Muñoz-Sanz et al. BMC Plant Biology (2017) 17:82
DOI 10.1186/s12870-017-1027-1

BMC Plant Biology

RESEARCH ARTICLE Open Access

Self-(in)compatibility in apricot germplasm is controlled by two major loci, *S* and *M*

Juan Vicente Muñoz-Sanz[1,2], Elena Zuriaga[1], Inmaculada López[1], María L. Badenes[1] and Carlos Romero[3*]

Abstract

Background: Apricot (*Prunus armeniaca* L.) exhibits a gametophytic self-incompatibility (GSI) system and it is mostly considered as a self-incompatible species though numerous self-compatible exceptions occur. These are mainly linked to the mutated S_C-haplotype carrying an insertion in the *S*-locus *F-box* gene that leads to a truncated protein. However, two *S*-locus unlinked pollen-part mutations (PPMs) termed *m* and *m'* have also been reported to confer self-compatibility (SC) in the apricot cultivars 'Canino' and 'Katy', respectively. This work was aimed to explore whether other additional mutations might explain SC in apricot as well.

Results: A set of 67 cultivars/accessions with different geographic origins were analyzed by PCR-screening of the *S*- and *M*-loci genotypes, contrasting results with the available phenotype data. Up to 20 *S*-alleles, including 3 new ones, were detected and sequence analysis revealed interesting synonymies and homonymies in particular with *S*-alleles found in Chinese cultivars. Haplotype analysis performed by genotyping and determining linkage-phases of 7 SSR markers, showed that the *m* and *m'* PPMs are linked to the same m_0-haplotype. Results indicate that m_0-haplotype is tightly associated with SC in apricot germplasm being quite frequent in Europe and North-America. However, its prevalence is lower than that for S_C in terms of frequency and geographic distribution. Structures of 34 additional *M*-haplotypes were inferred and analyzed to depict phylogenetic relationships and M_{1-2} was found to be the closest haplotype to m_0. Genotyping results showed that four cultivars classified as self-compatible do not have neither the S_C- nor the m_0-haplotype.

Conclusions: According to apricot germplasm *S*-genotyping, a loss of genetic diversity affecting the *S*-locus has been produced probably due to crop dissemination. Genotyping and phenotyping data support that self-(in)compatibility in apricot relies mainly on the *S*- but also on the *M*-locus. Regarding this latter, we have shown that the m_0-haplotype associated with SC is shared by 'Canino', 'Katy' and many other cultivars. Its origin is still unknown but phylogenetic analysis supports that m_0 arose later in time than S_C from a widely distributed *M*-haplotype. Lastly, other mutants putatively carrying new mutations conferring SC have also been identified deserving future research.

Keywords: Apricot, *Prunus*, Self-(in)compatibility, *S*-locus, *S*-alleles, Modifiers, *M*-locus

Background

Gametophytic self-incompatibility (GSI) is a system widely distributed in the plant kingdom [1] that prevents self-fertilization favoring outcrossing [2]. GSI specific recognition is under the control of a multi-allelic locus, termed *S*-locus, containing at least two linked genes: a pistil expressed S-RNase [3] and the pollen expressed *S*-locus *F-box* [4–6]. *S*-locus F-box proteins are thought to be components of E3 ubiquitin ligase complexes that recognize non-self *S*-RNases promoting their ubiquitination and degradation by the 26S proteasome proteolytic pathway [7, 8]. Recently, the collaborative model proposed in Solanaceae (*Petunia*) suggests that several F-box proteins are necessary to recruit *S*-RNases for degradation [9, 10]. This system seems to be extended to other plant families exhibiting GSI such as Rosaceae and particularly the Maloideae subfamily [10, 11]. However, *Prunus* does not seem to follow this model since knock-out of the *S*-locus F-box (*SFB*) gene leads to self-compatibility (SC) in contrast with the observations in Solanaceae. Reasons behind this behavior have been speculated for a long time [12, 13] but only recently evidence have been provided

* Correspondence: cromero@ibmcp.upv.es
[3]Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universidad Politécnica de Valencia-Consejo Superior de Investigaciones Científicas, Camino de Vera, 46022 Valencia, Spain
Full list of author information is available at the end of the article

supporting *S*-locus F-box like-2 protein as a 'general inhibitor' that detoxifies S-RNases unspecifically, unless affected by SFB [14].

In general, self-incompatibility (SI) trait predominates in stone fruits (*Prunus* genus) in accordance with the high degree of heterozygosity showed by many of these species. However, different 'degrees' of SC have been detected in this genus ranging from sweet cherry, almost strictly self-incompatible (with a few exceptions), or apricot, mainly self-incompatible (SC is restricted to European cultivars), to complete SC in peach [15]. SC is mostly related to mutations in the *S*-locus genes and mutations affecting both *S-RNases* and *SFBs* have been detected in many *Prunus* species [13, 16]. In apricot (*Prunus armeniaca* L.) the S_C-allele known to confer SC has been well characterized showing that a 358 bp insertion in the *SFB* gene leads to a putative truncated protein lacking the two essential 3′-hypervariable domains HVa and HVb [17]. Furthermore, the origin and dissemination of S_C has also been studied, identifying the non-mutated ancestor S_8-allele and detecting its presence in accessions from different geographic areas [18]. In fact, most *S*-genotyped self-compatible apricot cultivars have been shown to carry the S_C-allele [18–22].

Along with the *S*-locus specific products, other *S*-locus unlinked factors are also necessary for the GSI system to work. These factors known as 'modifiers' were firstly identified in Solanaceae [23]. Nevertheless, genetic evidence supporting modifiers have also been accumulated in other species including *Prunus spp*. [17, 24]. In apricot, pollen-part mutations (PPMs) conferring SC by putatively affecting modifiers have been identified in the Spanish local cultivar 'Canino' and the North-American one 'Katy'. Both PPMs were mapped at the distal end of chr.3 within the so-called *M*- and *M*'-loci, respectively [25, 26]. Given their distinct geographic origins and the significant genetic distance estimated between both cultivars, it was initially hypothesized that these two PPMs could have arisen independently but affecting the same gene [26].

In this context, we thought that additional *S*-locus unlinked mutations (and not just restricted to the *M*-locus) might also be present in apricot germplasm. Thus, in this study we have genotyped both the *S*- and *M*-loci in a wide set of apricot cultivars and accessions from different geographic origins, by using *SFB* and *S-RNase* introns and *M*-locus linked SSR markers. These data were combined with those of phenotyping to dissect SC causes and distribution in apricot germplasm.

Results

Self-incompatibility versus self-compatibility in apricot

A set of 67 apricot cultivars and accessions (hereinafter referred only as accessions) representing a wide range of geographic origins has been analyzed. Information about pedigree was only available for a few of them but it was quite useful to support genotyping data. Blooming time and self-(in)compatibility phenotypes were scored according to literature reports. Thus, apricot accessions were classified into 5 gross classes (early, mid-early, mid, mid-late and late) regarding the first phenotype and as self-incompatible or self-compatible regarding the second (Table 1).

Self-(in)compatibility phenotype data were completed with our own results when adult trees were available. In this regard, self-pollination tests were used to determine (8) or to confirm (13) self-(in)compatibility phenotypes but also to obtain progenies useful to search for *S*-locus unlinked mutations. Data suggest that 5 accessions are self-incompatible ('Aurora', 'Cow-2', 'Perla', 'Veecot' and 'Velázquez') while the remaining 16 show variable fruit-setting ranging from 0.5% ('Búlida') to 55% ('Ninfa'), being recorded as self-compatibles (Table 2). Some progeny could be finally obtained from all accessions of this latter group except from 'Búlida'. Those accessions with a progeny large enough were subsequently analyzed for non-*S*-locus mutations by *S*-genotyping embryos. The self-compatible controls ('Canino' and 'Katy'), already studied, and the S_C homozygote cultivar 'Galta Roja' were excluded. In summary, 40 out of the 67 accessions analyzed were classified as self-compatible, 16 as self-incompatible, 2 as male-sterile and the 9 remaining as undetermined (Table 3).

S-genotyping: identification of new *S*-alleles

All the 67 accessions were *S*-genotyped using primer pairs amplifying different fragments of the *S*-haplotype region: the first and second *S-RNase* introns as well as the 5′-UTR *SFB* intron (Table 3). *S*-alleles were determined by comparing intron size patterns with previously *S*-genotyped accessions and, when needed, sequence analyses supported these assignments. Fragment analyses of the PCR products obtained with 5 primer combinations allowed us to detect up to 20 different *S*-alleles. Seventeen out of them had already been reported in apricot (S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_9, S_{11}, S_{15}, S_{17}, S_{18}, S_{19}, S_{20}, S_{24} and S_C) and in the present work were designed basically according to the nomenclature established by Vilanova et al. [19], Halázs et al. [27] and Halázs [28], except for S_{20}, corresponding to the one described by Zhang et al. [29] and S_{24} reported by Wu et al. [30] and Gu et al. [31]. Another *S*-allele (present in 'Cow-1' and 'Cow-2') was found to have a previously non-described intron size pattern being named as S_{22}, according to its homology with the NCBI GenBank accession HM053569, corresponding to the S_{22}-*RNase* from the Chinese apricot cultivar 'Tuxiangbai'. Lastly, two additional *S*-alleles, preliminary suggested as S_V and

Table 1 Country of origin, pedigree, self-(in)compatibility and blooming time phenotypes of the analyzed apricot accessions

	Cultivar	Source[b]	Geog. area[c]	Country of origin[d]	Pedigree[e]	SI/SC[e]	Blooming time[e]
1	Alba	CAPA	WE	Italy	Unknown	?	Mid[f]
2	ASP[a]	FM	WE	Spain (V)	Unknown (isolated tree) [25]	?	Mid[f]
3	Aurora	CAPA	NA	USA	RR17–62 × NJA-13 [26]	SI [73]	Mid [74]
4	Bebecou	MAGRAMA	SE/NAf	Greece	Unknown [41]	SC [73]	Mid-early [41]
5	Bergeron	MAGRAMA	WE	France	Unknown [41]	SC [73]	Mid-late [41, 47]
6	Budapest	St. Istvan U	EE	Hungary	'Nancy' × ('Acme', 'Hungarian Best', 'Kései Rózsa') [16]	SC [16]	Mid-late [75]
7	Búlida	MAGRAMA	SE/NAf	Spain (M)	Unknown [41]	SC [73]	Mid [41]
8	Canino	IVIA	WE	Spain (V)	Unknown [41]	SC [73]	Mid [41]
9	Canino 9–7	IVIA	WE	Spain (V)	Clonal selection from Canino [27]	?	Mid [27]
10	Canino 14–4	IVIA	WE	Spain (V)	Clonal selection from Canino [27]	?	Mid [27]
11	Canino 14–6	IVIA	WE	Spain (V)	Clonal selection from Canino [27]	?	Mid [27]
12	Castlebrite	CEBAS	NA	USA	o.p. (Perfection × Castleton) [76]	SC [73]	Mid-late[f]
13	Castleton	CEBAS	NA	USA	Perfection × Newcastle [29]	SC [76]	Mid [41]
14	Ceglédi óriás	St. Istvan U	EE	Hungary	Unknown (local selection) [16]	SI [32]	Mid [35]
15	Colorao	CEBAS	SE/NAf	Spain (M)	Unknown [41]	↗ sterile [73]	Mid [41]
16	Corbató	MAGRAMA	WE	Spain (V)	Unknown [25]	?	Mid-late [74]
17	Cow-1	CAPA	WE	France	INRA	?	Mid [35]
18	Cow-2	CAPA	WE	France	INRA	?	?
19	Cristalí	FM	WE	Spain (V)	Unknown [25]	?	Mid [74]
20	Currot	IVIA	WE	Spain (V)	Unknown [25]	SC [73]	Early [74]
21	Dulcinea	CAPA	WE	Italy	Unknown (Toscana variety)	SC [73]	Mid [35]
22	Effect	St. Istvan U	EE	Ukraine	Krupnolodnyi o.p. [16]	SC [16]	Late [77]
23	Ezzine	CAPA	SE/NAf	Tunisia	INRAT	?	Early[f]
24	Fergani	St. Istvan U	EE	Former USSR	Unknown	?	?
25	Galta Roja[a]	CAPA	WE	Spain (V)	Unknown [25]	SC [73]	Mid-early [41]
26	GVV	FM	WE	Spain (V)	Unknown (isolated tree) [25]	?	?
27	Gandía	FM	WE	Spain (V)	Unknown [25]	?	Mid-early [74]
28	Gavatxet	FM	WE	Spain (V)	Unknown [25]	?	Mid [74]
29	Ginesta	IVIA	WE	Spain (V)	Unknown [25]	SC[e]	Mid-early [74]
30	Goldrich	IVIA	NA	USA	Sunglo × Perfection [76]	SI [73]	Mid [41, 47]
31	Gönci Magyar	St. Istvan U	EE	Hungary	Clone/hybrid? of Hungarian Best [16]	SC [16]	Mid-late [75]
32	Harcot[a]	IVIA	NA	Canada	[(Geneva × Naramata) × Morden 604] × NJA1(Phelps × Perfection) [76]	SI [73]	Mid [41]
33	Hargrand	St. Istvan U	NA	Canada	V51092 [(Reliable × o.p.) × o.p.] × NJA1 (Phelps × Perfection) [76]	SI [73]	Late[f]
34	Harlayne	IVIA	NA	Canada	V51092 [(Reliable × o.p.) × o.p.] × Sunglo [76]	SC [73]	Mid-late [41, 77]
35	Henderson	IVIA	NA	USA	Unknown [76]	SC?	Mid-late [41, 77]
36	Katy	IVIA	NA	USA	Zaiger's genetics (USA) [35]	SC [35]	Early[f]
37	Kech-pshar	St. Istvan U	EE	Uzbekistan	Unknown (local selection) [16]	?	?
38	K. Pozdnii[a]	St. Istvan U	EE	Ukraine	Unknown (chance seedling) [16]	SC [16]	Mid-late [75]
39	Lambertin-1	CEBAS	NA	USA	[Perfection × (Royal x Blush)] o.p. × (Perfection × Royal) [35]	SI [73]	Mid [35]
40	Lito	IVIA	SE/NAf	Greece	SEO × Tyrinthos [32]	SC [32]	Mid-late[f]
41	Manrí	FM	WE	Spain (V)	Unknown [25]	?	Mid-early [74]

