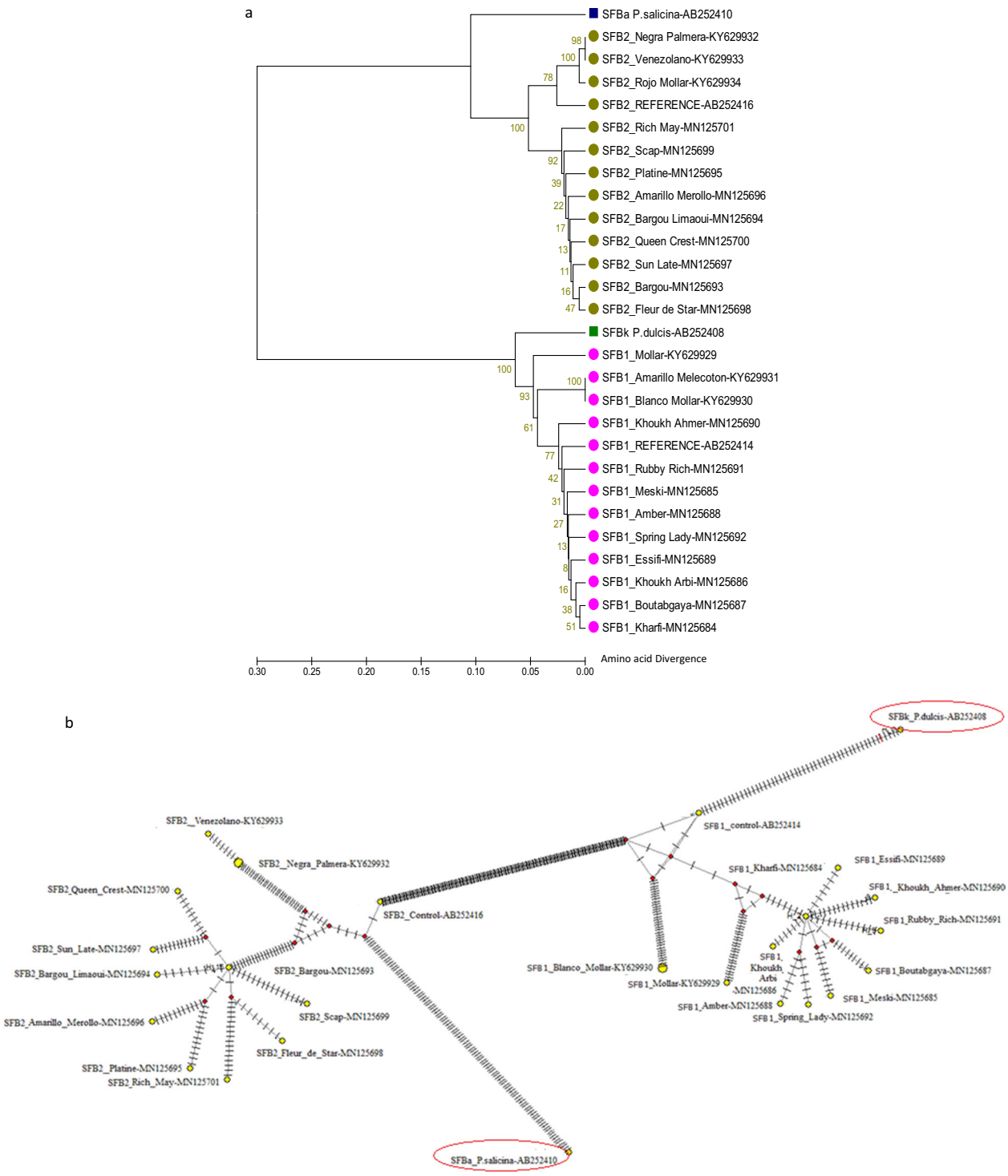

Fig. 6 Evolutionary relationships between the SFBp alleles and their ancestral copies. a UPGMA tree: the optimal tree with the sum of branch length = 1.16720195 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. The analysis involved 28 amino acid sequences. There were a total of 374 positions in the final dataset. **b** Minimum spanning network: the network was constructed using the median joining algorithm. The estimated number of mutations of the shortest tree = 651. The total number of taxa = 28. The total number of haplotypes is 26

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

the surface of the SFB protein, and, thus, be responsible for the discrimination between self and nonself *S*-RNase proteins[36,47]. For this reason, the most obvious structural difference between SFB1, SFB2 and their ancestral versions, involved the presence of α-helices at SFB1 HVb region and SFB2 V2, HVa, and HVb regions. A similar result was observed in five almond genotypes, where differences related to the conformation of secondary structure elements were found at the variable V4 and hypervariable RHV regions[28]. Nevertheless, further studies in peach are required to ascertain whether these structural differences mediate SC.

