ORIGINAL PAPER

Reproductive strategy and ploidy determine the genetic variability of *Sorbus aria*

Pedro A. Sosa · Edna A. González-González ·
Miguel A. González-Pérez · Agustín Naranjo-Cigala ·
Eduardo Carqué · Aurelio Acevedo

Abstract *Sorbus aria* (L.) Crantz (common whitebeam) from the Canary Islands has not been characterised genetically. We analysed the genetic variability of 184 individuals belonging to seven natural populations of *S. aria* from the Canarian Archipelago and the Iberian Peninsula. Our main aims were to obtain essential information to enable the exploration of the genetic relationship between populations from the Canary Islands and the Iberian Peninsula; to establish the existence of a spatial genetic structure and formulate appropriate management and conservation genetics strategies. Genetic variation was analysed using nine polymorphic microsatellite loci. The Canary Island populations (triploids) were found to have very low genetic variability and to be considerably differentiated from the populations from the peninsula (diploid and triploid), although with a connection to the Sierra Nevada population in the south of the Peninsula. This population, in turn, had many different genotypes, which is indicative of the existence of various origins. The level of genetic diversity was higher in all-diploid populations, which, in addition, presented a greater interpopulation gene flow, possibly the result of a prevalence of sexual reproduction. On the other hand, the triploid populations presented lower levels of genetic variability, with a significant degree of fixed heterozygosity, possibly due to asexual reproduction, mainly by apomixis. The reproductive biology and ploidy appear to be responsible for the levels of genetic variability in *S. aria*.

Keywords Apomixis · Agamospermy · Asexual reproduction · Canary Islands · Genetic variability · Iberian Peninsula · Microsatellites · Outcrossing · *Sorbus aria*

P. A. Sosa (✉) · E. A. González-González · M. A. González-Pérez
Departamento de Biología, Universidad de Las Palmas de Gran
Canaria, Las Palmas, Canary Islands, Spain
e-mail: psosa@dbio.ulpgc.es

A. Naranjo-Cigala
Departamento de Geografía, Universidad de Las Palmas de Gran
Canaria, Las Palmas, Canary Islands, Spain

E. Carqué
Grupo Tragsa, Carretera General del Norte, s/n Esquina Camino La
Colina, Local de Oficinas n° 1.38330 Guamasa, Santa Cruz de
Tenerife, Canary Islands, Spain

A. Acevedo
Tragsatec La Palma, C/Francisco Vega Monroy, N° 12, 38700 S/C La
Palma, Canary Islands, Spain

Introduction

The genus *Sorbus* L. (Rosaceae), which includes more than 250 species, is a well-known example of a complex plant group that includes diploid sexual outcrossing and self-incompatible species and a large number of polyploid species originated by hybridization, allopolyploidy and autopolyploidy (Nelson-Jones et al. 2002; Robertson et al. 2004, 2010; Chester et al. 2007; Gömory and Krajmerová 2008). Furthermore, many of these species are reproduced principally by asexual reproduction, through apomixis (hereafter synonymous with agamospermy), i.e. the production of "maternal" clonal seeds, both obligate and facultative (Robertson et al. 2010). Because apomictic plants are genetically identical from one generation to the next, each lineage has some of the characters of a true species, affording immediate, often total, reproductive isolation from parental taxa and maintaining distinctions from other apomictic lineages within the same genus (Aldasoro et al. 2004; Chester et al. 2007; Rasmussen and Kollmann 2008; Whitton et al. 2008; Del Río et al. 2009; Robertson et al. 2010). The taxonomical complexity of *Sorbus* has been widely noted, and various authors consider that apomictic polyploids are not equivalent to Linnean species and are described as

"microspecies" (Aldasoro et al. 2004; Nelson-Jones et al. 2002; Robertson et al. 2004, 2010).

Sorbus aria L. is included in the subgenus *Aria* that comprises not only the diploid sexual species (*S. aria* sensu stricto) but also a number of triploid and tetraploid species. The whole group is known as the *S. aria* aggregate and is often referred to as *S. aria* sensu lato (Nelson-Jones et al. 2002). *S. aria* aggregate covers a wide natural area throughout the whole of Europe and North Africa (Aldasoro et al. 2004). In the Iberian Peninsula, it extends through mountain areas between 600 and 2,200 m.a.s.l. above all in limestone areas (Aldasoro et al. 2004; Oria de Rueda et al. 2006; Del Río et al. 2009). The Canary Islands represent the most meridional distribution of the species. It has been described on just two islands, La Palma and Tenerife, located at more than 2,000 m.a.s.l, very isolated and with a very small number of individuals (Arechavaleta et al. 2010; Acevedo et al. 2011). At present, only 137 individuals have been described (Acevedo et al. 2011), 4 in Tenerife and 133 in La Palma; most of them within the limits of Teide and Caldera de Taburiente National Parks, respectively. There are no demographic, reproductive or genetic studies of this species in the Canary Islands or genetic studies in the south of Europe (Oria de Rueda et al. 2006). In fact, as far as we know, all genetic studies of *S. aria* have been carried out in the context of an intergenus comparison (Nelson-Jones et al. 2002; Chester et al. 2007; Gömory and Krajmerová 2008; Robertson et al. 2010). However, the relationships between populations located in the archipelago are unknown, as are their ploidy levels and biogeographical relationships with the continental populations.

In this study, we used codominant markers (microsatellites) to study the patterns of genetic diversity and genetic differentiation of *S. aria* from the Iberian Peninsula and the Canary Islands. Microsatellite markers are a powerful tool with which to assess vegetative reproduction because the probability that a number of sexually reproduced individuals have the same multilocus genotype, in such highly variable markers, is extremely low. The general aims of this study were as follows: (i) with respect to the archipelago, to investigate the level of genetic diversity of *S. aria* from the Canary Islands, determining the pattern of genetic differentiation and the existing genetic flow between populations where the species is located (the islands of La Palma and Tenerife); (ii) to obtain essential information to enable an exploration of the genetic relationship between populations of *S. aria* from the Canary Islands and the Iberian Peninsula; (iii) with respect to the Iberian Peninsula, to ascertain the genetic relationships between various locations separated by hundreds of kilometres; and (iv) to use molecular information to help in assessing the protection status of *S. aria* and to contribute to the formulation of appropriate management and conservation strategies in Spain's ,national parks, specially in the Canary Islands, where the species is threatened.