Table 1 Country of origin, pedigree, self-(in)compatibility and blooming time phenotypes of the analyzed apricot accessions *(Continued)*

42	Mari de Cenad	St. Istvan U	EE	Romania	Unknown [16]	SC [16]	?
43	Mariem	CAPA	SE/NAf	Tunisia	2nd generation (Bergeron × Ouardi) × (Carraut × Crossa-Raynaud) [35]	?	Mid [35]
44	Martinet	FM	WE	Spain (V)	Unknown [25]	?	Mid [74]
45	Mitger	IVIA	WE	Spain (V)	Unknown [25]	SC[e]	Mid [35]
46	Moniquí	CEBAS	SE/NAf	Spain (M)	Unknown [41]	SI [73]	Mid [41]
47	Ninfa	CAPA	WE	Italy	Ouardi × Tyrinthos [35]	SC [73]	Mid-early [47]
48	Orange Red	IVIA	NA	USA	Lasgerdi Mashhad × NJA2 [35]	SI [73]	Mid [41, 47]
49	Ouardi[a]	IVIA	SE/NAf	Tunisia	Canino × Hamidi [41]	SI [73]	Mid-early [41]
50	Palabras	IVIA	WE	Spain (V)	Unknown [25]	SC[e]	Mid-early [74]
51	Palau	IVIA	WE	Spain (V)	Unknown [25]	SC[e]	Mid-early [74]
52	Patterson	CEBAS	NA	USA	F_2 seedling Perfection × unknown [76]	SC [73]	Mid-late [76]
53	Perla	CAPA	WE	Italy	ANFIC (Italy)	?	Mid-early [35]
54	Portici[a]	FM	WE	Italy	Unknown [41]	SC [73]	Mid-early [41]
55	Rojo de Carlet	IVIA	WE	Spain (V)	Unknown [25]	SC[e]	Mid [74]
56	Rózsakajszi	St. Istvan U	EE	Hungary	Local selection Nagykórös [16]	SC [16]	Mid-late [75]
57	Sayeb[a]	CEBAS	SE/NAf	Tunisia	Canino × Hamidi [41]	SC [73]	Mid-early [41]
58	SEO[a]	IVIA	NA	USA	Unknown [76]	SI [73]	Mid-late [41]
59	Shalah[a]	St. Istvan U	EE	Armenia	Unknown	SC/SI? [73]	Late [77]
60	Stella	IVIA	NA	USA	Unknown [41]	SI [73]	Late [41, 77]
61	S. Mammut[a]	St. Istvan U	EE	Hungary	Unknown (local selection)	SI [73]	Mid [75]
62	Tadeo	IVIA	WE	Spain (V)	Unknown [25]	SC [73]	Mid-late [74]
63	Tirynthos	IVIA	SE/NAf	Greece	Unknown [41]	SC [73]	Early [41]
64	Trevatt	CEBAS	NA	Australia	Unknown [35]	SC [32]	Mid-late [41]
65	Veecot	IVIA	NA	Canada	o.p. from Reliable [76]	SI [73]	Mid [41]
66	Velázquez	MAGRAMA	SE/NAf	Spain (M)	Unknown	SI [73]	Late[f]
67	Xirivello	CAPA	WE	Spain (V)	Unknown [25]	?	Mid-late [74]

[a]Synonyms and acronyms: ASP '*Albaricoquero Sin Polen*'; Galta Roja 'Galta Roja de Mitger'; GVV 'Galta Vermella Valenciana'; Ouardi 'Priana'; Portici 'Portici-6'; Sayeb 'Beliana'; SEO 'Stark Early Orange' and Shalah 'Erevani'; K. Pozdnii 'Konservnyi Pozdnii'; S. Mammut 'Szegedi Mammut'
[b]Material sources (see Methods)
[c]Main geographic areas defined in this study: Western Europe (WE); Southern Europe/North Africa (SE/NAf); Eastern Europe (EE) and North America (NA)
[d]Accessions from Spain come from two different regions: Valencia (V) and Murcia (M). The first ones were assigned to WE and the second ones to SE/NAf
[e]References for pedigree are indicated when available. Reports for SI/SC and blooming time phenotypes are also indicated marking with (?) when no consistent data were found: García et al. [41]; Halázs et al. [27]; Della Strada et al. [76]; Halázs et al. [18]; Orero et al. [35]; Brooks and Olmo [32]; Russell [73]; Syrgianidis and Mainou [74]; Burgos et al. [47]; Badenes et al. [75]; Massai [77]; Nyujtó et al. [78]; Mehlenbacher et al. [36]
[f]IVIA's own data (see footnotes in Table 3)

S_X, were detected only once. The amplification of first *S-RNase* intron with the single primer pair SR1-F/SR1-R distinguished up to 13 *S*-alleles ranging from 260 (S_4) to 427 bp (S_{17}) (see Additional file 1: Table S1). Another primer combination (Pru-T2/SR1-R) allowed to amplify and distinguish two additional *S*-alleles S_6 and S_{24}. Exceptions were S_1 and S_7, and S_C and S_8 pairs having exactly the same fragment sizes and S_{19}, S_{20} and S_X that could not be PCR-amplified with any of the four different primer combinations tested (data not shown). Sizes for the second *S-RNase* intron were approximately determined by agarose gel electrophoresis ranging from 300 (S_{24}) to 2800 bp (S_C) and allowed to define 15 *S*-alleles with a single primer pair (Pru-C2/Pru-C6R). Exceptions were S_5 and S_6, and S_8 and S_C pairs as well as S_{19}, S_{20} and S_V sharing similar fragment sizes, and S_3 that could not be amplified with any of the two primer combinations used. Variability in size of the 5′-UTR *SFB* intron containing fragments ranged from 189 (S_{17}) to 210 bp (S_1) and facilitated the identification of 9 *S*-alleles. Exceptions were S_2 and S_{11}, S_{18} and S_{22}, and S_C and S_8 pairs sharing the same intron sizes and S_4, S_5, S_9, S_{19}, S_{20}, S_{24}, S_V and S_X which fragments could not be amplified. Altogether, combined data allowed us to distinguish unambiguously 18 out of the 20 *S*-alleles detected excluding S_V and S_X (see Additional file 1:

Table 2 Number of bagged flowers, fruit-setting percentage and SI/SC deduced phenotype in self-pollination tests

Cultivar/ Accession	Bagged flowers	Fruit-Setting	% Setting	Phenotype	Progeny[d]
Alba[a]	355	37	10.4	SC	1
Aurora	350	0	0	SI	–
Bebecou	760	108	14.2	SC	96
Búlida	200	1	0.5	SC?[b]	–
Canino	412	99	24.0	SC	99
Castlebrite	300	36	12.0	SC	2
Corbató[a]	320	52	16.3	SC	44
Cow-1[a]	200	8	4.0	SC	7
Cow-2[a]	315	0	0	SI	–
Cristalí[a]	200	24	12.0	SC	13
Dulcinea	850	111	13.1	SC	104
Ezzine[a]	350	160	45.7	SC	21
Galta Roja	775	106	13.7	SC	106
Katy	731	80	10.9	SC	80
Mariem[a]	450	11	2.4	SC?[c]	11
Ninfa	400	221	55.3	SC	12
Perla[a]	370	0	0	SI	–
Portici	850	63	7.4	SC	59
Tadeo	375	14	3.7	SC	5
Veecot	450	0	0	SI	–
Velázquez	270	0	0	SI	–

[a]Cultivars and accessions of unknown SI/SC phenotype
[b]In general, fruit-setting percentages below 2–3% should not be undoubtedly associated to SC since some degree of pollen contamination cannot be fully discarded
[c]S-genotyping of the progeny suggested some degree of non-self pollination and therefore SC could not be confirmed
[d]Number of individual plants obtained from the set fruits

Table S1). S_C and S_8-alleles could only be distinguished by PCR-amplifying a fragment containing the 358 bp insertion present in SFB_C and absent in SFB_8 [17, 18]. As a result, S_8 was only detected in four Hungarian cultivars ('Ceglédi óriás', 'Effect', 'Gönci Magyar' and 'Szegedi Mammut') while S_C was found in 38 cultivars from many different origins (Table 3). Some already identified *S*-alleles (S_3, S_{20} and S_{24}) were not present in previously *S*-genotyped control cultivars (Table 3). For instance, putative S_3 detected in 'Harlayne' and 'Henderson' has an estimated size for the first *S*-RNase intron that matches that of the 'Sunglo' S_3-allele identified by Vilanova et al. [19] and nucleotide sequences from the three cultivars aligned by CLUSTALW showed 99% identity (data not shown). Moreover, according to the pedigree 'Sunglo' is the 'Harlayne' male parent [32]. S_{20} could not be amplified for the first *S*-RNase intron but the estimated size as well as the nucleotide sequence (99% identity) (see Additional file 2: Table S2) of the second one matches almost perfectly that of the S_{20}-allele reported by Zhang et al. [29]. Interestingly, S_{19} could only be weakly amplified from 'Mari de Cenad' with PruC2/PruC6R primers and the fragment size observed in this work and that previously reported by Halázs et al. [20] is similar to S_{20}, suggesting that both *S*-alleles might be the same. Unfortunately, this fragment could not be finally sequenced to confirm this point. S_{24} was only detected in 'Ezzine' but the estimated sizes for the first and second *S*-RNase introns, and the nucleotide sequence of the first one (248 bp), matches that S_{24}-allele reported by Gu et al. [31] (see Additional file 2: Table S2).

