Running head: Pollen-part mutant analysis in apricot

Carlos Romero

Instituto Valenciano de Investigaciones Agrarias (IVIA). Ctra. Moncada-Náquera Km 5,5. 46113 Moncada (Valencia) Spain

Phone number: (34) 96 3424000 Ext 70166

Fax number: (34) 96 3424001

E-mail: cromero@ivia.es

Genetics, Genomics and Molecular Evolution

Self-compatibility of two *Prunus armeniaca* selections is associated with two pollen-part mutations of different nature

Santiago Vilanova[1a], María Luisa Badenes[1], Lorenzo Burgos[2], José Martínez-Calvo[1], Gerardo Llácer[1] and Carlos Romero[1b]

[1] Instituto Valenciano de Investigaciones Agrarias (IVIA). Ctra. Moncada-Náquera Km 5,5. 46113 Moncada (Valencia) Spain

[2] Departamento de Mejora de Frutales. CEBAS-CSIC, Apartado 164, 30100 (Murcia) Spain

This research was supported by a grant from the 'Ministerio de Ciencia y Tecnología' of Spain (AGL2001-1122-C02-02). S. Vilanova was funded by a fellowship from the 'Ministerio de Ciencia y Tecnología'.

[a] Present address: Departamento de Genética, IRTA Centre de Cabrils. Ctra. de Cabrils, s/n. 08348 Cabrils (Barcelona) Spain

[b] Corresponding author; e-mail cromero@ivia.es; fax number (34) 96 3424001

ABSTRACT

Loss of pollen-*S* function in *Prunus* self-compatible mutants has recently been associated with deletions or insertions in *S*-haplotype-specific F-box (*SFB*) genes. We have studied two self-compatible cultivars of apricot (*Prunus armeniaca* L.), 'Currot' ($S_C S_C$) and 'Canino' ($S_2 S_C$), sharing the naturally occurring S_C-haplotype. Sequence analysis showed that, whereas the S_C-*RNase* is unaltered, a 358 bp insertion is found in the SFB_C gene resulting in the expression of a truncated protein. The alteration of this gene is associated with self-incompatibility breakdown supporting previous evidence that point to *SFB* being the pollen-*S* gene of the *Prunus* self-incompatibility (*S*) locus. On the other hand, PCR-analysis of progenies derived from 'Canino' showed that pollen grains carrying the S_2-allele were also able to overcome the incompatibility barrier. However, alterations in the SFB_2 gene or evidence of pollen-*S* duplications were not detected. A new class of F-box genes encoding a previously uncharacterized protein with high sequence similarity (~62%) to *Prunus* SFB proteins was identified in this work but the available data rules them out of producing *S*-heteroallelic pollen and thus the cause of the pollen part mutation. These results suggest that the 'Canino' cultivar has an additional mutation, not linked to the *S* locus, that causes a loss of pollen-*S* activity when present in pollen. As a whole these findings support the proposal that the *S*-locus products besides other *S*-locus independent factors are required for gametophytic self-incompatibility in *Prunus*.

INTRODUCTION

Gametophytic self-incompatibility (GSI) is a widespread mechanism in the plant kingdom often controlled by a single multiallelic locus, termed the *S*-locus, that prevents inbreeding and promotes out-crossing (de Nettancourt, 2001). In the Solanaceae, Scrophulariaceae and Rosaceae the *S*-locus is considered to contain at least two linked genes, one encodes glycoproteins with ribonuclease activity in the pistils (S-RNases) (McClure et al., 1989; Boskovic and Tobutt, 1996; Xue et al., 1996) and the other is an F-box pollen-expressed gene (named *SLF* or *SFB*) (Lai et al., 2002; Entani et al., 2003; Ushijima et al., 2003; Sijacic et al., 2004).

In GSI, the RNase activity of S-RNases expressed in the styles is essential to arrest pollen growth by degrading pollen RNA (Huang et al., 1994) when the haploid pollen *S*-allele matches either of the two *S*-alleles of the diploid pistil. The recent identification of *S*-linked F-box genes as the pollen-*S* determinants of the GSI system in *Antirrhinum* (Lai et al., 2002), *Prunus* (Entani et al., 2003; Ushijima et al., 2003) and *Petunia* (Sijacic et al., 2004) has thrown some light on the underlying mechanism. The F-box proteins are involved in the ubiquitin/26S proteasome proteolytic pathway (Deshaies, 1999). Therefore, a system in which the non-self S-RNases, rather than be inhibited by the pollen-*S* factor, are marked to be degraded is more consistent (Ushijima et al., 2003; Ikeda et al., 2004). In *Antirrhinum* the pollen-expressed F-box protein $AhSLF_2$ has been already shown to interact with both self and non-self S-RNases (Qiao et al., 2004). These findings are compatible with the inhibitor model, where all S-RNases, regardless of their *S*-haplotype, enter into the pollen tube (Luu et al., 2000), and once there all of them are inhibited except the cognate S-RNase that degrades self-pollen RNA. How the cognate S-RNase is specifically protected to remain active is still unknown.

The spontaneous and induced self-compatible mutants have been extensively used to study the molecular basis of the GSI mechanism. Stylar part mutations in the *S*-locus have been reported in Solanaceae (Royo et al., 1994) and Rosaceae (Sassa et al., 1997) revealing that RNase activity of S-RNases is needed to inhibit pollen growth. Numerous pollen-part mutants (PPM) have been described in the Solanaceae and most of them are consistent with competitive interaction in which *S*-heteroallelic pollen fails to function in self-incompatibility (SI) (Thompson et al., 1991; Golz et al., 1999; Golz et al., 2001). In *Prunus*, the recent analysis of self-compatible mutant haplotypes has shown that *SFB* genes are defective, providing additional evidence that *SFB*s are the pollen-*S* genes in

GSI in *Prunus* (Ushijima et al., 2004; Sonneveld et al., 2005) and suggesting that mutations or deletions of the *SFB*s rather than *S*-allele duplications cause the pollen function breakdown in the Rosaceae. Finally, breakdown of SI has been also associated with mutations that may be affecting additional *S*-locus external factors in *Prunus avium* (Wünsch and Hormaza, 2004).

The degree of natural self-fertility varies between *Prunus* species. Peaches (*P. persica*) are normally self pollinated, apricots (*P. armeniaca*) partly self-compatible, and almonds (*P. dulcis*) and sweet cherries (*P. avium*) mostly self-incompatible (Watkins, 1976). Among the apricots, many early Spanish cultivars, genetically close related, are self-compatible (Burgos et al., 1993). Two of these apricot self-compatible mutants, 'Currot' (S_CS_C) and 'Canino' (S_2S_C), have been analyzed in this work. Loss of pollen-*S* function in the natural occurring S_C-haplotype has been associated with an insertion in the SFB_C allele that co-segregates with self-compatibility. In addition, this study provides evidence that the loss of function of an additional factor not linked to the *S*-locus is also involved in the breakdown of SI in the PPM 'Canino'. Finally, as a result of the pollen-*S* gene duplication analysis in both cultivars, an unusual type of F-box genes with high sequence similarity to *Prunus* SFB proteins and not tightly linked to the *S*-locus has been identified in apricot.

RESULTS

Genetic analysis of the self-compatible cultivar 'Currot' (S_CS_C)

Based on evidence obtained from stylar RNase analysis, PCR and pollen-tube growth tests on controlled crosses, the self-compatible apricot cultivar 'Currot' (S_CS_C) was suggested to be homozygous for the natural occurring S_C-haplotype (Alburquerque et al., 2002; Vilanova et al., 2005). *S-RNase* PCR-typing of the cross population 'Goldrich' (S_1S_2) x 'Currot' (S_CS_C) showed, as expected from the 'Currot' *S*-genotype, that the S_C-*RNase* is present in all the progeny (Supplemental Figure 1A) and grouped into two heterozygote classes S_1S_C and S_2S_C with an observed proportion 31:39 that fits with the expected segregation ratio 1:1 (Table 1). To analyze this progeny also for the S_C-*RNase* pollen counterpart, we cloned and sequenced an SFB_C fragment PCR-amplified with consensus primers from 'Currot'. Subsequently, we designed specific primers (RFBc-F/RFBc-R) to identify the SFB_C-allele by PCR. Supplemental Figure 1B

shows that the SFB_C allele is also present in all the 'Goldrich' x 'Currot' seedlings. The co-segregation of the S_C-*RNase* and SFB_C in the 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C) population (see Figures 3A and 3B) provided evidence of genetic linkage between the two genes.

Molecular analysis of the apricot S_C-haplotype

To test whether mutations or indels, affecting the putative S_C-*RNase* or S_C-haplotype specific *SFB* genes, were the cause of SI breakdown in the S_C-haplotype, we cloned and sequenced genomic DNA fragments, containing both genes and their adjacent regions, from the apricot self-compatible cultivar 'Currot' (S_CS_C). Figure 1A shows the putative genomic structure of the S_C-haplotype. Two fragments (~3.0 and ~5.0 Kbp, respectively) containing open reading frames (ORF) with homology to the known *SFB* and *S-RNase* genes (SFB_C and S_C-*RNase*) were identified. Several direct and inverted repeats along with microsatellites were localized mainly in the S_C-*RNase* intron. Physical distance between both genes and their relative positions are not yet definitively established but PCR segregation analysis suggests that they are genetically linked (see Figures 3A and 3B).

The putative apricot S_C-RNase shows the typical features of *Prunus* T2 type RNases with five conserved domains (C1, C2, C3, RC4 and C5) and one hypervariable region (RHV) (Supplemental Figure 2). The positions of the two S_C-*RNase* introns (260 and 2680 bp in size, respectively) are also identical to those of the other *Prunus S-RNases*. Amino acid identity of the S_C-RNase to other *Prunus* S-RNases ranged from 59.7 to 80.0% preserving the typical high allelic sequence diversity of the S-RNases (Table 2). The primary structure of the encoded apricot S_C-RNase is intact and, furthermore, motifs surrounding the histidine residues necessary for the RNase activity in the C2 and C3 domains (Kawata et al., 1989) are present. The only conserved potential N-glycosylation site (NxS/T consensus sequence) (Ishimizu et al., 1998) located at the Asn-142 according to the numbering for the *P. armeniaca* S_1-RNase (AY587561) was also identified (Supplemental Figure 2). Finally, RT-PCR analysis showed that the S_C-*RNase* is specifically expressed in style tissues but not in pollen or leaves (Figure 1D).