Materials and methods

Sample collection

A total of 184 individuals of *S. aria* belonging to seven natural populations were collected (Fig. 1 and Table 1). Four populations in the Iberian Peninsula were sampled, which corresponded to a representation of the geographical distribution of the species. All were located in high mountain habitats in the interior of Spanish national parks. An effort was made to include populations situated in both the north and south of the Iberian Peninsula, as well as sampling in those localities in the Iberian Peninsula closest to the Canarian Archipelago (Fig. 1 and Table 1). All samples were collected in Spanish national parks, which funded this research. The individuals were identified as *S. aria* according to the morphological characteristics of their leaves (laminas simple, entire, toothed or shallowly lobed, craspedodromous or camptodromous venation), previously identified as robust diagnostic characteristics (Aldasoro et al. 2004), and collected by specialist technicians from each national park between June 2007 and October 2008. The exact location for each sample studied was obtained (by GPS). They were mapped and located at a local scale. In Tenerife, the only four specimens described to date were sampled, three of which were found only a few metres from each other. In La Palma, two populations were studied: the first (La Palma 1) comprises an isolated and separated group located on terraces and slopes in the ravines of the national park of Caldera de Taburiente, and the second population (La Palma 2) is made up of a relatively dense nucleus of 64 specimens occupying an area of 25,000 m^2 in the Los Hombres ravine, north of the Caldera de Taburiente, at an altitude averaging 2,040 m.a.s.l. (Table 1). In Sierra Nevada, the samples were collected from isolated specimens, separated from each other by more than 25 m in different locations spanning an area of 57,000 km^2 of Sierra Nevada National Park. Due to technical problems, only six specimens from the Cabañeros population (CA) could be analysed. These specimens were collected from an area of approximately 1,500 m^2. In Aigüestortes (AI), the samples were collected in two adjacent areas separated by 7 km. Finally, in Picos de Europa (PE), the specimens were collected along a linear transect of 300 m, the greatest distance between specimens being 25 m (Table 1).

DNA extraction and microsatellite genotyping

DNA was extracted from silica gel-dried young leaves that were ground in a mixer mill (RETSCH MM 301, Haan, Germany) with two stainless steel balls in each microtube, following a modified 2× CTAB protocol (Doyle and Doyle 1987). About 150 μl of each total DNA sample was purified using a Genelute PCR Clean-Up Kit (Sigma-Aldrich, St. Louis, MO, USA) for further analysis.
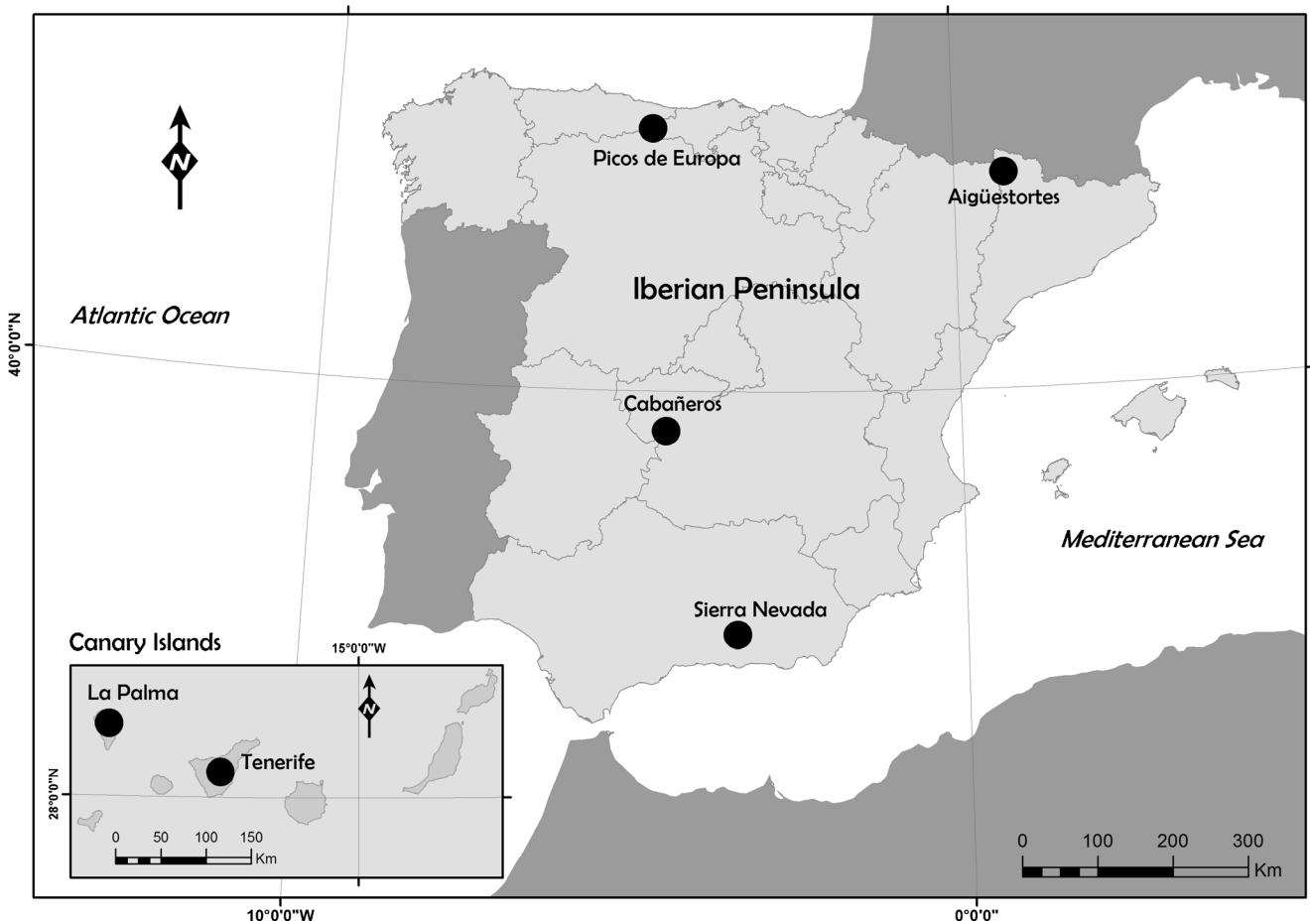
Fig. 1 Location of natural populations of *Sorbus aria* analysed