M-locus genotyping

The apricot *M*- and *M'*-loci were previously identified in cultivars 'Canino' and 'Katy' as carrying *S*-locus unlinked PPMs (*m* and *m'*, respectively) conferring SC [25, 26]. Up to 35 SSRs were found in the peach syntenic region of the *M*-locus (defined according to the 'Canino' genetic map) but just a five of them could be finally mapped due to several reasons (i.e. multiloci patterns, monomorphism, non-amplification, etc.). Two F_1 populations segregating for the PPMs obtained by crossing the self-incompatible cultivar 'Goldrich' with 'Canino' ('G × Ca') and 'Katy' ('G × K') were used for mapping [25, 26]. Three SSRs were mapped using 'G × Ca' (PGS3.71, PGS3.62 and PGS3.96) and two with 'G × K' (PGS3.22 and PGS3.23). In this work, linkage phases were determined by using JoinMap 3.0 software [33] and SSR-alleles in coupling were subsequently used to define haplotypes for the *M*-locus in 'Goldrich', 'Canino' and 'Katy'. These 5 SSR markers were also PCR-amplified from *M*- and *M'*-loci homozygote individuals derived from the self-pollination of 'Canino' [CC-77 (*MM*) and CC-67 (*mm*)] and 'Katy' [K06–17 (*m'm'*)] confirming estimated linkage phases and adding information to complete the haplotypes. Lastly, two additional SSRs identified from apricot genomic sequences located at the *M*-locus (AGS.20 and AGS.30) were also incorporated using the same procedure. SSR heterozygosity was estimated according to Nei [34] (Fig. 1; see Additional file 3: Table S3). As first finding, all SSR-alleles in coupling with the 'Canino' PPM *m* were found to be also in coupling with the 'Katy' PPM *m'*. In other words, both cultivars share the same pollen-part mutated haplotype designated hereinafter by m_0. As a whole, these results lead to define initially four different *M*-haplotypes (m_0, M_1, M_2 and M_3) according to the genotypes established for 'Canino', 'Katy' and the reference cultivar 'Goldrich' (Fig. 1).

These seven SSR markers were subsequently PCR-genotyped in the remaining accessions. Structures of up to 34 additional *M*-haplotypes were then statistically inferred (see Methods for details) while m_0, M_1, M_2 and M_3 were fully confirmed (see Additional file 4: Table S4). For the sake of simplicity and to facilitate the graphical representation of their relative frequencies, *M*-haplotypes were

Table 3 *S*- and *M*-haplotypes assigned to the apricot accessions analyzed in this study

	Cultivar	SI/SC	*S*-genot[d]	*M*-genot[e]		Cultivar	SI/SC	*S*-genot	*M*-genot
1	Alba	SC[b]	$S_1/\mathbf{S_C}$	M_{1-1}/M_3	35	Henderson	SC?	S_3/S_{17}	M_{2-0}/M_{16}
2	ASP	↗sterile[b,c]	$S_5/\mathbf{S_C}$	M_{4-0}/M_{5-1}	36	Katy[a]	SC/SC[b]	S_1/S_2	$M_3/\mathbf{m_{0-0}}$
3	Aurora[a]	SI	S_1/S_{17}	M_3/M_{11}	37	Kech-pshar[a]	?	S_{15}/S_Z	$M_?/M_?$
4	Bebecou	SC	$S_6/\mathbf{S_C}$	M_{4-0}/M_{4-2}	38	Konservnyi P. [a]	SC	$S_2/\mathbf{S_C}$	M_{4-1}/M_{8-1}
5	Bergeron[a]	SC	$S_2/\mathbf{S_C}$	M_{4-0}/M_{8-0}	39	Lambertin-1[a]	SI	S_1/S_2	M_{1-0}/M_{17}
6	Budapest[a]	SC	$S_2/\mathbf{S_C}$	M_{1-1}/M_{12}	40	Lito	SC[b]	$S_6/\mathbf{S_C}$	M_6/M_{10}
7	Búlida	SC	$S_5/\mathbf{S_C}$	M_{4-1}/M_{5-1}	41	Manrí	?	$\mathbf{S_C/S_C}$	$\mathbf{m_{0-0}/m_{0-0}}$
8	Canino[a]	SC/SC[b]	$S_2/\mathbf{S_C}$	$M_{1-0}/\mathbf{m_{0-0}}$	42	Mari de Cenad[a]	SC	$\mathbf{S_C}/S_{19}$	M_{8-0}/M_{18}
9	Canino 9–7	?	$S_2/\mathbf{S_C}$	$M_{1-0}/\mathbf{m_{0-0}}$	43	Mariem	SC?[b]	S_7/S_{20}	M_{1-0}/M_{8-2}
10	Canino 14–4	SC[b]	$S_2/\mathbf{S_C}$	$M_{1-0}/\mathbf{m_{0-0}}$	44	Martinet	?	S_2/S_2	$\mathbf{m_{0-0}/m_{0-1}}$
11	Canino 14–6	SC[b]	$S_2/\mathbf{S_C}$	$M_{1-0}/\mathbf{m_{0-0}}$	45	Mitger	SC[b]	$\mathbf{S_C/S_C}$	M_{5-0}/M_{5-0}
12	Castlebrite	SC	S_2/S_2	$M_3/\mathbf{m_{0-0}}$	46	Moniquí[a]	SI	S_2/S_6	M_{4-1}/M_{14-0}
13	Castleton	SC	S_1/S_2	$M_3/\mathbf{m_{0-0}}$	47	Ninfa	SC	$S_7/\mathbf{S_C}$	M_{7-0}/M_{10}
14	Ceglédi óriás[a]	SI	S_8/S_9	M_{12}/M_{13}	48	Orange Red	SI	S_6/S_{17}	M_{2-0}/M_{2-1}
15	Colorao[a]	↗sterile	$S_5/\mathbf{S_C}$	M_{4-0}/M_{4-0}	49	Ouardi[a]	SI	S_2/S_7	M_{1-0}/M_{7-0}
16	Corbató	SC[b]	S_2/S_5	$M_{5-0}/\mathbf{m_{0-0}}$	50	Palabras	SC/SC[b]	$\mathbf{S_C/S_C}$	$\mathbf{m_{0-0}/m_{0-0}}$
17	Cow-1	SC[b]	S_1/S_{31}	$M_3/\mathbf{m_{0-0}}$	51	Palau[a]	SC/SC[b]	$\mathbf{S_C/S_C}$	$\mathbf{m_{0-0}/m_{0-0}}$
18	Cow-2	SI[b]	S_{20}/S_{31}	$M_{1-0}/\mathbf{m_{0-0}}$	52	Patterson	SC	$S_1/\mathbf{S_C}$	M_{1-3}/M_3
19	Cristalí	SC[b]	$S_{20}/\mathbf{S_C}$	$M_{5-0}/\mathbf{m_{0-0}}$	53	Perla	SI[b]	S_2/S_{20}	M_{1-2}/M_{15-1}
20	Currot[a]	SC/SC[b]	$\mathbf{S_C/S_C}$	$\mathbf{m_{0-0}/m_{0-0}}$	54	Portici	SC/SC[b]	S_2/S_{20}	$M_{1-4}/\mathbf{m_{0-0}}$
21	Dulcinea	SC	$S_2/\mathbf{S_C}$	M_{7-3}/M_{14-1}	55	Rojo Carlet	SC[b]	$\mathbf{S_C/S_C}$	M_{4-0}/M_{5-0}
22	Effect[a]	SC	$S_8/\mathbf{S_C}$	M_{8-0}/M_{12}	56	Rózsakajszi[a,f]	SC	$S_2/\mathbf{S_C}$	M_{1-0}/M_{12}
23	Ezzine	SC[b]	$S_{24}/\mathbf{S_C}$	M_{1-0}/M_{7-1}	57	Sayeb[a]	SC	$S_7/\mathbf{S_C}$	M_{1-0}/M_{7-2}
24	Fergani	?	S_V/S_X	$M_?/M_?$	58	Shalah[a]	SC	S_5/S_{11}	M_{8-1}/M_{19}
25	GaltaRoja	SC/SC[b]	$\mathbf{S_C/S_C}$	M_{4-0}/M_{5-2}	59	SEO	SI	S_6/S_{17}	M_{2-0}/M_6
26	GW	?	$S_2/\mathbf{S_C}$	M_{4-0}/M_{5-2}	60	Stella	SI	S_6/S_{20}	M_9/M_9
27	Gandía	?	$\mathbf{S_C/S_C}$	$\mathbf{m_{0-0}/m_{0-0}}$	61	Szegedi M.	SI	S_8/S_9	M_{12}/M_{13}
28	Gavatxet	?	$S_{20}/\mathbf{S_C}$	$\mathbf{m_{0-0}/m_{0-0}}$	62	Tadeo	SC	$S_{20}/\mathbf{S_C}$	$M_{15-0}/\mathbf{m_{0-0}}$
29	Ginesta[a]	SC/SC[b]	$\mathbf{S_C/S_C}$	$\mathbf{m_{0-0}/m_{0-0}}$	63	Tirynthos	SC	$\mathbf{S_C/S_C}$	M_{10}/M_{10}
30	Goldrich[a]	SI/SI[b]	S_1/S_2	M_{1-0}/M_{2-0}	64	Trevatt	SC	$S_2/\mathbf{S_C}$	$M_{1-0}/\mathbf{m_{0-0}}$
31	Gönci Magyar[a]	SC	$S_8/\mathbf{S_C}$	M_{8-0}/M_{12}	65	Veecot	SI/SI[b]	S_2/S_{20}	M_{2-0}/M_3
32	Harcot[a]	SI/SI[b]	S_1/S_4	M_{1-0}/M_{2-2}	66	Velázquez	SI	S_5/S_{20}	M_{4-0}/M_{4-1}
33	Hargrand[a]	SI	S_1/S_2	M_{1-0}/M_{2-0}	67	Xirivello	?	$\mathbf{S_C/S_C}$	M_{4-0}/M_{4-1}
34	Harlayne	SC	S_3/S_{20}	M_{2-0}/M_9					

[a]Cultivars previously *S*-genotyped by Halázs et al. [18]; Vilanova et al. [19]; Halázs et al. [20]; Zuriaga et al. [26]; Halázs et al. [27]; Burgos et al. [42]; Alburquerque et al. [79]

[b]Own data on SI/SC phenotype obtained in this work (see Table 2). Additionally, SC had been observed in a set of accessions grown under insect-proof screen house at IVIA ('Rojo de Carlet', 'Mitger', 'Palabras', 'Palau', 'Currot', 'Ginesta', 'Canino 14–4' and 'Canino 14–6') showing moderate fruit-setting (not quantified) across several years

[c]Male-sterility in the ASP accession was indicated by shrunken pale anthers

[d]*S*-allele nomenclature is proposed according to Vilanova et al. [19]; Halázs et al. [27]; Zhang et al. [29]; Wu et al. [30]; Gu et al. [31] and Halázs et al. [28]. *S*-haplotype associated with SC (S_C) is written in bold

[e]*M*-haplotypes were named with two digits. The first one corresponds to the *M*-haplotype 'main class' and the second to the subtype. *M*-haplotype variants associated with SC (m_{0-0} and m_{0-1}) are written in bold. Haplotypes designated by $M_?$ could not be defined

[f]*S*-genotype determined for Rózsakajszi (S_2S_C) was not in agreement with that previously reported by Halázs et al. [18] (S_CS_C). Reasons for this discrepancy are still unknown

