

This is the peer reviewed version of the following article: **Martínez-López, V., García, C., Zapata, V., Robledano, F., De la Rúa, P. 2020. Intercontinental long-distance seed dispersal across the Mediterranean Basin explains population genetic structure of a bird - dispersed shrub. *Molecular Ecology*, 29: 1408-1420**, which has been published in final form at [<https://doi.org/10.1111/mec.15413>]. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.



DR. VICENTE MARTINEZ-LOPEZ (Orcid ID : 0000-0001-7800-0703)

DR. PILAR DE LA RUA (Orcid ID : 0000-0002-0058-1402)

Article type : Original Article

Title: Intercontinental long-distance seed dispersal across the Mediterranean Basin explains population genetic structure of a bird-dispersed shrub

Running title: Intercontinental long-distance seed dispersal

Authors: Martínez-López, V.^{1,2*}, García, C.^{3,4}, Zapata, V.², Robledano, F.² & De la Rúa, P.¹

¹*Department of Zoology and Physical Anthropology. Faculty of Veterinary. University of Murcia. Murcia, 30100 Spain;* ²*Department of Ecology and Hydrology. Faculty of Biology. University of Murcia. Murcia, 30100 Spain;* ³*Institute of Integrative Biology, Department of Evolution, Ecology, and Behaviour, University of Liverpool, Crown Street, Liverpool L69 7ZB, United Kingdom;* ⁴*CIBIO/InBIO-Universidade do Porto. Campus Agrário de Vairão, Rua Padre Armando Quintas, 4485-661, Vairão, Portugal.*

***Corresponding author:** e-mail: vicente.martinez2@um.es; phone: +34 86888 4985/ 8031; fax: +34 868 884906

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/mec.15413](https://doi.org/10.1111/mec.15413)

This article is protected by copyright. All rights reserved

Abstract

Long-distance dispersal (LDD) is a pivotal process for plants determining their range of distribution and promoting gene flow among distant populations. Most fleshy-fruited species rely on frugivorous vertebrates to disperse their seeds across the landscape. While LDD events are difficult to record, a few ecological studies have shown that birds move a sizeable number of ingested seeds across geographic barriers, such as sea straits. The foraging movements of migrant frugivores across distant populations, including those separated by geographic barriers, creates a constant flow of propagules that in turn shapes the spatial distributions of the genetic variation in populations. Here, we have analysed the genetic diversity and structure of 74 populations of *Pistacia lentiscus*, a fleshy-fruited shrub widely distributed in the Mediterranean Basin, to elucidate whether the Mediterranean Sea acts as a geographic barrier or alternatively whether migratory frugivorous birds promote gene flow among populations located on both sides of the sea. Our results show reduced genetic distances among populations, including intercontinental populations, and they show a significant genetic structure across an eastern-western axis. These findings are consistent with known bird migratory routes that connect the European and African continents following a north-southwards direction during the fruiting season of many fleshy-fruited plants. Further, Approximate Bayesian Analysis failed to explain the observed patterns as a result of historical population migrations at the end of Last Glacial Maximum. Therefore, anthropic and/or climatic changes that would disrupt the migratory routes of frugivorous birds might have genetic consequences for the plant species they feed upon.

Keywords: genetic clustering, genetic network, isolation by distance, migratory frugivorous birds, *Pistacia lentiscus*, wide distribution range

1. Introduction

The examination of the long-distance dispersal (LDD) of seeds is crucial to determine the demographic trends of plant populations, as well as their spatial genetic structures (SGS), and

distribution ranges (Nathan, 2006). Empirical LDD data are scarce owing to the inherent difficulty of tracking the fate of dispersed seeds (Nathan, 2006). However, molecular tools have provided ample evidence for the impact of LDD in: (i) expediting the demographic expansion of forests northwards as ice retreated at the end of the last glacial period (Hampe, Arroyo, Jordano, & Petit, 2003); (ii) clarifying the processes that underlie extreme bipolar disjunctions in plants, i.e., the biogeographic disjunction between the Arctic and the southernmost point of South America (Popp, Mirré, & Brochmann 2011); and (iii) explaining the distribution of plant species across biogeographic barriers, such as mountain ranges and sea straits (Dyer & Nason, 2004). Recent studies have evidenced that frugivorous birds mobilize propagules among distant populations hundreds of kilometres apart, including island and continental populations (Nogales, Heleno, Traveset, & Vargas, 2012; Viana, Gangoso, Bouten, & Figuerola, 2016a). Furthermore, a few studies have shown that long-distance migrant frugivorous birds play a crucial role in connecting plant populations across their distribution ranges, when their migratory movements overlap with the fruiting season of the plants they feed upon (Hampe et al., 2003). In the context of global climate change, bird migration routes are expected to collapse or shift due to warming (Both et al., 2009; Pulido & Berthold, 2010) or landscape fragmentation (Bairlein, 2016), consequently impacting the demographic trends, genetic connectivity, and the overall distribution range of animal-dispersed plants (Dirzo et al., 2014; Martínez-López, De la Rúa, Zapata, & Robledano, 2017; Parejo-Farnés, C., Robledo-Arnuncio, Albaladejo, Rubio-Pérez, & Aparicio, 2017).

To date, most of the studies assessing the SGS in plants across their entire distribution ranges have focussed on narrowly distributed plant species, as they typically have high levels of endemism, which make them particularly sensitive to the drivers of global change (e.g., González-Astorga et al., 2004; Medrano & Herrera, 2008; Binks et al., 2015). However, these studies are insufficient to fully understand the impact of the main drivers that underpin SGS in widely distributed plant species because: (i) the populations of widespread species are typically subjected to different environmental conditions across their distribution range (Lesica & Allendorf, 1995; Gapare, Aitken, & Ritland, 2005), including a variable level of human-driven disturbance; (ii) wide distribution ranges are shown to correlate with particular demographic and genetic traits (e.g. high levels of genetic diversity and low levels of population genetic differentiation; Hamrick & Godt, 1989; Hamrick & Godt, 1996; Eckert, Samis, & Loughheed, 2008); and (iii) the impact of LDD events is usually overlooked in studies with narrowly distributed species, but these events are crucial to explaining SGS in widely distributed plant species, particularly those distributed across

geographic barriers. Therefore, research on genetic variation including the full distribution ranges of widespread plant species is especially needed to understand the impact of LDD on the spatial distribution of genetic variation.

The Mediterranean Sea acts as a barrier to gene flow for a wide range of plant and animal populations located on both shores of the Mediterranean (e.g. Molina-Venegas, Aparicio, Lavergne, & Arroyo, 2015). However, studies assessing plant taxa migration between Europe and Africa through the Strait of Gibraltar found that it was not a barrier to the dispersal of those species with higher migration capabilities (Lavergne, Hampe, & Arroyo, 2013; Molina-Venegas et al., 2015). Current bird migratory routes were set at the end of the last glacial period (Bruderer & Liechti, 1999; Newton, 2007) and they connect the European and African continents, traversing distant populations of circum-Mediterranean endozoochorous plants (e.g., *Pistacia lentiscus* L., Figure 1). Seed mean retention times ranges between 1 and 4 h (Viana, Santamaría, & Figuerola, 2016b) and an average flight speed of 45 km/h has been described in passerine birds (Bloch & Bruderer, 1982). The flying distance covered by birds during their seasonal migratory movements would suffice to connect the European and African continents through different known routes across the Mediterranean Basin (Bruderer & Liechti, 1999; Fiedler & Davidson, 2012; and Bairlein, 2016). Thus, if frugivorous birds forage on fruiting trees during their migratory displacements, they are expected to mobilize seeds across distant populations, in turn, reducing the role of the Mediterranean Sea as a geographic barrier to gene flow.

The combination of both ecological information and genetic tools allows the testing of the hypothesis that the Mediterranean Sea acts as a barrier to gene flow, by comparing the observed genetic patterns with those expected under an isolation by distance (IBD) scenario, where the genetic distances between populations steadily increase with geographic distance (Wright, 1943). Following this approach, we herein analysed the spatial distribution of the genetic variation and the population genetic structure of a widely distributed fleshy-fruited shrub (*P. lentiscus*) across its distribution range, which includes the Mediterranean Basin and the Canary Islands. We applied microsatellites in combination with landscape genetic tools, to test whether the Mediterranean Sea acts as a barrier to gene flow among *P. lentiscus* populations located on opposite shores of the sea or, alternatively, whether the contemporary gene flow attributable to the migratory birds links the intercontinental populations. To that end, we first described the genetic diversity and structure of 74 populations encompassing the whole distribution range, and we tested for IBD patterns. Then,

we combined different landscape models (Bayesian clustering, discriminant analysis of principal components (DAPC), and genetic networks) to test whether known migratory routes of frugivorous birds explain the inferred genetic structure. Finally, we performed Approximate Bayesian Computations (ABC) to quantify the effects of the past demographic events (such as rapid colonization northwards at the end of the most recent glacial retreat *ca.* 20 k years ago) in shaping the observed spatial genetic patterns. We expected there to be reduced genetic distances among the populations encountered by frugivorous birds along their migratory routes that would promote contemporary gene flow among island and continental populations on both shores of the Mediterranean Sea.

