Population size, center–periphery, and seed dispersers’ effects on the genetic diversity and population structure of the Mediterranean relict shrub *Cneorum tricoccon*

Ana Lázaro-Nogal, Silvia Matesanz, Alfredo García-Fernández, Anna Traveset, Fernando Valladares

Abstract

The effect of population size on population genetic diversity and structure has rarely been studied jointly with other factors such as the position of a population within the species’ distribution range or the presence of mutualistic partners influencing dispersal. Understanding these determining factors for genetic variation is critical for conservation of relict plants that are generally suffering from genetic deterioration. Working with 16 populations of the vulnerable relict shrub *Cneorum tricoccon* throughout the majority of its western Mediterranean distribution range, and using nine polymorphic microsatellite markers, we examined the effects of periphery (peripheral vs. central), population size (large vs. small), and seed disperser (introduced carnivores vs. endemic lizards) on the genetic diversity and population structure of the species. Contrasting genetic variation ($H_E: 0.04–0.476$) was found across populations. Peripheral populations showed lower genetic diversity, but this was dependent on population size. Large peripheral populations showed high levels of genetic diversity, whereas small central populations were less diverse. Significant isolation by distance was detected, indicating that the effect of long-distance gene flow is limited relative to that of genetic drift, probably due to high selfing rates ($F_{IS} = 0.155–0.887$), restricted pollen flow, and ineffective seed dispersal. Bayesian clustering also supported the strong population differentiation and highly fragmented structure. Contrary to expectations, the type of disperser showed no significant effect on either population genetic diversity or structure. Our results challenge the idea of an effect of periphery *per se* that can be mainly explained by population size, drawing attention to the need of integrative approaches considering different determinants of genetic variation. Furthermore, the very low genetic diversity observed in several small populations and the strong among-population differentiation highlight the conservation value of large populations throughout the species’ range, particularly in light of climate change and direct human threats.

**KEYWORDS**
genetic diversity, islands, marginal populations, population size, relict plants, seed dispersal
Relict species are generally expected to show low levels of genetic diversity and strong population structure (Awad, Fady, Khater, Roig, & Cheddadi, 2014; Bauert, Kalin, Baltisberger, & Edwards, 1998; Ge et al., 2005) because they are usually composed of small, geographically isolated populations (Dobrowski, 2011; Hampe & Petit, 2005). However, experimental evidence has not always supported this generalization (see Peakall, Ebert, Scott, Meagher, & Offord, 2003; VandenBroeck et al., 2011). Low rates of evolution, genetic stability, and habitat reliability have been proposed as some of the factors that may influence the genetic patterns of relict species (Hampe & Petit, 2007). Information on genetic diversity and population structure—and their determinant factors—along a species’ distribution range has shown to be critical for designing conservation strategies for endangered relict species (Lesica & Allendorf, 1995).

Differences in genetic diversity and structure between peripheral and central populations have long been studied (e.g., Arnaud-Haond et al., 2006; Durka, 1999; Eckert, Samis, & Lougheed, 2008; Eckstein, O’Neill, Danihelka, Otte, & Köhler, 2006; Lammi, Siikamäki, & Mustajärvi, 1999; Pironon et al., 2016; Van Rossum, Vekemans, Gratia, & Meerts, 2003), but a clear pattern has not yet been found. It is often assumed that peripheral populations are small, isolated, and occur in ecologically marginal habitats where selection pressures are likely to be more intense (Brown, Stevens, & Kaufman, 1996; Eckert et al., 2008; Lawton, 1993; Lesica & Allendorf, 1995; Pulliam, 2000). Such populations can have low genetic diversity as a consequence of high inbreeding, genetic drift, and directional selection and may also show strong genetic structure due to reduced gene flow (Arnaud-Haond et al., 2006; Durka, 1999; Gapare & Aitken, 2005; Lammi et al., 1999; Schaal & Leverich, 1996). However, it is not known to what extent the effects of periphery are confined by those of population size. Populations established in peripheral areas can be smaller, equally large (see e.g., Dolan, 1994; Van Rossum et al., 2003) or even larger than central ones (see e.g., Kluth & Bruelheide, 2005), which could modulate the effect of periphery. Therefore, genetic diversity of peripheral populations could be expected to be similar to that of central ones. Accordingly, central distributions with small populations might have lower genetic diversity. Despite these implications, to our knowledge, the effects of population size on genetic diversity have been rarely studied in conjunction with other key factors, such as location within the species distribution range (i.e., center–periphery; see Pironon et al., 2016).

In the case of Mediterranean species, reproductive system, colonization success, and dispersal abilities have also been recognized as important aspects that could affect species’ evolution (Feliner, 2014) as well as shape their genetic diversity. For instance, species with long-range seed dispersal mechanisms, such as zoochory, often have higher within-population genetic diversity and lower population differentiation than species with limited dispersal (Avise, 2004; Hamrick, Murawski, & Nason, 1993; Vanden-Broeck et al., 2011). Therefore, the loss of local vertebrate frugivores can have detrimental genetic effects on plants that depend upon them (Babweteera & Brown, 2009; Christian, 2001). Seed dispersal disruptions can impact seed removal success and regeneration (Riera, Traveset, & García, 2002; Traveset, 1995b; Traveset, Gonzalez-Varo, & Valido, 2012) and may lead to the loss of genetic variation and inbreeding due to increased drift and clumping (Jordano et al., 2011). However, the genetic consequences of seed disperser loss have so far been poorly investigated (but see Calviño-Cancela et al., 2012; Pérez-Méndez, Jordano, & Valido, 2015), particularly in the face of other key aspects such as population size and center–periphery effects.