Analysis of selective forces at the *P. persica* SFB alleles

The analysis of polymorphism and divergence of the SFB1 allele as well as the MK test ($P_A/P_S \approx f_A/f_S$) suggest that the SFB1 allele seems to evolve neutrally. In fact, the null hypothesis of the MK test implied that the ratio of polymorphic sites at non-synonymous and synonymous sites (P_A/P_S) should equal the number of fixed differences at non-synonymous and synonymous sites (f_A/f_S)[38]. However, the negative values obtained by Tajima *D* statistic and Fu and Li's D^* and F^* tests in SFB1 sequences, suggest a recent selective sweep (which removed all the variation in the region), a recent population expansion, and/or an ongoing purifying selection because all tend to produce alleles at low frequency[50]. Previous works[32,48,51] have denied the notion of a selective sweep in the *Prunus* SFB gene. On the other hand, the higher negative *D* value of Tajima at synonymous than at non-synonymous sites ($Da1 < Ds1$) observed in SFB1 alleles, confirmed neither population expansion nor purifying selection. In fact, demographic factors, such as a population expansion, equally affect all genes and all regions of a gene which would create a homogeneous effect on all types of mutations and we would expect equal Tajima values of *D* statistic at synonymous and at non-synonymous sites ($Da < Ds$). However, selective forces, such as purifying selection, directly affect the genetic diversity at a target site and modify the genetic diversity within different sites[52]. Hence, purifying selection has a heterogeneous effect by eliminating deleterious non-synonymous mutations but a little or no effect on neutrally evolving synonymous changes and we would expect a *D* value higher at non-synonymous than at synonymous sites ($Da > Ds$)[50,52]. This result suggests that the SFB1 allele has diversified from its progenitor (SFBk) in a transition time between two different events, one resulting in a homogenous effect and the other resulting on a heterogeneous effect. In fact, a transition time between homogenous and heterogeneous processes (such population expansion and purifying selection respectively) would result in a period *when neither population expansion nor purifying selection can be detected* as described by Hahn et al.[50]. A similar result was observed by Hahn et al.[50] in an experimental system of T7 bacteriophage. At the statistical level, the slight difference between mutational classes was due to selective processes and the slight skew at synonymous mutations reflected demographic processes.

In SFB2, the relatively higher non-synonymous polymorphism ($\pi_A = 0.055$) suggested that non-synonymous replacements are more common than synonymous ones giving $\pi_A/\pi_S > 1$. However, the lower K_A/K_S ratio suggests that most of the non-synonymous mutations could be deleterious[53]. These results indicate that the deleterious mutations have not been fixed in the SFB sequences and, thus, created a strong non-synonymous nucleotide diversity such noted by Hughes[53]. Afterward, they would have been eliminated by a strong ongoing purifying selection. The signature of purifying selection was confirmed by the MK test ($f_A/f_S < P_A/P_S$), suggesting that a high proportion of non-synonymous changes were disadvantageous and, therefore, strongly affected polymorphism. In addition, the increased negative *D* value at non-synonymous sites observed in SFB2 ($Da > Ds$) revealed a signature of a heterogeneous effect and confirmed the signature of a strong ongoing purifying selection.

A hypothetical evolutionary scenario for peach self-incompatibility breakdown

In light of the obtained results, SFB2 allele diversified before SFB1, most probably soon after peach speciation. As reported by Ohta[54], at the time of speciation, the magnitudes of drift may increase and the selective force may change. In particular, after a strong bottleneck, such as that of peach in recent breeding history, the selection coefficients change through space and time[55]. In addition, Tao et al.[32] and Chen et al.[36] have reported a selection pressure for SC at the beginning of peach speciation. PPMs might preferentially be selected compared to pistil part mutants due to the higher number of pollen grains produced by each flower[36]. Yet, at the time of SFB2 diversification, purging of deleterious mutations was more efficient because of the smaller population size giving rise to reduced inbreeding depression and/or a decreased *S*-alleles number[13,56,57].

Due to selfing, we would expect a population expansion after the bottleneck. At the transition time between bottleneck and population expansion, SFB1 would have diversified from SFBk. In parallel, purifying selection would become increasingly efficient in purging deleterious changes getting closer to neutrality.