2. Materials and Methods

2.1 Study species and sampling design

P. lentiscus is a dioecious and anemophilous evergreen fleshy-fruited shrub, that is widely distributed across the Mediterranean Basin and is also found in the Canary Islands (Bolós & Vigo, 1990) (Figure 1), and flowers from March to May (Keynan et al., 1987; Milla, Castro-Díez, & Montserrat-Martí, 2010). *P. lentiscus* appeared in the Tertiary period, when the current Mediterranean Basin had a tropical climate, and is assumed to have evolved very little since then (Verdú & García-Fayos, 2002 and references therein). It is an important part of the Mediterranean shrublands, as it fruits in large amounts for long periods of time (from August till February), and these fruits are lipid-rich and fundamental to the diets of migratory birds (Verdú & García-Fayos, 2002). Therefore, *P. lentiscus* acts as a foundation species, *sensu* Whitham et al. (2006), by structuring the functioning of communities in which it is present (Herrera, 1984; Acosta-Rojas, Jiménez-Franco, Zapata-Pérez, De la Rúa & Martínez-López, 2019). Moreover, it has also been widely consumed by humans (Rouskas, 1996; Zamora, 1997; Piluzza, Viridis, Serralutzu, & Bullitta, 2015 and references therein).

The diets of numerous species of bird incorporate *P. lentiscus* fruits, most of which are small passerines (Jordano, 1989; Herrera, 1998). *Sylvia melanocephala* Gmelin, *Sylvia atricapilla* L., *Erithacus rubecula* L., and *Turdus merula* L. have been identified as the main consumers of *P. lentiscus* fruits in the Mediterranean environment (Herrera, 1984, 1998; Izhaki & Safriel, 1985;

Jordano, 1989). These four species have migrant and resident populations in the European continent; however, *S. melanocephala* and *T. merula* are considered to be mainly resident in the Mediterranean areas (e.g. Herrera, 1984), while *E. rubecula* and *S. atricapilla* show a higher mobility with important contingents of migrant populations that travel southwards from September to November (Leal, Monrós, & Barba, 2004; Arizaga, Alonso, & Barba, 2010) when the ripening peak of the *P. lentiscus* fruits occurs (Jordano, 1989). Thus, *E. rubecula* and *S. atricapilla* are the species with the greatest potential to facilitate LDD of *P. lentiscus* seeds.

The sampling design encompassed the entire distribution range of the plant species (Figure 1). A total of 1057 adult individuals were sampled from 2013 to 2015 (74 populations, Table S1). Plants sampled in each population were separated by at least 30 m to avoid clonality and reduce spatial autocorrelations. Leaf tissues were sampled and kept in paper envelopes with silica gel for drying until further examination in the lab.

2.2 DNA isolation and microsatellite genotyping

DNA was isolated from the dried leaves (0.1 g dry weight) using the NaOH protocol (Werner, Ros, & Guerra, 2002a). A set of 15 microsatellites designed for the species were used (six from Albaladejo et al., 2008 and nine from Motalebipour, Kafkas, Khodaeiaminjan, Çoban, & Gözel, 2016, see Table S2). PCR was performed in total volumes of 10 µl, containing 1 µl of template DNA; reagent concentrations and PCR profiles are shown in Table S2. The PCR products were electrophoretically separated on an ABI 3730 sequencer (Applied Biosystems, Foster City, CA, USA). Allele scoring was carried out manually with GeneMapper v3.7 software (Applied Biosystems). Genotypes were reviewed twice after scoring to reduce errors. We tested the performance of each marker by estimating the frequency of the null alleles with the Micro-Checker software, version 2.2.3 (Van Oosterhout, Hutchinson, Wills, & Shipley, 2004) (Table S2, null alleles by marker). The percentage of missing data was low (average 3.015 %, Table S2) and the overall average frequency of the null alleles per population was 0.104 ± 0.007 (see Table S1, null alleles by population). However, the null allele values were probably overestimated as the population sizes were limited for most of the locations (72 out of 74 populations contained < 30 individuals, Table S1).

2.3 Data analyses

Population genetic diversity estimates and AMOVA analysis

Data analyses were carried out using R v. 3.2.5 (R Development Core Team 2016). We used the package *DiveRsity* v. 1.9.90 (Keenan, McGinnity, Cross, Crozier, & Prodöhl, 2013) to estimate the basic population genetic diversity parameters, including the allelic richness (Ar) and rarefied allelic richness (Ar^*); expected, observed, and unbiased heterozygosity (H_e , H_o , and uH_e , respectively); and inbreeding coefficient (F_{is}). We computed the 95 % confidence intervals for Ar and F_{is} using 9999 iterations. We additionally estimated the frequency of the private alleles when implemented in GenALEx 6.5 (Peakall & Smouse, 2012). We tested for Hardy-Weinberg equilibrium (HWE) as implemented in the *DiveRsity* R package (Keenan et al., 2013), and genotypic linkage disequilibrium with the software Genepop (Raymond & Rousset, 1995). To search for a spatial trend in the distribution of the genetic diversity, we applied a Pearson's correlation between Ar^* and the frequency of private alleles, with the latitude and longitude of the sampled populations. We carried out an analysis of molecular variance (AMOVA) based on 999 replicates using GeneAlex 6.5 (Peakall & Smouse, 2012), to assess how the genetic diversity varied within and among the populations.

Spatial genetic structure across the whole distribution range

We first assessed the genetic differentiation across the populations (F_{st}) with the R package *hierfstat*, version 0.04.22 (Goudet, 2005), with confidence intervals and the significance of the results being calculated with 10 000 bootstrapped samples. We tested for IBD (Wright, 1943) by applying a Mantel test using the pairwise Nei's genetic and geographic distances (Nei, 1972) with the *adegenet* R package version 2.0.1 (Jombart, 2008). The geographic distance matrix (distances in kilometres) was calculated with the R package *geosphere* version 1.5-10 (Hijmans, 2019), using the Vincenty ellipsoid method to account for the curvature of the Earth (ellipsoid WGS84) (Vincenty, 1975). We also tested for IBD with the Euclidean longitudinal and latitudinal distances, to check for different patterns along the East-West and North-South axes. Furthermore, we tested the IBD according to the main bird migratory routes that traverse the Mediterranean Sea, by sorting the populations into four groups according to their migratory routes covering the distribution of the plant as follows (Figure 1): the Iberian route (IB, populations 1-31), the Balearic route (BL, populations 32-49), the Italian route (IT, populations 50-60), and the Eastern

route (EA, populations 61-74). The Eastern route includes two known migratory paths, one that goes through the Aegean (AG) Sea and a route that descended through Turkey (TK) (Figure 1), as we only had three populations overlapping with the TK route. For each test performed, we used 10 000 permutations to test the significance of the IBD patterns.

Genetic clustering

Recent studies have highlighted the importance of applying different clustering analyses to verify the number of genetic groups (Janes et al., 2017). Therefore we applied three types of clustering analysis: (i) a distance-based method called DAPC (Jombart, Devillard, & Balloux, 2010); (ii) a Bayesian clustering analysis (STRUCTURE, Pritchard, Stephens, & Donnelly, 2000); and (iii) an unweighted paired group method with arithmetic mean (UPGMA), based on Nei's genetic distance (Sokal & Sneath, 1963).

DAPC is a non-model-based method that builds genetic clusters from genetic data by synthesizing the original variables (alleles), and therefore, it does not assume HWE or linkage disequilibrium (Jombart et al., 2010). The most likely K value (i.e., the number of clusters) is selected by the *Bayesian information criterion* (BIC) and the optimal K value is identified through an accompanying decrease in the BIC (Jombart et al., 2010). STRUCTURE is a Bayesian model-based method that presumes HWE and linkage disequilibrium, and analyses the individual multilocus genotypes, assigning individuals to clusters according to their probability of assignment. The STRUCTURE analysis was performed with a burn-in period of 50 000 iterations, followed by 200 000 with MCMC, assuming an admixture model of the population structure and with K values ranging from 1 to 25 (10 runs per K value). The geographic location of the populations was not considered in the analysis. The final value of K was selected according to the Evanno method (Evanno, Regnaut, & Goudet, 2005), as implemented in Structure Harvester (Earl & vonHoldt, 2012). Although STRUCTURE generally performs well, its results can be misleading when the study system does not meet the general conditions of the underlying model (i.e. HWE and linkage disequilibrium). This is typically the case when the genetic structure varies gradually due to isolation by distance or because of strong environmental gradients (Lombaert, Guillemaud, & Deleury, 2018, and references therein). Thus, we have combined three genetic clustering methods with different methodological approaches to validate and strengthen our results. To

confirm the overall large-scale genetic structure (i.e., clusters), we performed an UPGMA as implemented in the R package “poppr” (Kamvar, Brooks, & Grünwald, 2015), based on 10 000 replications obtained by bootstrapping.

We additionally evaluated the effects of the geographical components on the SGS through a spatial Principal Component Analysis (sPCA hereafter), as implemented in the adegenet R package version 2.0.1 (Jombart, 2008). sPCA uses Moran’s index I (Moran, 1948, 1950) to detect SGS comparing allelic frequencies observed in one population to the values observed in the surrounding populations. The Moran’s index is positive when the populations at neighbouring places show similar values, and negative when close populations show dissimilar values, which denotes local and/or global genetic structures, respectively (Jombart, 2015). Significance of the spatial patterns was tested using 999 permutations, with the function `spca_randtest`, implemented in the package adegenet (Jombart, 2008), as described by Montano and Jombart (2017).

