Chapter XX

Spatial genetic structure of an explicit glacial refugium of maritime pine (Pinus pinaster Aiton) in southeastern Spain.

S. C. Gonzáles-Martínez, A. Gómez, J. S. Carrión, D. Agúndez, R. Alía and L. Gil

Keywords
genetic structure, glacial refugia, likelihood, Mediterranean Basin, conifers

Correspondence
E-mail: santiago@inia.es

Abstract  The Iberian Peninsula has been described as a glacial refugium for numerous organisms. In particular, there is evidence both from pollen records and genetic studies that shows the existence of Mediterranean conifers (Pinus halepensis Miller, Pinus pinaster Aiton) in southeastern Spain during the last glacial stage. Data from eight polymorphic allozyme markers were used to study the spatial genetic structure of 11 native populations of maritime pine, P. pinaster, in this region. Models of isolation-by-distance were adjusted to different groups of populations to test specific hypotheses about the role of mountain ranges in shaping the spatial genetic structure of maritime pine in southeastern Spain. In addition, pairwise gene interchange was analyzed using migration matrix models and maximum likelihood methods to make joint estimates of dispersal rates and population sizes. A complex pattern in the distribution of gene diversity was found, involving historical isolation due to geographic variables for particular populations. The role of mountain ranges in glacial refugia (i) reducing the risk of a population bottleneck by altitudinal migration in response to climatic change, and (ii) acting as geographical barriers to gene flow, were studied.
Introduction

Over the last decade, phylogeographic studies have revealed high genetic diversity and richness in southern Europe for most temperate species, whereas low genetic variation has been usually found in northern populations. The richness of southern populations is assumed to be the result of persistence of populations and accumulation of variation over several glacial stages (Hewitt 1996, 1999, Bennett 1997). Postglacial colonization followed concordant south to north migration routes in many species from Europe and the Pacific Northwest of the United States (Hewitt 2001). Fossil records in central and northern Europe have shown that, when typical migration rates for mesothermophilous species are invoked, there would not have been enough time for some species to arrive from their putative southern refugia (i.e., Reid’s paradox, Clark et al. 1998). Rapid, long-distance dispersal of small groups ahead of the main dispersing populations has been suggested (Hewitt 1999), but increasing evidence supports the existence of cryptic northern refugia in sheltered areas with suitable microhabitats (Steward and Lister 2001, Willis et al. 2000). Thus, current populations in central and northern Europe might be a result of the interaction between local or regional survival and continental-scale migration.

Hewitt (2001) suggested that in southern refugia, genomes are greatly subdivided geographically due to survival in disjunct locations without large geographical displacement. Coalescent simulations in grasshoppers have shown that glaciations promoted divergence among populations due to (i) drift associated with colonization of previous glaciated areas and (ii) differentiation among multiple allopatric glacial refugia (Knowles 2001). Analysis of spatial genetic structure within putative glacial refugia can provide relevant information about the distribution of gene diversity in stable populations and its causes. This is particularly true in long-lived organisms like forest trees (see Petit & Vendramin, this volume), where no genetic structure is usually found in recently colonized regions, partly because there have been an insufficient number of generations for the accumulation of variation and corresponding divergence between populations.

Major refugial areas have been identified in the Iberian Peninsula for a diversity of organisms (Taberlet et al. 1998; Santucci et al. 1998; Comes & Abbot
1998; Willis & Whittaker 2000, Gómez & Lunt, this volume). Recent studies showed a very high amount of genetic variation in forest tree populations of southern Iberia. Ferris et al. (1998) found three major cpDNA types in European white oaks (Quercus robur and Q. petraea) and suggested a postglacial migration pathway from the Iberian Peninsula based on the distribution of one of them. More recently, a consortium of 16 laboratories have studied chloroplast DNA variation in European white oaks (Quercus robur, Q. petraea, Q. pubescens, Q. frainetto, Q. faginea, Q. pyrenaica, Q. canariensis and Q. macranthera). Four of the six chloroplast lineages recognized so far in European white oaks were represented in the Iberian Peninsula and there was strong evidence for at least two major refugia in Spain (Petit et al. 2002; Olalde et al. 2002). Sinclair et al. (1999), in a wide-range study of mitochondrial variants of Scots pine, Pinus sylvestris L, observed within-population genetic variation in Iberian populations, whereas elsewhere in Europe, populations were fixed for one mitotype. Moreover, a population in southern Spain (Baza) showed a private mitotype. The Baza population is located close to Sierra Nevada, a region with a great level of endemism.

Mediterranean pines such as Aleppo pine, Pinus halepensis, and maritime or cluster pine, P. pinaster, have also shown high levels of diversity in southern Spain. These levels could be related to persistence during several glacial cycles (Salvador et al. 2000; González-Martínez et al. 2001; Gómez et al. 2001). Gene variation in Aleppo pine is clinally distributed from north to south, showing maximum levels of diversity in the extreme edges of its distribution (Agúndez et al. 1999). For maritime pine, chloroplast variation studies, including populations from most of the native range of the species showed three main refugial areas: the Atlantic coast of Portugal, southwestern Iberia, and Pantellaria and Sardinia in Italy (Ribeiro 2001 and references within; G.G. Vendramin pers. comm.). Burban & Petit (2003) found three maternal lineages in maritime pine using mtDNA-RFLP analysis (named western, eastern and Moroccan). The western mitotype might be associated with the Iberian glacial refugia and the eastern with those described for the Italian islands of Sardinia and Pantellaria.

