

Genome and population dynamics under selection and neutrality: an example of *S*-allele diversity in wild cherry (*Prunus avium* L.)

Ioannis Ganopoulos · Filippos Aravanopoulos ·
Anagnostis Argiriou · Athanasios Tsaftaris

Received: 28 November 2011 / Revised: 1 March 2012 / Accepted: 13 March 2012 / Published online: 31 March 2012
© Springer-Verlag 2012

Abstract Self-incompatibility, a common attribute of plant development, forms a classical paradigm of balancing selection in natural populations, in particular negative frequency-dependent selection. Under negative frequency-dependent selection population genetics theory predicts that the *S*-locus, being in command of self-incompatibility, keeps numerous alleles in equal frequencies demonstrating a wide allelic range. Moreover, while natural populations exhibit a higher within population genetic diversity, a reduction of population differentiation and increase of effective migration rate is expected in comparison to neutral loci. Allelic frequencies were investigated in terms of distribution and genetic structure at the gametophytic self-incompatibility locus in five wild cherry

(*Prunus avium* L.) populations. Comparisons were also made between the differentiation at the *S*-locus and at the SSR loci. Theoretical expectations under balancing selection were congruent to the results observed. The *S*-locus showed broad multiplicity (16 *S*-alleles), high genetic diversity, and allelic isoplethy. Genetic structure at the self-incompatibility locus was almost four times lower than at 11 nSSR loci. Analysis of molecular variance revealed that only 5 % of the total genetic variation concerns differentiation among populations. In conclusion, the wealth of *S*-allelic diversity found in natural wild cherry populations in Greece is useful not only in advancing basic population genetics research of self-incompatibility systems in wild cherry but also in the development of breeding programs.

Communicated by A. Dandekar

Ioannis Ganopoulos and Filippos Aravanopoulos contributed equally to this work.

I. Ganopoulos · A. Argiriou · A. Tsaftaris
Institute of Agrobiotechnology, CERTH,
Thermi 570 01 Thessaloniki, Greece

I. Ganopoulos · A. Tsaftaris
Department of Genetics and Plant Breeding,
Aristotle University of Thessaloniki,
Thessaloniki 54006, Greece

F. Aravanopoulos
Laboratory of Forest Genetics and Tree Breeding,
Faculty of Forestry and Natural Environment,
Aristotle University of Thessaloniki,
Thessaloniki 54124, Greece

A. Tsaftaris (✉)
Institute of Agrobiotechnology,
6th Km Charilaou Thermi Road,
Thermi 570 01 Thessaloniki, Greece
e-mail: tsaft@certh.gr

Keywords *Prunus avium* · Self-incompatibility ·
Balancing selection · *S*-alleles · *S*-RNase · Isoplethy ·
Population genetic structure

Introduction

Natural selection alters the distribution of allelic frequencies at target loci in finite populations in a different way compared to loci subject to other evolutionary forces like genetic drift, migration, and mutation (Hartl and Clark 2006). Typically, selection reduces variation within populations; nevertheless, the balancing selection class of models forms a general exception to this rule. Balancing selection results in stable non-neutral distributions of allelic frequencies, therefore in the maintenance of alleles in the population. One typical balancing selection model is negative frequency-dependent selection (NFDS; Maruyama and Nei 1981) and probably the most classical example of NFDS involves plant self-incompatibility (SI) systems (Wright 1939).

Outcrossing in plants is functionally favored by dedicated genetic systems (Barrett 2003) and almost half of plant species employ SI systems to limit inbreeding (Barrett 2002). RNase-based gametophytic self-incompatibility (GSI) is the most phylogenetically widespread system (~60 % of SI species, Franklin-Tong and Franklin 2003). Generally, a pollen grain sharing one of the self-incompatibility alleles (*S*-alleles) of the style fails to achieve fertilization, since the pollen tube cannot develop (Franklin-Tong and Franklin 2003). Self-fertilization is therefore impossible and mating between relatives is reduced.

NFDS arises in the genomic region associated with SI (Castric and Vekemans 2004) as any new incoming *S*-allele has a selective advantage in mating, which generates a high number of alleles within populations. This outcome is the product of the above sexual selection process that results in higher fitness for rare and novel alleles (Wright 1939). *S*-alleles are tightly linked both due to close physical association on the chromosome and reduced recombination rate in the area (Kamau et al. 2007). The second consequence of NFDS at the population level lies in the assumption that all *S*-alleles are selectively equivalent and hence selective strength acting on an *S*-allele only depends on its frequency. Therefore, selective equivalence among the different allelic specificities should lead to equal allelic frequencies at equilibrium (isoplethy hypothesis, i.e., all *S*-alleles should settle to an identical frequency, equal to the inverse of the number of *S*-alleles; Steiner and Gregorius 1999; Stoeckel et al. 2011; Wright 1939). Subsequently, the expected genetic signature of NFDS among populations is a reduction in population differentiation (compared to neutral loci) because of an increased effective migration rate at the genomic area associated with the *S*-locus (Glémin et al. 2005; Stoeckel et al. 2008). NFDS can be considered as an evolutionary force modulated by migration (Schierup 1998) and opposing to genetic drift, since this selection pressure tends to increase directly the frequency of rare alleles and thus limits the loss of alleles caused by drift (Wright 1939). Furthermore, if a plant species is characterized both by self-incompatibility and local dispersal an antagonism can exist between their evolutionary effects. This antagonism will be manifested differently in markers linked to a fairly large genomic area associated with the *S*-locus (due to reduced recombination near the *S*-locus; Kamau et al. 2007; Kawabe et al. 2006) than in markers that are unlinked.

Molecular data are now available to explore the impacts of ribonuclease (*S*-RNase) based GSI systems at a fine level because the pistil and pollen components of such systems have now been identified and the corresponding genes have been sequenced in some species (Newbigin et al. 2008).

Prunus avium L. (wild cherry), is a self-incompatible, multipurpose, scattered forest tree species used for its high quality timber, fruits and aesthetic value, which presents a

very good model to test NFDS and related hypotheses. Its GSI is determined by the multi-allelic *S*-locus that codes for *S*-RNases in the style (Bošković and Tobutt 1996) and for *S*-haplotype-specific F-box proteins in the pollen (Yamane et al. 2003). When the *S*-allele of the pollen grain matches one of the *S*-alleles of the style, these ribonucleases inhibit the growth of the pollen tube in the style. Various *S*-alleles have been cloned and shown to contain several highly conserved regions, while both allele-specific and consensus primers have been designed to amplify across each of the two introns (Sonneveld et al. 2001; Sonneveld et al. 2003). Both introns, especially the second one, present substantial length polymorphism, allowing most of the alleles to be distinguished according to the size of polymerase chain reaction (PCR) amplification products.