At the transition between bottleneck and population expansion, the effect of various amino acid changes became effectively neutral. Once diversified, SFB1 alleles behave like neutral alleles. Such status is called "The Nearly-Neutral Theory"[58], and represented a development of Kimura's Neutral Theory of molecular evolution. The

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

Nearly-Neutral Theory makes testable predictions that go beyond the mere null model of the Neutral Theory i.e., a theory of "no effect"[54]. Most of those predictions strongly support the results reported in this work. One of these predictions is that slightly deleterious variants will accumulate in a species that has undergone a severe bottleneck, and then they will have more chances of being purged by slightly purifying selection than being fixed by positive selection. In fact, unlike strictly neutral alleles, the fate of nearly neutral alleles depends on effective population size. Thus, when the effective population size gets larger for a long time, selection will decrease the frequency of slightly deleterious variants in the population and eventually eliminate them[54].

Association between demographic events and self-incompatibility breakdown

One question regarding the SC reversion in peach that remains unclear is whether SI–SC transition occurred before or after peach speciation and the results of this work do not allow solving this question. In fact, it is unclear whether the loss of SI in peach was due to the loss of function mutations (in which case the SC transition should have taken place before peach speciation) or to the impossibility to select a compensatory mutation due to purifying selection (in which case the SC transition should have taken place after peach speciation). Regardless of when the SC reversion took place in peach evolutionary history, our results showed that the loss of SI in peach was the result of an association between mutational events, bottleneck and population expansion.

A similar association between transition from SI to SC and strong bottleneck has been observed in several taxa. For example, in Brassicaceae, the loss of SI in *Capsella rubella* was associated with the split from its SI progenitor species *Capsella grandiflora* (50–100,000 years ago in relation to migration into glacial refuges) and a strong genetic bottleneck[11,13,59,60]. Also, the split between the SI *Leavenworthia alabamica* race a1 and the SC race a4 (150,000 years ago) was associated with a shift from SI to SC, with strong genome-wide genetic bottleneck and evolution of the selfing syndrome[12,13,61], although it is still unclear whether the SC transition occurred at the time of the split or later[13].

Conclusion

During evolution, peach has lost irreversibly the SI system. Our results show that no NFDS takes place at the *S*-locus and confirm a probable purifying selection system in SFB2, while SFB1 seems to evolve neutrally. The loss of the SI system in peach was due to the loss-of-function mutations in the C-terminal regions. In addition, the evolutionary history of *P. persica* played a crucial role to make complete and irreversible the loss of SI.

Materials and methods

Plant material

Twenty-four peach accessions from different origins were used in this work (Table 2). Ten of them were collected in Tunisia from growers' orchards and represent local endemic accessions; seven accessions were developed in the US and were collected from a Tunisian germplasm collection (Sodon); seven accessions were collected from farmer orchards in La Palma Island (Canary Islands, Spain).

Self-incompatibility phenotype analysis

To verify the SC phenotype of the 24 peach genotypes analyzed, self-pollinations were carried out in the laboratory. Two plum cultivars ["Bedri" (SeSh), known to be self-compatible, and "Cidre" (SaSe) considered as self-incompatible[61]] were used as controls that proof the success of this experimental approach in different species. Flowers of each genotype were collected at the balloon stage and pollen was obtained by manually removing and drying the anthers at room temperature during 24 h. The pollen was then sieved through a fine mesh and stored at −20 °C until required[62–64]. For pollinations, 15–20 flowers of each accession were collected at the balloon stage, 24 h before anthesis, emasculated and maintained on wet florist foam at laboratory temperature[64]. On the following day, the flowers were hand pollinated using a fine paintbrush. Three days later, the pollinated pistils were fixed in FAA [70% ethanol: acetic acid: formaldehyde (18: 1: 1, v/v/v)][62,63]. For microscope preparations, the fixed pistils were washed three times for 1 h with distilled water and left in 5% sodium sulfite at 4 °C. To soften the tissues, the pistils were autoclaved at 1 kg/cm^2 during 10 min in sodium sulfite[65], and stained with 0.1% (v/v) aniline blue in 0.1 N K_3PO_4 to stain callose[66]. Pollen tube growth in the style was observed under an Olympus BH2 microscope with UV.

SSR analysis

DNA extraction and amplification

Fresh young leaves were collected from each genotype and frozen. Genomic DNA was extracted using the cetyltrimethyl ammonium bromide (CTAB) method following the protocol described by Doyle and Doyle[67]. Extracted DNA was quantified using a Nanodrop 1000 and diluted to 10 ng/μL.