Maritime pine exhibits high genetic diversity in Spain, while Vendramin et al. (1998) found 34 different haplotypes in ten populations from Portugal, France, Italy and northern Africa, up to 69 haplotypes were recently found in
seven Spanish populations using the same cpSSR markers (our unpublished results). Moreover, an allozyme study of 32 Iberian populations showed that populations from southern Spain displayed the highest allelic richness in the Iberian Peninsula, including 82% of the total number of alleles (González-Martínez 2001). Population differentiation is relatively high in maritime pine from the Iberian Peninsula ($G_{ST} = 0.077$; Salvador et al. 2000) and weak, yet significant, fine-scale structure due to restricted gene dispersal has been found in a classical locality from central Spain (González-Martínez et al. 2002).

The primary aim of this work is to analyze the spatial genetic structure of maritime pine within one putative refugial area of southeastern Spain. The region under study is a physiographically complex, mountainous territory, so it represents an excellent model system to study the effect of mountain ranges as barriers to interpopulational gene flow. Finally, spatial analysis of gene diversity in maritime pine provides relevant information for genetic conservation of forest resources.

Materials and methods

Plant Material

Seeds were collected from 11 populations covering the native range of $P. pinaster$ in southern Spain (Figure 1 and Table 1). Six provenance regions (i.e., native locations of the species used as breeding units) have been delimited in southern Spain based on ecological and historical data (Alía et al. 1996). To test the provenance division as units for management and conservation practices, we sampled at least one population from each of them. One of the provenances, ‘Sierra de Segura-Alcaraz’, is one of the most important areas of maritime pine in southern Spain, covering more than 70,000 hectares. Seedlots from this provenance produce around 100-300 kg of seeds per year, and are used in plantations all around Spain. In this case, four populations were sampled to estimate within-population variation in ‘Sierra de Segura-Alcaraz’ breeding unit. Two other populations included in this study (‘Gaucín’ - MA2, and ‘Sierra de Oria’ - AL) are considered geographically marginal. In each stand, 2-3 cones
were collected from 80 trees, at least 50 m apart from each other. The material
analyzed was either 70-80 female gametophytes per population (four popula-
tions) or gametophytes and embryos of 35-40 seeds per population (seven pop-
ulations). The present study included four from the 12 populations analyzed by
Salvador et al. (2000) and five from those included in the study made by
González-Martínez et al. (2001).

**Table 1.** Geographical location and provenance of origin of the 11 studied Pinus pinaster pop-
ulations (The Arabidopsis Genome Initiative 2000 and other sources cited in the text).

<table>
<thead>
<tr>
<th>Population</th>
<th>Code</th>
<th>Provenance</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Altitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Riopar</td>
<td>AB</td>
<td>Sierra de Segura-Alcaraz</td>
<td>38° 28’ 05” N</td>
<td>2° 27’ 31” W</td>
<td>1200</td>
</tr>
<tr>
<td>Cazorla</td>
<td>J1</td>
<td>Sierra de Segura-Alcaraz</td>
<td>37° 55’ 05” N</td>
<td>2° 55’ 11” W</td>
<td>1100-1200</td>
</tr>
<tr>
<td>Siles</td>
<td>J2</td>
<td>Sierra de Segura-Alcaraz</td>
<td>38° 21’ 30” N</td>
<td>2° 34’ 30” W</td>
<td>800-1500</td>
</tr>
<tr>
<td>Segura</td>
<td>J3</td>
<td>Sierra de Segura-Alcaraz</td>
<td>38° 17’ 03” N</td>
<td>2° 26’ 12” W</td>
<td>800-1000</td>
</tr>
<tr>
<td>Estepona</td>
<td>MA1</td>
<td>Sierra Bermeja</td>
<td>36° 31’ 05” N</td>
<td>5° 07’ 11” W</td>
<td>500</td>
</tr>
<tr>
<td>Gaucín</td>
<td>MA2</td>
<td>Serranía de Ronda</td>
<td>36° 32’ 10” N</td>
<td>5° 17’ 56” W</td>
<td>500-700</td>
</tr>
<tr>
<td>Cómpeta</td>
<td>MA3</td>
<td>Sierra almijara-Nevada</td>
<td>36° 51’ 44” N</td>
<td>3° 53’ 33” W</td>
<td>1000-1500</td>
</tr>
<tr>
<td>Jubrique</td>
<td>MA4</td>
<td>Sierra Bermeja</td>
<td>36° 33’ 05” N</td>
<td>5° 10’ 41” W</td>
<td>1000</td>
</tr>
<tr>
<td>La Peza</td>
<td>GR</td>
<td>Sierra Almijara-Nevada</td>
<td>37° 16’ 26” N</td>
<td>3° 22’ 10” W</td>
<td>1400</td>
</tr>
<tr>
<td>Sierra de Oria</td>
<td>AL</td>
<td>Oria</td>
<td>37° 30’ 49” N</td>
<td>2° 20’ 11” W</td>
<td>1300</td>
</tr>
<tr>
<td>Moratalla</td>
<td>MU</td>
<td>Moratalla</td>
<td>38° 05’ 57” N</td>
<td>2° 11’ 37” W</td>
<td>1000-1200</td>
</tr>
</tbody>
</table>