For GSI systems, Wright (1939) stated that “under a moderate mutation/immigration rate of 10^{-3} , a population of 50 individuals should harbor about eight different *S*-alleles, whereas a population of 200 individuals would have about 15 *S*-alleles”. Empirical studies on *S*-allele diversity, employing either diallel crossing schemes or molecular-genetic approaches to assess *S*-genotypes, demonstrated that the examined *S*-allele distinction among populations is excessive (Busch and Schoen 2008; Castric and Vekemans 2004). *S*-allele genotyping in *P. avium* led to the rapid confirmation of *S*-alleles and incompatibility groups in many cultivars and genotypes, as well as in the identification of putative novel *S*-alleles through correlations with new PCR products (Choi et al. 2002; De Cuyper et al. 2005; Hauck et al. 2002; Sonneveld et al. 2003; Tao et al. 1999; Yamane et al. 2000; Zhou et al. 2002). *S*-allele typing has become a useful tool for biodiversity assessment in wild cherry populations, accessions (De Cuyper et al. 2005; Schueler et al. 2006), and germplasm collections (Gisbert et al. 2008; Marchese et al. 2007; Schuster et al. 2007). Up to 16 different *S*-alleles and 29 incompatibility groups have been reported thus far using different methods (Schuster et al. 2007).

Herein, we present results concerning allelic distribution and genetic structure at the gametophytic self-incompatibility locus in five wild cherry (*P. avium* L.) populations. We also used previously published data on eleven nuclear microsatellite (SSR) loci of the same populations (Ganopoulos et al. 2011) in a comparative analysis of microsatellite population genetic structure and differentiation at the *S*-locus and also data from the *S*-allele distribution in Greek (Ganopoulos et al. 2010) and other European (Boskovic and Tobutt 2001; De Cuyper et al. 2005) sweet cherry cultivars. The objectives of this study were: (1) to determine the allelic diversity of *S*-RNases by estimating numbers of alleles and allele frequencies at the *S*-locus for five natural wild cherry populations in Greece, (2) to compare *S*-locus genetic diversity of Greek natural populations to the genetic diversity of the Greek and

other western European cultivars, and (3) to investigate the occurrence of NFDS and its consequences regarding (a) the number of *S*-alleles and genetic diversity at the *S*-locus, (b) isoplethy, (c) *S*-locus neutrality, (d) migration rate, and (e) population differentiation.

Materials and methods

Study site and sampling design

Five natural populations, covering most of the northern Greek range of *P. avium*, were sampled (Table 1). Levels of sample size per population were influenced by census population numbers, the non-continuous distribution of wild cherry in nature and the need to avoid sampling clonal or filial structures. We collected young leaves from individual trees spaced at least 150 m apart, in order to avoid sampling related individuals. No individual was found to be a vegetative propagule of another according to an earlier nSSR analysis (Ganopoulos et al. 2011). The sample size of each population ranged from 15 to 20 individuals (mean 18.6); 93 individuals were sampled in total (Table 1). Collected leaves of each individual were stored at -80°C prior to DNA extraction. Total genomic DNA was isolated using the procedure described by Doyle and Doyle (1987). The DNA amount was quantified by a UV-spectrophotometer (Eppendorf Biophotometer, Hamburg, Germany). Samples were then diluted to a 20 ng/ μL working concentration.

Assessment of *S*-allele polymorphisms

Fluorescently labelled primers (IRD800 or IRD700) PaConsI-F and PaConsI-R2 (VBC-GENOMICS, Austria) were used in PCR to amplify regions encompassing the first of each *S-RNase* intron (Sonneveld et al. 2006). PCR products were electrophoresed and visualized on 6.5 % polyacrylamide gels using a LI-COR 4300 DNA analyzer. Results were analyzed using LI-COR SAGA^{GT} software. Amplification products, including the second intron of each *S-RNase* were generated using the primers PaConsII-F and PaConsII-R and

scored on agarose gels as described by Sonneveld et al. (2003). The 1 kb DNA ladder (Invitrogen) was loaded at the center of each gel. Twenty-one standard cultivars with known *S*-alleles were also included in the analysis as references (Ganopoulos et al. 2010).

S-allele frequency distribution and *S*-locus population genetic analysis

S-allele frequencies were calculated as a percentage of all alleles identified in 93 wild cherry genotypes. The software GenAlEx 6 (Peakall and Smouse 2006) was used. Corresponding calculations were performed on three large data-sets reported previously (Boskovic and Tobutt 2001; De Cuyper et al. 2005; Ganopoulos et al. 2010) for comparative purposes. To obtain population-wide estimates of *S*-allele diversity for populations in which not all individuals were sampled, the total number of *S*-alleles was estimated following Paxman (1963) as $a = n (1 - (1 - 2/n)^m)$, where “*a*” is the observed number of *S*-alleles, “*n*” is the effective number of *S*-alleles, and “*m*” is the number of plants sampled per population (Brennan et al. 2002; Holderegger et al. 2008). Paxman estimates were obtained from iteration and rounded to whole numbers. For comparative purposes, we have applied the Paxman (1963) procedure for all SSR loci as well.

To detect the presence of equal allelic frequencies in different populations (isoplethy hypothesis as expected in an infinite panmictic population), we tested the frequencies of *S*-alleles for equal distribution in each of the five populations studied using a modification by Davies of the chi-squared test by Mantel (1974), as described in Campbell and Lawrence (1981). As proposed by Mantel (1974), this statistic follows approximately a chi-squared distribution with $n_a - 1$ degrees of freedom, “ n_a ” being the number of alleles (Stoeckel et al. 2008).