Sequence analysis revealed a 358 bp insertion disrupting the putative SFB_C ORF at +904. The ends of the inserted sequence are two ~52 bp inverted repeats (~75% nucleotide identity) similar to the inverted terminal repeats (ITR) of transposable

elements (Figure 1B). Interestingly, BLASTN analysis (Altschul et al., 1990) revealed significant similarity between the insert and two sequences found in the EMBL/GenBank/DDBJ database, the peach BAC clone 28F08 (AC154900) and a partial coding sequence of the *Prunus salicina* gene for *S*n-*RNase* (AB093136) (82% and 77% nucleotide identities, respectively). Furthermore, the insertion also displayed significant similarity (52%) with the S_C-*RNase* 5′ flanking region located between -404 and –740 from the translational start codon. The inserted sequence leads to a frame shift and produces a premature stop codon at +904 in the SFB_C transcript (Figures 1A, 1B and 1C). Figure 1B also indicates the position of the putative intron associated with the 5' untranslated region (UTR) of *Prunus* SFB (Vaughan et al., 2006). Therefore, SFB_C transcript encodes a putative truncated protein lacking the last 75 amino acid residues of the C-terminal half including the HVa and HVb hypervariable domains (Figure 1C). No progenitor allele is available but the nucleotide sequence predicted for the hypothetical allele which is supposed to be the original SFB_C (OriSFB_C, according to the nomenclature established by Ushijima et al. (2004)), lacking the insert, encodes a typical *S*-locus F-box protein (Ikeda et al., 2004), with one F-box domain and four (Hyper)variable regions (V1, V2, HVa and HVb) (Figure 1C). OriSFB_C shows also the typical high sequence diversity to other *Prunus* SFBs with amino acid identities ranging from 65.1 to 78.9% (Table 2). Gene expression analysis performed by RT-PCR showed that SFB_C is expressed in pollen tissues but not in styles or leaves (Figure 1D).

The presence of the fragment inserted in SFB_C was confirmed by PCR-amplification in all apricot self-compatible cultivars sharing the S_C-haplotype used in this study (Figure 2). Because the S_C-*RNase* sequence is intact and an insertion leading to a premature stop codon in translation is found in the SFB_C, the S_C-haplotype is believed to be a PPM. As pollen-part mutants generated in the Solanaceae are mostly associated with *S*-locus duplications (Brewbaker and Natarajan, 1960; Golz et al., 1999; Golz et al., 2001), we tested this possibility in 'Currot' using flow cytometry and DNA blot analysis. In the flow cytometry analysis the peaks of nuclei isolated from 'Currot' and the control diploid plant ('Goldrich') were coincident pointing out that 'Currot' is a diploid. In addition, DNA blot analysis conducted with several restriction enzymes provided no evidence of genetic duplication (see below).

Genetic analysis of the self-compatible cultivar 'Canino' (S_2S_C)

The *S*-genotype of the self-compatible cultivar 'Canino' (S_2S_C) was previously determined by analysis of stylar ribonucleases (Alburquerque et al., 2002) and PCR (Vilanova et al., 2005). Initially, the origin of self-compatibility in this cultivar was associated with the natural occurring S_C-haplotype. Nevertheless, segregation analysis of *S*-alleles performed in different controlled crosses did not agree with this hypothesis. Figure 3A shows the *S-RNase* genotyping of the 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C) population determined by PCR-amplification. The *S*-genotypes of this population fell into four classes (S_1S_C : S_1S_2 : S_2S_2 : S_2S_C) (Table 1), instead of the two expected (S_1S_C : S_2S_C). Therefore, two unexpected *S*-genotype classes were obtained in this cross, S_1S_2 and S_2S_2, since pollen tubes carrying the S_2-allele from 'Canino' were supposed to be arrested in the style. Similarly, self-pollination of 'Canino' produced three *S*-genotype classes (S_2S_C : S_2S_2 : S_CS_C) including the unexpected S_2S_2 seedlings (Table 1), instead of the two expected (S_2S_C : S_CS_C). 'Pepito' (S_2S_C), a genetically close cultivar sharing the same *S*-genotype than 'Canino' (S_2S_C), was also tested for the SI breakdown not associated with the S_C-haplotype. However, in this case only the expected S_1S_C and S_2S_C genotypes derived from the 'Goldrich' (S_1S_2) x 'Pepito' (S_2S_C) cross were observed (Figure 3C). No pollen contamination was found by the detection of unrelated *S*-alleles in these populations. After self-pollination in the field, the four *S*-genotype classes obtained in the 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C) cross were shown to be self-compatible, setting at least 5% fruit in most seedlings and 2% fruit in nearly all seedlings (Table 3). It is noteworthy that the average fruit set decreases slightly in the S_1S_2 and S_2S_2 genotypes (Table 3). To confirm self-compatibility, two selected trees (GC-35 and GC-147) with a fruit set >10% and belonging to the S_1S_2 and S_2S_2 classes, respectively, were subjected to pollen-growth tests. Figure 4 shows that self-pollen growth of GC-147 (S_2S_2) as well as cross-pollination with the self-incompatible cultivar 'Goldrich' (S_1S_2) were successful. Similar results were obtained with GC-35 (S_1S_2) (data not shown).

SI of the cultivar 'Goldrich', previously determined by controlled cross-pollinations (Egea and Burgos, 1996), has been confirmed by pollen-growth tests in this work (Figure 4A). Thus, regardless of the S_C-haplotype, SI breakdown in the 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C) cross should be due to an alteration in the pollen of 'Canino'. This hypothesis is also supported by the *S*-genotype segregation resulting from the 'Canino' self-pollination. To test whether the pollen-part mutation in 'Canino' is linked or not to the *S*-locus we performed chi-square tests of the segregation ratios observed in

these two populations. The *S*-haplotypes of the 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C) progeny segregated into four classes (S_1S_C:S_2S_C:S_1S_2:S_2S_2) comprising 66:55:28:22 individuals, respectively (Table 1). In addition, self-pollination of 'Canino' (S_2S_C) produced three *S*-genotype classes (S_2S_C:S_2S_2:S_CS_C) comprising 53:11:35 individuals, respectively (Table 1). These observed ratios fit with the expected ratios for a model considering the pollen parent heterozygous for a pollen-part mutation unlinked to the *S*-locus (2:2:1:1 and 3:1:2, respectively) (Table 4) with χ^2 values of 2.98 and 2.20 (P=0.394 and P=0.333) (Table 1). On the contrary, if we consider pollen part mutations linked in coupling to the *S*-locus the expected ratios (1:1:1:1 and 2:1:1, respectively) do not fit with the observed data with χ^2 values of 31.31 and 12.13, respectively (P<0.0001 and P=0.002). Furthermore, *S*-allele segregation data of progenies derived from the self-pollination of S_1S_C and S_2S_C GC-seedlings revealed that both *S*-genotypes are able to produce homozygous (S_1S_1 or S_2S_2) individuals (Table 1). This result also seems to indicate that the loss of pollen-*S* function of 'Canino' is not tightly linked to the *S*-locus and therefore is inherited independently of the *S*-haplotype. In these cases segregations fit with the expected ratio for a single mutation unlinked to the *S*-locus (3:1:2) (Table 1). In agreement with this hypothesis, some GC-seedlings are not able to produce homozygous individuals, supposedly because they did not inherit the mutation, and their segregations fit with the expected ratio (1:1) for a non-mutated seedling (Table 1). To confirm these results 25 S_1S_C and 20 S_2S_C additional GC-seedlings have been tested and the observed segregations, 14:11 and 13:7 (mutated : non-mutated seedlings) respectively, fit with the expected ratio 1:1 with χ^2 values of 0.36 and 1.80 (P=0.548 and P=0.180).

Molecular analysis of the self-compatible cultivar 'Canino' (S_2S_C)

To test whether the growth of 'Canino' pollen tubes carrying the S_2-allele was due to mutations or indels affecting the S_2-haplotype, we cloned and sequenced genomic DNA fragments containing the *SFB*$_2$ and S_2-*RNase* genes and their flanking regions from the apricot cultivars 'Canino'(S_2S_C) and 'Pepito'(S_2S_C). The nucleotide sequence of the S_2-haplotype genomic region containing both genes in the self-incompatible cultivar 'Goldrich'(S_1S_2) was used as reference (Romero et al., 2004). No changes were detected in the nucleotide sequences of the two S_2-haplotype fragments cloned (~1.9

and ~2.2 Kbp, respectively) containing the ORFs of the *SFB*$_2$ and *S*$_2$-*RNase* genes among the three cultivars. Furthermore, gene expression analysis based on RT-PCR showed that both genes are expressed specifically in pollen and styles, respectively, in the three cultivars (Figure 5). Taken together, these findings indicate that neither the coding sequence nor the promoter functionality are altered in the 'Canino' *SFB*$_2$ and *S*$_2$-*RNase* genes.

Mutations that affect the pollen response are frequently associated with duplicated *S*-alleles in the Solanaceae (Golz et al., 2001). Thus, duplication of the *S*-locus F-box genes in the 'Canino' *S*$_2$-haplotype might also be the cause of pollen function breakdown. Genetic and molecular approaches were carried out to test this possibility. Firstly, as previously described for 'Currot', flow cytometry determined that 'Canino' is diploid discarding polyploidy as a possible origin of *S*-heteroallelic pollen. Afterwards, to look for evidence of genetic duplication, *SFB*-fragments PCR-amplified with consensus primers (SFBc-F and SFBc-R) from genomic DNA of 'Canino' (*S*$_2$*S*$_C$) and the self-compatible seedling GC-5 (*S*$_2$*S*$_2$) were cloned. Clones were analyzed for CAPS (Cleavage Amplified Polymorphism Sequence) digesting with the restriction enzymes *Eco*RI-*Nde*I, *Eco*RI-*Kpn*I and *Mva*I that recognize specific sites for cleavage in *SFB*$_2$ and *SFB*$_C$. Forty-three clones obtained from 'Canino' exhibited the restriction patterns predicted by the *SFB*$_2$ or *SFB*$_C$ sequences and twenty-six obtained from GC-5 the pattern predicted by *SFB*$_2$. Sequencing of six clones from each group showed that they indeed encoded SFB$_C$ and SFB$_2$. A different restriction pattern was found in three clones, named *FB*$_1$ (F-box like), two from 'Canino' and one from GC-5 (Supplemental Figures 3A and 3B). Digestions with *Kpn*I and *Mva*I confirmed these results (data not shown). The sequences obtained from these three clones were identical to each other and significantly similar to the *S*-locus F-box genes. Specific genomic PCR-amplification of this fragment showed that it is present in all the cultivars and GC-seedlings analyzed but, interestingly, it could not be amplified from BAC clones spanning a ~180 kb region containing the 'Goldrich' *S*-locus, suggesting that it is located outside of the *S*-locus (Supplemental Figure 3C). Genomic adjacent regions of this fragment were cloned and sequenced in 'Canino' and 'Goldrich' to obtain the complete F-box gene coding sequence. One and two allelic variants with no introns were found in 'Canino' and 'Goldrich', respectively and designated *ParFB*$_1$, *ParFB*$_2$ and *ParFB*$_3$ (*Prunus armeniaca* F-box protein 1, 2 and 3). The insertion of guanosine single nucleotides in the *ParFB*$_1$ and *ParFB*$_3$ (at +608 and +499, respectively) coding

sequences produce frame shifts leading to premature stop codons. On the contrary, *ParFB*$_2$ coding sequence was not disrupted and encoded a complete protein of 378 amino acids. ParFB amino acid deduced sequences shown the N-terminal F-box domain observed in all *Prunus* SFB proteins (including those encoded by which are supposed to be the original *ParFB*$_1$ (Ori*ParFB*$_1$) and *ParFB*$_3$ (Ori*ParFB*$_3$) lacking the inserted guanosines) (Supplemental Figures 4A and 4B). Amino acid sequence identity among ParFB proteins is very high (>99%) hence no typical SFB (Hyper)variable regions were found, but the level of identity between ParFB$_2$ and SFB proteins varies from 59% to 64%. RT-PCR assays (Supplemental Figure 4C) using *ParFB* specific primers showed that *ParFB* is strongly expressed in pollen but also in leaves and styles. Differential intensities of the bands may reflect stronger expression in pollen tissues although quantitative analysis should be performed to confirm this observation.