Forward and reverse primers specifically developed for *S. aria* by González-González et al. (2010) were used to amplify nine polymorphic microsatellite loci (SA-01, SA-02, SA-03, SA-06, SA-07, SA-08, SA-09, SA-14 and SA-19). Each

25 μL PCR reaction contained approximately 20 ng of DNA, 10 pmol of each primer as well as a PCR Master Mix (Reddy-Mix, ABgene, Surrey, UK) which included 0.625 units of Taq DNA polymerase, 75 mM Tris–HCl, 20 mM (NH4)$_2$SO$_4$,

Table 1 Genetic diversity indices obtained for nine microsatellite loci in *Sorbus aria* populations

Population	Code	Latitude	Longitude	Altitude	P	N	N_A	NPA	%PA	G	A	H_S
Tenerife	TF	341.571	3.122.681	2,354	3×	4	26	2	7.7	1	2.89	0.713
La Palma 1	LP1	218.509	3.185.157	2,267	3×	17	29	3	10.3	3	3.22	0.660
La Palma 2	LP2	219.124	3.184.805	2,263	3×	29	30	0	0.0	7	3.33	0.659
Canary Islands						50	35	7	20.0	-	4.22	0.677
Cabañeros	CA	363.705	4.377.333	1,300	3×	6	25	1	4.0	6	2.78	0.612
Aigüestortes	AI	317.335	4.719.635	1,682	2×	36	86	11	12.8	31	9.56	0.812
Picos de Europa	PE	345.737	4.778.456	1,181	2×	40	89	6	6.7	39	9.89	0.817
Sierra Nevada	SN	461.472	4.105.565	1,817	3×	52	65	8	12.3	20	7.22	0.755
Iberian Peninsula						134	135	36	26.7		14.78	0.860
Total						184	140		–		15.67	0.733

Latitude and longitude in UTM. Altitude in metres

P ploidy, N number of individuals sampled, N_A total number of different alleles seen across the loci, *NPA* number of private alleles, *%PA* percentage of private alleles, G number of different multilocus genotypes, A mean number of different alleles carried by each individual, averaged across loci, H_S expected heterozygosity within populations

0.01 % Tween 20, 2.5 mM of MgCl$_2$ and 0.2 mM of each dNTP. Amplifications were carried out using the following thermal cycling conditions: 3 min denaturation at 95°C; 35 cycles of 30 s denaturation at 95°C; 30 s at annealing temperature; and 90 s elongation at 72°C, followed by 5 min elongation at 72°C. These procedures are recorded in the Bank of Molecular Markers of Macaronesian Flora (www.banmac.ulpgc.es).

Ploidy

Amplified products were detected using an ABI 3130XL Genetic Analyzer, and fragment sizes were determined using GENESCAN v. 2.02 and GENOTYPER v. 1.1 (Applied Biosystems, Inc.). We identified allele peak profiles at each locus, and the presence of fragments was scored for each analysed individual. Since allele dosage determination in triploid individuals is complicated and difficult, we compared all multiloci genotypes following the method described by Robertson et al. (2010). Therefore, for each individual analysed in this study, microsatellite allele genotypes were determined for each of the nine loci. Ploidy level estimates were based on the maximum number of displayed alleles at a single locus. If three alleles were displayed for one locus, the individual was considered triploid. The number of different alleles detected in a triploid individual is often lower than its ploidy level. So, a triploid individual carrying alleles A and B may have two copies of either A or B and one of two different genotypes AAB or ABB. In this case, we included a missing value for this locus in the database. However, a triploid individual carrying only allele A will be AAA (Robertson et al. 2010). To estimate the number of triploids that have only one or two alleles, the known triploids were first used to calculate the proportion of triploids having three alleles at each microsatellite locus (pi) (Garner et al. 2008). This proportion was then used to calculate the probability that a triploid individual would receive three alleles at none of the nine loci $\left(\prod_{t=1}^{n} 1-\text{pi} = 5.17 \times 10^{-17}\right)$. This probability was then multiplied by the group size to determine the expected number of triploid not containing three alleles at any of the loci sampled (9.52×10^{-5}). So, we considered an individual as diploid when none of the nine loci analysed showed three alleles.

Statistical analysis

The quantification of genetic diversity for organisms with different ploidy (in this case, diploid and triploid; see "Results" section) can be difficult because some calculations are not included in the different software packages available. In our case, the most appropriate measure for studying genetic diversity in apomictic taxa is to compile and compare multilocus genotypes. As with other polyploid species of *Sorbus* (Robertson et al. 2010), we have included the following as measures: total number of different alleles seen across the loci (N_A); mean number of different alleles carried by each individual, averaged across loci (A); number of private alleles detected (NPA); percentage of private alleles (%NPA); number of different multiloci genotypes (G); and expected heterozygosity or gene diversity (H_S). All of these were calculated using GenoDive 2.0B22 software (Meirmans and Van Tienderen 2004), a programme specifically designed to handle polyploid data.

Recombination, hybridization, gene conversion and gene transfer all lead to histories that are not adequately modelled by a single tree, due to the imposition of a bifurcating (Neighbour-Joining) topology, which may be misleading (Bryant and Moulton 2004). In these cases, networks can provide a valuable tool for visualising a space of feasible trees. So, the genetic similarity matrices (Nei and Li 1985) were obtained with RAPDPLOT 2.3 (Black 1995), input into SplitsTree version 4.8 (Huson and Bryant 2006) and used to generate a Neighbour-Net network (Bryant and Moulton 2004) between the individuals analysed.