Standard population genetic estimates, i.e., the average number of alleles per locus (*A*), effective numbers of alleles (n_e), gene diversity (expected heterozygosity, H_E), observed heterozygosity (H_O), and Shannon's diversity index (*I*), were calculated using GenAlEx 6 (Peakall and Smouse 2006). We

Table 1 Geographic locality of *Prunus avium* natural populations studied

Population code	Geographic localities	Latitude	Longitude	Altitude (m)	Census population/sample size	H_E	F_{IS}	Observed/Paxman <i>S</i> -alleles number
VOR	Vorras, Imathia	21°.95′	40°.93′	1200	80/15	0.853	-0.138	12/13
HOL	Cholomontas, Chalkidiki	23°.52′	40°.42′	800	100/18	0.875	-0.050	11/12
KAT	Katafito, Drama	23°.69′	41°.35′	1000	900/20	0.861	-0.136	12/14
HAI	Hayntou, Xanthi	24°.66	41°.30′	1200	200/20	0.878	0.001	9/9
NIM	Nimphaea, Rodopi	25°.45′	41°.22′	700	400/20	0.764	-0.219	7/7

used the hierarchical rarefaction method available in HP-RARE (Kalinowski 2005) to calculate in addition to the above parameters allelic richness (AR). Pairwise F_{ST} estimates which characterize the level of population genetic structure for the S -locus was calculated using the Fstat vs. 2.9.3.2 software (Goudet 1995) based on the approach of Weir and Cockerham (1984). We have plotted H_E vs. F_{ST} to illustrate the performance of the S -locus compared to the nSSR loci. Historical gene flow (Nm) among populations was estimated indirectly based on the F_{ST} statistics.

The hierarchical distribution of genetic variation among and within populations for both markers was also characterized by an S -genotype-based analysis of molecular variance (AMOVA; (Excoffier et al. 1992; Michalakis and Excoffier 1996). We conducted a hierarchical AMOVA using the GENALEX 6 software (Peakall and Smouse 2006) as well. The tests were implemented using estimates of Φ_{ST} based on distances calculated from allelic data. Tests of significance were performed using 9,999 permutations within the total dataset.

Neutrality of a certain locus was tested by comparing the observed allele frequencies with the expected distribution of neutral alleles based on Ewens's sampling theory (Ewens

1972) using Watterson's homozygosity test (Watterson 1974). The Ewens–Watterson neutrality test was implemented in the software package Arlequin ver.3.0 (Excoffier et al. 2005) to test neutrality in S -RNase and microsatellite loci. The statistics for the test were calculated using the algorithm given by Manly (1985), using 1,000 simulated samples. Garrigan and Hedrick (2003) proposed the use of this test for neutral and the supposed selected loci in order to distinguish between demographic events and balancing selection.

Population genetics indices were compared to results from a neutral molecular codominant marker analysis (nSSR) of the same populations (Ganopoulos et al. 2011). A total of 11 nSSR loci have been employed by Ganopoulos et al. (2011) as part of a larger study. Nevertheless, one of these loci (EMPAS02) is closely linked to the Sc -locus (at 3.3 cM; Cachi and Wünsch 2011). As the S -locus genomic area is characterized by reduced recombination (Kamau et al. 2007), EMPAS02 was removed from the Ganopoulos et al. (2011) nSSR data set and population genetics indices were recalculated to obtain baseline estimated from truly neutral markers. Therefore, the following data sets were employed in comparative analyses: (a) S -locus, (b) nSRR (11 loci; Ganopoulos et al. 2011), (c) nSSR except EMPAS02 (10 loci).

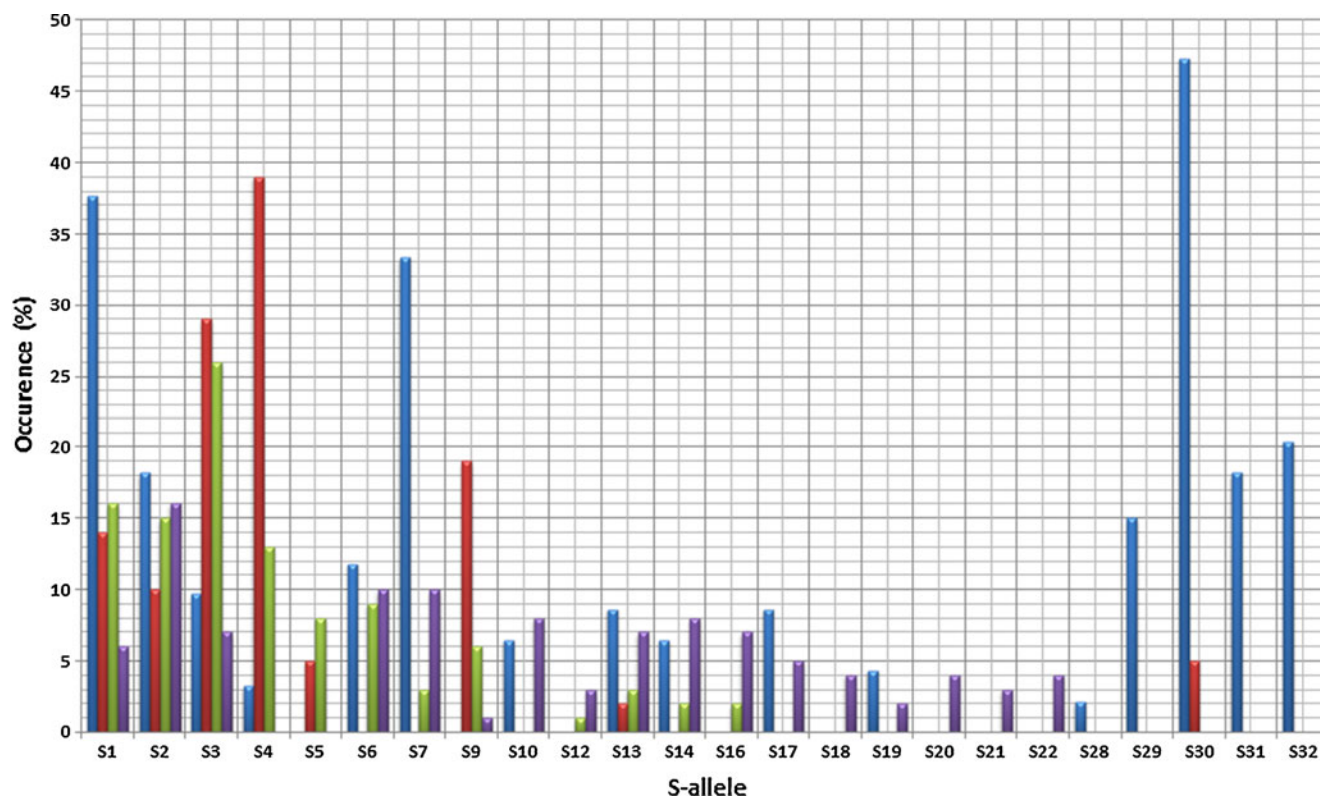


Fig. 1 Relative occurrence (percent) of different S-alleles in 93 wild cherry (*Prunus avium*) genotypes (blue) compared to the frequencies found by Ganopoulos et al. (2010) in 21 sweet cherry cultivars (red),

by Boskovic and Tobutt (2001) in 67 sweet cherry cultivars (green) and those reported by De Cuyper et al. (2005) in 65 wild cherry accessions (lilac)