To complement the CAPS analysis we studied the presence of genetic duplications by genomic DNA blot analysis performed on four cultivars ('Canino'(S_2S_C), 'Currot'(S_CS_C), 'Goldrich' (S_1S_2) and 'Pepito'(S_2S_C)) and four 'Goldrich' x 'Canino' seedlings using a 306 bp *SFB*$_1$ DNA fragment as a probe. Digestions with *Eco*RI and *Eco*RI-*Hind*III restriction enzymes resulted in single restriction fragments corresponding to *SFB*$_1$, *SFB*$_2$ and *SFB*$_C$-alleles (Figures 6A and 6B) according to the *S*-genotypes previously reported for each selection by genomic PCR (Vilanova et al., 2005). *SFB*$_2$ and *SFB*$_C$ alleles cross-hybridized slightly weaker to the *SFB*$_1$ probe than *SFB*$_1$. In addition, extra-bands common to all the analysed selections were detected in the two digestions. In the *Eco*RI digest the estimated size of the common band was ~6.2 kb (Figure 6A) but in the *Eco*RI-*Hind*III digest the extra-bands detected showed different sizes among the cultivars ~5.9 kb for 'Goldrich' and 'Pepito', ~6.1 kb for 'Currot' and ~6.6 kb for 'Canino'. Moreover, the four seedlings share both ~5.9 kb and ~6.6 kb restriction fragments in agreement with 'Goldrich' and 'Canino' being homozygous for these bands (Figure 6B). The re-hybridisation of the filters with a 707 bp *ParFB*$_1$ probe showed that these common bands correspond to *ParFB* alleles, because they hybridised specifically to this probe and the hybridisation signals corresponding to *SFB*-alleles were strongly reduced or undetected (Figures 6C and 6D). Results from genomic DNA blot analysis suggest that all selections analyzed carry *ParFB* homologous fragments.

DISCUSSION

A frame-shift mutation of the specific S_C-haplotype F-box gene is associated with self-compatibility in apricot

The presence of the S_C-RNase has been determined by stylar ribonuclease analysis and PCR in all apricot self-compatible cultivars studied to date (Burgos et al., 1998; Alburquerque et al., 2002; Vilanova et al., 2005). Moreover, it is absent in all self-incompatible cultivars analysed supporting the hypothesis of Burgos et al. (1998) that the *S_C-RNase* is associated with a self-compatible *S*-haplotype (S_C-haplotype). *S*-allele typing and self-compatibility determination of seedlings resulting from different crosses also confirmed this hypothesis and demonstrated that the *S_C-RNase* co-segregates with self-compatibility (Burgos et al., 1998; Alburquerque et al., 2002).

Results obtained in this work are consistent with *SFB_C*-allele disruption being the cause of SI breakdown associated with the S_C-haplotype. Indeed, sequence analysis showed that the *S_C-RNase* gene is unaltered whereas a 358 bp insertion is found in the middle of the putative *SFB_C* coding region. This insertion is located just upstream from the HVa hypervariable region and leads to a frame shift that produces transcripts of a defective SFB_C protein lacking the HVa and HVb hypervariable regions, suggesting that the S_C-haplotype is a PPM. RT-PCR analysis showed that *SFB_C* is expressed in pollen tissues but no data are available about the production of its truncated protein. However, even if the truncated protein is produced, it is unlikely to be functional, since the two C-terminal half hypervariable regions are lacking, and both are suggested to be essential for the function of SFB (Ikeda et al., 2004). This feature is common to two PPMs recently reported in *Prunus mume,* where a 6.8 kb insertion found in the middle of the *SFB^f* coding region leads to a frame shift and a premature stop codon (Ushijima et al., 2004), and *Prunus avium*, where the frame shift mutation is caused by a 4 bp deletion upstream from the HVa coding region of *$SFB^{4'}$* (Sonneveld et al., 2005; Ushijima et al., 2004).

The association between the self-compatibility phenotype and the loss of function of the haplotype-specific *SFB_C* gene in *Prunus armeniaca* supports previous evidence that *SFB* is the pollen-*S* gene in the GSI in *Prunus* (Entani et al., 2003; Ushijima et al., 2003; Sonneveld et al., 2005; Ushijima et al., 2004). In contrast with the *Prunus*, all well characterized PPMs in the Solanaceae result from duplications of an *S*-allele

leading to competitive interaction and the breakdown of the pollen SI response (Brewbaker and Natarajan, 1960; de Nettancourt, 2001). *S*-heteroallelic pollen is also the cause of self-fertility in some tetraploids (Stout and Chandler, 1941). The apricot is a diploid but occasional tetraploid mutants have been reported (Layne et al., 1996). Collectively, our data suggest that duplication of pollen *S*-alleles can be ruled out in the apricot PPMs studied in this work, reinforcing the conclusion that breakdown of pollen function in the apricot S_C-haplotype is due to the defective SFB_C protein.

In accordance with the original inhibitor model, defective SFBs have been suggested to lose the haplotype-specific interaction domain but retain the general interaction domain marking all S-RNases, including the cognate S-RNase, for subsequent degradation by the 26S proteasome pathway (Ushijima et al., 2004). However, Sonneveld et al. (2005) provided evidence of a self-compatible *SFB* knock-out mutant in *Prunus avium* suggesting that SFB can not have the role of general S-RNase inactivator, otherwise SFB loss of function would result in universal incompatible pollen lacking the S-RNase inhibitor mechanism. Results obtained by these authors support a specific role of the SFB proteins in the protection of the cognate S-RNase from the general inactivation mechanism in agreement with the two-components inhibitor model proposed by Luu et al. (2000). The analysis of the apricot S_C-haplotype has not provided evidence to support any of these two proposals.

A common evolutionary origin has been proposed for the Solanaceae, Rosaceae and Scrophulariaceae S-RNase based GSI systems (Igic and Kohn, 2001). Nevertheless, recent findings suggest differences at least between the pollen-*S* factor of the Rosaceae and Solanaceae. Firstly, *Prunus SFB* sequence allelic diversity is significantly higher than that found in *Antirrhinum AhSLF* or *Petunia PiSLF* genes (Zhou et al., 2003; Sijacic et al., 2004; Ikeda et al., 2004). Secondly, recent analysis of tetraploid sour cherry selections has showed that *S*-heteroallelic pollen retains its SI phenotype in *Prunus* (Hauck et al., 2006) unlike the phenomenon reported in the Solanaceae (Stout and Chandler, 1941) and Rosaceae (Crane and Lawrence, 1931) where polyploidy is a direct cause of self-compatibility as a result of competitive interaction. Finally, the absence of pollen-*S* deletions in the analysed Solanaceae PPMs prompted Golz et al. (2001) to suggest that pollen-*S* is required for pollen viability. However, several loss of pollen-*S* function self-compatible mutants have been found in *Prunus* besides those of apricot reported in this work, including a complete deletion of the *Prunus avium* SFB_3 (Sonneveld et al., 2005).

Loss of function of an *S*-locus external factor and SFB_C allele are co-responsible of self-incompatibility breakdown in 'Canino' (S_2S_C)

In the cross-pollination 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C), 'Canino' pollen tubes carrying the S_2-allele were supposed to be arrested in the styles of the self-incompatible cultivar 'Goldrich' (Egea and Burgos, 1996; Alburquerque et al., 2002). However, *S*-genotypes containing the S_2-allele derived from 'Canino' (S_1S_2 and S_2S_2) were identified in the progeny. Furthermore, *S*-genotyping of 'Canino' self-pollination progeny also showed homozygous S_2S_2 genotypes reinforcing the previous observation and suggesting that a loss of pollen-*S* function different from the mutant S_C-haplotype should be affecting 'Canino'. As described above, the loss of pollen-*S* function has been recently associated with deletions and mutations of the *SFB* genes in *Prunus avium* and *Prunus mume* self-compatible selections (Ushijima et al., 2004; Sonneveld et al., 2005). However, no mutations, insertions or deletions were found in the sequence analysis of the SFB_2 and S_2-*RNase* alleles of 'Canino' (S_2S_C), and their transcripts were detected specifically in pollen and styles, respectively.

Moreover, segregation ratios observed in the two controlled pollinations performed with 'Canino' (Table 1) fit with a model in which the pollen-part mutation is not linked to the *S*-locus. Indeed, all *S*-genotypes were found to be self-compatible by field self-pollination (Table 3) and the analysis of self-pollination progeny derived from S_2S_C and S_1S_C genotypes showed the presence of S_1 and S_2 homozygotes (Table 1). Thus, the pollen-part mutation seems to affect both S_1 and S_2 haplotypes indicating that it is inherited independently of the *S*-locus. Interestingly, the number of homozygous S_2 genotypes obtained in both controlled crosses is lower than that expected for a mutation not linked to the *S*-locus (Table 1). The average fruit set recorded from the progeny is consistent with that finding and seems to be lower in those *S*-genotypes containing the S_2-allele derived from 'Canino' (S_1S_2 and S_2S_2) (Table 3). A similar deviation was observed by Wünsch and Hormaza (2004) in the self-pollination of the sweet cherry PPM 'Cristobalina'. These authors suggest several reasons to explain this deviation such as, differences in the pollen competitive capacity to grow through the style, post-zygotic selection against homozygous embryos, or linkage between the dominant allele of the mutated factor and the *S*-locus. An additional possibility may be that the pollen-part mutation does not only break the SI rejection mechanism but also reduces pollen viability.

Pollen-part mutations can also arise from *S*-allele duplications located in a centric fragment, in a non-*S* chromosome or linked to the *S*-locus (Golz et al., 2001). In fact, *S*-genotype segregations observed in the 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C) cross and 'Canino' (S_2S_C) self-pollination might result from the loss of function of an additional factor outside *S*-locus or from duplicated pollen-*S* gene not linked in coupling with the *S*-locus. However, 'Canino' was confirmed as diploid (see above) and neither CAPS nor DNA blot analysis provided evidence of pollen-*S* genetic duplication. As discussed later, the *ParFB SFB*-like gene identified in this work was found in both self-compatible and self-incompatible cultivars, and therefore it can not be considered as a duplicated pollen-*S* gene involved in the SI breakdown. Together, these results seem to rule out competitive interaction resulting from *S*-heteroallelic pollen, hence, apart from the mutated S_C-haplotype, SI breakdown in 'Canino' could be due to a defective additional factor outside the *S*-locus.