The genetic differentiation between populations pair-wise was analysed using F_{ST} implemented in GenoDive 2.0B20 (Meirmans and Van Tienderen 2004). Also, principal coordinate analysis (PCoA) was performed to establish the genetic relationship between individuals from each population using GENALEX version 6.4 (Peakall and Smouse 2006).

In addition, population structure was also inferred using a Bayesian clustering procedure implemented in STRUCTURE 2.2 (Falush et al. 2007) to identify the origin of the sampled individuals and to assign them to different populations. The most likely value of K is assessed by comparing the likelihood of the data for different values of K. We assumed an admixture model of population and independent allele frequencies. Analysis consisted of 10^5 burn-in period replicated and a run length of 10^6 replicated. We conducted a series of five independent runs for each value of K between 1 and 10. Individuals were assigned to one cluster if their proportion of membership (qi) to that cluster was equal to or greater than 0.05.

Results

Ploidy of *S. aria*

Examination of the microsatellite profiles showed that most *S. aria* individuals have up to three alleles per locus. PCR products of three loci of three individuals were cloned in order to confirm the polyploidy status of the species. These analyses were described in a previous paper (González-González et al. 2010), and the cloned sequences confirmed the polyploidy

status of the plants. All the individuals from the Canary Island (LP1, LP2, TF), CA and Sierra Nevada (SN) populations were triploids (three alleles were detected in at least one locus). All individuals analysed from AI and PE had only one or two peaks in all loci, and these populations were estimated to be diploid (Table 1). We considered an individual as diploid when none of the nine loci analysed showed three alleles (see "Material and methods" section).

Genetic diversity level

One hundred forty alleles were recorded from the nine polymorphic microsatellites analysed in *S. aria*. At a population level, the average number of alleles ranged from 2.78 for the Cabañeros to 9.89 for the PE populations. The average number of alleles per locus was, in general, higher for the Iberian Peninsula ($A = 14.78$) than for the Canary Islands ($A = 4.22$). On the whole, expected heterozygosities (or gene diversity) were high, which ranged from $H_S = 0.612$ for CA to $H_S = 0.817$ for the population of PE (Table 1).

The population with the greatest number of private alleles was Aigüestortes (11). Only seven alleles out of 35 were exclusive from some of the populations in the Canary Islands (Table 1). We also found two private alleles from Tenerife, despite the majority of individuals showing the same multiloci genotype (Table 1). La Palma 2 (LP2) was the population with the least amount of private alleles (none). AI and PE shared 64 of the 112 alleles recorded in both populations, which shows their strong genetic connection.

The majority of the triploid populations showed a determined number of individuals that had the same multiloci genotype, although with a different frequency. Thus, the greatest percentage of individuals with the same multiloci genotype was detected in the Canary Island populations.

Genetic differentiation between populations

Genetic differentiation statistics F_{ST} (and its analogues) is appropriate to testing the gene flow because of its capacity to describe the genetic population structure and, therefore, the degree of isolation (or connection) between populations (González-Pérez et al. 2009a). Average genetic differentiation between populations (F_{ST}) was low although significant, ranging from 0.022 between the two La Palma populations and 0.284 between the Tenerife and Cabañeros populations (Table 2). The genetic differentiation between the Peninsular populations was almost double ($F_{ST}=0.115$) than that found between the Canarian populations ($F_{ST}=0.062$), increasing to a value of $F_{ST}=0.150$ when comparing the populations from the peninsula and the archipelago.

The split network computed from the similarity matrix displayed high levels of reticulation (Fig. 2). The network clearly separated the individuals from the Canary Islands and the Iberian Peninsula. Also, individuals from the island of La Palma were clustered together (Fig. 2). The network shows a clear separation of the Sierra Nevada individuals into three differentiated groups, one of which was found to be very close to the Canary Island populations. On the other hand, Aigüestortes and Picos de Europa were grouped very closely.

Consistent with these results, all the samples from the Canarian Archipelago were located separately in the same cluster in the principal coordinate analysis, with small differences between the specimens from Tenerife and La Palma (Fig. 3). The individuals from the northern Peninsula locations (Aigüestortes and Picos de Europa) were again clustered, which indicates their genetic closeness. However, the specimens from Sierra Nevada were divided into at least three subgroups: the first, formed of approximately half of the individuals, were very close to the Canarian populations (SN-1, Fig. 3), the second was separated considerably, forming one isolated group (SN-2), while the rest of the individuals (SN-3) were positioned more closely to the CA population (Fig. 3).

For the Bayesian analysis, using the total data set (184 individuals, 7 localities, 9 microsatellite loci) and $K=1$–10, the probability of the data was maximum with $K=5$ (Fig. 4). The first group included those individuals from the Canary Islands. The second inferred group clustered those individuals from CA. All individuals of *S. aria* from AI and PE could be attributed to cluster III, whereas samples from SN again showed signs of admixture between the inferred clusters IV and V, which exactly coincides with the results obtained in the PCoA and Neighbour-Net network.

Discussion

The genetic diversity of *S. aria* has been analysed in different parts of Europe (Chester et al. 2007; Robertson et al. 2010; Kucerová and Gömöry 2011) but always in the context of comparing it with other *Sorbus* species. It has never been studied on the southern limit of its distribution. Our results show that we are dealing with a *S. aria* complex, consisting of individuals and populations with different ploidy, and therefore with different evolutionary histories and even possibly with different reproductive behaviours, as has been described extensively in the *Sorbus* genus by many other authors (Nelson-Jones et al. 2002; Robertson et al. 2004; 2010; Oria de Rueda et al. 2006; Lepší et al. 2008; Del Río et al. 2009).