Ten loci on the sixth linkage group of peach were amplified: the second intron of the *S*-RNase gene (the female GSI component) of the *S*-locus and nine surrounding SSR markers (Table 2). PCR reactions were carried out in a volume of 20 µl, with 20 mMTris–HCl, pH 8.4, 50 mMKCl, 4 mM MgCl2, 0.1 mM of each dNTP, 0.2 µM of each primer, 40 ng of genomic DNA and 0.45 U of BioTaq™ DNA polymerase (Bioline, London, UK). PCR

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

reactions were run in an I-cycler (Bio-Rad Laboratories, Hercules, CA, USA) thermocycler using the following temperature cycles: an initial step of 2 min at 94 °C, 35 cycles of 45 s at 94 °C, 45 s at 57 °C, 1 min at 72 °C, and a final step of 5 min at 72 °C. The obtained fragments were analyzed by capillary electrophoresis with a Beckman Coulter GenomeLabGeXP^{TM} capillary DNA analysis system. Forward primers were labeled with a fluorescent dye on the 5-end. Samples were denaturalized at 90 °C for 120 s, injected at 2.0 kV for 30 s, and separated at 6.0 kV for 35 min.

Diversity analyses

To explore the genetic diversity, the number of observed alleles (*Ao*), the observed (*Ho*) and expected (*He*) heterozygosities and the power of discrimination (PD) of the *S*-locus and SSR loci were calculated with Arlequin ver. 3.5[68]. LD was estimated by computing squared allele frequency correlations (r^2) between each pair of alleles among the 10 loci with PowerMarker 3.25 software, considering unphased genotype data[69]. Since the distribution of allele frequencies may have an effect on the extent of LD, rare alleles whose frequencies are <5%, were excluded prior to further LD analyse[42].The distribution of *P* values per pairs of the ten tested loci was represented by heatmap using the Heatmapper.ca online server[70].

SFB gene analysis

The SFB gene was amplified using the general primers PsSFB-F_1 and PsSFB-R_1[71] with PCR specifications described in Abdallah et al.[47]. The amplified fragments were separated using a 1.5% agarose gel electrophoresis, stained with SYBRGreen and visualized with UV light. Fragment size estimation was done using a size standard (1 kb DNA Ladder; Invitrogen, Carlsbad, CA, USA). Amplification products were purified using the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) and sequenced with a Ready Reaction Big Dye Terminator Cycle Sequencing automated sequencer.

Datasets and sequence alignment

Nucleotide sequences were translated to amino acids using the DAMBE program[72], then DNA and peptide alignments were carried out using the accurate CLUSTALX algorithm version 1.64b[73] and minor adjustments were performed as described by Hammer et al.[74]. Database searches were performed using the National Center for Biotechnology Information's: Basic Local Alignment Search Tool (BLAST).

The consensus peptide sequences were obtained using the HIV Sequence Database (http://www.hiv.lanl.gov/). The protein secondary structure predictions were conducted using PSIPRED Version 4.01 (http://bioinf.cs.ucl.ac.uk/psipred)[75,76].

Sequence analysis

Diversity parameters and neutrality tests were calculated using DNAsp 5.1[77]. All alleles were included in the calculations and all sites with alignment gaps were eliminated.

Nucleotide diversity (π) was measured as the average number of pairwise nucleotide differences among all sequences[77]. Rates of synonymous and non-synonymous substitutions (π_S and π_A) in the coding regions were estimated via the approximate method of Nei and Gojobori[78] with the Jukes and Cantor correction[79] for multiple hits. The mean ratio of the number of non-synonymous substitutions per non-synonymous site (K_A) to the number of synonymous substitutions per synonymous site (K_S) was also calculated.

The MK test[38], Tajima statistic (D)[39], and Fu and Li's D* and F*[40] were calculated. In addition, Tajima's D was calculated in partitioned synonymous and non-synonymous data (D_S and D_A, respectively).

The evolutionary history was inferred using UPGMA method[80]. The evolutionary distances were computed using the Poisson correction method[81] and are in the units of the number of amino acid substitutions per site. The analysis involved the amino acid sequences. All ambiguous positions were removed for each sequence pair and evolutionary analyses were conducted in MEGA6[82].

The Minimum Spanning Network between SFB nucleotide sequences was run into the Network software ver.10.0.0.0using the Median joining algorithm[83]. The rho statistic, estimated by Network, measures the age of an ancestral node in mutational units. This mutational age was then converted into years by multiplication with the mutation rate[84].

Acknowledgements

This research was supported by the Tunisian "Ministère de l'Enseignement Supérieur et de la Recherche Scientifique" and "Ministerio de Economía y Competitividad—European Regional Development Fund, European Union (AGL2016-77267-R and PID2019-109566RB-I00). The authors would like to gratefully thank "Sodon collection" (Bou-Selem, Tunisia) and Tunisian farmers for kindly providing plant material.