Genetic networks

We built a weighted network, based on Reynolds genetic distances (Reynolds, Weir, & Cockerham, 1983), to depict the overall distributions of the genetic variations among the populations with the software EDENetwork (Kivelä, Arnaud-Haond, & Saramäki, 2015). Each node of the network represented a single population, with a size proportional to its betweenness centrality (i.e., the fraction of the shortest paths between pairs of nodes that pass through a given node), while link length, width and colour were proportional to the genetic distance between populations with shorter, wider and greener links showing low genetic distances. Overall, unlike pairwise-based metrics, these network metrics depict the genetic distances among populations, considering all sampled populations simultaneously. We tested for the statistical significance of the results by bootstrapping (i.e., resampling), by analysing the robustness of the betweenness centrality of the populations after re-sampling 85 and 50 % of the sampled populations for 1000 randomizations. We used Reynolds’ distances instead of Nei’s distances to statistically test the network, because the software requires raw genotype data to compute the Reynolds’ distance.

Nonetheless, a Nei's distances network was also computed with consistent results (Fig. S1a). Then, we conducted a modularity analysis with the R package igraph (Csardi & Nepusz, 2006), to detect substructures (modules) within the obtained network, where populations are more connected with other populations within their own module, than with populations outside their module. Modularity was estimated with the Walktrap algorithm (Pons & Latapy, 2005), that has been found to be accurate when detecting modules in small networks (less than 233 nodes) under different networks topologies (Yang, Algesheimer, & Tessone, 2016). The modularity was measured in the global network, and in the networks created according to the bird migratory routes. Groups in these networks were previously established in the IBD section.

Inference of ancestral relationship among populations

We applied Approximate Bayesian Computation (ABC) to test the influence of historical processes on the observed population genetic structure of *P. lentiscus*. This tool allows us to compare different competing scenarios describing ancestral relationships among the populations (e.g. divergence or migration), to explain the observed data. The approach is based on the generation of simulated datasets according to a range of values of demographic parameters that define different scenarios. We applied the ABC models to generate simulated data sets for the two competing scenarios, and to estimate their posterior probabilities that will be compared with the observed demographic parameters through summary statistics (e.g. mean population genetic diversity). These analyses were performed with DIYABC (Cornuet et al., 2008). We compared the two different competing demographic scenarios presumably describing the genetic divergence among the *P. lentiscus* populations across the Mediterranean Basin. The first scenario assumes that the populations from North Africa colonized the southern basin of the European region after the last glaciation, as previously described for numerous other plant species (Nieto-Feliner, 2014). After the historical colonization event, the populations on both sides of the Mediterranean Sea are thought to have remained genetically isolated (lack of gene flow across the Mediterranean Sea) (Fig. S2a). This model tests whether the observed pattern of lack of geographic barrier across the Mediterranean Basin could be explained by historical demographic expansions. Alternatively, the second model proposes that a single ancestral population from North Africa diverged into two isolated clusters located at the Eastern and Western extremes of the basin. Then, the individuals from both clusters would have colonized Europe through the Mediterranean Sea, and therefore, this would not act as a geographic barrier (Fig. S2b). This model tests whether the observed

genetic clustering could be already set in the ancestral population rather than the result of the lack of contemporary gene flow between eastern and western populations. We generated 1×10^6 simulated datasets for the 15 microsatellite loci, for each scenario, under a Generalized Stepwise Mutation model (GSM). Model details are included in Appendix S1. We retained the values of the following summary statistics: number of alleles, mean genetic diversity (Nei, 1978), mean size variance, pairwise genetic differentiation between groups (F_{st} ; Weir & Cockerham, 1984), shared allele distance between populations (Jin & Chakraborty, 1994), and genetic distance ($d\mu^2$) (Goldstein, Ruiz-Linares, Cavalli-Sforza, & Feldman, 1995). We tested whether the value of the summary statistics obtained for our study species overlapped with the posterior predictive distribution inferred for each hypothesized model by applying a Principal Component Analysis (PCA).

3. Results

3.1 Population genetic diversity

The number of alleles per locus varied from 5 to 43 (Table S2), with a total of 260 alleles detected across populations. Furthermore, for all populations, the unbiased expected heterozygosity (uH_e) ranged from 0.42 to 0.71 (average of 0.59 ± 0.06) and the rarefied allelic richness (Ar^*) varied from 2.29 to 3.74 (average of 2.81 ± 0.03). The frequency of private alleles ranged from 0 to 0.2 (Table S1). The inbreeding coefficient (F_{is}) was significant for 32 populations, 29 of which had positive values. In total, 45 out of the 74 populations showed signals of Hardy-Weinberg disequilibrium (Table S1), and linkage disequilibrium was significant for most pairs of loci (Table S3). The results showed a negative correlation between the frequency of private alleles and the rarefied allelic richness (Ar^*) with latitude ($\rho = -0.24$, p-value < 0.05 ; and $\rho = -0.23$, p-value < 0.05 , respectively) but a positive correlation with longitude ($\rho = 0.25$, p-value < 0.05 ; and $\rho = 0.59$, p-value < 0.001 , respectively) (Fig. S3 and Fig. S4). AMOVA analysis showed that the molecular variance was greater within populations (87 %) than among populations (13 %).

3.2 Genetic differentiation and population structure across the distribution range

Genetic differentiation (F_{st}) across the populations was 0.134 (0.118 - 0.204, p-value < 0.001) and the correlation between the genetic distances and the geographic distances was low, but significant

($\rho = 0.0001793$, $r^2 = 0.56$, $p\text{-value} < 0.001$) across the entire distribution range (Fig. S5), suggesting the presence of an IBD pattern ($p\text{-value} < 0.0001$). Similarly, an IBD pattern was also detected when latitudinal and longitudinal distances were analysed ($p\text{-value} < 0.0001$ in both cases). However, according to the coefficient of determination (r^2), the model that included longitude fitted better the genetic distances between the populations, in comparison to the model that included latitude ($\rho = 0.0001697$, $r^2 = 0.51$ and $\rho = 0.0002974$, $r^2 = 0.20$, respectively, both $p\text{-values} < 0.001$) (Fig. S5). The IBD was significant between the populations found in the IB and EA migratory routes, but not for the BL and IT routes (Table 1, Fig. S6).

All clustering analyses (DAPC, STRUCTURE, and UPGMA) concurred to show two main genetic clusters that broadly encompassed the eastern and western populations (Figure 2). The STRUCTURE and DAPC analyses showed two clusters ($K = 2$) with a western group that encompassed populations from the Canary Islands to the Sicily strait and an eastern group that included the populations from the Cyrenaica Peninsula to Israel (Figure 2a, b). Populations from the Balkans (populations 61 and 62) and Aegean Islands (populations 63 and 64) were mixed and were classified into both clusters. We report the assignment probability on an individual basis in Table S4 and the DeltaK values estimated by STRUCTURE in Fig. S7. In accordance with all previous results, the UPGMA phylogram placed all populations into two main groups that represented the eastern and western populations (Fig. S8). Results of sPCA pointed out the existence of a significant global genetic structure ($p\text{-value} < 0.001$) across populations with two main groups mirroring the East-West pattern described above (Figure 2c, Fig. S9). Clustering analysis within each of the identified groups did not detect any significant substructures within each group, although they both showed moderate levels of genetic structure ($F_{st} = 0.084$ and 0.054 for the western and eastern clusters, respectively, both $p\text{-value} < 0.001$)

The genetic network based on the Reynolds' genetic distances, also revealed two groups (Figure 3). The first included all populations from the western Mediterranean Basin (from the Strait of Gibraltar to the Strait of Sicily) and the Canary Islands, and the second contained the eastern Mediterranean locations. Populations from the Balkan Peninsula (Alyhi and Amarousion, 61 and 62 in the network) and the Aegean Islands (Amorgos and Lesvos islands, populations 63 and 64 in the network) scored the highest betweenness centrality in the network, and they acted as links between the two groups. The Sicilian populations (populations 58 and 59 in the network) and those from Linosa Island (population 60 in the network) also played an important role in

connecting the groups. These results were confirmed by the betweenness centrality analyses, which demonstrated the importance of the mentioned populations in the topology of the network (populations 61, 63, 45, 60, 58, and 62 were in the top five nodes of the 50 % bootstrap results, Figure 4). The network topology was robust with both 85 and 50 % bootstrapping (results of 50 % bootstrapping are shown in Fig. S1b). This overall structure defined by two large groups connected by the Aegean populations (61-64) was confirmed by a modularity analysis that identified three modules despite low modularity levels (0.048). Networks built with populations sorted by bird migratory routes showed low modularity with networks from routes IB, BL, and IT, composed by the main single module and EA network composed of the two main modules (Table S5) that roughly corresponded to the two migratory routes in the eastern Mediterranean (Figure 1, Table 1, Fig. S10).

3.3 Inference of ancestral relationship among populations

ABC analyses showed that both the hypothesized models (migration from Africa with subsequent genetic isolation hypothesis *vs.* ancestral East-West clustering coupled with gene flow, Fig. S2) were equally likely. The PCA explained the observed data in equal proportions for both models (Figure 5). In addition, some of the summary statistics showed low probabilities, which indicates that the proposed models are not suitable for explaining the observed data (Table S6). We also tested both hypothesized models with admixture events, but the results did not improve, compared to those obtained with the simple models (details in Fig. S11).