The Canary Island populations

The Canarian populations show low levels of genetic variability, which is reflected in all the indices used (Table 1). At least three specimens of *S. aria* from Tenerife (and possibly the

Table 2 F_{ST} values from pairwise comparisons between *Sorbus aria* populations

	Tenerife (TF)	La Palma 1 (LP1)	La Palma 2 (LP2)	Cabañeros (CA)	Aigüestortes (AI)	Picos Europa (PE)
La Palma 1 (LP1)	0.130					
La Palma 2 (LP2)	0.126	0.022				
Cabañeros (CA)	0.284	0.267	0.258			
Aigüestortes (AI)	0.111	0.131	0.139	0.098		
Picos Europa (PE)	0.105	0.126	0.131	0.093	0.023	
Sierra Nevada (SN)	0.169	0.165	0.161	0.156	0.104	0.091

All values are significant at $p < 0.001$ level

fourth individual also, but we have missing data in two loci) present the same multiloci genotype without any differences and only seven different genotypes were detected among the 46 individuals analysed from the island of La Palma. In addition, almost 80 % of these specimens from La Palma belong to only three different genotypes.

S. aria in the Canary Islands, with a mean expected heterozygosity value of 0.677 (H_S), is found in the described range for other autochthonous polyploid taxa from the Canaries (González-Pérez et al. 2009b). Moreover, in all cluster analyses performed (Figs. 2 and 3), the Canarian populations are grouped in a cluster clearly separated from the rest of the populations from the Iberian Peninsula. The low levels of genetic variability detected could be due to a combination of a founder effect, followed by the genetic drift incurred in the populations from both islands with an isolated and restricted distribution and a low number of individuals. All this is amplified by the possibility of mostly asexual reproduction, via agamospermy, consequence of its triploid character, which has resulted in an impoverished gene pool. These findings agree with those described in other polyploid populations of the complex *S. aria* and congeneric species (Robertson et al. 2004, 2010; Rasmussen and Kollmann 2008), as well as in other endemic trees in the Canary Islands (García-Verdugo et al. 2009).

The high numbers of individuals with the same multiloci genotype, together with the large number of shared alleles in both populations (76.9 % of the alleles found in Tenerife are
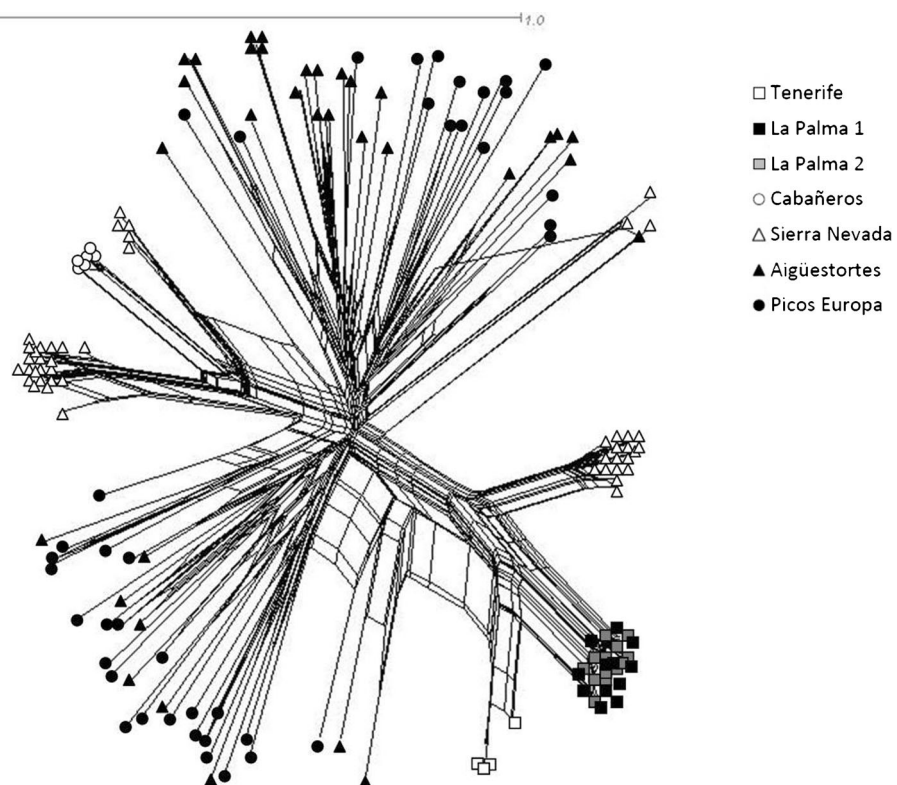

Fig. 2 Individual-based Neighbour-Net of *Sorbus aria*, based on pairwise genetic distances complementary to Nei and Li (1985). The *length* of the edge representing a given split is an indication of its support

Fig. 3 Principal coordinate analysis (PCoA) based on pairwise Nei and Li (1985) genetic similarity among *Sorbus aria* individuals. SN-1, SN-2 and SN-3 constitute the three genetic subgroups detected for Sierra Nevada population (see text for details)
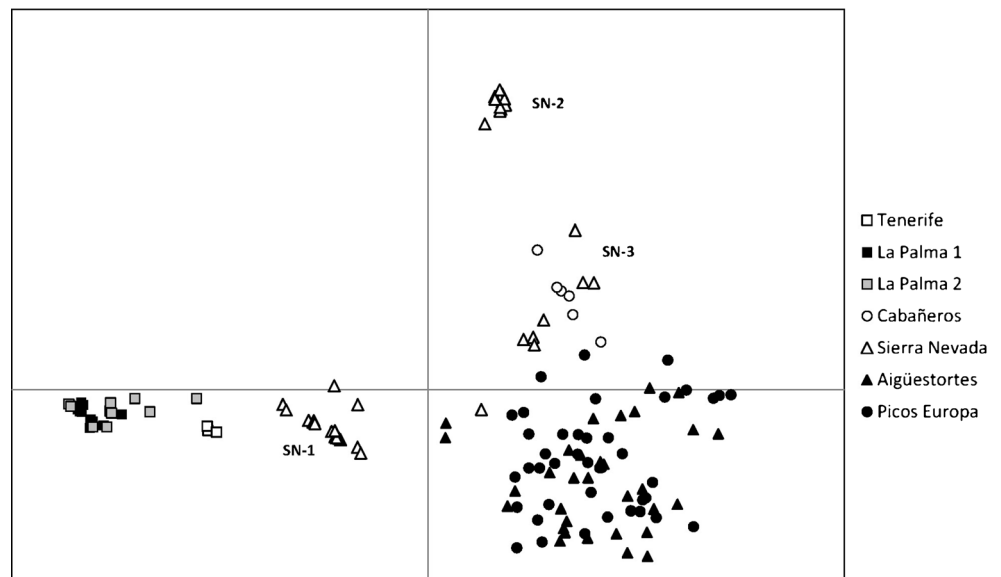
also found in La Palma), suggest that the *S. aria* populations from the Canarian Archipelago could have originated from a single colonisation event with very few effectives. The four individuals from the Teide National Park (Tenerife) could be the last remnants of a more numerous population of the species that was completely decimated and isolated. It would be feasible that the populations could have been affected by volcanic eruptions from the Teide volcano, such as Pico Viejo (17,000 years ago) or Montaña de La Cruz (3,500 years ago), resulting in a general loss of specimens in relatively recent times. However, as described above, at least three of these individuals present the same multiloci genotype, and it is also possible that these four trees are the lone stems of a genet originating from apomictic seeds produced by the same mother tree, which colonised this area recently.