Author details

[1]Faculté des Sciences de Tunis, Département Biologie, Université de Tunis El Manar, 2092 Tunis, Tunisie. [2]Laboratorio de Agrobiología Juan José Bravo Rodríguez (Cabildo Insular de La Palma), Unidad Técnica del Instituto de Productos Naturales y Agrobiología (IPNA-CSIC), 38700 S/C La Palma, Canary Islands, Spain. [3]Instituto de Hortofruticultura Subtropical y Mediterránea La Mayora (IHSM La Mayora-UMA-CSIC), 29750 Algarrobo-Costa, Malaga, Spain

Author contributions

D.A. performed the experiments, performed statistical analyses, developed the genetic analyses, and wrote the paper. G.B. discussed and corrected the content. V.P. provided some plant material and discussed the results. A.S.H. and

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

J.I.H. provided experimental instructions, supervised the work, and assisted in writing the paper.

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary Information accompanies this paper at (https://doi.org/10.1038/s41438-020-00392-z).

Received: 17 June 2020 Revised: 8 August 2020 Accepted: 12 August 2020
Published online: 01 October 2020

References

1. De Nettancourt, D. *Incompatibility and Incongruity in Wild and Cultivated Plants*. 2nd edn, (Springer, Berlin, 2001).
2. Voillemot, M. & Pannell, J. R. Maintenance of mixed mating after the loss of self-incompatibility in a long-lived perennial herb. *Ann. Bot.* **119**, 177–190 (2017a).
3. Igic, B., Bohs, L. & Kohn, J. R. Ancient polymorphism reveals unidirectional breeding system shifts. *PNAS* **103**, 1359–1363 (2006).
4. Igic, B., Lande, R. & Kohn, J. R. Loss of self-incompatibility and its evolutionary consequences. *Int. J. Plant. Sci.* **169**, 93–104 (2008).
5. Tsuchimatsu, T. et al. Patterns of polymorphism at the self-incompatibility locus in 1,083 *Arabidopsis thaliana* genomes. *Mol. Biol. Evol.* **34**, 1878–1889 (2017).
6. Takayama, S. A. & Isogai Self-incompatibility in plants. *Annu. Rev. Plant. Biol.* **58**, 467–489 (2005).
7. Voillemot, M. & Pannell, J. R. Inbreeding depression is high in a self-incompatible perennial herb population but absent in a self-compatible population showing mixed mating. *Ecol. Evol.* **7**, 8535–8544 (2017b).
8. Gaude, T., Glémin, S., Cabrillac, D. & Mignot, A. L'auto-incompatibilité chez les plantes à fleurs. *Med. Sci. (Paris)* **17**, 6–7 (2001).
9. Wright, S. I., Kalisz, S. & Slotte, T. Evolutionary consequences of self fertilization in plants. *Proc. R. Soc. Lon. [Biol].* **280**, 20130–20133 (2013).
10. Tang, C. L. et al. The evolution of selfing in *Arabidopsis thaliana*. *Science* **317**, 1070–1072 (2007).
11. Guo, Y. L. et al. Recent speciation of *Capsella rubella* from *Capsella grandiflora*, associated with loss of self-incompatibility and an extreme bottleneck. *PNAS* **106**, 5246–5251 (2009).
12. Busch, J. W., Joly, S. & Schoen, D. J. Demographic signatures accompanying the evolution of selfing in *Leavenworthia alabamica*. *Mol. Biol. Evol.* **28**, 1717–1729 (2011).
13. Vekemans, X., Poux, C., Goubet, P. M. & Castric, V. The evolution of selfing from outcrossing ancestors in Brassicaceae: what have we learned from variation at the S-locus? *J Evol. Biol.* **27**, 1372–1385 (2014).
14. Nei, M., Maruyama, T. & Chakraborty, R. Bottleneck effect and genetic variability in populations. *Evolution* **29**, 1–10 (1975).
15. Keller, S. R. & Taylor, D. R. History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecol. Lett.* **11**, 852–866 (2008).
16. Tedder, A., Ansell, S. W., Lao, X., Vogel, J. C. & Mable, B. K. Sporophytic self-incompatibility genes and mating system variation in Arabisalpina. *Ann. Bot.* **108**, 699–713 (2011).
17. Baker, H. G. Support for Baker's Law—as a rule. *Evolution* **21**, 853–856 (1967).
18. Chin, S. W., Shaw, J., Haberle, R., Wen, J. & Potter, D. Diversification of almonds, peaches, plums and cherries - molecular systematics and biogeographic history of *Prunus* (Rosaceae). *Mol. Phylogenetics. Evol.* **76**, 34–48 (2014).
19. Aguiar, B. et al. Convergent evolution at the gametophytic self-incompatibility system in *Malus* and *Prunus*. *Plos ONE* **10**, e0126138 (2015).
20. Yamane, H. & Tao, R. Molecular basis of self-(in)compatibility and current status of S-genotyping in Rosaceous fruit trees. *J. Jpn. Soc. Hort. Sci.* **78**, 137–157 (2009).
21. Zuriaga, E. et al. An S-locus independent pollen factor confers self-compatibility in 'Katy' apricot. *Plos ONE* **8**, e53947 (2013).
22. Yamane, H., Tao, R., Sugiura, A., Hauck, H. R. & Iezzoni, A. F. Identification and characterization of S-RNases in tetraploid sour cherry (*Prunus cerasus*). *J. Am. Soc. Hortic. Sci.* **126**, 661–667 (2001).