Genetic evidence of factors required for SI besides the specificity determinants has been widely reported in the Solanaceae (McClure et al., 2000). Indeed, similarly to the case described in this work, Ai et al. (1991) showed that the self-compatible cultivar of *Petunia hybrida* 'Strawberry Daddy' (S_OS_X) accumulates a non functional *S*-allele (S_O) and a recessive mutation in an additional factor necessary for SI. Moreover, mutations in modifier loci not related to the *S*-locus affecting the function of the pollen-*S* factor have been also suggested to explain SI breakdown in *Solanum tuberosum* (Thompson et al., 1991), *Petunia axillaris* (Tsukamoto et al., 2003) and more recently in *Prunus avium* (Wünsch and Hormaza, 2004). The hypothetical 'Canino' defective factor does not seem to affect the expression of *S*-locus genes or result in sterility suggesting that it belongs to the group 2 of modifier genes required for pollen rejection but with no wider role in pollination (McClure et al., 2000). If this is the case, according to the original RNase-mediated SI inhibitor model (Thompson and Kirch 1992) this factor could be involved in an unspecific-interaction preventing ubiquitination of the S-RNases, and therefore its loss of function would facilitate SI to be overcome. However, other possibilities can not be ruled out such as the requirement of this factor for the uptake of the S-RNases into the pollen tubes. Further research will be necessary to determine the SI function affected by the pollen-part mutation in 'Canino'.

ParFB a new type of F-box genes with sequence similarity to _SFB_

S-locus pollen expressed F-box genes with sequence similarity to *SFB* (or *SLF*) have been identified in *Prunus* (Entani et al., 2003; Ushijima et al., 2003), *Anthirrinum* (Zhou et al., 2003) and *Petunia* (Wang et al., 2003). In all these cases amino acid sequence identity shared to SFB (or SLF) is less than 60% (Tsukamoto et al., 2005). The SLFL F-box proteins found in the *S*-locus genomic region of *Prunus mume* and *Prunus dulcis* (Entani et al., 2003; Ushijima et al., 2003) exhibit high sequence identity among them (> 90%) and low amino acid identities (~25%) with the *Prunus S*-locus F-box proteins (SFB). More recently, *PaF1*, an *SFB*-like gene probably not linked to the *S*-locus, has been reported in *Petunia axillaris*. *PaF1* does not show any allelic polymorphism and unexpectedly PaF1 (*Petunia axillaris* F-box protein 1) shares ~90% sequence identity with PaSLF proteins (Tsukamoto et al., 2005). The function of all these *SFB*-like genes is presently unknown.

In this work we have identified a new type of F-box gene with sequence similarity to *SFB*, *ParFB* (*Prunus armeniaca* F-box gene). Three *ParFB* variants have been cloned and sequenced from two different apricot cultivars and they show very low allelic sequence polymorphism (~99% nucleotide sequence similarity), similarly to *Prunus* SLFL genes (Entani et al., 2003; Ushijima et al., 2003). Likewise, predicted amino acid sequences of these three F-box proteins exhibit low percentage identity (~25%) with *Prunus* SLFL (Entani et al., 2003; Ushijima et al., 2003) and *S*-locus linked F-box proteins identified in *Petunia inflata* (Wang et al., 2003) but a high sequence identity with *Prunus* SFB (~62% on average). Interestingly, two insertions comprising five and two amino acids were found in the middle of the putative ParFB regions corresponding to the SFB hydrophilic V1 and V2 domains, respectively. Similarly, a gap spanning three amino acids is found in the putative region of ParFB corresponding to HVb in the *S*-locus F-box gene. The V1, V2 and HVb domains of SFB have been shown to be under positive selection suggesting that these regions may be responsible for the discrimination between self and non-self *S*-RNases (Ikeda et al., 2004). Thus, it seems that whatever is the role of ParFB proteins, functions related to these putative domains could be changed or lost with regard to SFB proteins.

The isolation of a full-length *ParFB* coding sequence from the self-incompatible cultivar 'Goldrich' (S_1S_2) allows to discard its involvement in SI breakdown through competitive interaction. Moreover, the PPM 'Canino' is supposed to be homozygous for the disrupted variant $ParFB_1$ since it has been the only variant found in this cultivar. If this is the case *ParFB* would not be the mutated additional factor outside the *S*-locus

that causes SI breakdown in 'Canino'. Otherwise, the number of individuals belonging to each *S*-genotype obtained in the cross pollination 'Goldrich' (S_1S_2) x 'Canino'(S_2S_C) would have been the same (Table 1). On the other hand, *ParFB* genes share high sequence similarity with the pollen-*S* candidate *SFB* genes and are strongly expressed in pollen tissues but, at the same time, show very low allelic diversity and gene-expression is not pollen-specific. In addition, the analysis of their deduced amino acid sequences suggests that putative domains corresponding to those of SFB proteins involved in the specific recognition of the style counterpart are altered. Therefore, although the function of the ParFB proteins remains to be investigated these results would suggest a non-specific role, if any, in GSI.

Overall, findings of this study could be summarized in two main conclusions. On one side, the association between the partial loss-of-function mutation of the SFB_C and the 'Currot' self-compatibility phenotype reinforces previous evidence suggesting *SFB* as the pollen-*S* candidate in *Prunus*. On the other, genetic evidence has been provided supporting the involvement of an *S*-locus external factor in the *S*-RNase based GSI breakdown of the self-compatible cultivar 'Canino'. In addition, the *ParFB* gene identified in this work encodes a previously uncharacterized protein that shares high amino acid sequence identity to SFB but whose function is still unknown. Further research will be necessary to reveal the precise role of ParFB. In that respect, the study of natural self-compatible mutants has been shown as a useful tool to dissect the basis of GSI mechanism in *Prunus* where transgenic approaches present clear difficulties.

MATERIALS AND METHODS

Plant material

Eight apricot cultivars, 'Goldrich' (S_1S_2), 'Pepito' (S_2S_C), Colorao (S_5S_C) (Burgos et al., 1998), 'Currot' (S_CS_C), 'Canino' (S_2S_C), 'Beliana' (S_7S_C) (Alburquerque et al., 2002), 'Ginesta' (S_CS_C) and 'Palau' (S_CS_C) (Vilanova et al., 2005), and six progenies, consisting of 8 to 171 seedlings, derived from the crosses 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C) (designated as GC-), 'Goldrich' (S_1S_2) x 'Pepito' (S_2S_C) and 'Goldrich' (S_1S_2) x 'Currot' (S_CS_C) as well as from the self-pollination of 'Canino' (S_2S_C) and the seedlings GC-8, GC-10, GC-80 and GC-86 were used in this study. All these trees are maintained at the collections of the *Instituto Valenciano de Investigaciones Agrarias (IVIA)* in

609 Valencia (Spain) and at the *Departamento de Mejora de Frutales (CEBAS-CSIC)* in
610 Murcia (Spain).
611
612 **DNA extraction**
613
614 Five grams or two discs of leaves of each selection were collected and stored at
615 -80ºC before DNA isolation. Genomic DNA was extracted from leaf samples following
616 the method of Doyle and Doyle (1987). DNA quantification was performed by
617 comparison with a lambda DNA molecular weight marker (Promega, Madison, WI).
618
619 **Cloning and sequencing of the *S_C-RNase* and *SFB_C* genes from the cultivar**
620 **'Currot' (S_CS_C)**
621
622 A fragment of the *S_C-RNase* was PCR-amplified with the primers SRc-F (5´-CTC
623 GCT TTC CTT GTT CTT GC-3´) (Romero et al., 2004) and Pru-C5 (5´-TAC CAC
624 TTC ATG TAA CAA CTG AG-3´) (Tao et al., 1999) using 'Currot' (S_CS_C) genomic
625 DNA as template. Cycling conditions were as follows: an initial denaturing step of
626 94ºC for 1 min; 7 cycles of 94ºC for 5 s and 68ºC for 3 min; 30 cycles of 94ºC for 5 s,
627 60ºC for 1 min and 68ºC for 3 min; and a final extension of 68ºC for 10 min
628 (GeneAmp®PCR System 9700, Perkin-Elmer, Fremont, CA). Similarly, an
629 amplification product containing a fragment of the *SFB_C* gene was generated with
630 primers SFBc-F (5´-TCG ACA TCC TAG TAA GAC TAC CTG C-3´) and SFBc-R
631 (5´-ATT TCT TCA CTG CCT GAA TCG-3´) (Romero et al., 2004) using also 'Currot'
632 (S_CS_C) genomic DNA as template. In this case, amplifications were carried out using a
633 temperature profile with an initial denaturing of 94ºC for 2 min; 30 cycles of 94ºC for
634 30 s, 55ºC for 60 s and 72ºC for 1 min 30 s; and a final extension of 72ºC for 10 min
635 (GeneAmp®PCR System 9700, Perkin-Elmer, Fremont, CA).
636 PCR products were electrophoresed in 0.8% (w/v) agarose gel. Molecular sizes of
637 the amplified fragments were estimated using a 100-bp ladder (Invitrogen, Carlsbad,
638 CA). To obtain the complete sequences, adjacent 5´ and 3´ ends were isolated using the
639 Universal Genome Walker Kit (Clontech, Palo Alto, CA). Finally, these fragments
640 were purified from the agarose gels using the QIAquick Gel Extraction Kit (Qiagen,
641 Hilden, Germany) and cloned into the pGEM T-Easy vector (Promega, Madison, WI).
642 DNA sequences from four independent clones were determined automatically using an

643 ABI PRISM 377 DNA Sequencer (Applied Biosystems, Foster City, CA) and the
644 BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA)
645 following the manufacturer's instructions.
646
647 **Cloning and sequencing of the *S_2-RNase* and *SFB_2* genes from the cultivars**
648 **'Canino' (S_2S_C) and 'Pepito' (S_2S_C)**
649
650 Specific primers designed from the S_2-haplotype sequence identified in *Prunus*
651 *armeniaca* (Romero et al., 2004) were used for amplification of genomic fragments
652 containing the complete *S_2-RNase* [Sf-Hap2 (5′-CGC TAG AAA TCA AAG CCA
653 CAG -3′)/Sr-Hap2 (5′-GGC GTA AGC AAG TGG AAA AG -3′)] and *SFB_2* coding
654 sequences [FBf-Hap2 (5′-GCC CAA TTA CTT GGT CAC TG-3′)/FBr-Hap2 (5′-CAC
655 CCA CTT GAC TTG TCA GC-3′)] using 'Canino' (S_2S_C) and 'Pepito'(S_2S_C) genomic
656 DNA as templates. PCR conditions and methods for isolating, cloning and sequencing
657 these bands were the same used for the S_C-*RNase*.
658
659 **Sequence analysis, alignments and homology searches**
660
661 Sequences were assembled by BioEdit software (Hall, 1999). Primers were designed
662 using the online program Primer3 (Rozen and Skaletsky, 2000). The analysis of the *S*-
663 loci structure in the S_C-haplotype was performed using the following softwares:
664 EMBOSS (Rice et al., 2000), TROLL (Tandem Repeat Occurrence Locator) (Castelo et
665 al., 2002), GeneScan (Burge and Karlin, 1997) and GenomeScan (Yeh et al., 2001).
666 Alignment of the nucleotide and amino acidic sequences were carried out with
667 CLUSTALW (Thompson et al., 1994) and CLUSTALX (Thompson et al., 1997) programs,
668 respectively. Homology searches were performed with BLASTX and BLASTN (Altschul
669 et al., 1990).
670 The following sequences were used for the alignment of Figure 1:
671 EMBL/DDJB/GenBank accession numbers: *Par*-SFB_1 (AY587563), *Par*-SFB_2
672 (AY587562), *Par*-SFB_4 (AY587565) (Romero et al., 2004), *Par*-SFB_C (DQ422946)
673 (this work), *Pa*-SFB_3 (AB096857), *Pa*-SFB_6 (AB096858) (Yamane et al., 2003), *Pm*-
674 SFB_1 (AB092621), *Pm*-SFB_7 (AB092622) (Entani et al., 2003), *Pd*-SFBa (AB092966)
675 and *Pd*-SFBb (AB092967) (Ushijima et al., 2003).
676