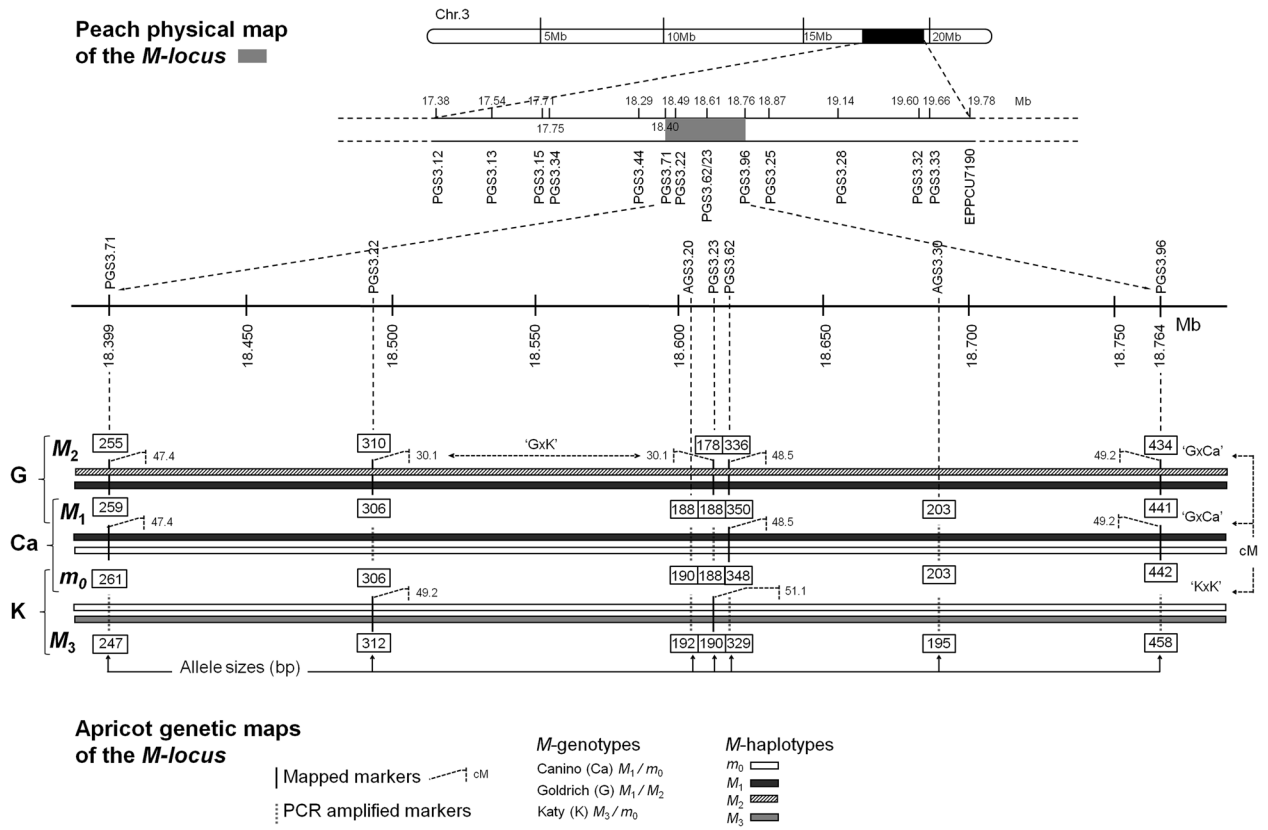

Fig. 1 Apricot *M*-locus haplotypes structure. The *peach* syntenic region at the distal end of chr. 3 (*black box*) comprising the *M*-locus (*grey box*) is zoomed twice and shows SSRs (*dashed lines*) PCR-amplified in apricot cultivars 'Goldrich' (G), 'Canino' (Ca) and 'Katy' (K). SSR positions in *peach* genome (Kb) and allele sizes (bp) determined in apricot are indicated. *White, black, diagonal striped* and *grey thick lines* represent apricot m_0, M_1, M_2 and M_3 haplotypes, respectively. SSR anchoring positions are shown in centimorgans (cM) (*boxed numbers*) according to the available mapping populations ('G × Ca', 'K × K' and 'G × K')

grouped into 'main classes' when they differ in no more than three SSR alleles (resulting in a total of 20 *M*-haplotype 'main classes': m_0 and M_1 to M_{19}). Thus, *M*-haplotypes were designed by two sub-indexes, the first one corresponds to the 'main class' and the second defines the subtype (see Additional file 4: Table S4). Structures of the *M*-haplotypes corresponding to the cultivars 'Kech-pshar' and 'Fergani' could not be inferred but according to the SSR allele sizes all four are distinct (data not shown). Phylogenetic reconstructions from Jaccard's and Bruvo's distances show that some 'main classes' clustered close together suggesting common ancestry such as, for instance, M_7 and M_8; M_2, M_4 and M_{16} or m_0, M_1 and M_{13} (Fig. 3). In this latter group, the nearest haplotypes to the majority haplotype m_{0-0} were M_{1-2}, M_{1-1} and M_{1-0} showing Bruvo's distance values of 0.18, 0.22 and 0.26, respectively.

Mutations conferring self-compatibility in apricot

The S_C-haplotype known to confer SC [17] was found in 38 accessions (being homozygous in 7) (Table 3). Thirty out of the 38 were self-compatible, 6 have undetermined phenotype and two were male-sterile. The m_0-haplotype associated with SC by Zuriaga et al. [25, 26] was detected in a total of 19 accessions (being homozygous in 8): 8 were previously shown to be self-compatible, 5 were undetermined and one was classified as self-incompatible ('Cow-2') since it did not produce any fruit after self-pollination (Tables 2 and 3). As a whole, all self-compatible accessions analyzed have at least one of these two haplotypes already known to confer SC (S_C and/or m_0) except for 'Harlayne', 'Henderson', 'Shalah' and 'Mariem'. The number of self-compatible haplotypes carried by each accession varied from 0 to 4. A total of 21 accessions do not carry neither S_C nor m_0, 26 carry only one self-compatible haplotype (corresponding to *S*- and *M*-genotypes S_C/x or m_0/x), 13 carry up to two (S_C/S_C, m_0/m_0 or S_C/x-m_0/x), 1 carry three (i.e. Gavatxet S_{20}/S_C-m_0/m_0) and 6 carry four (S_C/S_C-m_0/m_0) (Table 3).

Segregation of the *S*-genotypes in self-progenies obtained from two self-compatible cultivars carrying a single copy of the m_0-haplotype ('Portici' and 'Corbató') supported a mutation outside the *S*-locus as the cause for the phenotype (Table 4). Progenies from self-compatible cultivars ('Dulcinea' and 'Bebecou') not

carrying the m_0-haplotype were used as controls to confirm this finding (Table 4). Similar results were observed in other accessions (with fewer offspring) carrying ('Cow-1' and 'Cristalí') and not carrying ('Ezzine' and 'Ninfa') the m_0-haplotype (data not shown). Moreover, distortion ratios detected in the segregation of SSR markers tightly linked to the *M*-locus (AGS.20 and PGS3.23) point out that the mutation is located at the *M*-locus (Table 4).

Interestingly, *S*-genotypes segregation found in the 'Portici' progeny might also indicate the presence of another mutation affecting the S_2-haplotype, since the number of S_2S_2 genotypes was unexpectedly high (Table 4). Analysis of genomic DNA fragments containing the complete sequence of S_2-*RNase* and SFB_2 alleles from 'Portici' revealed only one mismatch (A/G) within SFB_2 located at position 1.296. This change leads to a non-synonymous substitution (lysine by arginine) in the hypervariable region HVb.

Lastly, *S*- and *M*-genotyping of 'Canino' selected clonal sibs (14–4, 14–6 and 9–7) [35] did not reveal differences with 'Canino' (Table 3), pointing out that S_C and *m* mutations are shared by all these four accessions.

Geographical distribution of self-(in)compatibility

For the sake of the analysis, cultivars and accessions studied in this work were grouped as belonging to four big geographic areas according to the country of origin, pedigree (Table 1) and data about dissemination [36–38]: North America (NA), Western Europe (WE), Eastern Europe (EE) and Southern Europe/North Africa (SE/NAf). Nonetheless, this classification is obviously arbitrary and exhibits some inconsistencies. This is particularly true for the WE and SE/NAf groups considered separately but clearly interrelated. Attending to geographic criteria, cultivars from Italy, France and Spain (Valencia region) were grouped into the first one while those from Tunisia, Greece and Spain (Murcia region) into the second. Spanish cultivars were divided into two subgroups that might reflect some differences according to [49]. Parentage relations among some cultivars from WE and SE/NAf groups (i.e. 'Canino' with 'Ouardi' and 'Sayeb') underscore the inaccuracy of this classification but, assuming these limitations, it produces more balanced groups without affecting data interpretation. Thus, SC phenotype is distributed across all four groups but frequencies varied from 2/19 (SI/SC) in WE to 9/16 in NA (Fig. 2). Similarly, blooming time types from early/mid-early to mid-late/late are present in all four groups, but frequency ranged from 9/24 in WE to 1/16 in NA (Table 1). In general it could be said that most early/mid-early classified accessions are self-compatible (13/15) but this phenotype is also generally found in many mid-late/late classified accessions (12/17). Therefore, even considering the reduced number of scored accessions and the limited phenotype data, it seems that there is no correlation between SC and early blooming in cultivated apricots.

Frequencies of the *S*-alleles also varied between the different regions analyzed. For instance, S_2 and S_C are present in all four groups but, on the contrary, S_8 and S_9 are only present in EE, S_3 in NA and S_7 in SE/NAf (Fig. 2). Regarding *S*-alleles distribution it is also important to take into account synonymies and homonymies. In the first case, BLASTN has revealed that some of the *S*-alleles initially found and numbered from European accessions have also been detected in Chinese cultivars but named differently. For instance, a 659 bp fragment containing the second intron of the S_{17}-*RNase* present in the North-American cultivar 'Aurora' (as well as in 'SEO', 'Orange Red' and 'Henderson') shows >99% identity with those from S_{44} and S_9-*RNases* identified from the Chinese cultivars 'Shailaiyulvke' and 'Xinshiji', respectively, suggesting these three *S*-alleles to be the same. Similar examples are reported for S_1, S_4, S_9, S_{12}, S_{16} and S_{20}-*RNases*. The other way around, BLASTN also points out homonymies. For example, S_9 derived from the Hungarian cultivar 'Ceglédi óriás' is different to that S_9 reported from the Chinese cultivar 'Xinshiji' (see Additional file 2: Table S2). BLASTN has also detected significant similarities between S_{13} reported in accessions from Armenia (including 'Shalah'), Turkey, Tunisia and Morocco [20–22]

Table 4 Segregation of *S-RNase* alleles in controlled self-pollinations

Cultivar *S*-genotype	Progeny *S*-genotypes			Total[a]	Ratio 1:1[b]	Ratio 1:2:1[b]	Ratio 1:1[c]
	A	H	B				
Bebecou (S_6/S_C)	0 (S_6/S_6)	33 (S_6/S_C)	45 (S_C/S_C)	96 (78)	1.85 (0.17)	53.8 (0)	n.d.[d]
Dulcinea (S_C/S_2)	36 (S_C/S_C)	37 (S_C/S_2)	0 (S_2/S_2)	104 (73)	0.34 (0.56)	37.9 (1e-8)	n.d.
Corbató (S_5/S_2)	10 (S_2/S_2)	17 (S_5/S_2)	9 (S_5/S_5)	44 (36)	—	0.17 (0.92)	0.17 (0.68) [24]
Portici (S_2/S_{20})	16 (S_2/S_2)	24 (S_2/S_{20})	4 (S_{20}/S_{20})	59 (44)	—	6.91 (0.03)	0.90 (0.35) [79]

[a]Number of individuals from the self-progeny and total analyzed (between parentheses) are indicated
[b]χ^2 and (*P*) values for *S*-genotypes expected ratios considering a single mutation unlinked (1:2:1) or linked to the *S*-locus (1:1)
[c]χ^2 and (*P*) values for PGS3.23 (in 'Corbató' self-progeny) and AGS.20 (in 'Portici' self-progeny) genotypes expected ratios considering a single mutation located at the *M*-locus (1:1). No homozygote for the SSR-allele in repulsion-phase linkage with the m_0-haplotype was detected in either case. [*n*] indicates number of individuals tested
[d]n.d. Not determined