Results and discussion

The frequency distribution of the *S*-alleles in wild populations

We detected a total of 16 *S*-alleles in the five wild cherry populations studied. We have used Vaughan et al. (2008) as a background reference for defining and assigning *S*-alleles. The number of *S*-alleles per population was moderately high, ranging from 7 to 12 with a mean of 10.20 ($s=2.17$) per population (Table 1, Fig. 1). The maximum number of 12 *S*-alleles was found in populations Katafito and Vorras. Even when a small population sample size was employed, ($n=15$, population Vorras), a large number of *S*-alleles was revealed (12 *S*-alleles). The numbers of *S*-alleles per population changed only slightly when adjusted for incomplete sampling (Paxman 1963) presenting an average of 11.00 ($s=2.91$; Table 1). A similar result was also observed when the Paxman (1963) procedure was applied to SSR loci (results not shown). According to NFDS predictions, a high number of *S*-alleles within populations are expected. Indeed, when the average number of *S*-alleles per population was compared to the corresponding value of neutral SSR markers (Ganopoulos et al. 2011), the number of alleles at

the *S*-locus was higher and the observed difference was statistically significant according to a *t* test ($t=2.963$, $p=0.031$). This result was also extended to the effective number of alleles values ($n_e=6.884$ at the *S*-locus, while $n_e=4.253$ at 11 SSR loci). The number of SI alleles per locus independent of sample size AR (the allelic richness) ranged from 5.810 (Nimphaea) to 8.360 (Hayntou) with an average of $AR=7.558$ compared to $AR=6.849$ for SSR markers. Although in a strict sense, true allelic richness can only be detected at the DNA sequence level, by and large the above results support Wright's (1939) prediction of high allelic richness at the *S*-locus. This outcome is in full congruence to most pertinent studies which report high allelic richness at the *S*-locus (Castric and Vekemans 2004; Lawrence 2000).

Overall allelic frequency ranged from 2.15 % (S_{28}) to 47.3 % (S_{30}). Out of a total of 16 *S*-alleles, six (37.5 %) were common to all wild cherry populations. Thirteen *S*-alleles were found in at least in two of the five populations used in this study. The most frequent alleles (S_7 , S_{17} , S_{29} , S_{30} , S_{31} , and S_{32}) were alleles found in all populations. Few *S*-alleles were private with low frequencies in the Cholomontas (two *S*-alleles) and the Katafito (one *S*-allele) populations, respectively (Fig. 2). The occurrence of *S*-alleles in natural Greek

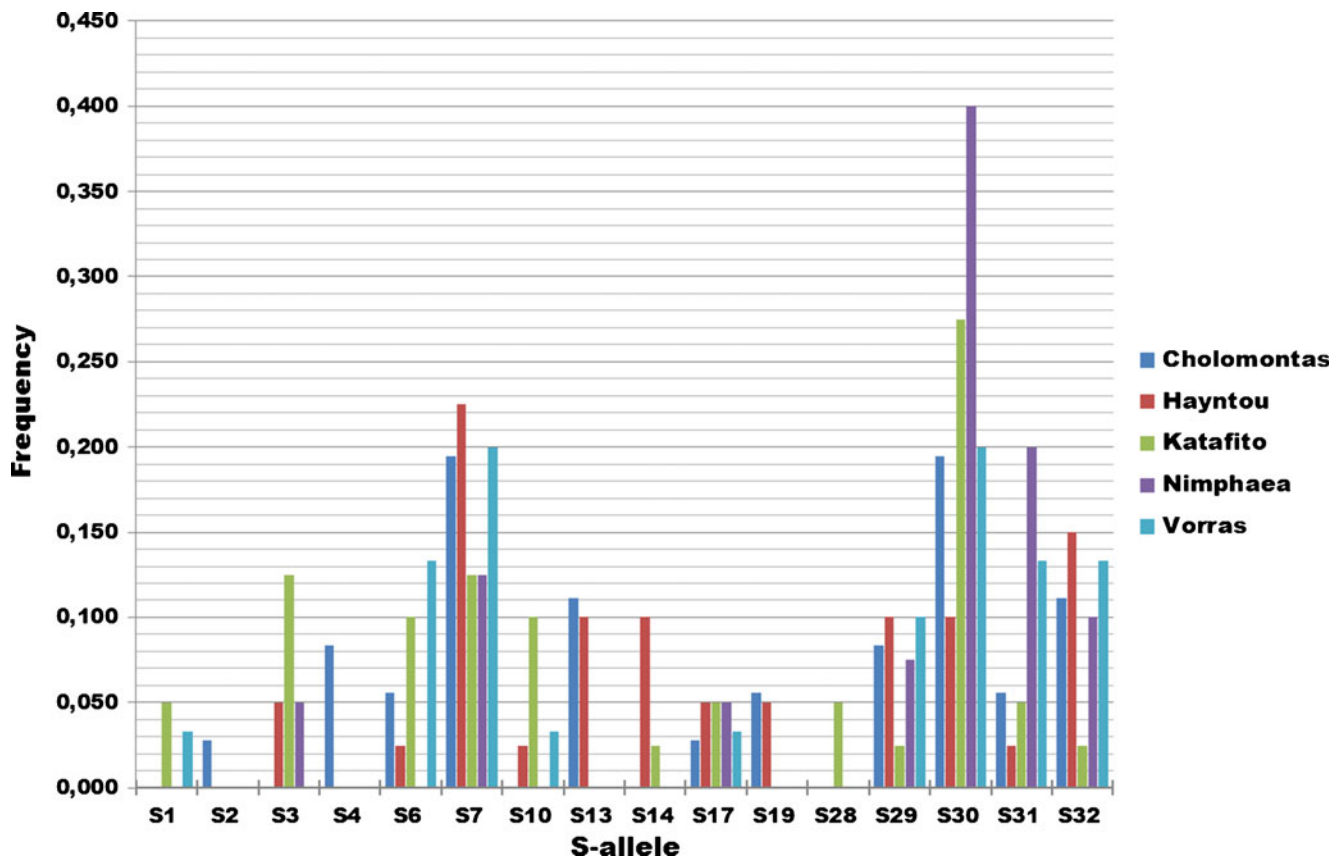


Fig. 2 *S*-allele frequencies in five Greek natural wild cherry (*Prunus avium*) populations