4. Discussion

Overall, our results have identified significant genetic structure in the populations of *P. lentiscus* in the Mediterranean Basin, with two distinctive genetic clusters located in the Eastern and Western ranges of the species distribution. The observed structure suggest that the Mediterranean Sea does not act as a geographical barrier for this species, as populations from both sides of the sea showed low levels of genetic distance. These findings are consistent with the hypothesis that migrating frugivorous birds connect the distant populations across the Mediterranean Sea, by mobilizing a sizeable number of seeds as they migrate between Europe and Africa. Further

analysis failed to find evidence that historical population expansions shaped the observed genetic patterns.

The SGS of the populations of *P. lentiscus* across its range of distribution reveals that the Mediterranean Sea does not act as a geographic barrier. All applied analyses concurred in finding two distinctive genetic clusters distributed across an East-West axis. If the Mediterranean Sea had imposed a barrier to gene flow, the genetic clusters would split along a North-South axis, with populations on both sides of the Mediterranean showing increased genetic distances. Previous studies assessing the spatial distributions of the genetic variations across the Mediterranean for this species have found mixed results, based on random amplification of polymorphic DNA (RAPDs) markers: while some reported a high genetic similarity among distant eastern and western populations (Barazani, Dudai, & Golan-Goldhirsh, 2003; Nahum, Inbar, Ne'eman, & Ben-Shlomo, 2008), others described significant genetic differences between populations from the Iberian Peninsula and the North of Africa (Werner, Sánchez-Gómez, Carrión-Vílches, & Guerra, 2002b). Overall, our study showed similar levels of genetic differentiation ($F_{st} = 0.134$), as those of previous studies with woody long-lived species (Hamrick, Godt, & Sherman-Broyles, 1992). By using a comprehensive sampling design and more informative molecular markers (SSRs), our study has described the spatial distribution of the genetic variation in greater detail, and has provided new insights into the processes that might have underpinned the observed structure. These findings highlight the importance of including populations across the entire distribution range when assessing the partitions of genetic variation in plant species (Medrano & Herrera, 2008; Hou & Lou, 2011).

Overall, our results showed that the genetic variation followed a predictable geographic trend that suggests a role of frugivorous birds in transporting ingested seeds across the Mediterranean when they migrate from Europe to Africa during the fruiting season of the species. Thus, the frequency of private alleles and Ar^* correlated negatively with latitude and positively with longitude, suggesting higher a genetic diversity when moving southwards and eastwards. Moreover, genetic distances between populations were better explained by the model that included longitudinal distance ($r^2 = 0.51$) than by the model that included latitudinal distance ($r^2 = 0.20$). These findings are congruent with the known migratory pathways followed by birds, including frugivores, across the Mediterranean Basin (Figure 1). Migrant frugivorous birds perform long-distance seed dispersal services when foraging upon fleshy-fruited plants, even overseas (Viana et al., 2016a).

Previous studies have also suggested that long-distance migratory bird movements are a key driver in fostering demographic expansion and shaping the genetic structure of bird-dispersed plants (Hampe et al., 2003; Nahum et al., 2008). Bird migration routes are determined by species-specific characteristics, such as food requirements, as well as climatic conditions (Newton, 2007), which might impose strong differences in the number of frugivorous birds migrating through each known migratory pathway (Figure 1). For instance, it was recorded that frugivorous birds, such as *E. rubecula* and *S. atricapilla*, migrate preferentially through the BL route rather than through the IB route (Migra programme, SEO/Birdlife, Spain (2012)). This differential use of migratory pathways by frugivorous birds might explain a significant IBD pattern found for populations distributed across the IB route (either including or excluding the population in the Canary Islands), but an absence of the IBD pattern among populations located along the BL route. In addition, migratory birds might also perform vagrant movements that deviate from the main migratory routes (Newton, 2007).

Besides the importance of contemporary long-distance seed dispersal events, the spatial distribution of the genetic variation is also known to be impacted by historical shifts in the distribution ranges (Hewitt, 2000). For example, the demographic expansion of plant populations northwards from the southern Mediterranean refugia after the Last Glacial Maximum (*ca.* 20 k years ago), reduced the genetic variation at the colonization front (Hampe et al., 2003), although some species show increased genetic diversity at intermediate latitudes, as a result of genetic admixtures from different refugia (Petit et al., 2003). Moreover, previous phylogeographic studies have described a divergence between populations located in the Eastern and Western regions of the Mediterranean Basin for many plant species (Nieto-Feliner, 2014). Therefore, the genetic patterns observed for *P. lentiscus* could also be explained by historical demographic processes. The decreased values of Ar^* and the frequency of private alleles northward, could result from a historical recolonization process from the southern populations after the last glacial retreat. Moreover, the high values of uH_e observed in the Aegean populations could be partially explained by an admixture process that would combine propagules from putative eastern and western refugia. However, the ABC models failed to find any statistical support for a demographic expansion northward from the southern populations, neither for an admixture process nor for ancestral eastern and western divergent populations. These findings suggest that the spatial distribution of the genetic variation in *P. lentiscus* is not fully explained by the historical

demographic processes, and therefore, contemporary long-distance seed dispersal remains a feasible driver with which to interpret our results. Future phylogeographic studies based on Next Generation Sequencing tools, such as deep sequencing, will allow us to fully understand the impact of historical events in shaping the current genetic patterns of this species.

Pollen dispersal patterns also shape the SGS of plants (Duminil et al., 2007). As anemophilous species, wind strength and direction during flowering time would determine pollen flow patterns in *P. lentiscus*. The prevailing winds in the Mediterranean Basin flow from the Northwest to the Southeast from March to May with a greater angular variance in the western than in the eastern regions (Soukissian, Karathanasi, Axaopoulos, Voukouvalas, & Kotroni, 2018). Even though pollen flow might account for low genetic distances among some populations at both shores of the Mediterranean Sea (e.g., Iberian Peninsula and Balearic populations), it does not explain the observed patterns for other populations (e.g., low genetic distance among Aegean and Libyan populations). Furthermore, most pollen dispersal distances for this species have been reported to remain under 500 m (Albaladejo, Guzmán, González-Martínez & Aparicio, 2012), while a recent study documents that pollen dispersal distances in wind-pollinated woody plants can reach up to 10 km (Souza, 2018 and references therein). These dispersal distances indicate that pollen flow contributes to the observed genetic patterns at the landscape level rather than across continents (Parejo-Farnés et al., 2017). Note that the shortest distance between Europe and Africa is *ca.* 14 km (Strait of Gibraltar). Other studies should include chloroplast markers (maternally inherited) to undoubtedly discern the relative contribution of pollen and seed dispersal to gene flow and the genetic patterns in *P. lentiscus*.

Besides all of the above-mentioned pollen and seed dispersal processes of *P. lentiscus*, ancient civilizations also traded products that were derived from this species, including seeds, across the Mediterranean Basin for millennia (Piluzza et al., 2015 and references therein; Alonso et al., 2016). Therefore, the observed genetic patterns might also be impacted to some extent by the historical maritime trade routes, as already hypothesized to explain the genetic patterns among the olive tree populations (di Rienzo et al., 2018). Active commerce by humans can blur the genetic footprints from natural dispersal processes, and this would explain the lack of statistical support for the tested historical (ABC) models.

Overall, this study shows that the Mediterranean Sea does not act as a geographic barrier among populations of *P. lentiscus*, a bird-dispersed plant. The observed genetic patterns suggest an important role for the migratory birds, acting as seed dispersal vectors among the distant populations on both shores of the Mediterranean. Therefore, any shift of the migratory bird routes would potentially impact the demographic and genetic patterns of the plant species dispersed by these frugivores. Recent studies (Bairlein, 2016) have pointed out that climate change, fragmentation, and poaching, threaten the populations and the migratory pathways of European birds. Climate change might decouple plant-frugivore interactions by shifting the fruiting phenology of the fleshy-fruited plants and/or interrupting the bird migration patterns, as previously reported (Zalakevicius, Bartkeviciene, Raudonikis, & Janulaitis, 2006; Both et al., 2009; Pulido & Berthold, 2010; Thackeray et al., 2010). These changes could have profound demographic and genetic consequences for bird-dispersed plants, by interrupting the arrival of the propagules and the connectivity patterns among the distant populations, potentially compromising the persistence of the plants that rely on birds to disperse their seeds.

Acknowledgements

We owe a debt to everyone who sent us samples for this study (Table S7), as it would have been impossible without these valuable contributions. We would like to thank the editor and the five anonymous reviewers, for their valuable comments and suggestions that have greatly improved this manuscript. We also thank M. Begon for kindly checking the English grammar and writing style. P. DIR. is supported by Project 19908/GERM/2015 of Regional Excellence (Fundación Seneca). V.M.L. was supported by a predoctoral FPU studentship (FPU13/05115). C.G. was supported by the Investigador FCT Programme (IF/01375/2012) and PLANTSHIFTS (PTDC/BIA-BIC/5233/2014 – POCI-01-0145-FEDER-016817) funded by national funds via FCT (Fundação para a Ciência e a Tecnologia, FCT) and FEDER funds through POCI-COMPETE. She is currently supported by the University of Liverpool (UK) Tenure-track Fellow program.