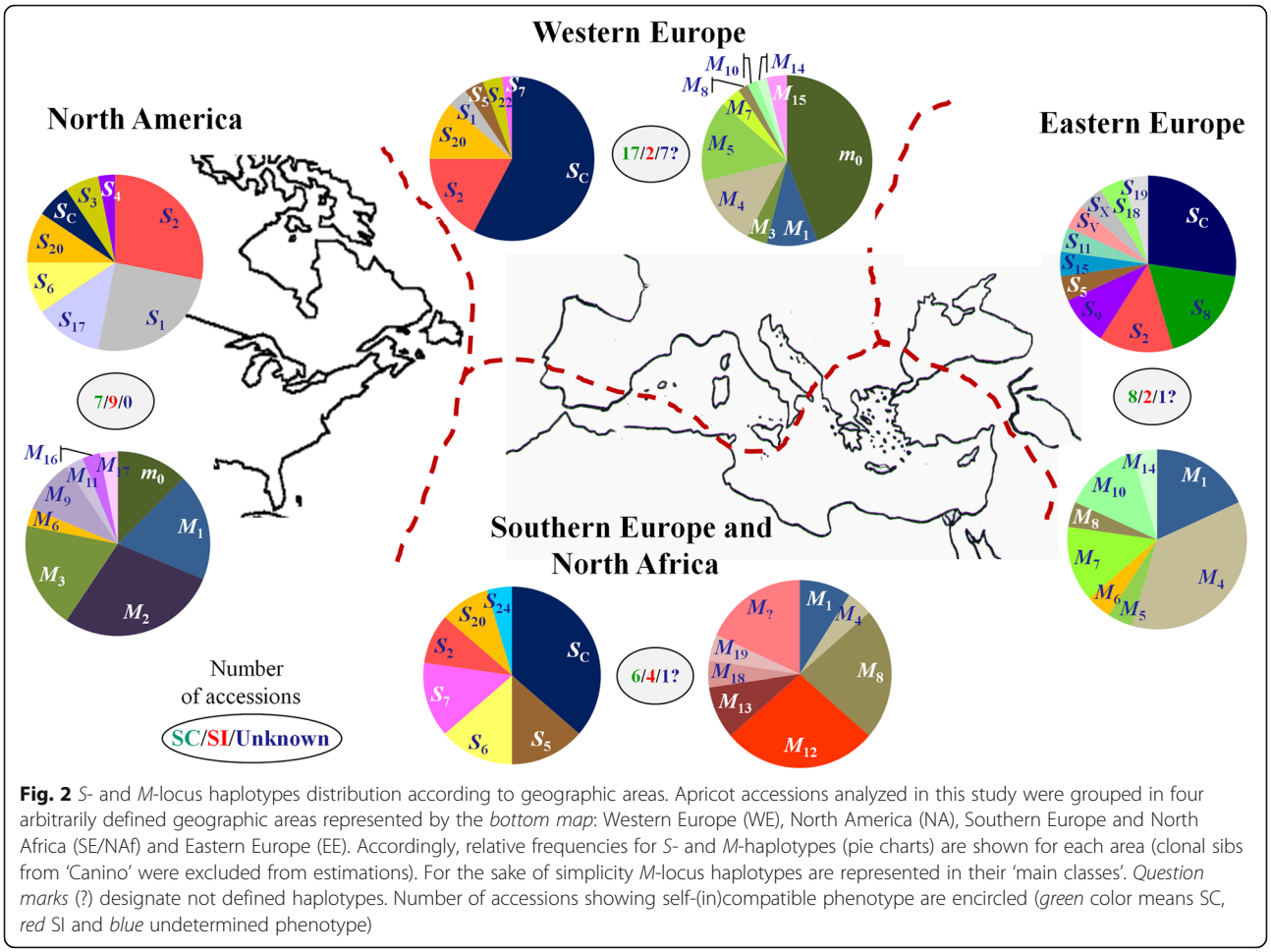

Fig. 2 *S*- and *M*-locus haplotypes distribution according to geographic areas. Apricot accessions analyzed in this study were grouped in four arbitrarily defined geographic areas represented by the *bottom map*: Western Europe (WE), North America (NA), Southern Europe and North Africa (SE/NAf) and Eastern Europe (EE). Accordingly, relative frequencies for *S*- and *M*-haplotypes (pie charts) are shown for each area (clonal sibs from 'Canino' were excluded from estimations). For the sake of simplicity *M*-locus haplotypes are represented in their 'main classes'. *Question marks* (?) designate not defined haplotypes. Number of accessions showing self-(in)compatible phenotype are encircled (*green* color means SC, *red* SI and *blue* undetermined phenotype)

and S_5, previously found in Southern Spanish cultivars [19], and among S_{12}, detected in Turkey and Tunisia [20, 22], S_{24} and S_4 (see Additional file 2: Table S2).

Distribution of the 38 inferred *M*-haplotypes is not uniform either. The M_1 'main class' is present in all four groups, m_0 is only found in WE and NA and others, such as M_{12} and M_2, were exclusive for EE and NA groups, respectively (Figs. 2 and 3).

Discussion

S-alleles and the apricot diffusion from Armenian and Chinese centers to Europe

Most apricot *S*-alleles identified in this work had been previously reported except for S_{22}, S_V and S_X [19, 27–30]. Geographical distribution of these *S*-alleles is heterogeneous including a few widely disseminated and others restricted to certain areas. These results are in agreement with previous works. For instance, S_2 and S_C had already been found in accessions from all countries analyzed in this work but also in Turkey [20], Morocco [21], Afghanistan [18], Iran [40], etc. On the contrary, S_1 is only present in NA and in a few accessions from WE. Meanwhile, S_{10} to S_{14}-alleles, suggested to be from Armenian origin by Halázs et al. [20], have been detected in Turkey, Tunisia, Morocco and in a few materials from EE but they were thought to be absent in WE [20–22, 41]. However, S_5 suggested to be the same that S_{13} connects (along with S_6) Armenian, Eastern-Turkish and Moroccan accessions [20–22] with Southern-Spanish accessions [19, 42] supporting the apricot Southwest-Mediterranean diffusion route, from the Irano-Caucasian gene pool, proposed by Bourguiba et al. [38].

The heterogeneous distribution of the apricot *S*-alleles is a consequence of the crop domestication and diffusion and this process has also produced a genetic bottleneck detected in neutral SSR markers [38]. Though not quantified this loss of genetic diversity can also be grossly appreciated at the *S*-locus. For instance, China is not only considered a center of origin but also a probable domestication center for the species [38] and at least 20 different *S*-alleles have been found in just 30 Chinese cultivars [29, 30]. This variability is considerably much higher than that detected in European accessions and, interestingly, only 5 of these *S*-alleles (S_9, S_{14}, S_{17}, S_{20} and S_{24}) seem to be represented in apricot cultivars outside China. In addition, other *S*-alleles detected in far-east

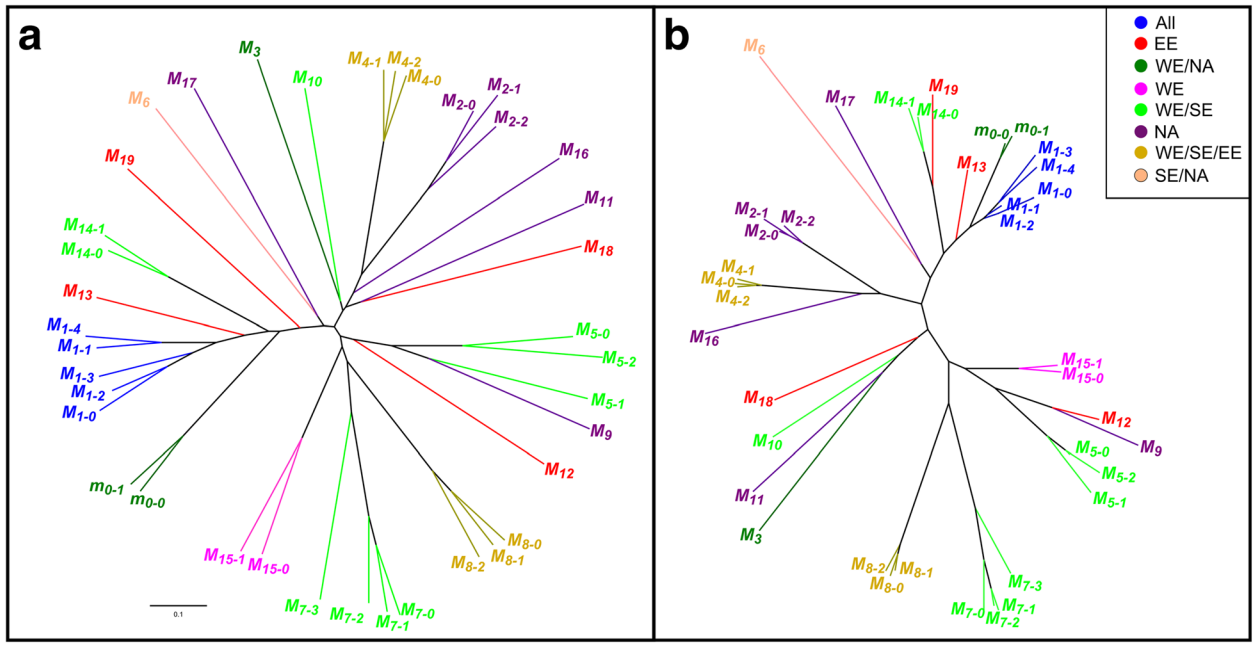

Fig. 3 Clustering analysis of apricot *M*-locus haplotypes based on genetic distances. **a** Clustering obtained by Neighbor-Joining algorithm using Jaccard's distance. **b** Clustering obtained by Neighbor-Joining algorithm using Bruvo's distance. Colors represent geographic areas where the distinct *M*-locus haplotypes were detected (see legend)

region cultivars such as S_{15}/S_{18} in 'Kech-pshar' (Uzbekistan), S_{10} in 'Harmat' (Armenia) [27] and S_V/S_X in 'Fergani' (former USSR) are not found in European accessions. *S*-alleles can therefore be a useful tool to study apricot dissemination but the establishment of a unified nomenclature would be advisable to favor this goal.

Pollen-part mutated *m*-haplotype origin and dissemination

The two *S*-locus unlinked PPMs (*m* and *m'*) conferring SC in apricot cultivars 'Canino' and 'Katy' were found to be located in an overlapping region at the distal end of chr. 3 [26]. This finding suggested that both PPMs might be the same. However, in addition to their different geographic origins, 'Canino' and 'Katy' exhibited a high value for the Nei's genetic distance (0.83) estimated on the basis of 85 SSR markers distributed across the whole genome [26]. Thus, initially it was hypothesized that these two PPMs originated independently but might affect the same gene. Nevertheless, a more detailed analysis of a set of SSR markers linked to the *M*-locus has allowed us to determine that haplotypes in coupling with the PPMs have the same structure in both cultivars. Thus, pollen-part mutated *m*- and *m'*-haplotypes previously associated with SC in 'Canino' and 'Katy' cultivars [25, 26] are now considered to be the same (m_0). Moreover, in this work the m_0-haplotype has been detected in 17 additional accessions (excluding 'Canino' clonal sibs) mainly Spanish (12 in total) but also from USA, Australia, France and Italy. Fifteen of them were confirmed as self-compatible (exceptions were 'Cow-2' described above as well as 'Gandía', 'Gavatxet', 'Manrí' and 'Martinet' with undetermined phenotype). The m_0-haplotype was frequently accompanied by the S_C-allele in confirmed self-compatible accessions (7 cases), which might suggest that mutations conferring SC tend to accumulate once the system is broken due to relaxed selective constraints [43]. Following this reasoning and considering the higher prevalence of the S_C-allele in confirmed self-compatible accessions (74%) compared with the m_0-haplotype (23%), it might also be speculated that PPM *m* arose in a self-compatible accession carrying S_C. However, against this hypothesis is the fact that the m_0-haplotype alone was also found in 6 confirmed self-compatible accessions and, therefore, further analyses are necessary to clarify this point. This latter group includes 'Portici' (S_2/S_{20}-M_{1-4}/m_{0-0}), previously shown to lack the S_C-haplotype but which causative mutation was still unknown [44].

Beside the m_0-haplotype, 37 additional *M*-haplotypes were identified by SSR analysis being grouped in 19 'main classes'. Microsatellite haplotype distances were analyzed by two alternative clustering methods to validate results. The first one relies on the proportion of shared alleles assuming independence and ignoring mutational processes which can bias distances, particularly when loci are highly polymorphic. This method was based on the similarity coefficient for binary data

developed by Jaccard [45]. The second takes into account the stepwise mutation model considering higher likelihood for small than for large changes in microsatellite repeat number and it is based on the Bruvo's distance [46]. Results obtained with both methods were equivalent but the second draws a more accurate phylogeny, where 'main classes' m_0, M_1 and M_{13} group close together suggesting a not too distant common ancestor. Remarkably, the closest haplotypes to m_{0-0} (M_{1-0}, M_{1-1} and M_{1-2}) are widely distributed geographically, while m_{0-0} seems to be restricted to NA and WE, and M_{13} to EE. The situation is different for the mutated S_C-allele present in all geographic areas studied [18–21] whereas the ancestor S_8-allele was only detected in Hungarian and Turkish cultivars [20]. Altogether, these results suggest that the m_0-haplotype arose later in time than S_C since its distribution is much more limited. It can also be hypothesized that the last common ancestor for m_0, M_1 and M_{13} spread throughout different regions and mutated somewhere in WE. Nevertheless, in general, it has not been observed a significant correlation between geographic distribution and clustering of the *M*-haplotypes (see Fig. 3), and this may be due to the frequent crosses between genotypes from different groups [39].