The Mediterranean basin is an important area for relict species, including those that emerged during the Tertiary (Palamarev, 1989). Many of these Tertiary relicts have evolved key traits (e.g., long life cycles, sprouting, seed bank development, and fleshy fruits for seed dispersion) to survive and enhance their performance (see revision in Rodríguez-Sánchez, Perez-Barrales, Ojeda, Vargas, & Arroyo, 2008). These adaptations, together with the climate oscillations during the Quaternary, the changes in sea level, and the tectonic movements around the Mediterranean Sea are some of the processes that have governed the complexity of the Mediterranean flora, making it a biodiversity hot spot (Myers, Mittermeier, Mittermeier, Da Fonseca, & Kent, 2000). However, these processes also complicate the formation of common patterns that could explain the biogeographic distributions of Mediterranean species and the test of general hypotheses (Feliner, 2014), such as the center (abundant)–periphery (marginal) rule (Pironon et al., 2016; Sagarin & Gaines, 2002).

*Cneorum tricoccon* L. (Figure 1) is a Mediterranean relict shrub of limited and regressive distribution (Lázaro-Nogal, Forner, Traveset, & Valladares, 2013; Traveset, 1995a). It is endemic to the western Mediterranean area, the Balearic Islands being their main distribution area. Isolated peripheral populations are found in the Iberian Peninsula, France, Sardinia, and Tuscany. *Cneorum tricoccon* has a tight seed dispersal mutualism with endemic lizards, but the introduction of carnivorous mammals in the Balearic Islands contributed to their extinction (particularly from the islands of Mallorca and Menorca) and, consequently, disrupted seed dispersal in this and other systems (Riera et al., 2002; Traveset, 1995b). Currently, in populations where lizards...
are extinct, some of those carnivorous mammals (mainly pine martens Martes martes L.) replace them to some extent as seed dispersers (Celedón-Neghme, Traveset, & Calviño-Cancela, 2013), but its consequences for the population genetics of C. tricoccon are still unknown.

Our aim in this study was to assess genetic diversity and structure of C. tricoccon populations encompassing the entire distribution range of the plant, using microsatellite markers. We hypothesized that there will be differences in genetic diversity between populations of similar size located in the center versus the periphery of the species range. Moreover, we predicted that populations where native dispersers are missing would also show different genetic patterns than those in which native dispersers still exist. We selected a set of 16 populations of contrasting population size, periphery, and type of seed disperser. Specifically, we addressed the following questions: (1) Are there differences in genetic diversity and structure among populations in relation to the population’s position in the distribution range (i.e., center vs. periphery) or population size (large vs. small)? (2) Are introduced mammals effective dispersers promoting greater genetic and structural diversity compared to the native lizards?

2 | METHODS

2.1 | Study species and population sampling

*Cneorum tricoccon* L. (Cneoraceae) is a perennial shrub approximately 1 m tall, although some individuals can reach up to 2 m. It is an andromonoecious, insect-pollinated, and self-compatible species (Traveset, 1995a). It represents a relic species of the Tertiary, which evolved under tropical climate conditions (Raven, 1973). It has been categorized as vulnerable (IUCN, International Union for Conservation of Nature) with human activities and climate change the major threats to the species’ survival. Its distribution has diminished in recent decades with some local extinctions (Traveset, Quintana, & Alcover, 2005). It inhabits the western Mediterranean area, with populations in the Balearic Islands considered the center of the species distribution range, and populations in the eastern and southern Iberian Peninsula, southeastern France, Sardinia, Giannutri, and Tuscany considered peripheral populations (Traveset, 1995a). In the Balearic Islands, seed dispersal is mostly mediated by endemic lizards (*Podarcis lilfordi*, *P. pityusensis*, and *P. siculus*). However, introduced carnivorous mammals (approximately 250 B.C.), such as pine martens (*M. martes*) or genets (*Genetta genetta* L.), are currently the main dispersers in localities where lizards are extinct (mainly Mallorca and Menorca; Traveset, 1995b; Celedón-Neghme et al., 2013).

A total of 399 individuals were sampled in 16 populations (Table 1, Figure 2) spanning the entire species’ distribution range. Selected populations covered contrasting levels of periphery (central vs. peripheral) and population size (large vs. small), and they differed in their main disperser (endemic lizards vs. introduced carnivorous mammals). To distinguish among peripheral and central populations, we followed a geographic selection criterion (Lesica & Allendorf, 1995; Pironon et al., 2016). We considered central populations as thriving populations located in the main species distribution (i.e., Balearic Islands) and peripheral as the other scant populations that appear in the mainland (Mediterranean coasts, Figure 2). We extensively sampled the extant populations located outside the main distribution area, and were able to include six of the seven extant peripheral populations in this study. To define the two levels of population size, we considered small populations those with <200 individuals and large populations those with >200 individuals (Frankham, 1996; Frankham, Briscoe, & Ballou, 2002), a relatively conservative threshold used in several studies (see e.g., Kery, Matthies, & Spillmann, 2000). Previous available field data and information on the species dispersal support our categorization (A. Traveset, personal communication). Complete information about the disperser type present in each population is available only for populations in the Balearic Islands. Consequently, we have tested the effect of this factor in a subsample of all populations, using a classification based on previous available data (Pérez-Mellado et al., 2008; Traveset, 1995b). Although there has not been any specific study on the dispersers in the peripheral populations, we know from personal observations and communications that carnivores, such as pine and stone martens, are the main frugivores in the mainland. Carnivore scats were occasionally found with *C. tricoccon* seeds in the Iberian Peninsula populations, and other studies have shown the importance of carnivores as seed dispersers for a large variety of fleshy fruited species (e.g., González-Varo, López-Bao, & Guitián, 2013).

We sampled 21–30 indiscriminately selected individuals within areas of about 1 km² per population. Leaves for DNA extraction (2–3) were collected and stored in bags with silica gel.