The two populations from La Palma, more numerous and with greater levels of genetic variability, also share many alleles, including various multiloci genotypes, and present a low differentiation coefficient between them ($F_{ST}=0.022$). Although this genetic closeness could be indicative of the high level of genetic flow between populations, we believe that is a reflection of the fact that the majority (or all) of the specimens of *S. aria* from La Palma form a part of a single population with a monophyletic origin. We believe that the existence of certain differences among multilocus genotypes could be associated with variation generated by mutation rather than independent colonisation events. Eighty percent of the specimens from La Palma belonged to only three different genotypes, a pattern that is much easier to explain by multiple mutation events, and enhanced by probable

Fig. 4 Bar plot for the proportion of coancestry inferred from Bayesian cluster analysis ($K=5$). Population codes as in Table 1
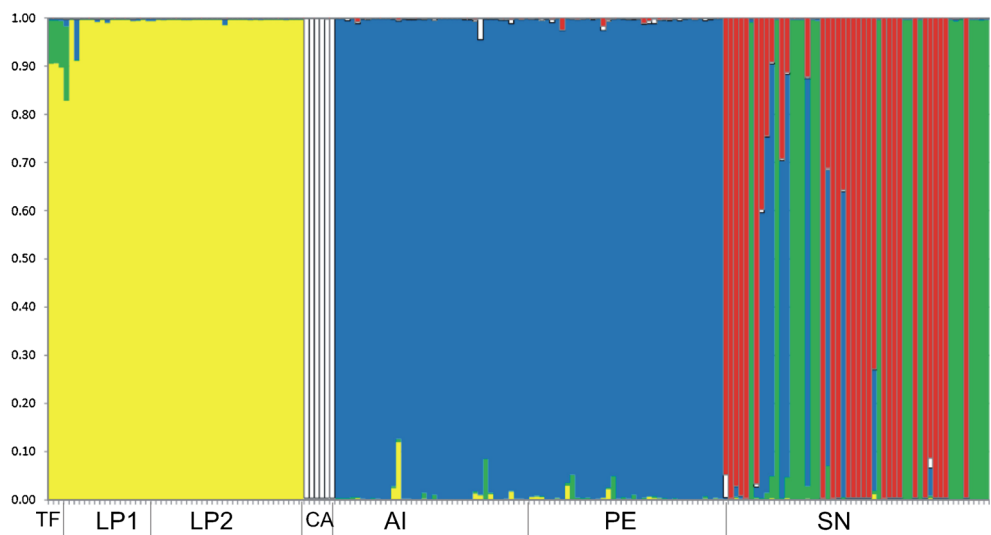
asexual reproduction in most if not all cases, by apomixis. Similar scenarios have been described widely for other *Sorbus* species (Robertson et al. 2004, 2010; Vít et al. 2012). We are probably, therefore, looking at a native non-endemic species shared between North Africa and the Mediterranean that colonised the Canary Islands in a single dispersal event.

What could be the origin of *S. aria* in the Canary Islands? We can dismiss the idea that the specimens from the archipelago come from a hybridization process between different species of the genus because other species of *Sorbus* have not been reported in the Canary Islands (Arechavaleta et al. 2010) nor, up to now, have related fossils been found. The flora of the Archipelago has a strong North African and Mediterranean component (Fernández-Palacios et al. 2011), and, indeed, *S. aria* is also found in the north of the African continent (Aldasoro et al. 2004), which makes this the most likely place of origin. Unfortunately, we were unable to obtain samples from this area to confirm this assertion. Nevertheless, we did find a significant relationship with one differentiated subgroup from the Sierra Nevada (SN-1 in Fig. 3), which shares 51.4 % of the alleles present in the Canarian populations. These results indicate that the Canarian populations could proceed from triploid specimens whose closest ancestor is found in the south of the Iberian Peninsula. So, *S. aria* could have taken refuge in the Canary Islands where the impact of the climate was buffered, maintaining a wide distribution within the Mediterranean basin, the Canary Islands and North Africa (Fernández-Palacios et al. 2011). Despite this, as we have already pointed out that samples from Morocco have not been studied, the results should be considered with caution.

Iberian Peninsula populations

Iberian Peninsula populations presented, in general, higher levels of genetic variability, with the exception of CA, probably due to the low number of samples analysed. Furthermore, the degree of genetic differentiation between the continental populations was variable, depending on the pairwise populations compared, the diploid populations (Picos de Europa and Aigüestortes) being much more closely related. Outcrossing perennial plant species generally exhibit higher levels of genetic variability and lower levels of population differentiation, indicating that the reproductive biology and mating system are the most important factors in determining the genetic structure of plant populations (Hamrick and Godt 1996; Duminil et al. 2007; Sosa et al. 2010).