References

- Acosta-Rojas D. C., Jiménez-Franco, M. V., Zapata-Pérez, V., De la Rúa, P. & Martínez-López, V. (2019). An integrative approach to discern the seed dispersal role of frugivorous guilds in a Mediterranean semiarid priority habitat. *PeerJ*, 7, e7609. doi: 10.7717/peerj.7609
- Albaladejo, R. G., Sebastiani, F., Aparicio, A., Buonamici, A., González-Martínez, S. C., & Vendramin, G. G. (2008). Development and characterization of eight polymorphic microsatellite loci from *Pistacia lentiscus* L. (Anacardiaceae). *Molecular Ecology Resources*, 8(4), 904-906. doi: 10.1111/j.1755-0998.2008.02110.x
- Albaladejo, R. G., Guzmán, B., González-Martínez, S. C., & Aparicio, A. (2012). Extensive pollen flow but few pollen donors and high reproductive variance in an extremely fragmented landscape. *PloS One*, 7(11), e49012. doi: 10.1371/journal.pone.0049012
- Alonso, N., Jordà, G. P., Rovira, N., & Reyes, D. L. (2016). Gathering and consumption of wild fruits in the east of the Iberian Peninsula from the 3rd to the 1st millennium BC. *Quaternary International*, 404, 69-85. doi: 10.1016/j.quaint.2015.07.021
- Arizaga, J., Alonso, D., & Barba, E. (2010). Patterns of migration and wintering of robins *Erithacus rubecula* in northern Iberia. *Ringing & Migration*, 25(1), 7-14. doi: 10.1080/03078698.2010.9674408
- Bairlein, F. (2016). Migratory birds under threat. *Science*, 354(6312), 547-548. doi: 10.1126/science.aah6215
- Barazani, O., Dudai, N., & Golan-Goldhirsh, A. (2003). Comparison of Mediterranean *Pistacia lentiscus* genotypes by random amplified polymorphic DNA, chemical, and morphological analyses. *Journal of Chemical Ecology*, 29(8), 1939-1952. doi: 10.1023/A:1024862614345
- Binks, R. M., Millar, M. A., & Byrne, M. (2015). Not all rare species are the same: contrasting patterns of genetic diversity and population structure in two narrow-range endemic sedges. *Biological Journal of the Linnean Society*, 114(4), 873-886. doi: 10.1111/bij.12465
- Bloch, R., & Bruderer, B. (1982). The air speed of migrating birds and its relationship to the wind. *Behavioral Ecology and Sociobiology*, 11(1), 19-24.

Bolós, O., & Vigo, J. (1990). Flora dels Països Catalans, Vol. 2. Editorial Barcino, Barcelona.

Both, C., Van Turnhout, C. A., Bijlsma, R. G., Siepel, H., Van Strien, A. J., & Foppen, R. P. (2009). Avian population consequences of climate change are most severe for long-distance migrants in seasonal habitats. *Proceedings of the Royal Society B: Biological Sciences*, 277(1685), 1259-1266. doi: 10.1098/rspb.2009.1525

Bruderer, B., & Liechti, F. (1999). Bird migration across the Mediterranean. In Proc Int Ornithol Congr, Durban. Johannesburg: BirdLife South Africa (pp. 1983-1999).

Cornuet, J. M., Santos, F., Beaumont, M. A., Robert, C. P., Marin, J. M., Balding, D. J., ... & Estoup, A. (2008). Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, 24(23), 2713-2719. doi: 10.1093/bioinformatics/btn514

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695(5), 1-9.

di Rienzo, V., Sion, S., Taranto, F., D'Agostino, N., Montemurro, C., Fanelli, V., ... & Zammit-Mangion, M. (2018). Genetic flow among olive populations within the Mediterranean basin. *PeerJ*, 6, e5260. doi: 10.7717/peerj.5260

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J., & Collen, B. (2014). Defaunation in the Anthropocene. *Science*, 345(6195), 401-406. doi: 10.1126/science.1251817

Duminil, J., Fineschi, S., Hampe, A., Jordano, P., Salvini, D., Vendramin, G. G., & Petit, R. J. (2007). Can population genetic structure be predicted from life-history traits? *The American Naturalist*, 169(5), 662-672. doi: 10.1086/513490

Dyer, R. J., & Nason, J. D. (2004). Population graphs: the graph theoretic shape of genetic structure. *Molecular Ecology*, 13(7), 1713-1727. doi: 10.1111/j.1365-294X.2004.02177.x

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359-361. doi:10.1007/s12686-011-9548-7

Eckert, C. G., Samis, K. E., & Loughheed, S. C. (2008). Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology*, 17(5), 1170-1188. doi: 10.1111/j.1365-294X.2007.03659.x

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14(8), 2611-2620. doi: 10.1111/j.1365-294X.2005.02553.x

Fiedler, W., & Davidson, S. (2012). Movebank: eine offene Internetplattform für Tierwanderungsdaten [Movebank: an open internet platform for animal movement data]. *Vogelwarte* 50(1): 15–20.

Gapare, W. J., Aitken, S. N., & Ritland, C. E. (2005). Genetic diversity of core and peripheral Sitka spruce (*Picea sitchensis* (Bong.) Carr) populations: implications for conservation of widespread species. *Biological Conservation*, 123(1), 113-123. doi: 10.1016/j.biocon.2004.11.002

Goldstein, D. B., Ruiz-Linares, A. R., Cavalli-Sforza, L. L., & Feldman, M. W. (1995). Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Sciences*, 92(15), 6723-6727.

González-Astorga, J., & Castillo-Campos, G. (2004). Genetic variability of the narrow endemic tree *Antirhea aromatica* Castillo-Campos & Lorence,(Rubiaceae, Guettardeae) in a tropical forest of Mexico. *Annals of Botany*, 93(5), 521-528. doi: 10.1093/aob/mch070

Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Resources*, 5(1), 184-186. doi: 10.1111/j.1471-8286.2004.00828.x

Hampe, A., Arroyo, J., Jordano, P., & Petit, R. J. (2003). Rangewide phylogeography of a bird-dispersed Eurasian shrub: contrasting Mediterranean and temperate glacial refugia. *Molecular Ecology*, 12(12), 3415-3426. doi: 10.1046/j.1365-294X.2003.02006.x

Hamrick, J. L., & Godt, M. J. (1989). Allozyme diversity in plant species. *Plant Population Genetics, Breeding and Genetic Resources*, 43-63.

Hamrick, J. L., Godt, M. J. W., & Sherman-Broyles, S. L. (1992). Factors influencing levels of genetic diversity in woody plant species. *New Forests*, 6(1), 95-124.

- Hamrick, J. L., & Godt, M. J. W. (1996). Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions: Biological Sciences*, 1291-1298. doi: 10.1098/rstb.1996.0112
- Herrera, C. M. (1984). A study of avian frugivores, bird-dispersed plants, and their interaction in Mediterranean scrublands. *Ecological Monographs*, 54(1), 1-23.
- Herrera, C. M. (1998). Long-term dynamics of Mediterranean frugivorous birds and fleshy fruits: a 12-year study. *Ecological Monographs*, 68(4), 511-538.
- Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405(6789), 907.
- Hijmans, R. J. (2019). geosphere: Spherical Trigonometry. R package version 1.5-10. <https://CRAN.R-project.org/package=geosphere>
- Hou, Y., & Lou, A. (2011). Population genetic diversity and structure of a naturally isolated plant species, *Rhodiola dumulosa* (Crassulaceae). *Plos One*, 6(9), e24497. doi: 10.1371/journal.pone.0024497
- Izhaki, I., & Safriel, U. N. (1985). Why do fleshy-fruit plants of the Mediterranean scrub intercept fall-but not spring-passage of seed-dispersing migratory birds? *Oecologia*, 67(1), 40-43.
- Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., & Andrew, R. L. (2017). The K= 2 conundrum. *Molecular Ecology*. 26(14), 3594-3602. doi: 10.1111/mec.14187
- Jin, L., & Chakraborty, R. (1994). Estimation of genetic distance and coefficient of gene diversity from single-probe multilocus DNA fingerprinting data. *Molecular Biology and Evolution*, 11(1), 120-127.
- Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403-1405. doi: 10.1093/bioinformatics/btn129
- Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, 11(1), 94. doi: 10.1186/1471-2156-11-94

Jombart, T. (2015). A tutorial for the spatial Analysis of Principal Components (sPCA) using adegenet 2.0.0. *Imperial College London*.

Jordano, P. (1989). Pre-dispersal biology of *Pistacia lentiscus* (Anacardiaceae): cumulative effects on seed removal by birds. *Oikos*, 55(375), D386.

Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, 6. doi: 10.3389/fgene.2015.00208

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, 4(8), 782-788. doi: 10.1111/2041-210X.12067

Keynan, N., Geller-Bernstein, C., Waisel, Y., Bejerano, A., Shomer-Ilan, A., & Tamir, R. (1987). Positive skin tests to pollen extracts of four species of *Pistacia* in Israel. *Clinical Allergy*, 17, 243-249.

Kivelä, M., Arnaud-Haond, S., & Saramäki, J. (2015). EDENetworks: A user-friendly software to build and analyse networks in biogeography, ecology and population genetics. *Molecular Ecology Resources*, 15(1), 117-122. doi: 10.1111/1755-0998.12290

Lavergne, S., Hampe, A., & Arroyo, J. (2013). In and out of Africa: how did the Strait of Gibraltar affect plant species migration and local diversification? *Journal of Biogeography*, 40(1), 24-36. doi: 10.1111/j.1365-2699.2012.02769.x

Leal, A., Monrós, J. S., & Barba, E. (2004). Migration and wintering of blackcaps *Sylvia atricapilla* in Eastern Spain. *Ardeola*, 51(2), 345-355.