2.2 | DNA extraction, microsatellite analysis, and fragment scoring

Total genomic DNA was extracted from 20 mg of dry leaf tissue using the DNaseasy Plant Mini Kit (Qiagen, Valencia, CA, USA) following the manufacturer’s protocol. Its concentration and purity were quantified on a NanoDrop spectrophotometer (NanoDrop Products, Wilmington, DE, USA).

We genotyped each sample at nine microsatellite loci known to be polymorphic across the species, described in Garcia-Fernandez, Lazaro-Nogal, Traveset, and Valladares (2012): Ctric 00490, Ctric 02925, Ctric 07615, Ctric 14301, Ctric 06384, Ctric 10195, Ctric 19884, Ctric 15341, and Ctric 09344. An M13 tail was added to one of the PCR primers (Schuelke, 2000). PCRs were performed in a 10 µl mix containing 4.1 µl of RNase-free water, 1× Taq buffer (Biotools, Madrid, Spain), 2 mmol/L MgCl₂, 0.25 µmol/L of each dNTP, 0.04 µmol/L of the forward primer with M13 tail, 0.16 µmol/L of the reverse primer and 0.16 µmol/L of the universal fluorescently labeled M13 primer ([6-FAM, VIC, PET, and NED]), 0.5 U of Taq polymerase (Biotools), and 10 ng of template DNA. The PCR program consisted of one step of 4 min at 94°C followed by 30 cycles each of 30 s at 95°C, 45 s at 56°C, and 45 s at 72°C. Another eight cycles were then performed, consisting of 30 s at 95°C, 45 s at 53°C, and 45 s at 72°C, followed by a final step of 10 min at 72°C. A S1000 Thermal Cycler (Bio-Rad laboratories, Richmond, CA, USA) was used. Each
TABLE 1  Population code, location, and characterization of the 16 sampled Cneorum tricoccon populations

<table>
<thead>
<tr>
<th>Code</th>
<th>Population</th>
<th>Region</th>
<th>Lat</th>
<th>Long</th>
<th>Insularity</th>
<th>Pop Size</th>
<th>Disperser</th>
<th>Habitat</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA</td>
<td>S’Arboçar</td>
<td>Mallorca</td>
<td>39.678</td>
<td>2.544</td>
<td>LIs</td>
<td>L</td>
<td>Mammals</td>
<td>Oak forest</td>
</tr>
<tr>
<td>LL</td>
<td>Lluc</td>
<td>Mallorca</td>
<td>39.789</td>
<td>2.865</td>
<td>LIs</td>
<td>L</td>
<td>Mammals</td>
<td>Oak forest</td>
</tr>
<tr>
<td>BL</td>
<td>Cap Blanc</td>
<td>Mallorca</td>
<td>39.479</td>
<td>2.738</td>
<td>LIs</td>
<td>L</td>
<td>Mammals</td>
<td>Maquis</td>
</tr>
<tr>
<td>CO</td>
<td>Pla de Corona</td>
<td>Ibiza</td>
<td>39.044</td>
<td>1.333</td>
<td>SIs</td>
<td>L</td>
<td>Lizards</td>
<td>Maquis</td>
</tr>
<tr>
<td>TG</td>
<td>Tagomago</td>
<td>Ibiza</td>
<td>39.035</td>
<td>1.644</td>
<td>SIs</td>
<td>L</td>
<td>Lizards</td>
<td>Maquis</td>
</tr>
<tr>
<td>BB</td>
<td>Cap de Barbaria</td>
<td>Formentera</td>
<td>38.642</td>
<td>1.39</td>
<td>SIs</td>
<td>L</td>
<td>Lizards</td>
<td>Maquis</td>
</tr>
<tr>
<td>CA</td>
<td>Cabrera</td>
<td>Cabrera</td>
<td>39.150</td>
<td>2.953</td>
<td>SIs</td>
<td>S</td>
<td>Lizards</td>
<td>Maquis</td>
</tr>
<tr>
<td>CC</td>
<td>Conillera de Cabrera</td>
<td>Cabrera</td>
<td>39.182</td>
<td>2.962</td>
<td>SIs</td>
<td>S</td>
<td>Lizards</td>
<td>Maquis</td>
</tr>
<tr>
<td>DR</td>
<td>Dragonera</td>
<td>Dragonera</td>
<td>39.587</td>
<td>2.329</td>
<td>SIs</td>
<td>L</td>
<td>Lizards</td>
<td>Maquis</td>
</tr>
<tr>
<td>ME</td>
<td>Sa Mesquida</td>
<td>Menorca</td>
<td>39.917</td>
<td>4.287</td>
<td>LIs</td>
<td>S</td>
<td>Mammals</td>
<td>Maquis</td>
</tr>
</tbody>
</table>

Peripheral populations

<table>
<thead>
<tr>
<th>Code</th>
<th>Population</th>
<th>Region</th>
<th>Lat</th>
<th>Long</th>
<th>Insularity</th>
<th>Pop Size</th>
<th>Disperser</th>
<th>Habitat</th>
</tr>
</thead>
<tbody>
<tr>
<td>GI</td>
<td>Gianuttri</td>
<td>Italy</td>
<td>42.255</td>
<td>11.099</td>
<td>SIs</td>
<td>L</td>
<td>?</td>
<td>Maquis</td>
</tr>
<tr>
<td>CR</td>
<td>Cap de Creus</td>
<td>NE Spain</td>
<td>42.322</td>
<td>3.318</td>
<td>Con</td>
<td>S</td>
<td>?</td>
<td>Maquis</td>
</tr>
<tr>
<td>FI</td>
<td>Fitou</td>
<td>France</td>
<td>43.54</td>
<td>3.785</td>
<td>Con</td>
<td>L</td>
<td>?</td>
<td>Maquis</td>
</tr>
<tr>
<td>MO</td>
<td>Montpellier</td>
<td>France</td>
<td>42.892</td>
<td>2.978</td>
<td>Con</td>
<td>L</td>
<td>?</td>
<td>Maquis</td>
</tr>
<tr>
<td>CG</td>
<td>Cerro Gordo</td>
<td>S Spain</td>
<td>36.741</td>
<td>−3.777</td>
<td>Con</td>
<td>S</td>
<td>?</td>
<td>Maquis</td>
</tr>
<tr>
<td>MA</td>
<td>Fuente de Maro</td>
<td>S Spain</td>
<td>36.759</td>
<td>−.85</td>
<td>Con</td>
<td>S</td>
<td>?</td>
<td>Maquis</td>
</tr>
</tbody>
</table>

Lat, latitude; Long, longitude; LIs, large island; SIs, small island; Con, continent; L, large (>200); S, small (<200); ?, incomplete information on dispersers.