RT-PCR

Total RNA was extracted from leaves, pollen grains and styles of the cultivars 'Currot' (S_CS_C), 'Canino' (S_2S_C), 'Pepito' (S_2S_C) and 'Goldrich' (S_1S_2) following the method of Salzman et al. (1999) with modifications. All RNAs were treated with DNase I (Roche Diagnostics, Basel, Switzerland). The cDNAs were synthesized by ThermoScript™ RT-PCR System (Invitrogen, Carlsbad, CA) with oligo-d(T) primer. The obtained cDNAs were used as a template for PCRs with gene-specific primer sets RSc-F (5′-GTG TTT CAT TAT GAG CAC TAG ATC-3′) and RSc-R (5′-TTA ATG TCA ACG TTA TTC CAG C-3′) for *S*$_C$*-RNase*, RS2-F (5′-ATG AGC ACT GGT GAT GGA AC-3′) and RS2-R (5′-ACG GAG TGC AGG ATC AGT TC-3′) for *S*$_2$*-RNase*, RFBc-F (5′-GAG GAG TGC TAC AAA CTA AGC-3′) and RFBc-R (5′-ACC CCT ATG ATG TTC CAA AG-3′) for *SFB*$_C$, RFB2-F (5′-TGA ACG TCA GAA CGA CAC TG-3′) and RFB2-R (5′-ACC CTT ATA ATG CTG CCA AG-3′) for *SFB*$_2$, FB-F (5′-TAC GAA AAC TAC GAG GAC TAC-3′) and FB-R (5′-AAG CAT CAT CTT TGT GGA CG-3′) for *ParFB*s, and UBI-F (5′-CTC CTC TGA CAC CAT CGA CAA-3′) and UBI-R (5′-CAT AGG TCA ACC CAC ACT TG-3′) for the ubiquitin gene (Entani et al., 2003). PCRs were performed as described by Romero et al. (2004).

Genomic PCRs for *S*-allele typing and *SFB*$_C$ specific amplification

S-allele typing of populations and cultivars was performed by PCR-amplification of the *S-RNase* first intron with the primer pair SRc-F (5′-CTC GCT TTC CTT GTT CTT GC-3′) (Romero et al., 2004) and SRc-R (5′-GGC CAT TGT TGC ACA AAT TG-3′) (Vilanova et al., 2005), following the protocol described by Vilanova et al. (2005).

SFB$_C$ specific PCR-amplification from genomic DNA was performed with the primer pair RFBc-F/RFBc-R. Two additional primers, RFBc-F and SFBins-R (5′-TCA AGA ACT TGG TTG GAT TCG-3′), designed from the consensus sequence of the *Prunus SFB* alleles (Romero et al., 2004) were used to amplify the *SFB*$_C$ insertion from genomic DNA of several apricot cultivars.

Pollination and pollen-tube growth test

One hundred and fifteen trees derived from the cross 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C) were tested for self-compatibility by self-pollination in the field. Before anthesis, insect-proof bags were put over several branches, containing approximately 200 to 250 flower buds in total per tree, to prevent cross-pollination. Subsequent fruit set was recorded and fruits collected about three months later. Following the same procedure 'Canino' was self-pollinated to obtain a population of study.

Pollen-tube growth tests were performed for self-pollinations of 'Goldrich' (S_1S_2) and two seedlings derived from the cross 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C), GC-35 (S_1S_2) and GC-147 (S_2S_2). Moreover, cross-pollinations between 'Goldrich', as female parent, and GC-35 or GC-147, as male parents, were also evaluated. Branches with about 30 flowers in balloon stage, for each combination, were placed in beakers with a 5% (w/v) sucrose solution and the flowers were emasculated immediately. Branches were maintained in a chamber at a controlled temperature of 24ºC. Pistils were self- or cross-pollinated 24 h later and, 72 h after pollination, pistils were harvested in a 5% formaldehyde fixing solution at 40%, 5% acetic acid, and 90% ethanol at 70% (FAA). Pistils were washed in distilled water three times, for 1 h each, to remove the FAA and then placed in the autoclave for 30 min at 100 KP in 5% (w/v) sodium sulphite solution to soften the pistils. The pistils were then stained for 24 h with 0.1% (v/v) aniline blue in 0.1N potassium phosphate, pH 8.7. The epidermis was removed and the pistils squashed for observation with an Olympus BH2 microscope (Olympus, Tokyo, Japan) with a BH2-RFL-T2 ultraviolet light source, using an Osram HBO 100 W/2 high-pressure mercury lamp (Osram GmbH, Berlin-Munich, Germany).

Ploidy level determination

Ploidy level was determined using the *Partec CyStain UV precise P* reagent kit for nuclei extraction and DNA staining of nuclear DNA from plant tissues. Approximately 0.5 cm^2 leaf tissue was chopped using a sharp razor blade in 400 µl extraction buffer and filtered through a *Partec 50 µm CellTrics* disposable filter. Samples were then incubated for 60 seconds in the staining solution and analyzed in the *Partec* flow cytometer Ploidy Analyzer PA (Partec, Münster, Germany) in the blue fluorescence channel.

Genomic DNA blot analysis

744

745 Genomic DNA was digested with two restriction enzymes (*Eco*RI and *Hind*III)
746 (Roche Diagnostics, Basel, Switzerland), electrophoresed on 0.8% agarose gels and
747 blotted onto Hybond N nylon membranes (Amersham Biosciences, Piscataway, NJ).
748 Blots were performed using approximately 3 μg of DNA per track. An SFB_1 alkali-
749 labile digoxygenin DIG-labeled (306 bp long) probe was synthesized by PCR from the
750 apricot BAC DNA clone 102/D9 (Vilanova et al., 2003) using the primer pair SFB-5F
751 (5'-TAG GAC CCC TCC AAT GAG C-3') and SFBc-R (5´-ATT TCT TCA CTG CCT
752 GAA TCG-3´). The SFB_1 probe corresponds to a *Prunus SFB* conserved region located
753 between V1 and V2 variable domains. Filters were hybridized overnight with the DIG-
754 labeled probe at 65ºC and washed at low-stringency conditions (twice at room
755 temperature in 2xSSC, 0.1%SDS for 5min., and twice at 68ºC in 0.5xSSC, 0.1%SDS for
756 15 min.). Hybridization signals were detected following the manufacturer's
757 recommendations (Roche Diagnostics, Basel, Switzerland) using Kodak XAR films.
758 Filters were then washed for removing the probe and re-hybridized following the same
759 procedure with a $ParFB_1$ 707 bp long DIG-labeled probe synthesized by PCR from the
760 apricot DNA clone C14 using the primer pair SRc-F (5´-CTC GCT TTC CTT GTT
761 CTT GC-3´) and SFBc-R (5´-ATT TCT TCA CTG CCT GAA TCG-3´). Detection was
762 performed as described above.
763

764 **CAPS analysis**
765

766 *SFB* coding sequence fragments were non-specifically PCR-amplified with the
767 consensus primers SFBc-F and SFBc-R (Romero et al., 2004) using 'Canino' (S_2S_C) and
768 GC-5 (S_2S_2) genomic DNA as templates. PCR-products were cloned and subsequently
769 digested. PCR conditions and methods for isolating and cloning these fragments into the
770 pGEM T-Easy vector (Promega, Madison, WI) were the same described above for
771 SFB_C. Clones were digested with the restriction enzymes *Eco*RI-*Nde*I, *Eco*RI-*Kpn*I and
772 *Mva*I (Roche Diagnostics, Basel, Switzerland). Restriction products were
773 electrophoresed on 2% agarose gels, and stained with ethidium bromide.
774

775 **Cloning and sequencing of the *ParFB* genes from 'Goldrich' (S_1S_2) and 'Canino'**
776 **(S_2S_C)**
777

778 *ParFB* fragments amplified from 'Canino' and Goldrich' using the primers FB-F and
779 SFBc-R (Romero et al., 2004) were cloned into the pGEM T-Easy vector (Promega,
780 Madison, WI) as described above. Twelve clones from each cultivar were sequenced
781 falling into three variants ($ParFB_1$ from 'Canino'; $ParFB_2$ and $ParFB_3$ from
782 'Goldrich'). PCR conditions and methods for isolating, cloning and sequencing these
783 fragments and the adjacent 5′ and 3′ ends were the same used for SFB_C.
784
785 **Accession numbers**
786
787 Sequence data from this article can be found in the EMBL/GenBank data libraries
788 under accession numbers DQ422943, DQ422944, DQ422945, DQ422946 and
789 DQ422947.
790
791 **Supplemental data**
792
793 Supplemental Figure 1. Segregation analysis of *S*-alleles of the 'Goldrich' x 'Currot'
794 progeny.
795 Supplemental Figure 2. Alignment of the amino acidic sequences of S-RNases from
796 several *Prunus* species carried out with CLUSTALX (Thompson et al., 1997). Sequences
797 used were the following: EMBL/DDJB/GenBank accession numbers: *Par*-S_1
798 (AY587561), *Par*-S_2 (AY587562), *Par*-S_4 (AY587564), *Par*-S_C (DQ422947), *Pa*-S_3
799 (AB010306), *Pa*-S_6 (AB010305), *Pm*-S_1 (AB101438), *Pm*-S_7 (AB101439), *Pd*-Sa
800 (AB026836) and *Pd*-Sb (AB011469).
801 Supplemental Figure 3. Search of *SFB* genetic duplications in the cultivar 'Canino'
802 (S_2S_C) by CAPS analysis with *Eco*RI-*Nde*I.
803 Supplemental Figure 4. Molecular characterization of *ParFB* alleles. Sequences used
804 for the alignments were: *Par*-SFB_1 (AY587563), *Par*-SFB_2 (AY587562), *Par*-SFB_4
805 (AY587565) (Romero et al., 2004), *Par*-FB_1 (DQ422943), *Par*-FB_2 (DQ422944) and
806 *Par*-FB_3 (DQ422946) (this work).
807
808 **ACKNOWLEDGEMENTS**
809
810 Authors thank Dr. Wim Deleu for his helpful revision of this manuscript, to José
811 Juarez for his help in the ploidy level determination and to all the members of the Fruit

812 Trees lab for their contributions. All experiments described in this paper comply with
813 the current laws of Spain.
814
815 **LITERATURE CITED**
816