23. Watari, A. et al. A low transcriptional level of Se-RNase in the Se-haplotype confers self-compatibility in Japanese plum. *J. Am. Soc. Hort. Sci.* **132**, 396–406 (2007).
24. Hauck, N. R., Yamane, H., Tao, R. & Iezzoni, A. F. Accumulation of nonfunctional S-haplotypes results in the breakdown of gametophytic self-incompatibility in tetraploid *Prunus*. *Genetics* **172**, 1191–1198 (2006).
25. Entani, T. et al. Comparative analysis of the self-incompatibility (S-) locus region of *Prunus mume*: Identification of a pollen-expressed-box gene with allelic diversity. *Genes Cells* **8**, 203–213 (2003).
26. Vilanova, S., Romero, C., Burgos, L., Llacer, G. & Badenes, M. L. Identification of self-(in)compatiblity alleles in apricot (*Prunus armeniaca* L.) by PCR and sequence analysis. *J. Am. Soc. Hort. Sci.* **130**, 893–898 (2005).
27. Fernández i Martí, A., Gradziel, T. M. & Socias i Company, R. Methylation of the Sf locus in almond is associated with S-RNase loss of function. In Kodad, O., López-Fran cos, A., Rovira, M., Socias i Company R. (eds). *XVI GREMPA Meeting on Almonds and Pistachios*. Options Méditerranéennes: Série A. Séminaires Méditerranéens **119**, 127–130 (CIHEAM, Zaragoza, 2016).
28. Shashi, N. Goonetilleke et al. Variation among S-locus haplotypes and among stylar RNases in almond. *Sci. Rep.* **10**, 583 (2020).
29. Yu, Y. et al. Genome re-sequencing reveals the evolutionary history of peach fruit edibility. *Nat. Commun.* **9**, 5404 (2018).
30. Zheng, Y., Crawford, G. W. & Chen, X. Archaeological evidence for peach (*Prunus persica*) cultivation and domestication in China. *PLos ONE* **9**, e106595 (2014).
31. Cao, Ke et al. Two-stage transcriptional reprogramming in *Saccharomyces cerevisiae* for optimizing ethanol production from xylose. *Metab. Eng.* **24**, 150–159 (2014).
32. Tao, R. et al. Self-compatible peach (*Prunus persica*) has mutant versions of the S haplotypes found in self-incompatible *Prunus* species. *Plant Mol. Biol.* **63**, 109–123 (2007).
33. Hegedus, A. & Halasz, J. Self-incompatibility in plums (*Prunus salicina* Lindl., *Prunus cerasifera* Ehrh. and *Prunus domestica* L.). A minireview. *J. Hortic. Sci.* **12**, 137–140 (2006).
34. Hanada, T. et al. Two novel self-compatible S haplotypes in peach (*Prunus persica*). *J. Jpn. Soc. Hortic. Sci.* **83**, 203–213 (2014).
35. Chao, Gu, Wang, Lu, Korban, SchuylerS. & Han, Yuepeng Identification and characterization of S-RNase genes and S-genotypes in *Prunus* and *Malus* species. *Can. J. Plant Sci.* **95**, 1–13 (2015).
36. Chen, Q. et al. SLFL genes participate in the ubiquitination and degradation reaction of S-RNase in self-compatible peach. *Front. Plant Sci.* **9**, 227 (2018).
37. Stoeckel, S. et al. Heterozygote excess in a self-incompatible and partially clonal forest tree species *Prunus avium* L. *Mol. Ecol.* **15**, 2109–2118 (2006).
38. McDonald, J. H. & Kreitman, M. Adaptive protein evolution at the Adh locus in Drosophila. *Nature* **351**, 652–654 (1991).
39. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
40. Fu, Y. & Li, W. Statistical tests of neutrality of mutations. *Genetics* **133**, 693–709 (1993).
41. Wright, S. The distribution of self-sterility alleles in populations. *Genetics* **24**, 538–552 (1939).
42. Arunyawat, U., Capdeville, G., Decroocq, V. & Mariette, S. Linkage disequilibrium in French wild cherry germoplasm and worldwide sweet cherry germoplasm. *Tree Genet. Genomes* **8**, 737–755 (2012).
43. Guo, Y. L., Zhao, X., Lanz, C. & Weigel, D. Evolution of the *S*-locus region in Arabidopsis relatives. *Plant Physiol.* **157**, 937–946 (2011).
44. Edh, K., Widen, B. & Ceplitis, A. Molecular population genetics of the SRK and SCR self-incompatibility genes in the wild plant species *Brassica cretica* (Brassicaceae). *Genetics* **181**, 985–995 (2009).
45. Leducq, J. B. et al. Self-Incompatibility in Brassicaceae: identification and characterization of SRK-Like sequences linked to the S-locus in the tribe Biscutelleae. *Genet. Sex* **4**, 983–992 (2014).
46. Kamau, E. & Charlesworth, D. Balancing selection and low recombination affect diversity near the self-incompatibility loci of the plant *Arabidopsis lyrata*. *Curr. Biol.* **15**, 1773–1778 (2005).
47. Abdallah, D., Baraket, G., Ben Tamarzizt, H., Ben Mustapha, S. & Salhi-Hannachi, A. Identification, evolutionary patterns and intragenic recombination of the gametophytic self incompatibility pollen gene (SFB) in Tunisian *Prunus* Species (Rosaceae). *Plant Mol. Biol. Rep.* **34**, 339–352 (2016).
48. Nunes, M. D. S., Santos, R. A. M., Ferreira, S. M., Vieira, J. & Vieira, C. P. Variability patterns and positively selected sites at the gametophytic self-incompatibility