FIGURE 2  Distribution of Cneorum tricoccon (black areas) and location (white dots) of the populations sampled. Letter codes correspond to those listed in Table 1. Codes in italics refer to peripheral populations.
reaction was checked for successful amplification by running 3 μl of PCR product of six indiscriminately selected individuals per population and microsatellite locus on a 1% agarose gel stained with SYBR Safe (Invitrogen, Carlsbad, CA, USA). PCR products (1 μl) were diluted with 9.1 μl of a loading mixture containing 9 μl of HiDi Formamide and 0.1 μl GeneScan LIZ500 internal size standard (Applied Biosystems, Foster City, CA, USA) and analyzed on an automated DNA sequencer (ABI PRISM 3730 Genetic Analyser, Applied Biosystems) in Unidad de Genómica, Parque Científico de Madrid, Spain. Amplified fragment lengths were scored with GENEMARKER v. 2.4.0 (Softgenetics LLC, State College, PA, USA).

2.3 | Data analysis

2.3.1 | Genetic diversity within populations

Genetic diversity indices, including P, percentage of polymorphic loci; A_e, mean number of effective alleles (1/2p_i^2), where p_i is the frequency of the ith allele for the population; H_o, observed heterozygosity (number of heterozygotes/N, where N is the number of individuals per population); H_e, expected heterozygosity (1 – 2p_i^2); F_is, fixation index (1 – (H_o/H_e)), and the number of private alleles, were estimated using GENALEX v. 6.41 (Peakall & Smouse, 2006). F_is was estimated using INEst 2.0 software (Chybicki & Burczyk, 2009) that corrected for the excess of homozygosity due to the effects of null alleles and genotyping errors (50 × 10^5 Markov chain Monte Carlo [MCMC] iterations, burn-in = 50,000, thinning = 50). Deviation from the Hardy–Weinberg equilibrium (HW) was evaluated at each population with the MCMC approximation (dememorization = 10,000, batches = 100, iterations per batch = 10,000) of Fischer’s test implemented in GENEPOP V. 4.1 (Rousset, 2008). The same software was used to check for genotypic linkage disequilibrium (LD) between pairs of loci within each population using the log-likelihood ratio G statistic (dememorization = 10,000, batches = 100, iterations per batch = 10,000).

2.3.2 | Population structure and differentiation

To determine population differentiation, we calculated pairwise F_st values (Weir & Cockerman, 1984) using FreeNA, which implements a correction to provide accurate estimation of F_st in the presence of null alleles (Chapuis & Estoup, 2007). A Mantel test (Legendre & Legendre, 1998) was performed to check for the correlation between genetic and geographical distances among populations (isolation by distance [IBD]). The matrix of genetic differentiation was calculated with pairwise F_st values, and the matrix of geographical (Euclidean) distances was generated with PASSAGE v. 2 (Rosenberg & Anderson, 2011). The Mantel test was carried out also with the PASSAGE software, using 10,000 permutations. This was repeated using the matrix of logarithm transformation of the distance among populations, which could normalize the distribution.

Population genetic structure in central and peripheral regions was examined by means of analysis of molecular variance (AMOVA, Excoffier, Smouse, & Quattro, 1992) computed in ARLEQUIN. We examined the distribution of genetic variation at three hierarchical levels: (1) among regions (center vs. periphery), (2) among populations within regions, and (3) within populations. We thus computed three fixation indices: F_C, variation among regions, F_SC (among populations within regions), and F_ST (among all populations). The significance of the variance components was tested using nonparametric permutation procedures, with 50,000 permutations.

Genetic structure was also analyzed using Bayesian clustering methods implemented in STRUCTURE v. 2.3.1. (Pritchard, Stephens, & Donnelly, 2000), comparing the results obtained with AMOVA with a different statistical approach. STRUCTURE assumes a model in which there are K genetic clusters (where K is unknown), where each K is characterized by a set of allele frequencies at each locus. Individuals are probabilistically assigned to one or more clusters. Our analyses were based on an admixture ancestral model with correlated allele frequencies (Falush, Stephens, & Pritchard, 2003) for a range of K values starting from two to the number of populations plus two (i.e., 18). The proportion of membership of each individual and population to the inferred K cluster was then calculated. We performed 12 independent runs for each K, using a burn-in period and a run length for the MCMC of 5 × 10^5 and 10^6 iterations, respectively. We used HARVESTER (Earl & vonHoldt, 2011) to extract the relevant data from STRUCTURE and to determine the number of clusters most appropriate for the interpretation of our data. First, HARVESTER calculates the mean log probability of the data for each K. Second, it calculates ΔK following the method described in Evanno, Regnaut, and Goudet (2005). ΔK is an ad hoc parameter which estimates the rate of change in the log probability of the data between the successive K values. We then used CLUMPP v. 1.2.2 (Jakobsson & Rosenberg, 2007) to combine results from the 12 runs at each K, using the Greedy option (for K < 9) and the LargeKGreedy option (for K ≥ 10). Membership in clusters was graphically represented using DISTRUCT v. 1.1 (Rosenberg, 2004). Additionally, a Bayesian analysis was conducted using the R package GENELAND (Guillot, Mortier, & Estoup, 2005), which also uses MCMC algorithms to perform clustering analyses with the option to include nonspatial as well as spatial models. In this case, we performed spatial clustering; MCMC iterations were set at 10^5, thinning at 100, and the burn-in period of 100. The number of K to be tested was set at 1–18, and five replicates for each possible K were run. The best result was chosen, based on the highest average posterior probability. Finally, a principal coordinates analysis (PCoA) based on FST pairwise values was performed in GENALEX to provide further insight into population differences.