Aigüestortes and Picos de Europa exhibited the highest indices of genetic variability (Table 1), possibly due to their ability to recombine genotypes between individuals. The highest average number of alleles at each locus was detected in these populations, and almost every individual (with the exception of a few specimens) has a unique multilocus

genotype (Table 1). Moreover, both populations were in Hardy-Weinberg equilibrium (data not shown) as predicted for outcrossing taxa. They also share a high number of alleles (74.4 %) and show the lowest genetic differentiation index of the established comparisons in the Peninsula (F_{ST}=0.023; Table 2), and all the cluster analyses (Neighbour-Net (Fig. 2), PCoA (Fig. 3) and Bayesian analysis (Fig. 4)) indicate that they are genetically very closely related.

The individuals of these populations are diploid (*S. aria* sensu stricto) and therefore outcrossing, as described by numerous authors (Nelson-Jones et al. 2002; Oria de Rueda et al. 2006; Del Río et al. 2009; Robertson et al. 2010). High capacity for long distance gene flow by both pollen dispersal by generalist insects and seed dispersal carried out by passerine birds and fruit-spreading by mammals have been widely described in other diploid species of *Sorbus* (Del Río et al. 2009; Oria de Rueda et al. 2006; Angelone et al. 2007; Klein et al. 2008; Robertson et al. 2010).

However, the high genetic diversity detected in the Sierra Nevada seems, a priori, incompatible with clonal reproduction, considering that its individuals are triploid and contain genomes in which individual loci are not free to recombine. The association between polyploidy and apomixis has been widely noted, and it has been suggested that polyploidy acts either as a stimulus for apomixis, is a prerequisite for its maintenance or both (Soltis et al. 2004; Whitton et al. 2008; Hörandl 2010; Robertson et al. 2010). In addition, the specimens from the Sierra Nevada were found to be strongly genetically differentiated; the value of genetic differentiation between the subgroups SN-1, SN-2 and SN-3 (F_{ST}=0.240) being considerably greater than all the pairs of comparisons made between populations. In addition, there is no geographic separation between these individuals, and many of the specimens from the genetic sub-groups are separated by only a few metres.

What these results, Network, PCoA and Bayesian analysis (Figs. 1, 2 and 3), show is the polyphyletic nature of the polyploids within the Sierra Nevada population. It is possible that this population has been driven primarily by a series of interspecific hybridisations and backcrosses between closely related taxa, with new genotypes being fixed and propagated via apomictic reproduction (20 different multilocus genotypes were identified in 52 individuals analysed).

The observation supporting multiple origins for these subgroups, rather than mutation in the Sierra Nevada population, is that the individuals belonging to each of the three subgroups differ considerably in the alleles of the majority loci, a pattern which is much easier to explain by independent hybrid origins than by multiple mutation events. However, the differences within each subgroup are few; for example, subgroups SN-1 and SN-3 were made up of only five and nine different multilocus genotypes, respectively (data not shown). The result is a scenario derived from a dynamic evolutionary

complex which may proceed from multiple hybridizations and backcrosses between closely related taxa, where *S. aria* is present, a finding that has been described on numerous occasions in the *Sorbus* genus (Aldasoro et al. 2004; Nelson-Jones et al. 2002; Robertson et al. 2010; Vít et al. 2012). Determining whether or not these distinct genetic pools are different taxa (or species) is not the objective of this work, but it opens new doors for further study.

The delimitation of the species in *S. aria* sensu lato is extremely difficult and is not clearly distinguished by morphological features (Aldasoro et al. 2004; Lepší et al. 2008). The consideration that the agamosperm microspecies could be considered an actual species has been discussed by different authors (Aldasoro et al. 2004; Lepší et al. 2008), and there are yet other authors who separate the products of hybridization events as independent taxa (Lepší et al. 2008).

Conclusions and implications for conservation

In conclusion, we have described the divergent behaviour of the polyploid (and possibly apomictic) populations *S. aria* sensu lato in the south of the Iberian Peninsula and the Canary Islands from that of the diploid populations (with possible sexual reproduction) in the north of the Iberian Peninsula.

The occurrence of only triploid individuals in populations from the Canary Islands and the large proportion of genetically identical samples on both the islands of Tenerife and La Palma make it likely that the reproductive processes of this species in the Archipelago involve apomixis (agamospermy). However, an analysis between parents and their offspring will be needed to reveal if seed production utilises obligate apomixis in every population or whether, in addition to apomixis, populations also resort to sexual reproduction (facultative apomixis). If facultative apomixis occurs in the Canarian populations, interpopulation reinforcement would be feasible within La Palma, where a certain degree of genetic diversity has been found. This would be achieved through controlled crosses or by the use of seedlings from germplasm banks with different multilocus genotypes, avoiding the collection of clone specimens. In the case of Tenerife, and given that there is a high number of clonal specimens, it would be very important to initiate a reproductive process.

In this study, basing our conclusions exclusively on molecular results, we found three genetically differentiated entities with different gene pools in the south of the Iberian Peninsula (the Sierra Nevada population). For this population, we could be revealing the presence of different microspecies of *Sorbus*, which are extremely difficult to distinguish by morphological differences, which must be considered in a possible conservation programme.

With this information, we hope to contribute to the development of a more accurate conservation policy for *S. aria* in the Spanish national parks that both takes into account its evolutionary past and facilitates its future evolution.