Table 2 Genetic diversity at the *S*-locus and 11 neutral SSR loci in *Prunus avium*

Data set	<i>A</i>	n_e	<i>I</i>	H_e	H_o	AR	F_{ST}	<i>E</i>	N_m	Isoplethy	Neutrality test
VOR	9	6.818	2.020	0.853	1.000	6.833	–	0.919	–	Equal 0.165 (ns)	–
HOL	11	8.000	2.220	0.875	0.944	6.939	–	0.926	–	Equal 0.738 (ns)	–
KAT	12	7.207	2.211	0.861	1.000	6.500	–	0.890	–	Equal 0.720 (ns)	–
HAI	12	8.163	2.267	0.878	0.900	7.519	–	0.912	–	Equal 0.623 (ns)	–
NIM	7	4.233	1.672	0.764	0.950	6.455	–	0.859	–	Unequal $p < 0.001$	–
Only SI	10.2	6.884	2.078	0.846	0.958	7.558	0.035	0.901	6.754		$F(p) < 0.001$
SSR without EMPaS02	7.12	4.246	1.603	0.739	0.685	6.040	0.092	0.816	0.977	–	$F(p) = 0.209$ – 0.374 Neutral
SSR (Ganopoulos et al. 2011)	7.2	4.253	1.609	0.74	0.680	6.849	0.097	0.815	1.190	–	$F(p) = 0.207$ – 0.374 Neutral

wild cherry genotypes was compared to those found in 21 sweet cherry cultivars of wide cultivation in Greece (Ganopoulos et al. 2010). A total of 10 *S*-alleles (S_6 , S_7 , S_{10} , S_{14} , S_{17} , S_{19} , S_{28} , S_{29} , S_{30} , S_{32}) were present only in the wild populations and two *S*-alleles (S_5 , S_9) were present only in cultivars, while six *S*-alleles (S_1 , S_2 , S_3 , S_4 , S_{13} , S_{30}) were present in both (Fig. 1). Furthermore, results were compared to data from 67 cultivars examined by Boskovic and Tobutt (2001) and from 65 Belgian wild cherry accessions originating from natural populations (De Cuyper et al. 2005; Fig. 1). A total of seven *S*-alleles were common to the Boskovic and Tobutt (2001) study, while ten alleles were common to the De Cuyper et al. (2005) study. Alleles S_1 occurred at high frequency in the Greek wild cherry genotypes, although they were rare in the De Cuyper et al. (2005) survey. Nevertheless, the frequencies of alleles S_1 and S_2 were similar to the Boskovic and Tobutt (2001) study (Fig. 1). Five *S*-alleles found in the Greek wild cherry populations (S_{28} , S_{29} , S_{30} , S_{31} , S_{32}) have not been detected in European cultivars and natural accessions by Boskovic and Tobutt (2001) and De Cuyper et al. (2005). However, they were recently found in wild cherry genotypes of western Europe and have been characterized as novel alleles by Vaughan et al. (2008). On the contrary, alleles S_5 , S_9 , S_{12} , S_{16} , S_{18} , S_{20} , S_{21} , and S_{22} were absent in the Greek populations studied; these belong to well-known international sweet cherry cultivars (Boskovic and Tobutt 2001), as well as to eastern sweet cherry genotypes (Ipek et al. 2011) and are present in the Belgian wild cherry population (De Cuyper et al. 2005) in addition to being present in the Greek sweet cherry cultivars (Ganopoulos et al. 2010).

Comparative analyses of genetic diversity and gene flow in natural wild cherries and orchard cultivars in Northern Greece have indicated that these groups are very well differentiated and realized genetic introgression is practically

absent (Ganopoulos et al. 2012). Cultivars do not appear to be crossbreeding with wild trees due to significant disjunction in flowering time and a considerable geographic distance between the two groups (Ganopoulos et al. 2012). The genetic base of the cultivar collection is restricted compared to natural populations probably as a result of prolonged and strong anthropogenic selection (Ganopoulos et al. 2012). Nevertheless, the presence of private *S*-alleles in cultivars compared to natural populations indicates that at least some cultivars have not been derived from local origin.

The tests for equal *S*-allele distribution within each population confirmed the presence of isoplethy in all populations, but one or in 80 % of the populations studied (Table 2). According to NFDS predictions, equal allelic frequencies at equilibrium are expected at the *S*-locus in natural populations (Wright 1939). The isoplethy hypothesis was supported in *Pyrus pyraeaster* (Holderegger et al. 2008); nevertheless, it was not confirmed in *P. avium* (Schueler et al. 2006; Stoeckel et al. 2008) and in *Prunus lannesiana* (Kato and Mukai 2004). Different behavior of populations within a species is not uncommon and has been observed before in *P. avium* (Stoeckel et al. 2008), while deviations from isoplethy have been explained on theoretical grounds

Table 3 Pairwise F_{ST} values at the *S*-locus for the five *Prunus avium* natural populations

Population	Cholomontas	Hayntou	Katafito	Nimphaea
Hayntou	0.0063			
Katafito	0.0202	0.0331		
Nimphaea	0.0342	0.0666	0.0254	
Vorras	0.0063	0.0066	0.0091	0.0211