2.3.3 | Effects of center–periphery, population size, and disperser type on genetic variation and population structure

To assess the effects of the study factors on within-population genetic diversity and population structure, we compared diversity indices (A, H_o, H_e, F_is) and differentiation (F_SC) among groups using FSTAT v. 2.9.3.3 (Goudet, 1995). Populations were grouped (see Table 1)
according to periphery (central vs. peripheral), population size (large vs. small), and disperser type (lizard vs. mammal), and significance was assessed based on 10,000 permutations.

3 | RESULTS

3.1 Genetic diversity within populations and comparisons between populations

A total of 80 alleles were scored from the nine microsatellites in the 399 individuals analyzed, an average of 8.89 alleles per locus. The number of alleles per locus ranged from 6 to 16. Genetic diversity varied significantly across populations (summarized data for genetic diversity indices are shown in Table 2). These differences were not related to population location or disperser type; that is, there were no significant differences in the genetic diversity indices when peripheral versus central populations were compared or when populations were grouped according to disperser type (endemic lizards vs. introduced mammals) (Table 3). Conversely, we found an effect of population size on genetic diversity; \( A, H_O, \) and \( H_E \) were lower and \( F_{ST} \) was larger in small compared to large populations (Table 3).

The percentage of polymorphic loci (\( P \)) was high in central populations with the exception of CA and ME, in which 89% of the loci were monomorphic. Conversely, \( P \) was low (33%–56%) in all peripheral populations, except in FI, where all loci were polymorphic. The lowest mean number of alleles per locus (\( A \)) was found in CA and ME (1.11), whereas SA had the largest value (4.11). The mean number of effective alleles (\( A_e \)) showed similar results, ranging from 1.0 (in CA) to 2.82 (in SA). Observed (\( H_O \)) and expected (\( H_E \)) heterozygosities tended to be higher in central populations than in peripheral populations (with the exception of CA, ME, and FI), but differences were not significant (Table 3). Observed heterozygosities (\( H_O \)) ranged from 0.004 (CA) to 0.366 (SA), whereas expected heterozygosities (\( H_E \)) ranged from 0.004 (CA) to 0.485 (LL). The inbreeding coefficient (\( F_{IS} \)) varied among populations from 0.185 (SA) to 0.887 (CA) and also showed a (not-significant) tendency to be higher in peripheral populations, with the exception of CA and ME, which showed the highest \( F_{IS} \) values. We found a total of 20 private alleles, present in 10 of the 16 populations. The number of private alleles per population ranged from one (SA, CO, CR, MA, CG) to five (DR). All populations showed significant departures from HW equilibrium toward heterozygote deficiency (\( p < .001 \) across loci and populations). No consistent LD was found between any pairwise comparisons across loci and populations (data not shown).

3.2 Genetic structure and population differentiation

All pairwise \( F_{ST} \) values of differentiation among populations were highly significant and generally very high (Table S1), ranging from 0.153, between MA and CG, to 0.977 between CA and ME, with an overall \( F_{ST} \) of 0.556 ± 0.179 (mean ± SD). Pairwise \( F_{ST} \) values were higher than