References

Acevedo A, Diaz M, Mentrel D, Calero C (2011) Seguimiento de las poblaciones naturales de *Sorbus aria* (L.) Crantz en el Parque Nacional de Taburiente. Parque Nacional de la Caldera de Taburiente. Gobierno de Canarias. Unpublished report

Aldasoro JJ, Aedo C, Garmendia FM, Pando de la Hoz F, Navarro C (2004) Revision of *Sorbus* subgenera Aria and Torminaria (Maloideae–Rosaceae). Syst Bot Monogr 69:148

Angelone S, Hilfiker K, Holderegger R, Bergamini A, Hoebee SE (2007) Regional population dynamics define the local genetic structure in *Sorbus torminalis*. Mol Ecol 16:1291–1301

Arechavaleta M, Rodríguez S, Zurita N, García A (2010) Lista de especies silvestres de Canarias. Hongos, plantas y animales terrestres. Gobierno de Canarias, Santa Cruz de Tenerife

Black IVWC (1995) FORTRAN programs for the analysis of RAPD-PCR markers in populations. Colorado State University, Colorado

Bryant D, Moulton V (2004) Neighbour-net: an agglomerative method for the construction of phylogenetic networks. Mol Biol Evol 21:255–265

Chester M, Cowan RS, Fay MF, Rich TCG (2007) Parentage of endemics *Sorbus* L. (Rosaceae) species in the British Isles: evidence from plasmid DNA. Bot J Linn Soc 154:291–304

Del Río J, Martínez de Azagra A, Oria de Rueda JA (2009) Ecología del paisaje del género *Sorbus* L. en la Península Ibérica y en Baleares. Ecología 22:25–44

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Duminil J, Fineschi S, Hampe A, Jordano P, Salvini D, Vendramín GG, Petit RJ (2007) Can population genetic structure be predicted from life-history traits? Am Nat 169:662–672

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data, dominant markers and null allele. Mol Ecol Notes 7:574–578

Fernández-Palacios JM, de Nascimento L, Otto R, Delgado JD, García-del-Rey E, Ramón-Arévalo J, Whittaker RJ (2011) A reconstruction of Paleo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. J Biogeogr 38:226–246

García-Verdugo C, Fay MF, Granado-Yela C, Rubio de Casas R, Balaguer L, Besmard G, Vargas P (2009) Genetic diversity and differentiation processes in the ploidy series of *Olea europea* L.: a

multiscale approach from subspecies to insular populations. Mol Ecol 18:454–467

Garner SR, Madison BN, Bernier NJ, Neff BD (2008) Juvenile growth and aggression in diploid and triploid Chinook salmon *Oncorchynchus tshawytscha* (Walbaum). J Fish Biol 73:169–185

Gömory D, Krajmerová D (2008) Spatial structure of a natural mixed topodeme of subalpine *Sorbus* taxa. Acta Soc Bot Pol 77:305–311

González-González EA, González-Pérez MA, Rivero E, Sosa PA (2010) Isolation and characterization of microsatellite loci in *Sorbus aria* (Rosaceae). Conserv Genet Resour 2:341–343

González-Pérez MA, Sosa PA, Rivero E, González-González EA, Naranjo A (2009a) Molecular markers reveal no genetic differentiation between *Myrica rivas-martinezii* and *M. faya* (Myricaceae). Ann Bot 103:79–86

González-Pérez MA, Lledó MD, Lexer C, Fay M, Marrero M, Bañares A, Carqué E, Sosa PA (2009b) Genetic diversity and differentiation in natural and reintroduced populations of *Bencomia exstipulata* and comparisons with *B. caudata* (Rosaceae) in the Canary Islands: an analysis using microsatellites. Bot J Linn Soc 160:429–441

Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. Philos Trans R Soc 351:1291–1298

Hörandl E (2010) The evolution of self-fertility in apomictic plants. Sex Plant Reprod 23:73–86

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23:254–267

Klein EK, Desassis N, Oddou-Muratorio S (2008) Pollen flow in the wildservice tree, *Sorbus torminalis* (L.) Crantz. IV. Whole interindividual variance of male fecundity estimated jointly with the dispersal kernel. Mol Ecol 17:3323–3336

Kucerová V, Gömöry D (2011) Reticulate phylogeny in the genus *Sorbus*: the case of *Sorbus haljamovae* Bernátová et Májovský. Folia Oecol 38:169–175

Lepší M, Vít P, Lepší P, Blublík K, Suda J (2008) *Sorbus milensis*, a new hybridgenous species from Northwestern Bohemia. Preslia 80:229–244

Meirmans PG, Van Tienderen PH (2004) GenoType and GenoDive: two programs for analysis of genetic diversity of asexual organisms. Mol Ecol Notes 4:792–794

Nei M, Li WH (1985) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci U S A 76:5269–5273

Nelson-Jones EB, Briggsm D, Smith AG (2002) The origin of intermediate species of the genus *Sorbus*. Theor Appl Genet 105:953–963

Oria de Rueda JA, de Azagra M, Paredes A, Álvarez Nieto A (2006) Botánica forestal del género *Sorbus* en España. Investig Agrar: Sist Recur For 15:166–186

Peakall R, Smouse PE (2006) GENALEX 6.4. Genetic Analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Rasmussen KK, Kollmann J (2008) Low genetic diversity in small peripheral populations of a rare European tree (*Sorbus torminalis*) dominated by clonal reproduction. Conserv Genet 9:1533–1539

Robertson A, Newton AC, Ennos RA (2004) Multiple hybrid origins, genetic diversity and population genetic structure of two endemic *Sorbus* taxa on the Isle of Arran, Scotland. Mol Ecol 13:123–134

Robertson A, Rich TCG, Allen AM, Houston L, Roberts C, Bridle JR, Harris SA, Hiscock SJ (2010) Hybrization and polyploidy as drivers of continuing evolution and speciation in *Sorbus*. Mol Ecol 19:1675–1690

Soltis DE, Soltis PS, Tate JA (2004) Advances in the study of polyploidy since plant speciation. New Phytol 161:173–191

Sosa PA, González-Pérez MA, Moreno C, Clarke JB (2010) Conservation of the endangered endemic *Sambucus palmensis* Link (Sambucaceae) from the Canary Islands. Conserv Genet 11:2357–2368

Vít P, Lepŝí M, Lepŝí P (2012) There is no diploid apomictic among Czech *Sorbus* species: a biosystematics revision of *S. eximia* and discovery of *S. barrandienica*. Preslia 84:71–96

Whitton J, Sears CJ, Baack EJ, Otto SP (2008) The dynamic nature of apomixis in the angiosperm. Int J Plant Sci 169:169–182