817 **Ai Y, Kron E, Kao T-H** (1991) S-alleles are retained and expressed in a self-
818 compatible cultivar of *Petunia hybrida*. Mol Gen Genet **230**: 353-358
819 **Alburquerque N, Egea J, Pérez-Tornero O, Burgos L** (2002) Genotyping apricot
820 cultivars for self-(in)compatibility by means of RNases associated with *S* alleles.
821 Plant Breed **121**: 343-347
822 **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ** (1990) Basic local
823 alignment search tool. J Mol Biol **215**: 403-410
824 **Brewbaker JL, Natarajan AT** (1960) Centric fragments and pollen part mutation of
825 incompatibility alleles in *Petunia*. Genetics **45**: 699-704
826 **Boskovic R, Tobutt KR** (1996) Correlation of stylar ribonuclease zymograms with
827 incompatibility alleles in sweet cherry. Euphytica **90**: 245-250
828 **Burge C, Karlin S** (1997) Prediction of complete gene structures in human genomic
829 DNA. J Mol Biol **268**: 78-94
830 **Burgos L, Berenguer T, Egea J** (1993) Self- and cross-compatibility among apricot
831 cultivars. Hortsci **28**(2): 148-150
832 **Burgos L, Pérez-Tornero O, Ballester J, Olmos E** (1998) Detection and inheritance
833 of stylar ribonucleases associated with incompatibility alleles in apricot. Sex. Plant
834 Reprod **11**: 153-158
835 **Castelo AT, Martins W, Gao GR** (2002) TROLL-Tandem Repeat Occurrence Locator
836 Bioinformatics **18**(4): 634-636
837 **Crane MB, Lawrence, WJC** (1931) Sterility and incompatibility in diploid and
838 polyploid fruits. J. Genet. **24**: 97-107
839 **De Nettancourt D** (2001) Incompatibility and incongruity in wild and cultivated
840 plants. Springer, Berlin, pp 1-215
841 **Deshaies RJ** (1999) SCF and Cullin/Ring H2-based ubiquitin ligases. Annu Rev Cell
842 Dev Biol **15**: 435-467
843 **Doyle JJ, Doyle JL** (1987) A rapid isolation procedure for small quantities of fresh leaf
844 tissue. Phytochem. Bull **19**: 11-15

Egea J, Burgos L (1996) Detecting cross-incompatibility of three North American apricot cultivars and establishing the first incompatibility group in apricot. J Am Soc Hort Sci **121**: 1002-1005

Entani T, Iwano M, Shiba H, Che FS, Isogai A, Takayama S (2003) Comparative analysis of the self-incompatibility (S-) locus region of *Prunus mume*: identification of a pollen-expressed F-box gene with allelic diversity. Genes Cells **8**: 203-213

Golz JF, Su V, Clarke AE, Newbigin E (1999) A molecular description of mutations affecting the pollen counterpart of the *Nicotiana alata S* locus. Genetics **152**: 1123-1135

Golz JF, Oh HY, Su V, Kusaba M, Newbigin E (2001) Genetic analysis of *Nicotiana* pollen-part mutants is consistent with the presence of an *S*-ribonuclease inhibitor at the S locus. Proc Natl Acad Sci USA **98**: 15372-15376

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser **41**: 95-98

Hauck NR, Yamane H, Tao R, Iezzoni AF (2006) Accumulation of non-functional S-haplotypes results in the breakdown of gametophytic self-incompatibility in tetraploid Prunus. Genetics (in press)

Huang S, Lee HS, Karunanandaa B, Kao TH (1994) Ribonuclease activity of *Petunia inflata* S proteins is essential for rejection of self-pollen. Plant Cell **6**: 1021-1028.

Igic B, Kohn JR (2001) Evolutionary relationships among self-incompatibility RNases. Proc Natl Acad Sci USA **98**: 13167-13171

Ikeda K, Igic B, Ushijima K, Yamane H, Hauck NR, Nakano R, Sassa H, Iezzoni AF, Kohn JR, Tao R (2004) Primary structural features of the S haplotype-specific F-box protein, SFB, in *Prunus*. Sex Plant Reprod **16**: 235-243

Ishimizu T, Shinkawa T, Sakiyama F, Norioka S (1998) Primary structural features of rosaceous S-RNases associated with gametophytic self-incompatibility. Plant Mol Biol **37**: 931-941

Kawata Y, Sakiyama F, Hayashi F, Kyogoku Y (1989) Identification of two essential histidine residues of ribonuclease T_2 from *Aspergillus oryzae*. Eur J Biochem **187**: 255-262

Lai Z, Ma W, Han B, Liang L, Zhang Y, Hong G, Xue Y (2002) An F-box gene linked to the self-incompatibility (*S*) locus of *Antirrhinum* is expressed specifically in pollen and tapetum. Plant Mol Biol **50**: 29-42

879 **Layne REC, Bailey CH, Hough LF** (1996) Apricots. *In* J Janick, JN Moore, eds, Fruit
880 Breeding: Tree and Tropical Fruits, Vol II, John Wiley & Sons, New York, pp 89
881 **Luu DT, Qin XK, Morse D, Cappadocia M** (2000) S-RNase uptake by compatible
882 pollen tubes in gametophytic self-incompatibility. Nature **407**: 649-651
883 **McClure BA, Cruz-García F, Beecher B, Sulaman W** (2000) Factors affecting inter-
884 and intra-specific pollen rejection in *Nicotiana*. Ann Bot **85**: 113-123
885 **McClure BA, Haring V, Ebert PR, Anderson MA, Simpson RJ, Sakiyama F,**
886 **Clarke AE** (1989) Style self-incompatibility gene products of *Nicotiana alata* are
887 ribonucleases. Nature **342**: 955-957
888 **Qiao, H, Wang H, Zhao L, Zhou J, Huang J, Zhang Y, Xue Y** (2004) The F-box
889 protein AhSLF-S_2 physically interacts with S-RNases that may be inhibited by the
890 ubiquitin/26S proteasome pathway of protein degradation during compatible
891 pollination in *Antirrhinum*. Plant Cell **16**: 571-581
892 **Rice P, Longden I, Bleasby A** (2000) EMBOSS: The European Molecular Biology
893 Open Software Suite. Trends Genet **16**(6): 276-277
894 **Roalson EH, McCubbin AG** (2003) S-RNases and sexual incompatibility: structure,
895 functions, and evolutionary perspectives. Mol Phylogenet Evol **29**: 490-506
896 **Romero C, Vilanova S, Burgos L, Martínez-Calvo J, Vicente M, Llácer G, Badenes**
897 **ML** (2004) Analysis of the *S*-locus structure in *Prunus armeniaca* L. Identification
898 of *S*-haplotype specific *S-RNase* and *F-box* genes. Plant Mol Biol **56**: 145-157
899 **Royo J, Kuntz C, Kowyama Y, Anderson M, Clarke AE** (1994). Loss of a histidine
900 residue at the active site of *S*-locus ribonuclease is associated with self-compatibility
901 in *Lycopersicon peruvianum*. Proc Natl Acad Sci USA **91**: 6511-6514
902 **Rozen S, Skaletsky HJ** (2000) Primer3 on the WWW for general users and for
903 biologist programmers. *In* S Krawetz, S Misener, eds, Bioinformatics Methods and
904 Protocols: Methods in Molecular Biology, Humana Press, Totowa, NJ, pp 365-386
905 **Salzman RA, Fujita T, Zhu-Salzman K, Hasegawa PM, Bressan RA** (1999) An
906 improved RNA isolation method for plant tissues containing high levels of phenolic
907 compounds or carbohydrates. Plant Mol Biol Reptr **17**: 11-17
908 **Sassa H, Hirano H, Nishio T, Koba T** (1997) Style-specific self-compatible mutation
909 caused by deletion of the S-RNase gene in Japanese pear (*Pyrus serotina*). Plant J
910 **12**: 223-227

Sijacic P, Wang X, Skirpan AL, Wang Y, Dowd PE, McCubbin AG, Huang S, Kao T-h (2004) Identification of the pollen determinant of S-RNase-mediated self-incompatibility. Nature **429**: 302-305

Sonneveld T, Tobbutt KR, Vaughan SP, Robbins TP (2005) Loss of pollen-*S* function in two self-compatible selections of *Prunus avium* is associated with deletion/mutation of an *S* haplotype-specific F-box gene. Plant Cell **17**: 37-51

Stout AB, Chandler C (1941) Change from self-incompatible to self-compatible accompanying change from diploid to tetraploidy. Science **94**: 118

Tao R, Yamane H, Sugiura A (1999) Molecular typing of S-alleles through identification, characterization and cDNA cloning for S-RNases in sweet cherry. J Am Soc Hort Sci **124**(3): 224-233

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res **25**: 4876-4882

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res **22**: 4673-4680

Thompson RD, Kirch HH (1992) The *S* locus of flowering plants: When self-rejection is self-interest. Trends Genet **8**: 381-387

Thompson RD, Uhrig H, Hermsen JGT, Salamini F, Kaufmann H (1991) Investigation of a self-compatible mutation in *Solanum tuberosum* clones inhibiting S-allele activity in pollen differentially. Mol Gen Genom **226**: 283-288

Tsukamoto T, Ando T, Takahashi K, Omori T, Watanabe H, Kokubun H, Marchesi E, Kao T-h (2003) Breakdown of self-incompatibility in a natural population of *Petunia axillaris* caused by loss of pollen function. Plant Physiol **131**: 1903-1912

Tsukamoto T, Ando T, Watanabe H, Marchesi E, Kao T (2005). Duplication of the *S*-locus F-box gene is associated with breakdown of pollen function in an *S*-haplotype identified in a natural population of self-incompatible *Petunia axillaris*. Plant Mol Biol **57**: 141-153

Ushijima K, Sassa H, Dandekar AM, Gradziel TM, Tao R (2003) Structural and transcriptional analysis of the self-incompatibility locus of almond: Identification of

944 a pollen-expressed F-box gene with haplotype-specific polymorphism. Plant Cell **15**:
945 771-781
946 **Ushijima K, Yamane H, Watari A, Kakehi E, Ikeda K, Hauck NR, Iezzoni AF, Tao**
947 **R** (2004) The *S* haplotype-specific F-box protein gene, *SFB*, is defective in self-
948 compatible haplotypes of *Prunus avium* and *P. mume*. Plant J **39**: 573-586
949 **Vaughan SP, Russell K, Sargent DJ, Tobutt KR** (2006) Isolation of *S*-locus F-box
950 alleles in *Prunus avium* and their application in a novel method to determine self-
951 incompatibility genotype. Theor Appl Genet (in press)
952 **Vilanova S, Romero C, Abernathy D, Abbott AG, Burgos L, Llácer G, Badenes ML**
953 (2003) Construction and application of a bacterial artificial chromosome (BAC)
954 library of *Prunus armeniaca* L. for the identification of clones linked to the self-
955 incompatibility locus. Mol Gen Genom **269**: 685-691
956 **Vilanova S, Romero C, Burgos L, Llácer G, Badenes ML** (2005) Identification of
957 Self-(in)compatibility Alleles in Apricot (*Prunus armeniaca* L.) by PCR and
958 Sequence Analysis. J Am Soc Hort Sci **130**(6): 893-898
959 **Wang Y, Wang X, McCubbin AG, Kao T-h** (2003) Genetic mapping and molecular
960 characterization of the self-incompatibility (*S*) locus in *Petunia inflata*. Plant Mol
961 Biol **53**: 565-580
962 **Watkins R** (1976) Cherry, plum, peach, apricot and almond. *In* NW Simmonds, ed,
963 Evolution of crop plants, Longman, London-New York, pp 242-247
964 **Wünsch A, Hormaza JI** (2004) Genetic and molecular analysis in Cristobalina sweet
965 cherry, a spontaneous self-compatible mutant. Sex. Plant Reprod **17**: 203-210
966 **Xue Y, Carpenter R, Dickinson HG, Coen ES** (1996) Origin of allelic diversity in
967 *Antirrhinum S* locus RNases. Plant Cell **8**: 805-814
968 **Yamane H, Ikeda K, Ushijima K, Sassa H, Tao R** (2003) A pollen-expressed gene
969 for a novel protein with an F-box motif that is very tightly linked to a gene for S-
970 RNase in two species of cherry, *Prunus cerasus* and *P. avium*. Plant Cell Physiol **44**:
971 764-769
972 **Yeh RF, Lim LP, Burge CB** (2001) Computational inference of homologous gene
973 structures in the human genome. Genome Res **11**: 803-816
974 **Zhou J, Wang F, Ma W, Zhang Y, Han B, Xue Y** (2003) Structural and
975 transcriptional analysis of *S*-locus F-box genes in *Antirrhinum*. Sex. Plant Reprod
976 **16**: 165-177
977