SC and genotyping: uncovering accessions carrying unanalyzed mutations

Part of the materials analyzed in this work could not be self-pollinated because adult trees were not available. However, though phenotype could not be directly assessed in these cases, according to the almost perfect association observed between SC and S_C/m_0-alleles it should be expected that most, if not all, accessions carrying whatever of these two alleles were self-compatible. However, some incongruities between phenotype and genotype need to be highlighted. For instance, in spite of having the m_{0-0} haplotype self-pollination of 'Cow-2' (S_{20}/S_{31}-M_{1-0}/m_{0-0}) did not produce any fruit. This behavior could be due to fruit-set problems, as exemplified by the self-compatible cultivar 'Búlida' [47] which setting was also nearly null. Alternately, it may also be possible that m_{0-0} in 'Cow-2' does not carry the expected PPM but further analyses are needed to check this point. On the contrary, four cultivars classified as self-compatible do not carry neither the S_C nor the m_{0-0} haplotypes: 'Mariem' (S_7/S_{20}-M_{1-0}/M_{8-2}), 'Shalah' (S_5/S_{11}-M_{8-1}/M_{19}), 'Harlayne' (S_3/S_{20}-M_{2-0}/M_9) and 'Henderson' (S_3/S_{17}-M_{2-0}/M_{16}). Again, erroneous phenotyping can lead to discrepancies. In fact, 'Mariem' progeny could not be fully confirmed as derived from self-pollination (Table 2), 'Shalah' was reported as self-compatible by Burgos et al. [47] but scored as self-incompatible by Halázs et al. [20] and information describing 'Henderson' as self-compatible just comes from nursery catalogs. So, phenotype can be questioned in all these three cultivars. However, SC in 'Harlayne' is fully confirmed since it has been tested by several independent sources [47, 48]. Regarding genotypes, all the *S*-alleles carried by these cultivars can be frequently found in self-incompatible accessions and therefore expected to be 'functional'. Something similar happens with M_{1-0} and M_{2-0} haplotypes but M_{8-1} and M_{8-2} belong to the M_8-haplotype 'main class' always detected in self-compatible accessions (a total of 7 in this work and mostly East-European). However, no progenies are available and therefore it could not be tested whether M_8 is mutated (non-functional) or not. As for M_{16} and M_{19}, they were detected only once in the set of accessions analyzed. Overall, there are still no evidence supporting the nature of the mutation/s conferring SC in these cases and neither S/M-loci unlinked mutations nor other mutations affecting *S*- and *M*-loci can be discarded.

Lastly, genetic analysis suggests the presence of a SNP mutation within SFB_2 HVb region in 'Portici' that could also be associated with SC. It could be speculated that a single non-synonymous change within a SFB hypervariable region might alter its specificity, since these domains (strongly hydrophobic and under positive selection) were already suggested to have a role in the specific recognition of S-RNases [49]. Interestingly, this change has also been found in 'Katy' SFB_2 where a similar distortion was detected in the self-progeny [26]. However, further research is needed to validate this observation and other possibilities cannot be discarded (i.e. mutations affecting pollen viability).

Forces selecting for self-compatibility in apricot

SC is therefore quite common in apricot but distribution is not uniform across geographic areas. SI is the prevalent phenotype in three out of the four major eco-geographical groups for apricot (centers of origin): Central Asian, Irano-Caucasian and Dzhungar-Zailij and also in the later proposed Chinese group [36]. In fact, studies about Chinese apricots do not report self-compatible cultivars [29, 30, 50]. On the contrary, SC predominates in the European group but some disequilibrium can also be observed. Among the materials analyzed in this work, SI is frequent in commercial North-American cultivars (two thirds) but unusual in West and East-European countries (one fifth). This might point out a non-European SI ancestral donor for North-American cultivars apricots as previously suggested for PPV resistance trait [51].

It is generally accepted that apricot genetic diversity decreases from east to south-west [36] and, in this context, it is questionable whether SC might be one of the causes. Recent works suggest that a substantial part of this loss is independent of the selection impact due to the domestication bottleneck, being more related to apricot diffusion routes [38]. In general, self-fertility would be advantageous

under unfavorable pollination conditions including early blooming [52] but this correlation has not been observed in this study. However, we cannot discard that SC was originally selected from early-flowering apricot types. In fact, in cherry (*Prunus avium*) only a very few cultivars are reported to be naturally self-compatible and these include two exceptionally early-flowering cherries, 'Cristobalina' and 'Kronio' [53, 54]. Moreover, they were selected independently since each one carry a different mutation conferring SC, 'Cristobalina' in a non-S-locus pollen modifier [24] and 'Kronio' in the SFB_5 allele [54]. On the other hand, in spite of the benefits linked to SC for growing stone fruits (increased yield, removing of interspersed pollinators, etc.) it was not usually considered as a target trait. Some morphological and agronomic traits are putatively linked to the *S*- (fruit shape) and *M*-loci (polycarpel and flower color) in *Prunus* [55] but it seems unlikely that SC was brought in by linkage drag with these particular traits. Thus, SC in apricot is more likely the indirect result of selective breeding on major traits, such as ecological adaptation [36], and only recently it has become a breeding objective by itself. In any case, SC must have had an effect on reducing diversity since S_CS_C homozygote cultivars (potentially resulting from self-pollination events) are frequently found in most European countries.

In this work we have confirmed two independent PPMs (S_C and *m*) as the main causes for SC in apricot. According to [16] a total of 29 mutations conferring SC have been identified in *Prunus*. Most of them affect *S*-locus (27) and only two correspond to mutated modifiers, the apricot mutated m_0-haplotype and a non *S*-locus PPM in sweet cherry [17, 24]. Indeed, these two are located at the distal end of chr.3 but it is still unknown whether they affect different genes [26]. Within the supposedly complex GSI mechanism it is noteworthy this low number of mutations affecting modifiers. One possible explanation is that other modifiers participating in the control of the GSI system may be redundant and when mutated do not confer SC. Furthermore, loss of function of some factors might lead to SI as may be predicted for SLFL2 [14]. Lastly, in spite of the high number of accessions already evaluated in all *Prunus* species, it can also be argued that a more exhaustive screening might reveal novel mutations. Indeed, in the light of the results of this work, it is conceivable that additional mutations conferring SC may remain 'hidden' in the large amount of unanalyzed plant material.

Conclusions

It has been shown that apricot *S*-allele geographical distribution is heterogeneous. In addition, it seems that the loss of genetic diversity, hypothesized for this crop during the dissemination process, affected not only neutral markers but also the *S*-locus. Indeed, only a few *S*-alleles present in self-incompatible cultivars from China (the domestication center of the species) can be found in European cultivars. However, S_C, one of the most diffused *S*-alleles and the main cause of SC in apricot has not been detected in China. Apart from the PPM associated to S_C, other *S*-locus unlinked PPMs conferring SC have been found in apricot cultivars 'Canino' and 'Katy'. At odds with the initial hypothesis, we have shown that both are associated with the same m_0-haplotype, in spite of the genetic distance between the two cultivars. Moreover, this haplotype has been surprisingly detected in other Spanish, French, Italian and North-American accessions. Its origin is still a matter of speculation but the phylogenetic analysis supports that m_0 arose later in time than S_C from a widely distributed *M*-haplotype. As a whole, self-(in)compatibility in apricot seems to rely mainly on the *S* and *M* loci. However, other self-compatible mutants putatively carrying different mutations have been identified in this work. These results deepen our knowledge on the self-(in)compatibility trait in stone fruits and open new questions that deserve future research.

Methods

Plant Material and self-pollination test

Sixty seven apricot cultivars and accessions from diverse geographic origins were used in this study (Table 1). Most are currently kept at the collection of the *Instituto Valenciano de Investigaciones Agrarias* (IVIA) in Valencia (Spain). Part of this collection was kindly provided by *Frutales Mediterráneo S.A.* (FM) company, by the *Consellería de Agricultura, Pesca y Alimentación* (CAPA) and by the *Ministerio de Agricultura, Alimentación y Medio Ambiente* of Spain (MAGRAMA). Other samples, already used in previous works [56], were provided by the *Departamento de Mejora y Patología Vegetal del* CEBAS-CSIC in Murcia (Spain) and by the University of St. Istvan (Budapest, Hungary).

Trees from different cultivars and accessions (Table 2) were tested for SC by self-pollination in the field. Before anthesis, insect-proof bags were put over several branches, containing approximately 200–250 flower buds in total per cultivar *ad minimum*, to prevent cross pollination. Subsequent fruit set was recorded and fruits collected about 3 months later. Seed-derived embryos were dissected from the rest of the seed tissue and stored at −20 °C.

DNA extraction

Two leaf discs were collected from each accession, frozen in liquid N_2 and stored at −80 °C before DNA isolation. Genomic DNA was extracted following the method of [57] with slight modifications: the ratio of fresh leaf tissue/CTAB buffer was reduced to 10 mg/250 μl and nucleic acids were recovered by precipitation with 1×

volume of isopropanol. DNA quantification was performed by NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE) and integrity was checked on 1% agarose gel by comparison with lambda DNA (Promega, Madison, WI, USA). Embryo DNA was extracted by incubating for 10 min at 95 °C with 20 μL of TPS (100 mM Tris-HCl, pH 9.5; 1 M KCl; 10 mM EDTA) isolation buffer [58].

S-locus genotyping

Apricot *S*-alleles were identified by PCR-amplifying fragments comprising first and second introns of the *S-RNase* as well as the 5′-UTR *SFB* intron, respectively. Genomic DNA isolated from cultivars and accessions listed in Table 1 was used as PCR template. PCRs were performed in a final volume of 20 μL containing 1 × DreamTaq buffer, 0.2 mM of each dNTP, 20 ng of genomic DNA and 1 U of DreamTaq DNA polymerase (Thermo Fisher Scientific, Waltham, MA, USA). Previously developed primers designed from conserved regions of *Prunus armeniaca S-RNase* genomic sequences, SRc-F and SRc-R [19, 58] and from *Prunus avium S-RNase*-cDNA sequences, Pru-T2 and Pru-C2R [59], were used to amplify the first intron (see Additional file 5: Table S5) in all four possible combinations. The amplification was carried out using a temperature profile with an initial denaturing of 95 °C for 2 min; 30 cycles of 95 °C for 45 s, 52 °C for 1 min and 72 °C for 1 min 30 s; and a final extension of 60 °C for 30 min (UNO96, VWR, Radnor, PA, USA). Each PCR was performed by the procedure of [60] using three primers: the specific forward primer with M13(−21) tail at its 5′ end at 0.4 mM, the reverse primer at 0.8 mM, and the universal fluorescent-labeled M13(−21) primer at 0.4 mM. Allele lengths were determined using an ABI Prism 3130 Genetic Analyzer with the aid of GeneMapper software, version 4.0 (Applied Biosystems, Foster City, CA, USA).

The second intron was amplified using two sets of primers designed from *Prunus avium S-RNase*-cDNA sequences [61]: Pru-C2/Pru-C4R and Pru-C2/Pru-C6R (see Additional file 5: Table S5). PCRs were performed using 0.25 μM of each primer and the program previously described by [62] to amplify long PCR products. PCR products were electrophoresed in 0.8% (*w*/*v*) agarose gels using 1 x TBE (89 mM Tris, 89 mM boric acid, and 2 mM EDTA (pH 8.0)) buffer, stained with RedSafe Nucleic acid Staining Solution (iNtRON Biotechnology, Korea) and visualized under UV light. Molecular sizes of amplified fragments were estimated using GeneRuler 100 bp Plus DNA ladder (Thermo Fisher Scientific).