<table>
<thead>
<tr>
<th>Population code</th>
<th>N</th>
<th>P</th>
<th>A</th>
<th>( A_e )</th>
<th>( H_O )</th>
<th>( H_E )</th>
<th>( F_{IS} )</th>
<th>Private alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Central populations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SA</td>
<td>30</td>
<td>89</td>
<td>4.11</td>
<td>2.82</td>
<td>0.366</td>
<td>0.476</td>
<td>0.155*</td>
<td>1</td>
</tr>
<tr>
<td>LL</td>
<td>25</td>
<td>100</td>
<td>3.67</td>
<td>2.33</td>
<td>0.182</td>
<td>0.485</td>
<td>0.651</td>
<td>3</td>
</tr>
<tr>
<td>BL</td>
<td>25</td>
<td>100</td>
<td>3.89</td>
<td>2.21</td>
<td>0.142</td>
<td>0.477</td>
<td>0.731*</td>
<td>0</td>
</tr>
<tr>
<td>CO</td>
<td>25</td>
<td>100</td>
<td>3.78</td>
<td>2.21</td>
<td>0.273</td>
<td>0.482</td>
<td>0.476*</td>
<td>1</td>
</tr>
<tr>
<td>TG</td>
<td>25</td>
<td>89</td>
<td>3.33</td>
<td>1.95</td>
<td>0.178</td>
<td>0.387</td>
<td>0.550*</td>
<td>0</td>
</tr>
<tr>
<td>BB</td>
<td>25</td>
<td>100</td>
<td>2.89</td>
<td>1.8</td>
<td>0.089</td>
<td>0.408</td>
<td>0.768*</td>
<td>2</td>
</tr>
<tr>
<td>CA</td>
<td>25</td>
<td>11</td>
<td>1.11</td>
<td>1</td>
<td>0.004</td>
<td>0.004</td>
<td>0.887 (8)*</td>
<td>0</td>
</tr>
<tr>
<td>CC</td>
<td>25</td>
<td>67</td>
<td>2.11</td>
<td>1.83</td>
<td>0.209</td>
<td>0.354</td>
<td>0.453 (3)*</td>
<td>0</td>
</tr>
<tr>
<td>DR</td>
<td>25</td>
<td>100</td>
<td>2.89</td>
<td>1.66</td>
<td>0.173</td>
<td>0.366</td>
<td>0.555*</td>
<td>5</td>
</tr>
<tr>
<td>ME</td>
<td>23</td>
<td>11</td>
<td>1.11</td>
<td>1.04</td>
<td>0.039</td>
<td>0.032</td>
<td>0.865 (8)*</td>
<td>0</td>
</tr>
<tr>
<td><strong>Peripheral populations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GI</td>
<td>21</td>
<td>56</td>
<td>1.89</td>
<td>1.22</td>
<td>0.027</td>
<td>0.132</td>
<td>0.795 (4)*</td>
<td>3</td>
</tr>
<tr>
<td>CR</td>
<td>25</td>
<td>44</td>
<td>1.56</td>
<td>1.27</td>
<td>0.032</td>
<td>0.144</td>
<td>0.780 (5)*</td>
<td>1</td>
</tr>
<tr>
<td>FI</td>
<td>25</td>
<td>100</td>
<td>3.56</td>
<td>1.97</td>
<td>0.236</td>
<td>0.437</td>
<td>0.481*</td>
<td>3</td>
</tr>
<tr>
<td>MO</td>
<td>25</td>
<td>56</td>
<td>1.89</td>
<td>1.39</td>
<td>0.102</td>
<td>0.207</td>
<td>0.540 (4)*</td>
<td>0</td>
</tr>
<tr>
<td>MA</td>
<td>25</td>
<td>33</td>
<td>1.67</td>
<td>1.12</td>
<td>0.036</td>
<td>0.079</td>
<td>0.591 (6)*</td>
<td>1</td>
</tr>
<tr>
<td>CG</td>
<td>25</td>
<td>56</td>
<td>1.78</td>
<td>1.14</td>
<td>0.04</td>
<td>0.085</td>
<td>0.556 (4)*</td>
<td>1</td>
</tr>
</tbody>
</table>

N = number of individuals sampled; P = proportion of polymorphic loci; A = mean number of alleles per locus (allelic richness); \( A_e \) = mean number of effective alleles; \( H_O \) = observed heterozygosity; \( H_E \) = expected heterozygosity; \( F_{IS} \) = inbreeding coefficient and number of private alleles. The number of monomorphic loci is shown in parentheses.

*Significant departures from Hardy–Weinberg equilibrium (\( P ≤ 0.001 \)). See text for details on statistics. Letter codes are listed in Table 1.
0.5 in ≥70% of the comparisons. In general, FST values were lower between nearby populations (e.g., MA and CG). Accordingly, we detected significant IBD between populations, either using the matrix of linear Euclidean distances (FST = 0.34; p = 0.039) or the logarithm of the distances (FST = 0.37; p = 0.008; Fig. S1). Hierarchical AMOVA (Table 4) showed a high proportion of genetic variation among (57.3%) and within populations (40.1%), while a very small fraction of the variation was due to differences among regions (central vs. peripheral) (2.52%).

Results from Bayesian clustering conducted by STRUCTURE and strictly interpreted using the method of Evanno et al., 2005 would suggest that two genetic clusters are sufficient for interpretation of our data (K = 2, Figure 3). However, we focus on the results with K = 6 and K = 15 because there are two secondary large peaks in AK at K = 6 and K = 15 that could explain a secondary substructure (Figure 3). Moreover, larger values of K are consistent with the observed high among-population differentiation revealed by AMOVA and by high pairwise FST values. We thus show results for K = 2-6 to provide a comprehensive understanding of population structure (Figure 4). In the K = 6 solution, all populations contained admixed individuals; that is, an individual was assigned to different genetic clusters (Figure 4). However, several populations were assigned to a predominant cluster; for example, SA and LL were included mainly in the green cluster; CO, TG, BB, DR, BL, and CA in the red one; CG and MA in the purple one. In general, nearby populations were assigned to the same cluster (e.g., SA and LL; CG and MA); however, there were also close populations included in different clusters (e.g., CA and CC; FI and MO). Central and peripheral populations were generally assigned to different clusters. In the K = 15 solution, all populations were assigned to different clusters, except for the nearest populations (CG and MA) that were included in the same cluster (Fig. S2).

The clustering obtained with K = 6 was in general consistent with results from a PCoA based on FST distances (Fig. S3). The PCoA also revealed the strong differentiation of CA and ME to other populations, which was only observed at larger K values in the Bayesian analysis. Furthermore, these results were comparable to the second Bayesian clustering method used, GENELAND, which allowed inclusion of the geographical location of the populations. Again, 15 clusters were found as the most plausible scenario for explaining the data structure. This approach assigned each population to different clusters, except for CG and FM that were grouped together, which clearly supports the observed high genetic structure.

4 | DISCUSSION

While several studies have shown reduced genetic diversity in peripheral compared with central populations (Arnaud-Haond et al., 2006; Durka, 1999; Eckert et al., 2008; Eckstein et al., 2006; Lammi et al., 1999), our results showed an alternative pattern where other interacting factors may override the presumed effects of the center–periphery hypothesis. Cneorum tricoccon populations located at range margins tended to show lower genetic diversity (P, A, H0, H1) and higher inbreeding (FIS) than central populations, although differences were not significant. In contrast, population size was significantly

---

**Table 3** Effects of center–periphery, population size, and disperser type on genetic diversity and population structure of Cneorum tricoccon