Lesica, P., & Allendorf, F. W. (1995). When are peripheral populations valuable for conservation? *Conservation Biology*, 9(4), 753-760. doi: 10.1046/j.1523-1739.1995.09040753.x

Lombaert, E., Guillemaud, T., & Deleury, E. (2018). Biases of STRUCTURE software when exploring introduction routes of invasive species. *Heredity*, 1. doi: 10.1038/s41437-017-0042-1

- Martínez-López, V., De la Rúa, P., Zapata, V. M., & Robledano, F. (2017). Ecological and genetic consequences of fragmentation in a semiarid Mediterranean urban forest. *Urban Ecosystems*, 20, 1161-1168. doi: 10.1007/s11252-017-0668-z
- Medrano, M., & Herrera, C. M. (2008). Geographical structuring of genetic diversity across the whole distribution range of *Narcissus longispathus*, a habitat-specialist, Mediterranean narrow endemic. *Annals of Botany*, 102(2), 183-194. doi: 10.1093/aob/mcn086
- Milla, R., Castro-Díez, P., & Montserrat-Martí, G. (2010). Phenology of Mediterranean woody plants from NE Spain: synchrony, seasonality, and relationships among phenophases. *Flora-Morphology, Distribution, Functional Ecology of Plants*, 205(3), 190-199. doi: 10.1016/j.flora.2009.01.006
- Molina-Venegas, R., Aparicio, A., Lavergne, S., & Arroyo, J. (2015). The building of a biodiversity hotspot across a land-bridge in the Mediterranean. *In Proceedings of the Royal Society B*, 282(1813), 20151116. doi: 10.1098/rspb.2015.1116
- Montano, V., & Jombart, T. (2017). An Eigenvalue test for spatial principal component analysis. *BMC Bioinformatics*, 18(1), 562. doi: 10.1186/s12859-017-1988-y
- Moran, P. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society B*, 10, 243-251.
- Moran, P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37, 17-23.
- Motalebipour, E. Z., Kafkas, S., Khodaeiaminjan, M., Çoban, N., & Gözel, H. (2016). Genome survey of pistachio (*Pistacia vera* L.) by next generation sequencing: Development of novel SSR markers and genetic diversity in *Pistacia* species. *BMC Genomics*, 17(1), 998. doi: 10.1186/s12864-016-3359-x.
- Nahum, S., Inbar, M., Ne'eman, G., & Ben-Shlomo, R. (2008). Phenotypic plasticity and gene diversity in *Pistacia lentiscus* L. along environmental gradients in Israel. *Tree Genetics & Genomes*, 4(4), 777-785. doi: 10.1007/s11295-008-0150-4
- Nathan, R. (2006). Long-distance dispersal of plants. *Science*, 313(5788), 786-788. doi: 10.1126/science.1124975

- Nei, M. (1972) Genetic distances between populations. *American Naturalist*, 106, 283-292.
- Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89(3), 583-590.
- Newton, I. (2007). The Migration Ecology of Birds. *Academic Press*, London, UK.
- Nieto-Feliner, G. N. (2014). Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. *Perspectives in Plant Ecology, Evolution and Systematics*, 16(5), 265-278.
- Nogales, M., Heleno, R., Traveset, A., & Vargas, P. (2012). Evidence for overlooked mechanisms of long-distance seed dispersal to and between oceanic islands. *New Phytologist*, 194(2), 313-317. doi: 10.1111/j.1469-8137.2011.04051.x
- Parejo-Farnés, C., Robledo-Arnuncio, J. J., Albaladejo, R. G., Rubio-Pérez, E., & Aparicio, A. (2017). Effects of habitat fragmentation on parental correlations in the seed rain of a bird-dispersed species. *Tree Genetics & Genomes*, 13(1), 17. doi: 10.1007/s11295-017-1100-9
- Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research an update. *Bioinformatics*, 28, 2537e2539. doi: 10.1093/bioinformatics/bts460
- Petit, R. J., Aguinagalde, I., de Beaulieu, J. L., Bittkau, C., Brewer, S., Cheddadi, R., ... & Mohanty, A. (2003). Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, 300(5625), 1563-1565. doi: 10.1126/science.1083264
- Piluzza, G., Viridis, S., Serralutzu, F., & Bullitta, S. (2015). Uses of plants, animal and mineral substances in Mediterranean ethno-veterinary practices for the care of small ruminants. *Journal of Ethnopharmacology*, 168, 87-99. doi: 10.1016/j.jep.2015.03.056i
- Pons, P., & Latapy, M. (2005). Computing communities in large networks using random walks. In International symposium on computer and information sciences (pp. 284-293). Springer, Berlin, Germany.
- Popp, M., Mirré, V., & Brochmann, C. (2011). A single Mid-Pleistocene long-distance dispersal by a bird can explain the extreme bipolar disjunction in crowberries (*Empetrum*). *Proceedings of the National Academy of Sciences*, 108(16), 6520-6525. doi: 10.1073/pnas.1012249108

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945-959.

Pulido, F. & Berthold, P. (2010). Current selection for lower migratory activity will drive the evolution of residency in a migratory bird population. *Proceedings of the National Academy of Sciences*, 107(16), 7341-7346. doi: 10.1073/pnas.0910361107

R Development Core Team (2016) R: a language an environment for statistical computing. R foundation for statistical computing. Vienna, URL <http://www.R-project.org>

Raymond, M. & Rousset, F. (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86, 248-249.

Reynolds, J., Weir, B. S., & Cockerham, C. C. (1983). Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics*, 105(3), 767-779.

Rouskas, D. (1996). Conservation strategies of *Pistacia* genetic resources in Greece. In: Paludosi, S., Caruso, T. & Barone, E. (eds.) Taxonomy, distribution, conservation and uses of *Pistacia* genetic resources. *IPGRI*, Rome.

SEO/BirdLife. (2012). Análisis preliminar del banco de datos de anillamiento de aves del Ministerio de Agricultura, Alimentación y Medio Ambiente, para la realización de un atlas de migración de aves de España. *SEO/BirdLife-Fundación Biodiversidad*. Madrid.

Sokal, P.R. & Sneath, P.H.A. (1963). Principles of numerical taxonomy. *W. H. Freeman and Company*, San Francisco.

Soukissian, T., Karathanasi, F., Axaopoulos, P., Voukouvalas, E., & Kotroni, V. (2018). Offshore wind climate analysis and variability in the Mediterranean Sea. *International Journal of Climatology*, 38(1), 384-402. doi: 10.1002/joc.5182

Souza, F. B. D. (2018). Historic pollen and seed dispersal in fragmented populations of *Cariniana estrellensis* (Raddi) Kuntze and *Cariniana legalis* (Mart.) Kuntze. Thesis. Universidade Estadual Paulista, São Paulo, Brazil.

Thackeray, S. J., Sparks, T. H., Frederiksen, M., Burthe, S., Bacon, P. J., Bell, J. R., ... & Clutton-Brock, T. I. M. (2010). Trophic level asynchrony in rates of phenological change for marine,

freshwater and terrestrial environments. *Global Change Biology*, 16(12), 3304-3313. doi: 10.1111/j.1365-2486.2010.02165.x

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P., & Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4(3), 535-538. doi: 10.1111/j.1471-8286.2004.00684.x

Verdú, M., & García-Fayos, P. (2002). Reproductive ecology of *Pistacia lentiscus* L. (Anacardiaceae): an evolutionary anachronism in the Mediterranean shrubland. *Revista Chilena de Historia Natural*, 75(1), 57-65.

Viana, D. S., Gangoso, L., Bouten, W., & Figuerola, J. (2016a). Overseas seed dispersal by migratory birds. *In Proceedings of the Royal Society B*, 283(1822), 20152406. doi: 10.1098/rspb.2015.2406

Viana, D. S., Santamaría, L., & Figuerola, J. (2016b). Migratory birds as global dispersal vectors. *Trends in Ecology & Evolution*, 31(10), 763-775. doi: 10.1016/j.tree.2016.07.005

Vincenty, T. (1975). Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. *Survey Review*, 23(176), 88-93.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 1358-1370.

Werner, O., Ros, R. M., & Guerra, J. (2002a). Direct amplification and NaOH extraction: two rapid and simple methods for preparing bryophyte DNA for polymerase chain reaction (PCR). *Journal of Bryology*, 24(2), 127-131. doi: 10.1179/037366802125000980

Werner, O., Sánchez-Gómez, P., Carrión-Vilches, M. A., & Guerra, J. (2002b). Evaluation of genetic diversity in *Pistacia lentiscus* L. (Anacardiaceae) from the southern Iberian Peninsula and North Africa using RAPD assay. Implications for reforestation policy. *Israel Journal of Plant Sciences*, 50(1), 11-18.

Whitham, T. G., Bailey, J. K., Schweitzer, J. A., Shuster, S. M., Bangert, R. K., LeRoy, C. J., ... & Fischer, D. G. (2006). A framework for community and ecosystem genetics: from genes to ecosystems. *Nature Reviews Genetics*, 7(7), 510-523. doi:10.1038/nrg1877

Wright, S. (1943). Isolation by distance. *Genetics*, 28, 114-138.