The 5′-UTR *SFB* intron was also amplified using the degenerate primer pair (F-BOX5′A/F-BOXintronR) developed by Vaughan et al. [63] from sweet cherry sequences (see Additional file 5: Table S5). PCR components, thermo-cycler conditions and detection procedure were identical to that described above for the first *S-RNase* intron. Two additional primers (RFBc-F/SFBins-R), designed from the consensus sequence of the *Prunus SFB* alleles [17], were used to amplify the SFB_C insertion from genomic DNA of several apricot cultivars in order to distinguish S_C and S_8-alleles (see Additional file 5: Table S5).

M-locus genotyping

Seven SSR markers comprised within (or flanking) the *M*-locus were genotyped: PGS3.71, PGS3.22, PGS3.62, PGS3.23 and PGS3.96 [25, 26] and AGS-20 and AGS-30 (this work) (see Additional file 3: Table S3). SSR amplifications were performed in a GeneAmp® PCR System 9700 thermal cycler (Perkin–Elmer, Freemont, CA, USA) in a final volume of 20 μL, containing 1 × DreamTaq buffer, 0.2 mM of each dNTP; 20 ng of genomic DNA and 1 U of DreamTaq DNA polymerase (Thermo Fisher Scientific). Regarding primers, PCRs were performed as described above by the procedure of [60]. The following temperature profile was used: 94 °C for 2 min, then 35 cycles of 94 °C for 45 s, 50–60 °C for 1 min, and 72 °C for 1 min and 15 s, finishing with 72 °C for 5 min. Allele lengths were determined using an ABI Prism 3130 Genetic Analyzer with the aid of GeneMapper software, version 4.0 (Applied Biosystems).

Sequence analysis of PCR products containing *S-RNase* introns

PCR products containing the first and second introns of S_{20} and S_{24}-*RNase* alleles previously obtained from genomic DNA of 'Ezzine' (S_CS_{24}), on one side, and from cultivars 'Cow-2' ($S_{20}S_{22}$), 'Harlayne' (S_3S_{20}), 'Cristalí' ($S_{20}S_C$), 'Gavatxet' ($S_{20}S_C$), 'Mariem' ($S_{20}S_7$), 'Perla' ($S_{20}S_2$), 'Portici' ($S_{20}S_2$), 'Stella' ($S_{20}S_6$), 'Tadeo' ($S_{20}S_C$), 'Veecot' ($S_{20}S_2$) and 'Velázquez' ($S_{20}S_5$), on the other, were sequenced to check their identities. Similarly, PCR products associated with S_3, S_{17} and S_{22}-*RNase* alleles obtained from cultivars 'Harlayne' (S_3S_{20}), 'Henderson' (S_3S_{17}), 'Orange Red' (S_6S_{17}), 'SEO' (S_6S_{17}), 'Cow-1' (S_1S_{22}) and 'Cow-2' ($S_{20}S_{22}$) were also sequenced. Primer combinations SRc-F/SRc-R and Pru-C2/Pru-C4R were used in all cases for the first and second introns respectively except for S_{24} (Pru-T2/SRc-R) and S_{22} (Pru-C2/Pru-C6R). PCR products were electrophoresed in 0.8% or 2% (*w*/*v*) agarose gels (second or first intron, respectively) stained with RedSafe Nucleic acid Staining Solution (iNtRON Biotechnology). Molecular sizes of the amplified fragments were estimated using GeneRuler 100 bp Plus DNA ladder (Thermo Fisher Scientific). Fragments were extracted and purified from the agarose gels using the Zymoclean Gel DNA Recovery kit (Zymo Research, Irvine, CA, USA). Sequences were determined

automatically using an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) and the Big Dye Terminator Cycle Sequencing Kit v3.1 (Applied Biosystems) following the manufacturer's instructions. Sequences were assembled and edited with the Staden package v.2.0 [64]. Multiple alignments were made with CLUSTALW v.1.82 [65] and edited in BioEdit v.7.0.9 [66]. Homology searches were performed against the NCBI Genbank database using the BLASTN program [67].

S_2-locus sequence analysis

Specific primers designed from the apricot S_2-haplotype sequence [17] were used to amplify genomic fragments containing the complete S_2-*RNase* (Sf-Hap2/Sr-Hap2) and SFB_2 coding sequences (FBf-Hap2/FBr-Hap2) (see Additional file 5: Table S5) using 'Portici' (S_2S_{20}) genomic DNA as template. PCR conditions and methods for isolating and sequencing (using the three additional primers SFBc-F, FBF5 and FBF6) these bands (see Additional file 5: Table S5) were the same reported above for fragments containing the *S-RNase* second intron.

Clustering analysis

Reference m_{0-0}, M_{1-0}, M_{2-0} and M_3-haplotypes were established using genetic maps from 'G × Ca', 'K × K' and 'G × K' populations through the automatic determination of linkage phases by JoinMap 3.0 [33]. Remaining *M*-haplotypes were inferred from SSR genotypes by comparing with the references and confirmed by using the EM algorithm [68] implemented in PowerMarker V3.25 software [69]. When available pedigree was used to confirm assignments. Similarities between *M*-haplotypes were estimated by using Jaccard's similarity coefficient [45] through the Phylip 3.62 package [70] and Bruvo's genetic distance [46] through a hand-made phyton script. Bootstrapped data matrices were obtained using the Phyltools 1.32 software [71] and used to test the stability of the neighbor joining trees constructed with the Phylip 3.62 package [70]. HyperTree software [72] was used to visualize the obtained trees.

Additional files

Additional file 1: Table S1. Fragment sizes (bp) determined for *S-RNase* and *SFB* allele introns. Fragment sizes for the first *S-RNase* and the 5′-UTR *SFB* introns were exactly determined by an ABI PRISM 3100 Genetic Analyzer and those for the second *S-RNase* intron were approximately estimated from 0.8% agarose gels. References reporting *S*-allele molecular sizes for the first time are provided. (DOCX 15 kb)

Additional file 2: Table S2. Main results of homology searches performed by BLASTN against the NCBI Genbank database using *Prunus armeniaca S*-RNase sequences. Query sizes, Query cover %, Total score, E-value and Identity % are indicated. Putative names were assigned to the *S*-alleles according to the fragment sizes and sequence analyses. (DOCX 18 kb)

Additional file 3: Table S3. Characteristics of SSR primers developed from peach (PGS) and apricot genome sequences (AGS) located at the *M*-locus. Repeat motifs, Primer sequences, Scaffold positions, ORF (Prupe), Size range, N° of alleles and Heterozygosity (in the set of accessions analyzed) are indicated. (DOCX 14 kb)

Additional file 4: Table S4. Allele sizes (bp) for the SSR markers comprised into the *M*-haplotypes. 'Main class' as well as the subtype are indicated for every *M*-haplotype. Absolute frequency in the set of accessions analyzed is also provided. (DOCX 18 kb)

Additional file 5: Table S5. Primers used in this study. Sequences and references are indicated. (DOCX 14 kb)

Abbreviations

EE: Eastern Europe; GSI: Gametophytic Self-Incompatibility; NA: North America; PPM: Pollen-Part Mutation; SC: Self-Compatibility; SE/NAf: Southern Europe/North Africa; SFB: *S*-locus F-box; SI: Self-Incompatibility; WE: Western Europe

Acknowledgements

This work was supported by two grants from the Ministerio de Economia y Competitividad del Gobierno de España (AGL19018-2010) and the Instituto Nacional de Investigaciones Agrarias (RF2011-00020-C02-02). The authors want to thank Laura Ramirez and Mati González for their technical assistance and to Antonio Monforte and Mario Fares for helpful discussion on the manuscript. We also acknowledge support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI).

Funding

This work was supported by two grants from the *Ministerio de Economia y Competitividad del Gobierno de España* (AGL19018–2010) and the *Instituto Nacional de Investigaciones Agrarias* (RF2011–00020-C02–02).

Availability of data and materials

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

JVMS performed most of the genotyping experiments and part of self-pollination tests and analyzed the data. EZ analyzed the data and developed most of the bioinformatic and statistical analyses. IL contributed to plant sampling and genotyping experiments. MLB contributed materials, reagents and analysis tools, analyzed part of the data and revised the manuscript. CR performed part of the genotyping experiments and self-pollination tests, analyzed the data, designed and coordinated the study and drafted the manuscript. All authors read and approved the final manuscript.

Author's information

When this work started CR worked for the IVIA but since 2014 CR holds a position as CSIC researcher at the IBMCP. JVMS is not currently working for the IVIA either.

Competing interests

The authors declare that they do not have competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Publisher's Note

Author details

[1]Fruit Tree Breeding Department, Instituto Valenciano de Investigaciones Agrarias (IVIA), CV-315, Km. 10,7., 46113 Moncada, Valencia, Spain. [2]Department of Biochemistry, University of Missouri, 117 Schweitzer Hall,

65211 Columbia , MO, USA. [3]Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universidad Politécnica de Valencia-Consejo Superior de Investigaciones Científicas, Camino de Vera, 46022 Valencia, Spain.

Received: 23 September 2016 Accepted: 7 April 2017
Published online: 26 April 2017