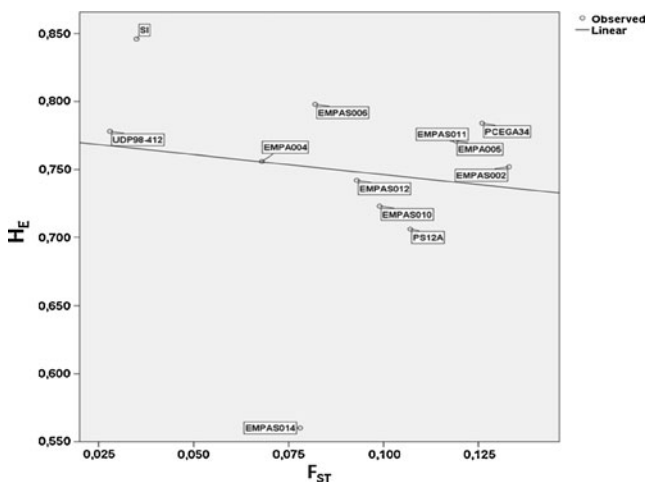


Fig. 3 Plot of expected heterozygosity (H_E) and fixation index F_{ST} for the *S*-locus and the nSSR loci

(Campbell and Lawrence 1981; Muirhead 2001). In two reviews of empirical studies, Lawrence (2000) reported no departure from the identical allelic frequencies hypothesis in 16 out of 19 sampled populations from 12 species with GSI, whilst Castric and Vekemans (2004) reported that of 11 species with GSI, a highly significant rejection of the isoplethy hypothesis was obtained in two cases. Further results of mixed nature have been presented for perennial woody angiosperms in recent studies. Overall, it appears that, although allelic frequency tests have low statistical power because of generally low sample sizes, most species with GSI support the equal allelic frequencies hypothesis as this study reported. The prevalence of isoplethy provides further support regarding the occurrence of NFDS in the Greek wild cherry populations studied.

Genetic diversity and marker neutrality

Besides the large number of *S*-alleles, high levels of diversity at the *S*-locus (Table 2) were observed. Gene diversity was $H_E=0.846$ at the *S*-locus (Table 2), higher than the respective value at 11 SSR loci ($H_E=0.740$; Ganopoulos et al. 2011). When the EMPAS02 nuclear microsatellite locus which is closely linked to the *S*-locus was removed from the analysis, neutral marker diversity of the remaining ten nSSR loci was even lower ($H_E=0.685$, Table 2). These results are in agreement to the generally high *S*-allele diversity which is

a distinct feature of the *S*-locus in self-incompatible plant species (Lawrence 2000). At the population level the average heterozygosity was highest in the Hayntou ($H_E=0.878$) and lowest in the Nimphaea ($H_E=0.764$) populations. High levels of genetic diversity at the *S*-locus were indicated especially in populations Katafito and Hayntou (Table 2). All pairwise F_{ST} (0.0063–0.0342) values were highly significant ($p<0.001$; Table 3). It can be therefore deduced that the levels of genetic diversity at the *P. avium S*-locus are in good agreement with theoretical models of *S*-locus population dynamics.

The Ewens–Watterson homozygosity test (Watterson 1974) revealed a strong deviation from neutrality for the *S*-locus ($p<0.001$; Table 2). On the other hand, for nSSR loci, no departure from neutrality was detected ($0.207<F(p)<0.374$; Table 2). Similar results were found by Schueler et al. (2006) also in *P. avium* (in a comparison of the *S*-locus and nSSR loci), as well as by Garrigan and Hedrick (2003) in a comparison of MHC and SSR loci in sheep. Especially for selectively neutral loci that are in synteny with the *S*-locus in *P. avium* the above results may be regarded as supporting the finding of rapid linkage disequilibrium decay in this species (Arunyawat et al. 2012). The sharp difference between the results of the *S*-locus neutrality test and the nSSR loci neutrality tests in this analysis provides an additional clear evidence of the impact of NFDS in the wild cherry populations studied.

Gene flow and genetic structure

The elevated number of *S*-alleles in the wild cherry populations potentially indicates a high effective gene flow among populations for the genomic region associated with the GSI system. Indeed, gene flow based on data from the RNase-based gametophytic self-incompatibility alleles was $N_m=6.754$. This value is almost six times higher than the corresponding value estimated from neutral marker (nSSR) data for the same populations and individuals ($N_m=1.190$; Ganopoulos et al. 2011). Thus, substantial historical gene flow is clearly suggested at the *S*-locus. Theory predicts a greater rate of effective migration of *S*-alleles, which are under balancing selection, compared to alleles of neutral molecular markers (Muirhead 2001; Castric and Vekemans 2004). Consequently, our results are also in accordance to NFDS expectations regarding elevated gene flow as estimated by *S*-allele analysis.

Table 4 *S*-genotype-based analysis of molecular variance (AMOVA) for the five *Prunus avium* populations

Source of variation	d.f.	SSD	Variance component	Total variance (%)	Φ statistics	<i>p</i> value
Among populations	4	6.026	0.039	5 %		
Within populations	88	68.200	0.775	95 %	0.048	<0.001
Total	92	74.226	0.814	100 %		

Another finding which is consistent both with the above results and with theoretical predictions concerns population differentiation at the *S*-locus. The *S*-locus F_{ST} value of 0.035 was strikingly low compared to the high degree of population differentiation observed at neutral loci in this species ($F_{ST}=0.097$ for 11 nSSRs; Ganopoulos et al. 2011). This result can be readily visualized in Fig. 3 where H_E and F_{ST} are plotted for all loci. When the closely linked to the *S*-locus EMPaS02 nSSR locus was removed the magnitude of the difference was not substantially altered ($F_{ST}=0.092$). Therefore, population differentiation was found to be approximately three times lower in the *P. avium* genomic region compared to the rest of the genome. This outcome is in accordance to Stoeckel et al. (2006, 2008) who found the F_{ST} at the *S*-locus to be three times lower than F_{ST} at seven microsatellite loci among three *P. avium* populations from France. Similarly, wild pear (*P. pyraster*) populations exhibited a three times lower F_{ST} at the *S*-locus ($F_{ST}=0.011$) than at five microsatellite loci ($F_{ST}=0.039$; Holderegger et al. 2008). In addition, (Glemin et al. 2005) found a population differentiation at the *S*-locus accounting for half of that at 11 nSSR loci in self-incompatible *Brassica insularis*.

Additional support for the above findings was obtained through the use of *S*-locus AMOVA. AMOVA showed that approximately 95 % of the total variation at the *S*-locus was partitioned within populations, while, among population differentiation is associated with approximately only 5 % of the total variation ($\Phi_{ST}=0.048$; $P=0.001$; Table 4). A lower genetic structure at the *S*-locus when compared with marker loci seems to be a general feature of plant SI systems (Glemin et al. 2005; Brennan et al. 2002; Kamau et al. 2007). In general, our results show clearly a lower genetic structure at the *S*-locus compared to the genetic structure of the same populations in neutral loci in congruence to the relevant literature (Kamau et al. 2007).