Yang, Z., Algesheimer, R., & Tessone, C. J. (2016). A comparative analysis of community detection algorithms on artificial networks. *Scientific Reports*, 6, 30750. doi: 10.1038/srep30750

Zalakevicius, M., Bartkeviciene, G., Raudonikis, L., & Janulaitis, J. (2006). Spring arrival response to climate change in birds: a case study from eastern Europe. *Journal of Ornithology*, 147(2), 326-343. doi: 10.1007/s10336-005-0016-6

Zamora, M.C. (1997). Aprovechamientos forestales en la comarca del campo de Cartagena durante la Edad Media. *Revista electrónica de Geografía y Ciencias Sociales. Scripta Nova*, 13. Universidad de Barcelona [ISSN 1138-9788] URL: <http://www.ub.edu/geocrit/sn-13.htm> (Accessed the 1st of May 2018).

Data accessibility

- Information about microsatellites (Genebank accession) (Table S2)
- Populations information (Table S1)

Authors contribution

The study was conceived by V.M.L., F.R., and P.DIR. The sampling design and collection materials were organised by V.M.L. with the support of F.R., P.DIR., and V.Z. Lab work was performed by V.M.L. Data analysis was done by V.M.L. with the assistance of C.G. and P. DIR. The first draft of the manuscript was written by V.M.L. Comments from V.M.L., G.C, F.R., and P.DIR. contributed to the final version of the manuscript.

Figures

Figure 1. Distribution range of *P. lentiscus* (dark grey shadow) (based on Bolós & Vigo 1990) and a detailed photograph of the *P. lentiscus* fruits (upper left). Orange lines depict the main bird migratory routes from Europe to Africa based on Bruderer & Liechti (1999), Fiedler & Davidson (2012), and Bairlein (2016). In the text, we refer to the migratory routes as: Iberian route (IB), Balearic route (BL), Italian route (IT), Aegean route (AG), and Turkish route (TK). For some analyses AG and TK are combined due to a low number of sampled populations to create an Eastern route (EA). Black dots indicate the locations of the sampled populations for this study, with the ID of each population indicated by a number (see Table S1).

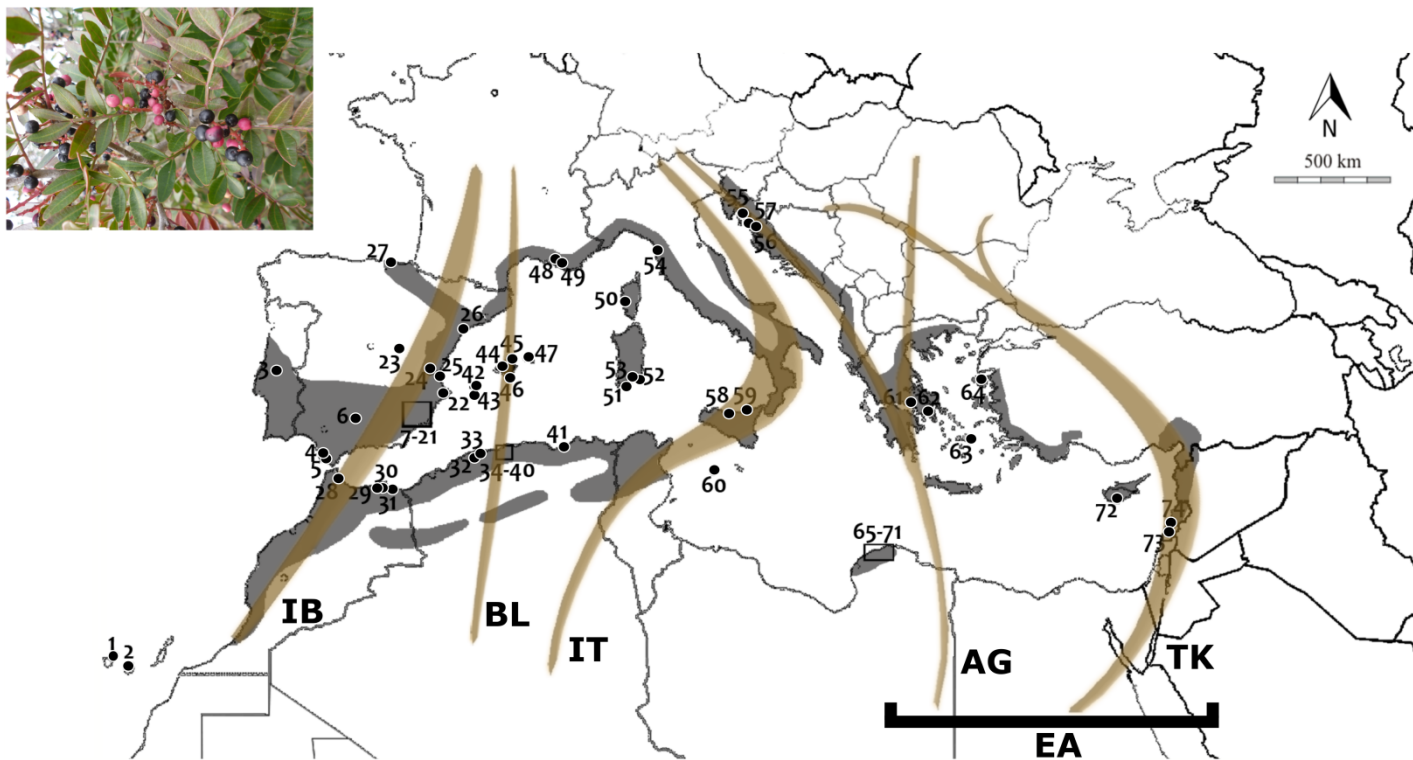


Figure 2. (a) Bar plot showing the STRUcTURE results (blue = group 1; red = group 2). Numbers refer to population IDs (Table S1). (A = Canary Islands, B = Iberian Peninsula, C = North Africa, D = Balearic Islands, E = Southeast France, F = Corsica and Sardinia, G = Italian Peninsula, H = Adriatic Sea, I = Sicily, J = Linosa Island, K = Balkans, L = Aegean Islands, M = Cyrenaican Peninsula, N = Cyprus, O = Middle East); (b) Discriminant Analysis of Principal Components (DAPC) for all populations (western populations in blue and eastern populations in red); (c) Spatial

Principal Component Analysis (sPCA) with dots depicting the location of sampled populations and the colours expressing the eigenvalues of the global structure retained in the first axis reported in Fig. S9. (Map from Map data © 2018 Google, ORION-ME).

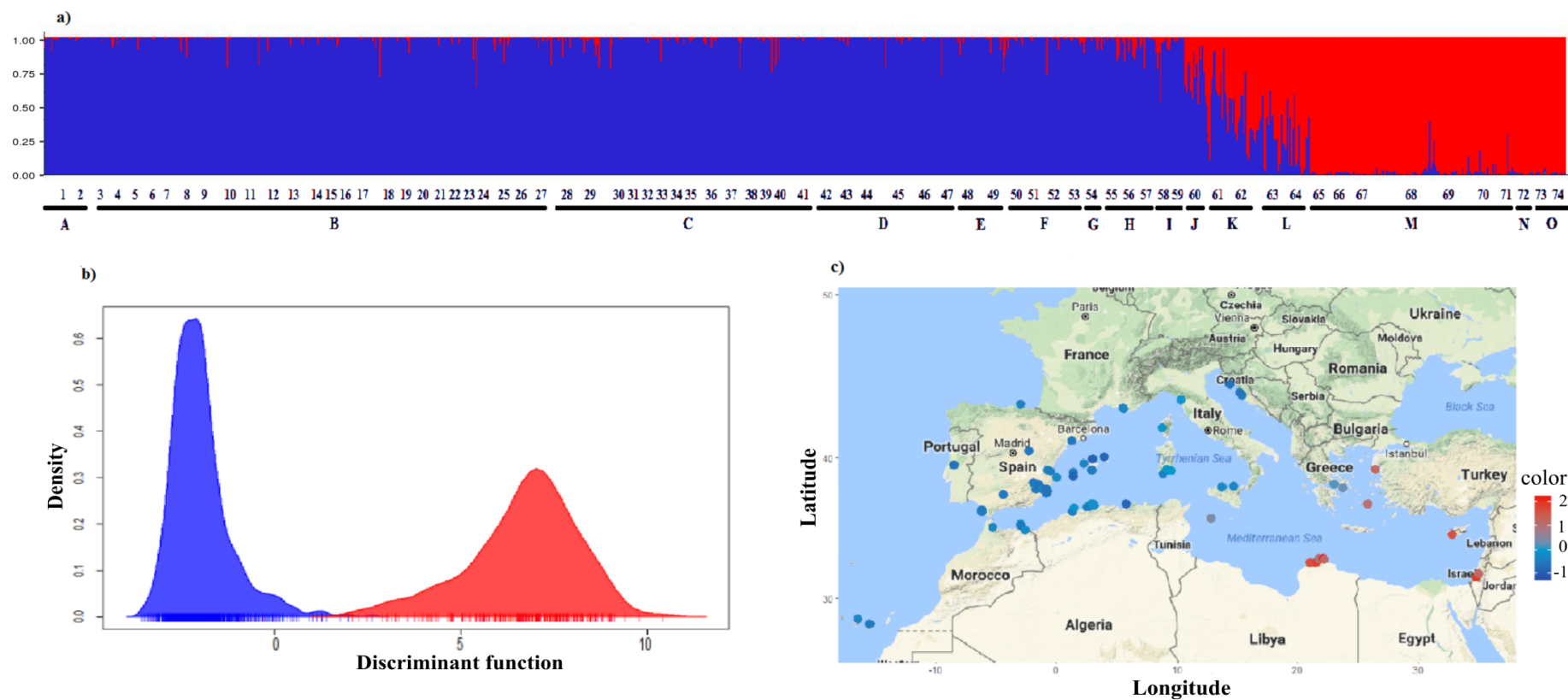


Figure 3. Genetic network obtained based on the Reynolds' genetic distances for all sampled populations. The nodes represent the sampled populations, numbered by their IDs (see Table S1). The size of each node is proportional to the betweenness centrality of the population in the network. The length, the width and the colour of the links are proportional to the genetic distances between the pairwise populations, with shorter, thicker and greener links indicating lower genetic distance and long, thin and blue links indicating increased values of genetic distances. The visualization of the network is based on the spring-charge-algorithm of Himmeli (see <http://www.finndiane.fi/software/Himmeli> for more details).