References

1. Igic B, Kohn JR. Evolutionary relationships among self-incompatibility RNases. Proc Natl Acad Sci U S A. 2001;98:13167–71.
2. De Nettancourt D. Incompatibility and incongruity in wild and cultivated plants. Berlin: Springer-Verlag; 2001.
3. McClure BA, Haring V, Ebert PR, Anderson MA, Simpson RJ, Sakiyama F, Clarke AE. Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. Nature. 1989;342:955–7.
4. Lai Z, Ma W, Han B, Liang L, Zhang Y, Hong G, Xue Y. An F-box gene linked to the self-incompatibility (*S*) locus of *Antirrhinum* is expressed specifically in pollen and tapetum. Plant Mol Biol. 2002;50:29–42.
5. Sijacic P, Wang X, Skirpan AL, Wang Y, Dowd PE, McCubbin AG, Huang S, Kao T-h. Identification of the pollen determinant of S-RNase-mediated self-incompatibility. Nature. 2004;429:302–5.
6. Ushijima K, Sassa H, Dandekar AM, Gradziel TM, Tao R, Hirano H. Structural and transcriptional analysis of the self-incompatibility locus of almond: Identification of a pollen-expressed F-box gene with haplotype-specific polymorphism. Plant Cell. 2003;15:771–81.
7. Hua Z, Kao T-h. Identification and characterization of components of a putative *Petunia S*-locus F-box-containing E3 ligase complex involved in S-RNase based self-incompatibility. Plant Cell. 2006;18:2531–53.
8. Huang J, Zhao L, Yang Q, Xue Y. AhSSK1, a novel SKP1-like protein that interacts with the S-locus F-box protein SLF. Plant J. 2006;46:780–93.
9. Kubo K, Entani T, Tanaka A, Wang N, Fields AM, Hua Z, Toyoda M, Kawashima S, Ando T, Isogai A, Kao T, Takayama S. Collaborative non-self recognition system in S-RNase-based self-incompatibility. Science. 2010;330:796–9.
10. Fujii S, Kubo K, Takayama S. Non-self- and self-recognition models in plant self-incompatibility. Nat Plants. 2016;2:16130.
11. Kakui H, Kato M, Ushijima K, Kitaguchi M, Kato S, Sassa H. Sequence divergence and loss-of-function phenotypes of S locus F-box brothers (SFBB) genes are consistent with non-self recognition by multiple pollen determinants in self-incompatibility of Japanese pear (*Pyrus pyrifolia*). Plant J. 2011;68:1028–38.
12. Matsumoto D, Tao R. Distinct self-recognition in the *Prunus* S-RNase-based gametophytic self-incompatibility system. Hort J. 2016; doi:10.2503/hortj.MI-IR06.
13. Tao R, Iezzoni AF. The S-RNase-based gametophytic self-incompatibility system in *Prunus* exhibits distinct genetic molecular features. Sci Hort. 2010; 124:423–33.
14. Matsumoto D, Tao R. Recognition of a wide-range of S-RNases by S locus F-box like 2, a general-inhibitor candidate in the Prunus-specific S-RNase-based self-incompatibility system. Plant Mol Biol. 2016;91:459–69.
15. 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. 2009;78:137–57.
16. Hegedűs A, Lénárt J, Halász J. Sexual incompatibility in Rosaceae fruit tree species: molecular interactions and evolutionary dynamics. Biol Plant. 2012; 56:201–9.
17. Vilanova S, Badenes ML, Burgos L, Martínez-Calvo J, Llácer G, Romero C. Self-compatibility of two apricot selections is associated with two pollen-part mutations of different nature. Plant Physiol. 2006;42:629–41.
18. Halász J, Pedryc A, Hegedűs A. Origin and dissemination of the pollen-part mutated S_C haplotype which confers self-compatibility in apricot (*Prunus armeniaca*). New Phytol. 2007;176:792–803.
19. Vilanova S, Romero C, Llácer G, Badenes ML. Identification of self-(in)compatibility alleles in apricot by PCR and sequence analysis. J Am Soc Hortic Sci. 2005;130:893–8.
20. Halász J, Pedryc A, Ercisli S, Yilmaz KU, Hegedűs A. *S*-genotyping supports the genetic relationships between Turkish and Hungarian apricot germplasm. J Amer Soc Hort Sci. 2010;135:410–7.
21. Kodad O, Hegedűs A, Socias i Company R, Halász J. Self-(in)compatibility genotypes of Moroccan apricots indicate differences and similarities in the crop history of European and North African apricot germplasm. BMC Plant Biol. 2013;13:196.
22. Lachkar A, Fattouch S, Ghazouani T, Halasz J, Pedryc A, Hegedűs A, Mars M. Identification of self-(in)compatibility *S*-alleles and new cross-incompatibility groups in Tunisian apricot (*Prunus armeniaca* L.) cultivars. J Hortic Sci Biotechnol. 2013;88:497–501.
23. McClure BA, Cruz-García F, Beecher B, Sulaman W. Factors affecting inter- and intra-specific pollen rejection in *Nicotiana*. Ann Bot. 2000;85:113–23.
24. Wünsch A, Hormaza JI. Genetic and molecular analysis in Cristobalina sweet cherry, a spontaneous self-compatible mutant. Sex Plant Reprod. 2004;17:203–10.
25. Zuriaga E, Molina L, Badenes ML, Romero C. Physical mapping of a pollen modifier locus controlling self-incompatibility in apricot and synteny analysis within the Rosaceae. Plant Mol Biol. 2012;79:229–42.
26. Zuriaga E, Muñoz-Sanz JV, Molina L, Gisbert AD, Badenes ML, Romero C. An S-locus independent pollen factor confers self-compatibility in "Katy" apricot. PLoS One. 2013;8:e53947.
27. Halász J, Hegedüs A, Hermán R, Stefanovits-Bányai E, Pedryc A. New self-incompatibility alleles in apricot (*Prunus armeniaca* L.) revealed by stylar ribonuclease assay and S-PCR analysis. Euphytica. 2005;145:57–66.
28. Halász J. Molecular background of the S-locus controlled self-incompatibility in apricot, PhD Diss. Budapest, Hungary: Corvinus University of Budapest; 2007.
29. Zhang L, Chen X, Chen X, Zhang C, Liu X, Ci Z, Zhang H, Wu C, Liu C. Identification of self-incompatibility (*S*-) genotypes of Chinese apricot cultivars. Euphytica. 2008;160:241–8.
30. Wu J, Gu C, Zhang SL, Zhang SJ, Wu HQ, Heng W. Identification of *S*-haplotype specific *S-RNase* and *SFB* alleles in native Chinese apricot (*Prunus armeniaca* L.). J Hortic Sci Biotechnol. 2009;84:645–52.
31. Gu C, Wu J, Du YH, Yang YN, Zhang SL. Two different *Prunus SFB* alleles have the same function in the self-incompatibility reaction. Plant Mol Biol Rep. 2013;31:425–34.
32. Brooks RM, Olmo HP. The Brooks and Olmo Register of Fruit and Nut Varieties. 3rd ed. Alexandria VA: ASHS Press; 1997.
33. Van Ooijen JW, Voorrips RE. JoinMap®3.0, Software for the calculation of genetic linkage maps. Wageningen, The Netherlands: Plant Research International; 2001.
34. Nei M. Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci U S A. 1973;70:3321–3.
35. Orero G, Cuenca J, Romero C, Martínez-Calvo J, Badenes ML, Llácer G. Selection of seedling rootstocks for apricot and almond. Acta Hort (ISHS). 2004;658:529–33.
36. Mehlenbacher SA, Cociu V, Hough LF. Apricots. In: Moore JN, Ballington JR, editors. Genetic Resources of Temperate Fruit and Nut Crops. Wageningen, The Netherlands: International Society for Horticultural Science; 1990. p. 65–107.
37. Faust M, Surányi D, Nyujtó F. Origin and dissemination of apricot. Hort Rev. 1998;22:225–66.
38. Bourguiba H, Audergon JM, Krichen L, Trifi-Farah N, Mamouni A, Trabelsi S, D'Onofrio C, Asma BM, Santoni S, Khadari B. Loss of genetic diversity as a signature of apricot domestication and diffusion into de Mediterranean Basin. BMC Plant Biol. 2012;12:49–64.
39. Martin C, Herrero M, Hormaza JI. Molecular characterization of apricot germplasm from an old stone collection. PLoS One. 2011;6:e23979.
40. Hajilou J, Grigorian V, Mohammadi SA, Nazemmieh A, Romero C, Vilanova S, Burgos L. Self- and cross-(in)compatibility between important apricot cultivars in northwest Iran. J Hort Sci Biotech. 2006;81:513–7.
41. García S, Sánchez-Capuchino JA, Dalmau F, Casanova R, Salazar D. Taxonomía pomológica de los albaricoqueros valencianos (I). Agrícola Vergel. 1985;4:538–51.
42. Burgos L, Pérez-Tornero O, Ballester J, Olmos E. Detection and inheritance of stylar ribonucleases associated with incompatibility alleles in apricot. Sex Plant Reprod. 1998;11:153–8.
43. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 1980;16:111–20.
44. Dondini L, De Franceschi P, Sansavini S. Gametophytic incompatibility in pome and stone fruits: genes controlling *S*-locus. Acta Hort (ISHS). 2012;967:67–75.
45. Jaccard P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Bull Soc Vaudoise Sci Nat. 1901;37:547–79.
46. Bruvo R, Michiels NK, D'Souza TG, Schulenburg H. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. Mol Ecol. 2004;13:2101–6.
47. Burgos L, Alburquerque N, Egea J. Review. Flower biology in apricot and its implications for breeding. Span J Agric Res. 2004;2:227–41.
48. Milatović D, Nikolić D, Fotirić-Akšić M, Radović A. Testing of self-(in)compatibility in apricot cultivars using fluorescence microscopy. Acta Sci. Pol. Hortorum Cultus. 2013;12:103–13.

49. Ikeda K, Igic B, Ushijima K, Yamane H, Hauck NR, Nakano R, Sassa H, Iezzoni AF, Kohn JR, Tao R. Primary structural features of the *S* haplotype specific F-box protein, SFB, in *Prunus*. Sex Plant Reprod. 2004;16:235–43.
50. Jie Q, Shupeng G, Jixiang Z, Manru G, Huairui S. Identification of self-incompatibility genotypes of apricot (*Prunus armeniaca* L.) by S-allele-specific PCR analysis. Biotechnol Lett. 2005;27:1205–9.
51. Zhebentyayeva TN, Reighard GL, Lalli D, Gorina VM, Krška B, Abbott AG. Origin of resistance to plum pox virus in Apricot: what new AFLP and targeted SSR data analyses tell. Tree Genet Genomes. 2008;4:403–17.
52. Lloyd DG. Self- and cross-fertilization in plants. II. The selection of self-fertilization. Int J Plant Sci. 1992;153:370–80.
53. Cachi AM, Hedhly A, Hormaza JI, Wünsch A. Pollen tube growth in the self-compatible sweet cherry genotype, 'Cristobalina', is slowed down after self-pollination. Ann Appl Biol. 2014;164:73–84.
54. Marchese A, Boskovic RI, Caruso T, Raimondo A, Cutuli M, Tobutt KR. A new self-compatibility haplotype in the sweet cherry 'Kronio' S_5' attributable to a pollen-part mutation in the SFB gene. J Exp Bot. 2007;58:4347–56.
55. Dirlewanger E, Graziano E, Joobeur T, Garriga-Calderé F, Cosson P, Howad W, Arús P. Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proc Natl Acad Sci. 2004;101:9891–6.
56. Romero C, Pedryc A, Munoz V, Llacer G, Badenes ML. Genetic diversity of different apricot geographical groups determined by SSR markers. Genome. 2003;46:244–52.
57. Doyle JJ, Doyle JL. A rapid isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11–5.
58. Thomson D, Henry R. Single-step protocol for preparation of plant tissue for analysis by PCR. BioTechniques. 1995;19:394–400.
59. Romero C, Vilanova S, Burgos L, Martínez-Calvo J, Vicente M, Llácer G, Badenes ML. Analysis of the *S*-locus structure in *Prunus armeniaca* L. Identification of *S*-haplotype specific *S-RNase* and *F-box* genes. Plant Mol Biol. 2004;56:145–57.
60. Schuelke M. An economic method for the fluorescent labelling of PCR fragments. Nat Biotechnol. 2000;18:233–4.
61. Tao R, Yamane H, Sugiura A. Molecular typing of *S*-alleles through identification, characterization and cDNA cloning for S-RNases in sweet cherry. J Am Soc Hortic Sci. 1999;124:224–33.
62. Sonneveld T, Tobutt KR, Robbins TP. Allele-specific PCR detection of sweet cherry self-incompatibility (S) alleles S1 to S16 using consensus and allele-specific primers. Theor Appl Genet. 2003;107:1059–70.
63. Vaughan SP, Russell K, Sargent DJ, Tobutt KR. Isolation of S-locus F-box alleles in *Prunus avium* and their application in a novel method to determine self-incompatibility genotype. Theor Appl Genet. 2006;112:856–66.
64. Bonfield J: Staden package, version 2.0. http://staden.sourceforge.net (2004). Accessed 16 September 2016.
65. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22:4673–80.
66. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser. 1999;41:95–8.
67. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
68. Excoffier L, Slatkin M. Maximum-Likelihood estimation of molecular haplotype frequencies in a diploid population. Mol Biol Evol. 1995;12:921–7.
69. Liu K, Muse SV. PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics. 2005;21:2128–9.
70. Felsenstein J. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Seattle: Department of Genome Sciences, University of Washington; 2005.
71. Buntjer JB. Phylogenetic computer tools (PhylTools). Version 1.32 for Windows. Wageningen: Laboratory of Plant Breeding, Wageningen University; 1997.
72. Bingham J, Sudarsanam S. Visualizing large hierarchical clusters in hyperbolic space. Bioinformatics. 2000;16:660–1.
73. Russell D. The stonefruit cultivar system (A database of worldwide stonefruit cultivars and rootstocks). Queensland, Australia: Department of Primary Industries; 1998.
74. Syrgianidis GD, Mainou AC. Deux nouvelles varieties d'abricotier resistances à la maladie à virus de la sharka (Plum pox) issues de croissements. In: Audergon JM, editor. Deuxièmes rencontres sur l'abricotier, Avignon (France). Luxembourg: Commission des Communautés européennes (EUR 15009R); 1993. p. 135–7.
75. Badenes ML, Martínez-Cavo J, García-Carbonell S, Villarrubia D, Llácer G. Descripción de variedades autóctonas valencianas de albaricoquero. Valencia: Generalitat Valenciana (Consellería de Agricultura, Pesca y alimentación) Series Divulgación Técnica; 1997.
76. Della Strada G, Pennone F, Fideghelli C, Monastra F, Cobianchi D. Monografía di cultivar di albicocco. Roma: Ministerio dell'Agricoltura e delle Foreste. Direzione Generale della Produzione Agricola; 1989.
77. Massai R. Variability of apricot cultivars traits inside the 'List of recommended fruit varieties' project. Acta Hort (ISHS). 2010;862:129–36.
78. Nyujtó F, Brozic S, Nyéki J, Brozic S. Flowering and fruit set in apricot varieties grown in Hungary, and combination of varieties within the plantation. Acta Hort (ISHS). 1982;121:159–65.
79. Alburquerque N, Egea J, Pérez-Tornero O, Burgos L. Genotyping apricot cultivars for self-(in)compatibility by means of RNases associated with S alleles. Plant Breed. 2002;121:343–7.