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026

pollen SFB gene in a wild self-incompatible *Prunus spinosa* (Rosaceae) population. *New Phytol.* **172**, 577–587 (2006).
49. Zeng, B. et al. Identification of a novel SBP1-containing SCF*SFB* complex in wild dwarf almond (*Prunus tenella*). *Front. Genet.* **10**, 1019 (2019).
50. Hahn, M. W., Rausher, M. D. & Cunningham, C. W. Distinguishing between selection and population expansion in an experimental lineage of bacteriophage T7. *Genetics* **161**, 11–20 (2002).
51. Shimizu, K. K. et al. Darwinian selection on a selfing locus. *Science* **306**, 2081–2083 (2004).
52. Le Corre, V., Roux, F. & Reboud, X. DNA polymorphism at the FRIGIDA gene in Arabidopsis thaliana: extensive nonsynonymous variation is consistent with local selection for flowering time. *Mol. Biol. Evol.* **19**, 1261–1271 (2002).
53. Hughes, A. L. Near neutrality: leading edge of the neutral theory of molecular evolution. *Ann. NY. Acad. Sci.* **1133**, 162–179 (2008).
54. Ohta, T. Near-neutrality in evolution of genes and gene regulation. *PNAS* **99**, 16134–16137 (2002).
55. Nei, M. Bottlenecks, genetic polymorphism and speciation. *Genetics* **170**, 1–4 (2005).
56. Brennan, A. C., Harris, S. A. & Hiscock, S. J. The population genetics of sporophytic self-incompatibility in *Senecio squalidus* L. (Asteraceae): the number, frequency, and dominance interactions of *S* alleles across its British range. *Evolution* **60**, 213–224 (2006).
57. Pujol, B., Zhou, S. R., Sanchez-Vilas, J. & Pannell, J. R. Reduced inbreeding depression after species range expansion. *PNAS* **106**, 15379–15383 (2009).
58. Ohta, T. The nearly neutral theory of molecular evolution. *Ann. Rev. Ecol. Sys.* **23**, 263–286 (1992).
59. St Onge, K. R., Kallman, T., Slotte, T., Lascoux, M. & Palme, A. E. Contrasting demographic history and population structure in *Capsella rubella* and *Capsella grandiflora*, two closely related species with different mating systems. *Mol. Ecol.* **20**, 3306–3320 (2011).
60. Brandvain, Y., Slotte, T., Hazzouri, K. M., Wright, S. I. & Coop, G. Genomic identification of founding haplotypes reveals the history of the selfing species *Capsella rubella*. *PLoS Genet.* **9**, e1003754 (2013).
61. Chantha, S. C., Herman, A. C., Platts, A. E., Vekemans, X. & Schoen, D. J. Secondary evolution of a self-incompatibility locus in the Brassicaceae genus *Leavenworthia*. *PLos Biol.* **11**, e1001560 (2013).
62. Abdallah, D. et al. Analysis of self-incompatibility and genetic diversity in diploid and hexaploid plum genotypes. *Front. Plant Sci.* **10**, 3389 (2019).
63. Guerra, M. E., Rodrigo, J., Lopez-Corrales, M. & Wünsch, A. *S*-RNase genotyping and incompatibility group assignment by PCR and pollination experiments in Japanese plum. *Plant Breed.* **128**, 304–311 (2009).
64. Guerra, M. E. & Rodrigo, J. Japanese plum pollination: a review. *Sci. Hortic.* **197**, 674–686 (2015).
65. Jefferies, C. J. & Belcher, A. R. A fluorescent brightener used for pollen tube identification in vivo. *Stain Technol.* **49**, 199–202 (1974).
66. Linskens, H. F. & Esser, K. Uber einespezifischeanfarbung der pollen schlaucheimgriffelunddiezahl der kallosepfropfennachselbstung und fremdung. *Naturwissenschaften* **44**, 16 (1957).
67. Doyle, J. J. & Doyle, J. L. Isolation of DNA from fresh plant tissue. *Focus* **12**, 13–15 (1987).