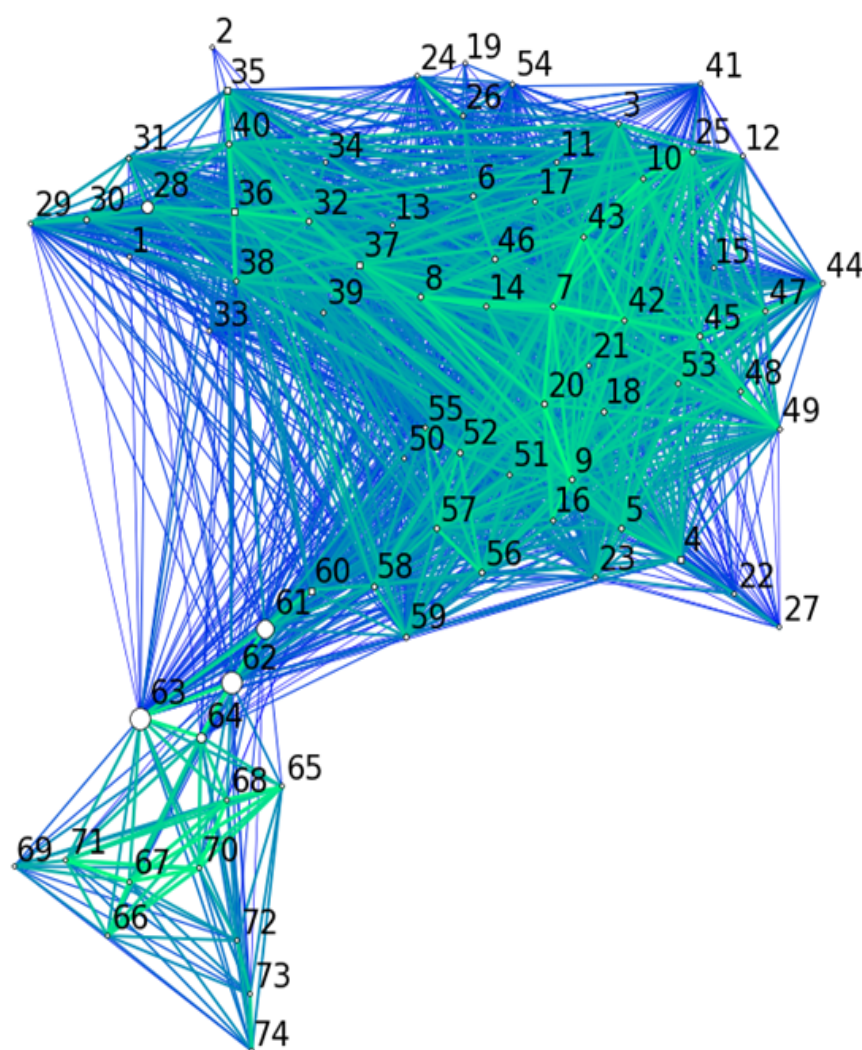


Figure 4. Number of times that a given population showed a high value of betweenness centrality (top 5) with the 50 % bootstrap analysis (see population codes in Table S1) (61 and 62 = Alyhi and Amaraouion, Balkan Peninsula; 63 = Amorgos Island, Aegean Islands; 45 = Majorca, Balearic Islands; 60 = Linosa Island; 58 = Palermo, Sicily; 7 = Tazona, Iberian Peninsula; 55 and 57 = Neresine and Diklo, Adriatic Sea; and 36 = Tipaza, North of Africa).

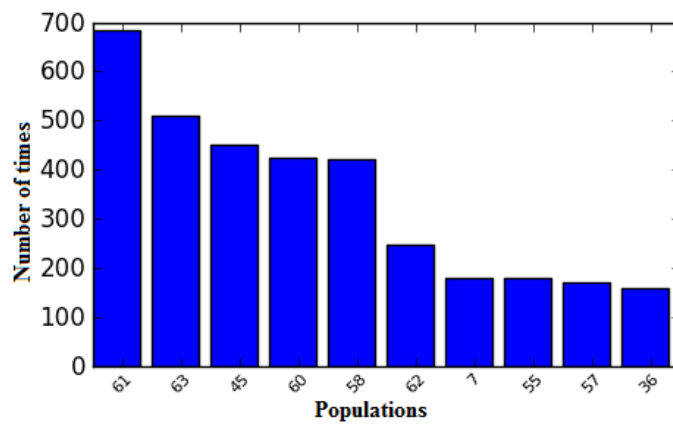


Figure 5. Model checking for the two hypothesized demographic models (Fig. S2): (a) Model 1 tests for a demographic expansion from ancestral populations in Africa to Europe at the end of the last glaciation followed by genetic isolation; (b) Model 2 tests for an ancestral population from North Africa strongly structured into an East and West clustered that expanded over Europe. The figure shows the results of a Principal Component Analysis (PCA) in the space of summary statistics with the 1% of the closest simulated data sets (10,000 in our case) and 1,000 data sets simulated from the posterior distribution parameters. If the model fits the observed data well, the yellow dot should be located in the middle of the cluster generated from the posterior predictive distribution data sets (green filled dots).

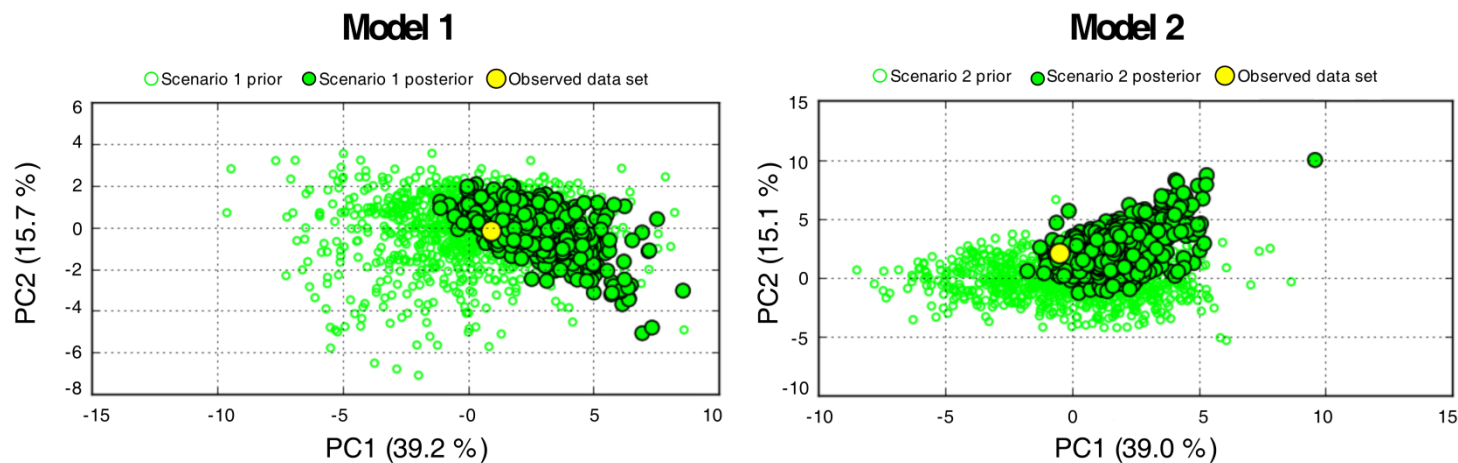


Table 1. Isolation by distance (IBD) patterns, statistical correlation between genetic and geographic distances, and the value of the modularity of the genetic networks (i.e. substructure/division inside the networks) obtained for populations sorted by migration routes between Europe and Africa (Figure 1) (Significance of results; NS = p-value > 0.05; * = p-value < 0.05; ** = p-value < 0.01; *** = p-value < 0.001). Migratory route names: Iberian route (IB), Balearic route (BL), Italian route (IT), and Eastern route (EA). Population IDs are indicated in Table S1.

Migratory route	Populations	No. of <i>P. lentiscus</i> individuals	IBD	Genetic distance vs geographic distance	Modularity
IB	1-31	420	***	$r^2=0.47^{***}$	0.0017
BL	32-49	243	NS	$r^2=0.05^{**}$	0.0023
IT	50-60	143	NS	$r^2=0.01^{NS}$	0.01
EA	61-74	251	**	$r^2=0.25^{***}$	0.11
All populations	1-74	1057	***	$r^2=0.56^{***}$	0.048