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>df</th>
<th>Sum squares</th>
<th>Variance components</th>
<th>% Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among regions (central vs. peripheral)</td>
<td>1</td>
<td>124.115</td>
<td>0.082</td>
<td>2.52 ns</td>
</tr>
<tr>
<td>Among populations within regions</td>
<td>14</td>
<td>1,321.098</td>
<td>1.866</td>
<td>57.34***</td>
</tr>
<tr>
<td>Within populations</td>
<td>782</td>
<td>1,021.394</td>
<td>1.306</td>
<td>40.14***</td>
</tr>
</tbody>
</table>

**Fixation indices**

| FST  | 0.599 | ***     |
| FSC  | 0.588 | ***     |
| FCT  | 0.025 | .127    |

The significance of the variance components was tested using nonparametric permutation procedures with 50,000 permutations. *** P<0.001

---

**Table 4** Population genetic structure inferred by hierarchical analysis of molecular variance and Weir and Cockerham F-statistics estimates: FCT, variation among regions (peripheral vs. central populations); FSC, among populations within regions; and FST, among all populations

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>df</th>
<th>Sum squares</th>
<th>Variance components</th>
<th>% Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among regions (central vs. peripheral)</td>
<td>1</td>
<td>124.115</td>
<td>0.082</td>
<td>2.52 ns</td>
</tr>
<tr>
<td>Among populations within regions</td>
<td>14</td>
<td>1,321.098</td>
<td>1.866</td>
<td>57.34***</td>
</tr>
<tr>
<td>Within populations</td>
<td>782</td>
<td>1,021.394</td>
<td>1.306</td>
<td>40.14***</td>
</tr>
</tbody>
</table>

**Fixation indices**

| FST  | 0.599 | ***     |
| FSC  | 0.588 | ***     |
| FCT  | 0.025 | .127    |

The significance of the variance components was tested using nonparametric permutation procedures with 50,000 permutations. *** P<0.001
associated with most genetic diversity indices. It is generally predicted that central populations are larger than peripheral ones (Eckert et al., 2008; Lawton, 1993; Sagarin & Gaines, 2002). However, C. tricoccon contains several large populations at the range margins and small ones in the center. The large peripheral populations (e.g., Fl) maintain high levels of genetic diversity, despite their geographic isolation. Similarly, small populations located in the center of the species distribution (e.g., CA, ME) are genetically depauperated, showing even lower genetic diversity than peripheral populations. Our results support the stronger role of population size in determining the genetic diversity of populations (Ellstrand & Elam, 1993) and, more importantly, its potential to counteract the effect of periphery. Furthermore, our results add insight into scenarios where the central–marginal hypothesis may not hold due to insularity and isolation effects (Pironon et al., 2016).

The several methods used to assess population structure showed high genetic differentiation among C. tricoccon populations. According to hierarchical AMOVA, 57% of the variation was found among populations, and mean \( F_{ST} \) value was 0.56, much higher than the average of \( F_{ST} = 0.24 \) for plant species with mixed breeding systems (Nyborg, 2004). Concurrent with the strong population structure, we also found a pattern of IBD; that is, closer populations were genetically more similar than populations farther apart (see Awad et al., 2014; Eckstein et al., 2006). Spatial distribution of genetic variation in plants is the outcome of several factors such as mating system, gene flow, genetic drift, and natural selection (Petit & Hampe, 2006). Our results suggest that gene flow among populations is limited, likely due to a combination of different factors, such as high selfing rates, populations with high kinships coefficients between parents (Hirao, 2010), and limited pollen and seed dispersal. In general, outcrossing species maintain most genetic variation within populations, while selfing species harbor comparatively higher variation among populations (Avise, 2004). Cneorum tricoccon is a self-compatible species, and, thus, its marked population structure and high inbreeding coefficients suggest that selfing (and/or crosses between parents with moderate–high kinship) could be more frequent than outcrossing in current conditions, likely due to the limited pollinator visits observed in many populations (A. Traveset, personal communication). Limited gene flow among populations is also in agreement with the strong genetic drift observed in certain populations (e.g., CA and ME), where many alleles were fixed. Finally, as has been documented in other long live Tertiary relicts, gene flow may have been interrupted during some historic periods, increasing isolation and therefore \( F_{ST} \) values (Dubreuil, Riba, & Mayol, 2008; Rodríguez-Sánchez, Guzman, Valido, Vargas, & Arroyo, 2009; Rodríguez-Sánchez et al., 2008).

As a Tertiary relict, C. tricoccon survived during several Quaternary climate oscillations and, as such, its phylogeographic history may also have altered its genetic patterns (Rodríguez-Sánchez et al., 2008).

![FIGURE 3](a) Number of genetic clusters (K) detected by Evanno et al. (2005). Mean log probability of the data for the 12 STRUCTURE runs at each K. Error bars are SD; (b) \( \Delta K \), rate of change in the log probability of data between successive K values.

![FIGURE 4](Population structure inferred by Bayesian cluster analyses (STRUCTURE) for 399 Cneorum tricoccon individuals from 16 central and peripheral populations. Results for K (number of clusters) ranging from 2 to 6. Each individual (grouped by population) is represented by a vertical bar. Letter codes correspond to the population listed in Table 1.)
The high level of genetic differentiation, the presence of a marked spatial structure, and the isolation-by-distance pattern found must be interpreted in the complex phylogeographic context of the Mediterranean basin (Feliner, 2014; Rodríguez-Sánchez et al., 2008). *Cneorum tricoccon* has experienced marine transgression during the glacial and interglacial periods of the Quaternary, which greatly influenced the Balearic Islands (Vesica et al., 2000). Furthermore, the Messinian Salinity Crisis that almost desiccated the Mediterranean Sea, the recurrent connections, and disconnections with Northern Africa vegetation (Rodríguez-Sánchez et al., 2008) or the existence of microrefugia for relict species in the Mediterranean coasts (MoEco; Dubreuil et al., 2008; Lumaret et al., 2002) may have also influenced current species’ distributions (Hewitt, 2011). All these processes have fostered both expansion pulses—such as the several contacts of the Balearic Islands with the Iberian Peninsula (Garnatje, Pérez-Collazos, Pellicer, & Catalán, 2013) or the cohesion of Mallorca and Cabrera in a single land mass (Vesica et al., 2000)—as well as isolation processes of the populations that colonized the extremes of the distribution (such as MA, CG, or GI).