Conclusions

By analyzing *S*-locus allelic data derived from five natural populations of wild cherry (*P. avium*) and comparing population genetics parameters between *S*-locus data and neutral molecular marker data, we detected many of the signatures expected from a genomic region under negative frequency-dependent selection. These are: (a) a high number of alleles and allelic richness at the *S*-locus, (b) prevalence of isoplethy, (c) high *S*-allele genetic diversity, (d) a strong departure from neutrality at the *S*-locus, (e) a greater rate of effective migration of *S*-alleles compared to alleles of neutral molecular markers, (f) a lower population genetic structure at the *S*-locus. These results indicate that negative frequency-dependent selection is a major evolutionary selective force

at the genomic area associated with the *S*-locus in accordance to the predictions of Wright (1939).

In practical terms, these results indicate the absence of restrictions concerning mate availability in the relatively small natural wild cherry populations studied. The lack of available data regarding the composition of *S*-alleles in Greek natural populations is now largely resolved. The determination of *S*-locus genotypes is paramount in designing crosses in breeding programs and our results are useful for the proper selection of plus trees. Furthermore, they can also contribute to the advance of population genetics and self-incompatibility research in wild cherry.

Acknowledgments This research has been co-financed by the European Union (European Social Fund—ESF) and Hellenic national funds through the Operational Program “Education and Lifelong Learning” of the National Strategic Reference Framework (NSRF)—Research Funding Program: Heraclitus II. Continuous support for the Institute of Agrobiotechnology/CERTH from the General Secretariat of Research and Technology of Greece is also acknowledged.

References

- Arunyawat U, Capdeville G, Decroocq V, Mariette S (2012) Linkage disequilibrium in French wild cherry germplasm and worldwide sweet cherry germplasm. *Tree Genet Genom*. doi:10.1007/s11295-011-0460-9
- Barrett SCH (2002) The evolution of plant sexual diversity. *Nat Rev Genet* 3:274–284
- Barrett SCH (2003) Mating strategies in flowering plants: the outcrossing–selfing paradigm and beyond. *Philos T Roy Soc B* 358:991–1004
- Bošković R, Tobutt KR (1996) Correlation of stylar ribonuclease zymograms with incompatibility alleles in sweet cherry. *Euphytica* 90:245–250
- Boskovic R, Tobutt KR (2001) Genotyping cherry cultivars assigned to incompatibility groups, by analysing stylar ribonucleases. *Theor Appl Genet* 103:475–485
- Brennan AC, Harris SA, Tabah DA, Hiscock SJ (2002) The population genetics of sporophytic self-incompatibility in *Senecio squalidus* L. (Asteraceae) I: S allele diversity in a natural population. *Heredity* 89:430–438
- Busch JW, Schoen DJ (2008) The evolution of self-incompatibility when mates are limiting. *Trends Plant Sci* 13:128–136
- Cachi AM, Wünsch A (2011) Characterization and mapping of non-S gametophytic self-compatibility in sweet cherry (*Prunus avium* L.). *J Exp Bot* 62:1847–1856
- Campbell JM, Lawrence MJ (1981) The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. II. The number and frequency of S-alleles in a natural population. *Heredity* 46:81–90
- Castric V, Vekemans X (2004) Plant self-incompatibility in natural populations: a critical assessment of recent theoretical and empirical advances. *Mol Ecol* 13:2873–2889
- Choi C, Ta R, Andersen RL (2002) Identification of self-incompatibility alleles and pollen incompatibility groups in sweet cherry by PCR based S-allele typing and controlled pollination. *Euphytica* 123:9–20
- De Cuyper B, Sonneveld T, Tobutt KR (2005) Determining self-incompatibility genotypes in Belgian wild cherries. *Mol Ecol* 14:945–955

- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11–15
- Ewens WJ (1972) The sampling theory of selectively neutral alleles. *Theor Popul Biol* 3:87–112
- Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform Online* 1:47–50
- Franklin-Tong VE, Franklin FCH (2003) The different mechanisms of gametophytic self-incompatibility. *Philos T Roy Soc B* 358:1025–1032
- Ganopoulos IV, Argiriou A, Tsaftaris AS (2010) Determination of self-incompatible genotypes in 21 cultivated sweet cherry cultivars in Greece and implications for orchard cultivation. *J Hortic Sci Biotech* 85:444–448
- Ganopoulos I, Aravanopoulos F, Argiriou A, Kalivas A, Tsaftaris A (2011) Is the genetic diversity of small scattered forest tree populations at the southern limits of their range more prone to stochastic events? A wild cherry case study by microsatellite-based markers. *Tree Genet Genom* 7:1299–1313
- Ganopoulos IV, Aravanopoulos FA, Tsaftaris A (2012) Genetic differentiation and gene flow between wild and cultivated *Prunus avium*: an analysis of molecular genetic evidence at a regional scale. *Plant Biosyst* (accepted)
- Garrigan D, Hedrick PW (2003) Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution* 57:1707–1722
- Gisbert AD, Badenes ML, Tobutt KR, Llacer G, Romero C (2008) Determination of the S-allele composition of sweet cherry (*Prunus avium* L.) cultivars grown in the southeast of Spain by PCR analysis. *J Hortic Sci Biotech* 83:246–252
- Glémin S, Gaude T, Guillemin M-L, Lourmas M, Olivieri I, As M (2005) Balancing selection in the wild: testing population genetics theory of self-incompatibility in the rare species *Brassica insularis*. *Genetics* 171:279–289
- Goudet J (1995) Fstat (version 1.2); a computer program to calculate F-statistics. *J Hered* 86:485–486
- Hartl D, Clark A (2006) *Principles of Population Genetics*, edn. Sinauer Associates, Inc, Sunderland
- Hauck NR, Yamane H, Tao R, Iezzoni AF (2002) Self-compatibility and incompatibility in tetraploid sour cherry (*Prunus cerasus* L.). *Sex Plant Reprod* 15:39–46
- Holderegger R, Häner R, Csencsics D, Angelone S, Hoebee SE (2008) S-allele diversity suggests no mate limitation in small populations of a self-incompatible plant. *Evolution* 62:2922–2928
- Ipek A, Gulen H, Akcay ME, Ipek M, Ergin S, Eris A (2011) Determination of self-incompatibility groups of sweet cherry genotypes from Turkey. *Genet Mol Res* 10(1):253–260
- Kalinowski ST (2005) Hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes* 5:187–189
- Kamau E, Charlesworth B, Charlesworth D (2007) Linkage disequilibrium and recombination rate estimates in the self-incompatibility region of *Arabidopsis lyrata*. *Genetics* 176:2357–2369
- Kato S, Mukai Y (2004) Allelic diversity of *S-RNase* at the self-incompatibility locus in natural flowering cherry populations (*Prunus lannesiana* var. *speciosa*). *Heredity* 92:249–256
- Kawabe A, Hansson B, Forrest A, Hagenblad J, Charlesworth D (2006) Comparative gene mapping in *Arabidopsis lyrata* chromosomes 6 and 7 and *A. thaliana* chromosome IV: evolutionary history, rearrangements and local recombination rates. *Genet Res* 88:45–56
- Lawrence MJ (2000) Population genetics of the homomorphic self-incompatibility polymorphisms in flowering plants. *Ann Bot* 85:221–226
- Manly BFJ (1985) *The Statistics of Natural Selection on Animal Populations*. Chapman and Hall, London
- Mantel N (1974) Some reasons for not using the Yates continuity correction on 2×2 contingency tables: comment and a suggestion. *Amer Statist Assoc* 69:378–380
- Marchese A, Tobutt KR, Raimondo A, Motisi A, Boskovic RI, Clarke J, Caruso T (2007) Morphological characteristics, microsatellite fingerprinting and determination of incompatibility genotypes of Sicilian sweet cherry cultivars. *J Hortic Sci Biotech* 82:41–48
- Maruyama T, Nei M (1981) Genetic variability maintained by mutation and overdominant selection in finite populations. *Genetics* 98:441–459
- Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* 142:1061–1064
- Muirhead CA (2001) Consequences of population structure on genes under balancing selection. *Evolution* 55:1532–1541
- Newbigin E, Paape T, Kohn JR (2008) *RNase*-based self-incompatibility: puzzled by pollen S. *Plant Cell* 20:2286–2292
- Paxman GJ (1963) The maximum likelihood estimation of the number of self-sterility alleles in a population. *Genetics* 48:1029–1032
- Peakall R, Smouse P (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295
- Schierup MH (1998) The number of self-incompatibility alleles in a finite, subdivided population. *Genetics* 149:1153–1162
- Schueler S, Tusch A, Scholz F (2006) Comparative analysis of the within-population genetic structure in wild cherry (*Prunus avium* L.) at the self-incompatibility locus and nuclear microsatellites. *Mol Ecol* 15:3231–3243
- Schuster M, Flachowsky H, Kohler D (2007) Determination of self-incompatible genotypes in sweet cherry (*Prunus avium* L.) accessions and cultivars of the German Fruit Gene Bank and from private collections. *Plant Breeding* 126:533–540
- Sonneveld T, Robbins TP, Boskovic R, Tobutt KR (2001) Cloning of six cherry self-incompatibility alleles and development of allele-specific PCR detection. *Theor Appl Genet* 102:1046–1055
- Sonneveld T, Tobutt KR, Robbins TP (2003) Allele-specific PCR detection of sweet cherry self-incompatibility (S) alleles S1 to S16 using consensus and allele-specific primers. *Theor Appl Genet* 107:1059–1070
- Sonneveld T, Robbins TP, Tobutt KR (2006) Improved discrimination of self-incompatibility S-RNase alleles in cherry and high throughput genotyping by automated sizing of first intron polymerase chain reaction products. *Plant Breeding* 125:305–307
- Steiner W, Gregorius H-R (1999) Incompatibility and pollen competition in *Alnus glutinosa*: evidence from pollination experiments. *Genetica* 105:259–271
- Stoeckel S, Grange J, Fernandez-Manjarres J, Bilger I, Frascaria-Lacoste N, Mariette S (2006) Heterozygote excess in a self-incompatible and partially clonal forest tree species—*Prunus avium* L. *Mol Ecol* 15:2109–2118
- Stoeckel S, Castric V, Mariette S, Vekemans X (2008) Unequal allelic frequencies at the self-incompatibility locus within local populations of *Prunus avium* L.: an effect of population structure? *J Evolution Biol* 21:889–899
- Stoeckel S, Klein EK, Oddou-Muratorio S, Musch B, Mariette S (2011) Microevolution of S-allele frequencies in wild cherry populations: respective impacts of negative frequency dependent selection and genetic drift. *Evolution* (in press)

- Tao R, Yamane H, Sugiura A, Murayama H, Sassa H, Mori H (1999) Molecular typing of S-alleles through identification, characterization and cDNA cloning for *S-RNases* in sweet cherry. *J Am Soc Hort Sci* 124:224–233
- Vaughan SP, Boskovic RI, Gisbert-Climent A, Russell K, Tobutt KR (2008) Characterisation of novel S-alleles from cherry (*Prunus avium* L.). *Tree Genet Genom* 4(3):531–541
- Watterson GA (1974) The sampling theory of selectively neutral alleles. *Adv Appl Probab* 6:463–488
- Weir B, Cockerham C (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–1370
- Wright S (1939) The distribution of self-sterility alleles in populations. *Genetics* 24:538–552
- Yamane H, Tao R, Murayama H, Ishiguro M, Abe Y, Soejima J, Sugiura A (2000) Determining S- genotypes of two sweet cherry (*Prunus avium* L.) cultivars, ‘Takasago (Rockport Bigarreau)’ and ‘Hinode (Early Purple)’. *J Japan Soc Hort Sci Hort Res* 69:29–34
- Yamane H, Ikeda K, Ushijima K, Sassa H, Tao R (2003) A pollen-expressed gene for a novel protein with an F-box motif that is very tightly linked to a gene for *S-RNase* in two species of cherry, *Prunus cerasus* and *P-avium*. *Plant Cell Physiol* 44:764–769
- Zhou L, Kappel F, MacDonald R, Hampson C, Bakkeren G, Wiersma PA (2002) Determination of S-genotypes and self-fertility of sweet cherry in Summerland advanced selections. *J Am Pomol Soc* 56:173–179