68. Excoffier, L., Laval, G. & Schneider, S. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinform.* **1**, 47–50 (2005).
69. Li, X. W. et al. Peach genetic resources: diversity, population structure and linkage disequilibrium. *BMC Genet.* **14**, 84 (2013).
70. Sasha, B. et al. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res.* **10**, 1093 (2016).
71. Zhang, S. L., Huang, S. X., Kitashiba, H. & Nishio, T. Identification of S-haplotype-specific F-box gene in Japanese plum (*Prunus salicina* Lindl). *Sex. Plant Reprod.* **20**, 1–8 (2007).
72. Xia, X. & Xie, Z. DAMBE: software package for data analysis in molecular biology and evolution. *J. Hered.* **92**, 371–373 (2001).
73. Thompson, J., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. The ClustalX window interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids. Res.* **25**, 4876–4882 (1997).
74. Hammer, L. B., Kossek, E. E., Yragui, N. L., Bodner, T. E. & Hanson, G. C. Development and validation of a multidimensional measure of family support supervisor behaviors (FSSB). *J. Manag.* **35**, 837–856 (2009).
75. Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**, 195–202 (1999).
76. Buchan, D. W. A. & Jones, D. T. The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Res.* **47**, 402–407 (2019).
77. Rozas, J., Sánchez-Del Barrio, J. C., Messeguer, X. & Rozas, R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496–2497 (2003).
78. Nei, M. & Gojobori, T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**, 418–426 (1986).
79. Jukes, T. H. & Cantor, C. R. Evolution of protein molecules. in *Mammalian Protein Metabolism* (ed. MUNRO, H.) 21–32 (Academic Press, New York, 1969).
80. Sneath, P. H. A. & Sokal, R. R. Numerical taxonomy. (Freeman, San Francisco, 1973).
81. Zuckerkandl, E. & Pauling, L. Evolutionary divergence and convergence in proteins. In *Evolving Genes and Proteins* (eds Bryson, V. & Vogel, H.J.) pp. 97–166 (Academic Press, New York, 1965).
82. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
83. Bandlet, H.-J., Forster, P. & Röhl, A. Median joining networks for inferring intraspecific phylogenies. *Mol. Bio. Evol.* **16**, 37–48 (1999).
84. Morral, N. et al. The origin of the major cystic fibrosis mutation (delta F508) in European populations. *Nat. Genet.* **7**, 169–175 (1994).
85. Dirlewanger, E. et al. Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). *Theor. Appl. Genet.* **105**, 127–138 (2002).
86. Sosinski, B. et al. Characterization of microsatellite markers in peach [*Prunus persica* (L.) Batsch]. *Theor. Appl. Genet.* **101**, 421–428 (2000).
87. Aranzana, M. J., Garcia-Mas, J., Carbo, J. & Arús, P. Development and variability analysis of microsatellite markers in peach. *Plant Breed.* **12**, 184–184 (2002).
88. Testolin, R. et al. Microsatellite DNA in peach (*Prunus persica* L. Batsch) and its use in fingerprinting and testing the genetic origin of cultivars. *Genome* **43**, 512–520 (2000).
89. Sonneveld, T., Tobutt, K. R., Robbins, T. P. & Allele-specific, P. C. R. detection of sweet cherry self incompatibility (S) alleles S1 to S16 using consensus and allele-specific primers. *Theor. Appl. Genet.* **107**, 1059–1070 (2003).

Downloaded from academic.oup.com/hr/article/doi/10.1038/s41438-020-00392-z/6445679 by guest on 03 September 2026