Factors related to seed dispersal and insularity can also account for the high divergence observed. *Cneorum tricoccon* has a tight seed dispersal mutualism with lizards. The existence of natural dispersers may also be considered as a biological criterion to define optimal and marginal populations (Pironon et al., 2016). Lizards are known to be highly territorial and thus do not move seeds to long distances as do other vertebrates, such as birds or mammals. Furthermore, several *C. tricoccon* populations are restricted to small islands and islets and, consequently, are strongly isolated from each other. Past fragmentation leading to geographical isolation, and limited pollen and seed dispersal may collectively explain the high genetic differentiation among populations (Awad et al., 2014; Ge et al., 2005; Van Rossum et al., 2003). In populations of *C. tricoccon* where lizards are not present, introduced carnivorous mammals (e.g., pine martens and genets) may replace the native lizards as seed dispersers (Riera et al., 2002; Traveset, 1995b) although little is known about the implications for genetic diversity and population structure of this seed dispersal disruption. As pine martens and genets have extensive home ranges—up to 900 ha in the case of pine martens males (Clevenger, 1993)—higher gene flow may be expected in populations with carnivorous mammals as seed dispersers than those where lizards are still present, and therefore, low differentiation could be expected among these populations (Kirkpatrick & Barton, 1997). Surprisingly, no differences in genetic diversity and structure were found among populations with different dispersers in the Balearic Islands. These results could be explained by the habitat preferences of the introduced mammals. In a previous study, Celedón-Neghme et al. (2013) showed that mammals replace lizards as seed dispersers only in populations occurring in the understory of holm oak forests. Therefore, gene flow mediated by introduced mammals is probably restricted to populations close to forested mountain areas instead of coastal maquis, the typical habitat of *C. tricoccon*. The low number of sampled populations occurring in holm oak forest (SA and LL) may account for the absence of significant differences in genetic diversity and population structure among populations with different dispersers. In other Tertiary relics, where natural dispersers have disappeared, other species have occupied their role with important consequences for the species dispersion patterns (e.g., *Laurus nobilis*, Hampe, 2003).

The contrasting levels of genetic variation observed among populations can help identify valuable populations for conservation purposes. Assuming that genetic variation estimated with neutral molecular markers can reflect overall genetic diversity and be used as an indicator of variation in quantitative traits, several large populations with high genetic variation and different genetic makeup (e.g., FI, BL, DR) can be considered for further investigations into their conservation value. Accordingly, the low genetic variation found in some populations (e.g., FM, CG, GI, CR, CA, and ME) may compromise their long-term viability. In particular, two populations (CA and ME) showed extremely low genetic variation, in terms of heterozygosity and polymorphism. Both populations are small and are geographically isolated in islands, where genetic drift can increase the fixation of deleterious alleles and outcrosses may be carried out between related parents. Therefore, reduced individual fitness and increased inbreeding depression can be expected in both populations (Kimura, Maruyama, & Crow, 1963; Lynch & Gabriel, 1990; Oakley & Winn, 2012). However, these populations are located at the dry edge of the species’ distribution range, and strong selection pressures related to water availability may be selecting drought-adapted phenotypes, as has been suggested for marginal populations of other species (Kirkpatrick & Barton, 1997; Lesica & Allendorf, 1995), including Tertiary relics such as *Ramonda myconi* (Muller et al., 1997). An experimental study with *C. tricoccon* adds support to this hypothesis, showing that dry-edge populations (e.g., CA) may be locally adapted, as they exhibited multiple functional traits that favored drought tolerance (Lázaro-Nogal et al., 2015). It is then likely that low genetic diversity and high divergence in these populations are explained by past strong directional selection combined with isolation-driven low gene flow and strong genetic drift. Conservation of these *C. tricoccon* populations may be especially important in light of climate change as they may preserve genetic combinations that are important for adaptation to drier conditions (Hampe & Petit, 2005; Lesica & Allendorf, 1995).

In conclusion, our results highlight that, despite geographic isolation and important gene flow limitations related to pollen and seed dispersal mechanisms, population size can override, at least partially, the effects of geographical periphery. Integrative approaches taking into account different factors that determine genetic diversity and population structure, such as the position of a population in the species’ distribution range, size, the presence of effective seed and pollen dispersers, and the stability of local adaptations, are necessary to develop effective conservation strategies.

**ACKNOWLEDGMENTS**

A L-N was supported by JAE-predoc fellowship from the International Laboratory of Global Change (LINCGlobal). Funding was provided by the Spanish Ministry for Innovation and Science with the grants...
Consolider Montes (CSD2008_00040) and VULGLO (CGL2010-22180-C03-03) provided to FV, and the Ministry of Economy and Competitiveness, grant ECOMETAS (CGL2014-53840-REDT) and Junta de Andalucía Excellence project to Juan Arroyo and AT. Leaf collection from all sampled populations was partly possible thanks to previous projects to AT and Alfredo Valido. We also thank David Montero for his valuable assistance in processing material for DNA extractions, and Alastair Plant for linguistic assistance.

CONFLICT OF INTEREST
None declared.

AUTHOR CONTRIBUTIONS
A. L.-N., FV, and AT conceived the ideas; A L-N collected the data; A L-N, A G-F, and SM analyzed the data; and A L-N, A G-F, and SM led the writing with contributions from all authors.

REFERENCES


SUPPORTING INFORMATION
Additional Supporting Information may be found online in the supporting information tab for this article.