978 **FIGURE LEGENDS**

979

980 **Figure 1.-** Genetic analysis of the apricot S_C-haplotype. **A)** Genomic structure of the
981 apricot *S*-locus region in the S_C-haplotype. Grey boxes represent ORFs for the *S-RNase*
982 and *SFB* alleles. The physical distance between both genes and their relative positions
983 have not been established yet. Direct and inverted repeats and microsatellites identified
984 in the genomic sequences are also indicated. The broken line represents the inserted
985 fragment, black boxes the ~52 bp inverted repeats and lined box the sequence
986 corresponding to the OriSFB_C below the insertion. **B)** Nucleotide and the deduced
987 amino acid sequence of the SFB_C allele. The first nucleotide of the translational
988 initiation codon is indicated (+1). The nucleotide sequence indicated by lower case
989 correspond to the 5' flanking region of SFB_C. The broken line is underlining the
990 putative intron within the 5´ UTR found in Japanese apricot (Ushijima et al., 2004) and
991 sweet cherry (Vaughan et al., 2006). A 358 bp DNA fragment inserted at position 904
992 to 1261 produces a frame shift leading to a premature stop codon in translation. Inverted
993 repeats located at both ends of the insertion are underlined. The nucleotide sequence
994 below the stop codon is in italics. Residues highlighted in grey correspond to the
995 original intact OriSFB_C. *Arrows* indicate positions and directions of primers used in
996 Supplemental Figure 1B, Figures 1D and 3B (RFBc-F/RFBc-R) and Figure 2 (RFBc-
997 F/SFBins-R). **C)** Alignment of the deduced amino acid sequences of SFBs from apricot
998 *Par* (*Prunus armeniaca*), sweet cherry *Pa* (*P. avium*), Japanese apricot *Pm* (*P. mume*)
999 and almond *Pd* (*P. dulcis*). *Asterisks* indicate conserved sites, *dots* conservative
1000 substitutions and *dashes* gaps. Residues different from the consensus are *highlighted* in
1001 *black boxes*. F-box and (Hyper)variable (V1, V2, HVa and HVb) regions (Ikeda et al.,
1002 2004) are *boxed*. **D)** RT-PCR analysis of the S_C-*RNase* and SFB_C gene expression in
1003 pollen (P), styles (S) and leaves (L) of the apricot cultivars 'Currot' (S_CS_C) and 'Canino'
1004 (S_2S_C). *Ubiquitin* (*Ubi* - positive control) genes.

1005

1006 **Figure 2.-** Genomic PCR-amplification of several apricot cultivars for the SFB_C
1007 insertion. The primers used (RFBc-F/SFBins-R) were designed from the consensus
1008 sequence of the *Prunus SFB* alleles (Romero et al., 2004). Samples are as follows: (Cu)
1009 'Currot' (S_CS_C), (Gi) 'Ginesta' (S_CS_C), (P) 'Palau' (S_CS_C), (Pe) 'Pepito' (S_2S_C), (Ca)
1010 'Canino' (S_2S_C), (B) 'Beliana' (S_7S_C), (Co) 'Colorao' (S_5S_C) and (G) 'Goldrich' (S_1S_2).

1011

Figure 3.- Segregation analysis of *S*-alleles of the 'Goldrich' x 'Canino' and 'Goldrich' x 'Pepito' progenies. **A**) PCR-amplification of apricot genomic DNA with consensus primers (SRc-F/SRc-R) for the first *S-RNase* intron. **B**) SFB_C specific PCR-amplification performed with RFBc-F/RFBc-R primers. Samples in (A) and (B) are as follows: (G) 'Goldrich' (S_1S_2); (Ca) 'Canino' (S_2S_C); 16 seedlings derived from the 'Goldrich' x 'Canino' cross. **C**) *S-RNase* allele fragments PCR-amplified with SRc-F/SRc-R primers in the 'Goldrich' x 'Pepito' progeny. Samples are as follows: (G) 'Goldrich' (S_1S_2); (P) 'Pepito' (S_2S_C); 8 seedlings derived from the 'Goldrich' x 'Pepito' cross.

Figure 4.- Pollen-tube growth after controlled-pollination in two apricot selections. Photographs show the style of pistils. **A)** Self-pollination in 'Goldrich' (S_1S_2) **B)** Self-pollination in seedling GC-147 (S_2S_2). **C)** Cross-pollination between 'Goldrich' and seedling GC-147 (S_2S_2). Pollen-tubes of self-incompatible cultivar 'Goldrich' are arrested at the upper part of the pistil (A) but those of the two compatible pollinations complete their development in the pistil and reach the ovary tissue (B and C).

Figure 5.- RT-PCR analysis of the S_2-*RNase* and SFB_2 gene expression in pollen (P), styles (S) and leaves (L) of the apricot cultivars 'Goldrich' (S_1S_2), 'Canino' (S_2S_C) and 'Pepito' (S_2S_C). *Ubiquitin* (*Ubi* - positive control) genes.

Figure 6.- DNA blot analysis of apricot self-compatible and incompatible cultivars. **(A, B)** Blots of genomic DNA hybridized with an SFB_1 probe. **(C, D)** Blots of genomic DNA re-hybridized with a $ParFB_1$ probe. Samples were Cu, 'Currot' (S_CS_C), G, 'Goldrich' (S_1S_2), Ca, 'Canino' (S_2S_C), GC-1, seedling GC-1 (S_1S_C), GC-13, seedling GC-13 (S_2S_C), GC-6, seedling GC-6 (S_1S_2), GC-5, seedling GC-5 (S_2S_2), and 'Pepito' (S_2S_C). DNAs were digested with two restriction enzymes *Eco*RI (A and C) and *Eco*RI-*Hind*III (B and D).

Table 1.- Segregation of the *S-RNase* alleles in the progenies of controlled field crosses and self-pollinations. *S*-genotypes were determined by PCR. Observed *S-RNase* genotypes, expected segregation ratios and chi-square values (χ^2) obtained for each population are indicated.

Seed parent (*S*-genotype)	**Pollen parent (*S*-genotype)**	***S*-genotypes observed**						**Total**	**Exp. ratio**	**χ^{2d} (*P*-value)**
		S_1S_C	S_2S_C	S_1S_2	S_2S_2	S_1S_1	S_CS_C			
Goldrich (S_1S_2)	Currot (S_CS_C)	31	39	--	--	--	--	70	1:1	0.91 (0.339)
Goldrich (S_1S_2)	Canino (S_2S_C)	66	55	28	22	--	--	171	2:2:1:1[b]	2.98 (0.394)
Canino (S_2S_C)	Canino (S_2S_C)	--	53	--	11	--	35	99	3:1:2[b]	2.20 (0.333)
GC-8 (S_2S_C)[a]	GC-8 (S_2S_C)	--	14	--	4	--	6	24	3:1:2[b]	0.83 (0.659)
GC-10 (S_2S_C)[a]	GC-10 (S_2S_C)	--	14	--	--	--	10	24	1:1[c]	0.67 (0.414)
GC-80 (S_1S_C)[a]	GC-80 (S_1S_C)	15	--	--	--	3	6	24	3:1:2[b]	1.50 (0.472)
GC-86 (S_1S_C)[a]	GC-86 (S_1S_C)	13	--	--	--	--	11	24	1:1[c]	0.17 (0.683)

[a] Seedlings derived from the cross 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C)

[b] Expected ratios for a single mutation unlinked to the *S*-locus

[c] Expected ratios for non-mutated GC-seedlings

[d] Observed ratios do not differ significantly from expected at $P < 0.05$ in any case

Table 2.- Amino acid sequence identities (%) among *Prunus S*-locus genes. The upper half shows identities between the *S-RNases* and the lower half between the *SFB*s.

	Prunus avium		***Prunus dulcis***		***Prunus mume***		***Prunus armeniaca***			
	Pa-S3	***Pa-S6***	***Pd-Sa***	***Pd-Sb***	***Pm-S1***	***Pm-S7***	***Par-S1***	***Par-S2***	***Par-S4***	***Par-Sc***
Pa-S3	---	72.8	57.7	73.8	72.2	67.3	73.1	75.6	74.6	74.1
Pa-S6	80.0	---	55.8	80.0	74.3	74.4	77.4	71.8	71.6	77.9
Pd-Sa	67.2	68.3	---	57.7	58.3	55.9	59.1	54.9	57.9	59.7
Pd-Sb	77.1	78.7	67.3	---	72.5	71.7	81.8	73.8	75.1	80.0
Pm-S1	82.1	81.1	69.3	80.6	---	75.7	75.0	68.8	70.9	75.8
Pm-S7	77.7	79.6	69.5	78.6	80.6	---	72.1	69.0	67.7	75.5
Par-S1	77.1	80.6	67.6	75.3	81.1	78.6	---	77.5	73.7	77.3
Par-S2	78.1	79.7	67.2	79.3	82.1	79.3	79.8	---	73.3	75.9
Par-S4	79.8	80.6	70.8	79.6	83.0	82.5	80.1	79.0	---	75.5
Par-Sc	77.1	78.9	65.1	78.7	78.9	78.0	77.7	77.1	78.5	---

Table 3.- Evaluation of self-compatibility in seedlings derived from the cross 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C) by recording fruit set after field self-pollination.

***S*-genotype**	**No of seedlings evaluated**	**No and (%) of seedlings with fruit set > 2%**	**No and (%) of seedlings with fruit set > 5%**	**Average fruit set (%)**
S_1S_C	47	46 (98%)	42 (89%)	19.2
S_2S_C	38	36 (95%)	34 (89%)	18.0
S_1S_2	16	15 (94%)	12 (75%)	11.7
S_2S_2	14	13 (93%)	10 (71%)	8.3
Total	115	110	98	

Table 4.- Expected gamete and seedling genotypes formed from the outcross 'Goldrich' (S_1S_2) x 'Canino' (S_2S_C) and the selfing of 'Canino' (S_2S_C) considering 'Canino' heterozygous for a pollen-part mutation unlinked to the *S*-locus (*Mm*)

Female 'Goldrich' (S_1S_2 *MM*) /**Male** 'Canino' (S_2S_C *Mm*)	$\boldsymbol{S_2M}^b$	$\boldsymbol{S_2m}$	$\boldsymbol{S_CM}$	$\boldsymbol{S_Cm}^b$
$\boldsymbol{S_1M}$	**X**[a]	S_1S_2 *Mm*	S_1S_C *MM*	S_1S_C *Mm*
$\boldsymbol{S_2M}$	**X**	S_2S_2 *Mm*	S_2S_C *MM*	S_2S_C *Mm*
Female/Male 'Canino' (S_2S_C *Mm*)	$\boldsymbol{S_2M}^b$	$\boldsymbol{S_2m}$	$\boldsymbol{S_CM}$	$\boldsymbol{S_Cm}^b$
$\boldsymbol{S_2M}$	**X**	S_2S_2 *Mm*	S_2S_C *MM*	S_2S_C *Mm*
$\boldsymbol{S_2m}$	**X**	S_2S_2 *mm*	S_2S_C *Mm*	S_2S_C *mm*
$\boldsymbol{S_CM}$	**X**	S_2S_C *Mm*	S_CS_C *MM*	S_CS_C *Mm*
$\boldsymbol{S_Cm}$	**X**	S_2S_C *mm*	S_CS_C *Mm*	S_CS_C *mm*

[a] Pollen incompatibility

[b] If the pollen-part mutation was linked in coupling with S_2 the S_2M and S_Cm gametes from 'Canino' would not be formed

Figure 1

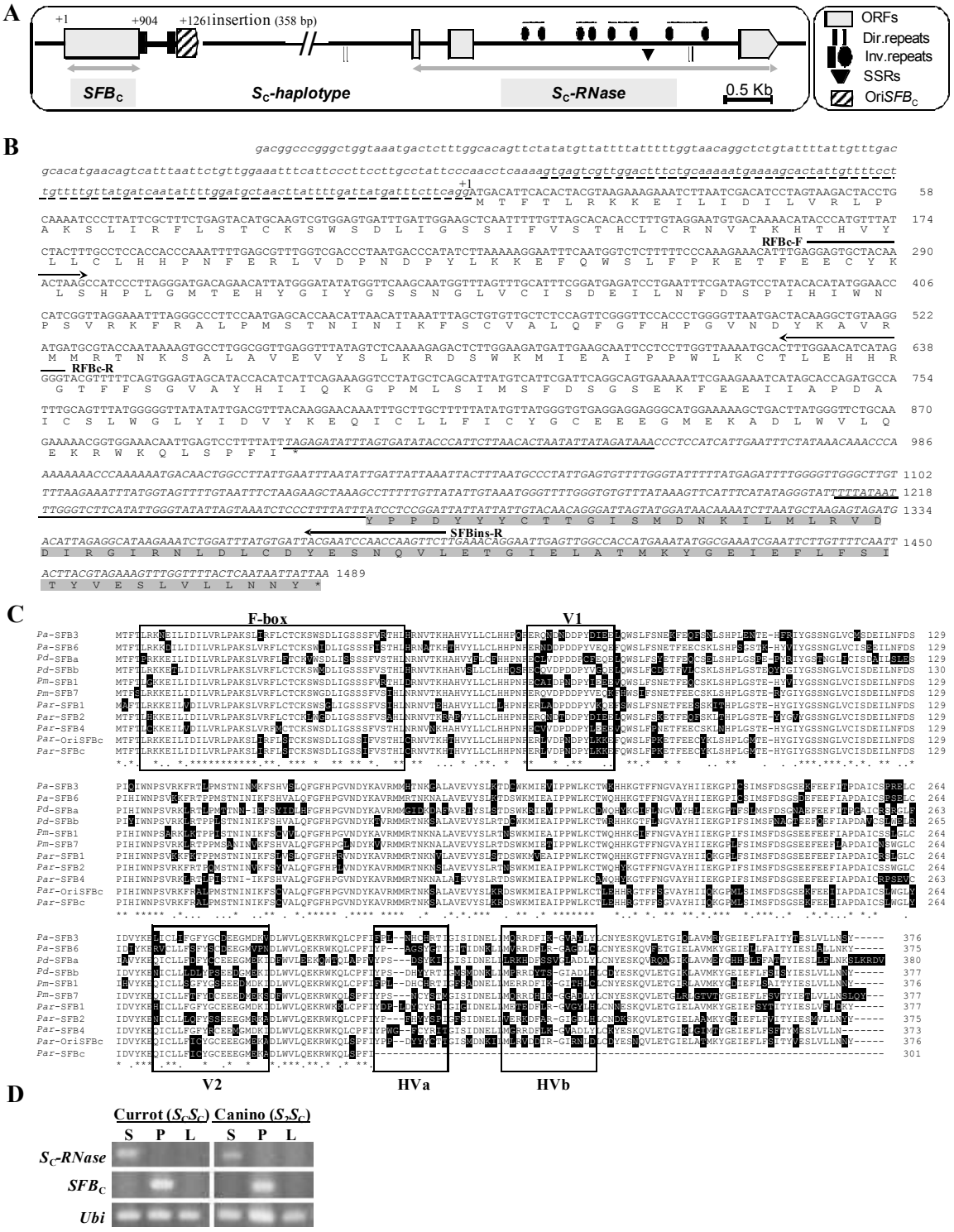

A
+1
+904
+1261insertion (358 bp)
SFBc
Sc-haplotype
Sc-RNase
0.5 Kb
ORFs
Dir.repeats
Inv.repeats
SSRs
OriSFBc
B
RFBc-F
RFBc-R
SFBins-R
C
F-box
V1
V2
HVa
HVb
D
Currot (Sc-Sc)
Canino (Sc-Sc)
S P L S P L
Sc-RNase
SFBc
Ubi

Figure 2

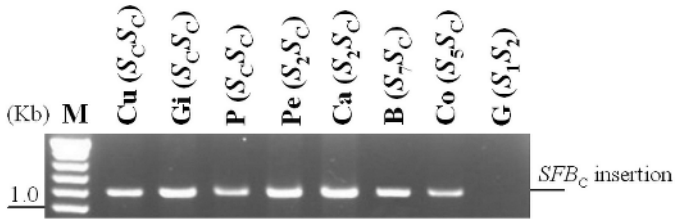


Figure 3

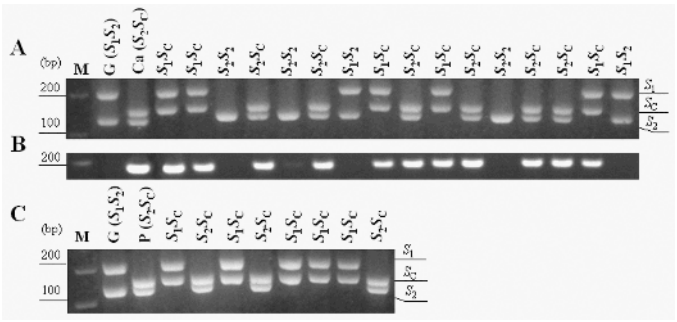


Figure 4

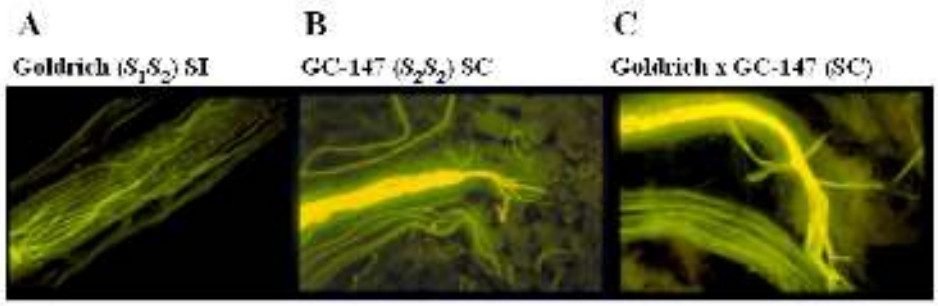


Figure 5

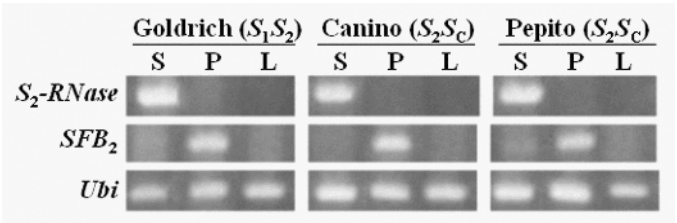
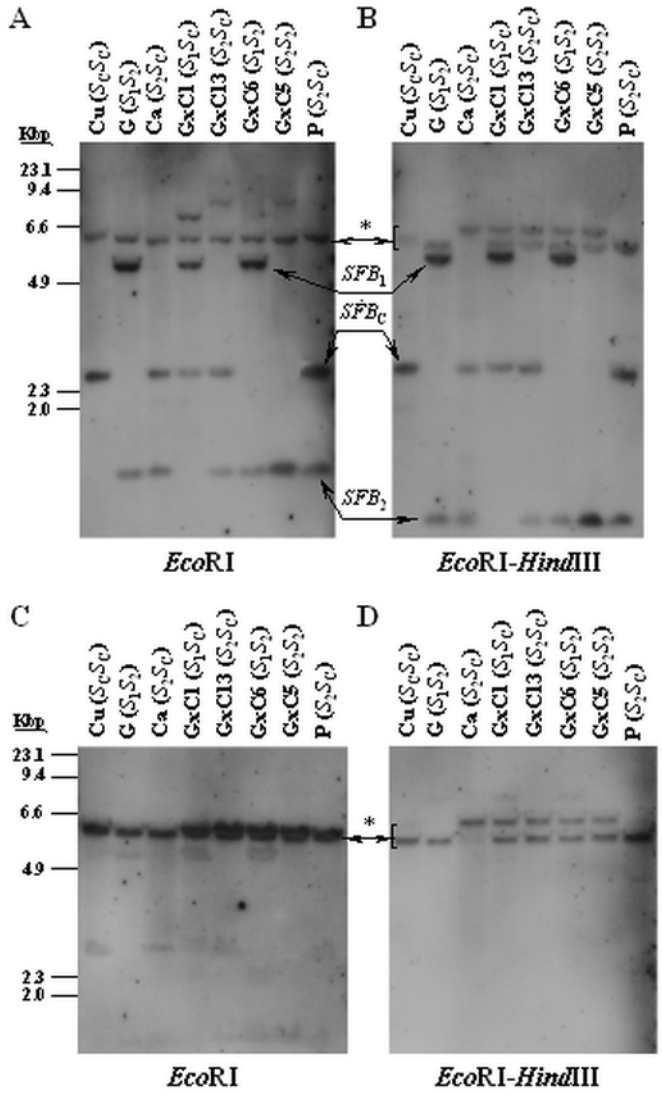


Figure 6


A
B
C
D
Cu (S_cS_c)
G (S_1S_2)
Ca (S_2S_c)
GxC1 (S_1S_c)
GxC13 (S_2S_c)
GxC6 (S_1S_2)
GxC5 (S_2S_2)
P (S_2S_c)
Kbp
23.1
9.4
6.6
4.9
2.3
2.0
*
SFB_1
SFB_c
SFB_2
EcoRI
EcoRI-HindIII