**Figure 1.** Physiography of southeastern Iberia and location of P. pinaster populations.
Molecular Markers

Seeds were stored at 4°C in a dry environment until enzyme extraction. For enzyme extraction, seeds were soaked on filter paper at room temperature for 24 h and germinated in a Petri dish in an incubator. Enzymes were extracted from 3-4 mm long radicles, following Conkle et al. (1982). Different studies dealing with allozyme variation in maritime pine from the Iberian Peninsula have shown that only a limited number of markers are significantly polymorphic, even when more than 20 allozyme loci have been successfully analyzed in this species (Castro 1989; Salvador 1997). Thus, only loci known to be polymorphic in the Iberian Peninsula at the 95% level were used. Horizontal starch-gel electrophoresis of seven allozymes encoded by eight loci was performed. The enzyme systems and the scored loci were: isocitrate dehydrogenase (*Idh*; EC 1.1.1.42), malate dehydrogenase (*Mdh*-2 and *Mdh*-3; EC 1.1.1.37), phosphoglucose isomerase (*Pgi*-2; EC 5.3.1.9), acid phosphatase (*Acph*; EC 3.1.3.2), glutamate dehydrogenase (*Gdh*; EC 1.4.1.3), glutamate-oxalacetate transaminase (*Got*-2; EC 2.6.1.1), and leucine aminopeptidase (*Lap*; EC 3.4.11.1). Genetic interpretation of enzyme systems and staining methods can be found in Castro (1989) and Salvador (1997). All loci included in this work showed to be neutral under the conditions of the Ewens-Watterson neutrality test (see Hartl & Clark 1997, pp 298-300).

Data Analysis

Spatial genetic structure was analyzed using the isolation-by-distance approach of Rousset (1997). In two-dimensional habitats, this method involves the regression of $F_{st}/(1-F_{st})$ estimates for pairs of populations on the logarithm of distance. Rousset (1997) showed that under isolation-by-distance, the variables considered show a linear relationship, the slope of which is inversely correlated with the product of effective population density and the second moment of parental axial distance ($4D\pi\sigma^2$). The parameter $\sigma^2$ is a measure of the speed at which two lineages descending from a common ancestor depart from each other in space (Rousset 2001). The absence of a pattern of isolation-by-distance (null slope of the regression, $D\sigma^2$ infinite) was tested by an exact permutation procedure using Genepop vs. 3.3 (M. Raymond & F. Rousset, ISEM, Université
de Montpellier 2, France). We computed the correlation between pairwise $F_{st}/(1-F_{st})$ and the logarithm of distance in two cases: (i) within the ‘Sierra de Segura-Alcaraz’ provenance; and (ii) using all the populations. The first analysis was done in order to study the spatial genetic structure among populations of the same seed collection region. In the second case, we used two types of distances: straight geographic distances and the length of the shortest pathway below 800 meters above sea level (m.a.s.l.). Nowadays populations of maritime pine are usually distributed at medium altitudes. During the ice ages, displacements of this species across some of the high mountains that separate populations in this region may have been difficult. Our hypothesis is that any spatial structure would be more easily detected using distances along low altitude pathways.

Gene interchange between pairs of populations in a subset of six populations (one from each provenance region) was analyzed using migration matrix models and maximum likelihood methods. The estimation process used an expansion of the coalescent theory that included migration. Sampling of genealogies has been done using a Markov chain Monte Carlo approach and the Metropolis-Hastings algorithm (see description in Chib & Greenberg 1995). Two runs with 10 short chains (200 trees used out of 4,000 sampled) and 2 long chains (2,000 trees used out of 40,000 sampled) were done. As the correlation between the two runs was high (> 80%), we present only the average values. Migrate version 1.1. software (Beerli 1997-2001) has been used to perform this analysis (see Beerli & Felsenstein 2001 for a detailed description).

**Results**

Maritime pine showed a marginally significant spatial structure but only when the length of the shortest pathway below 800 meters above sea level was considered ($b = 0.00816; p = 0.09$). In this case, the slope of the correlation between $F_{st}/(1-F_{st})$ and the logarithm of the distance provided an indirect estimation of $4D\pi\sigma^2$ for dispersal in two dimensions ($4D\pi\sigma^2 = 122.55$). No genetic spatial structure was found within the ‘Sierra de Segura-Alcaraz’ provenance, nor
when straight geographic distances between all pairs of populations were used. The correlation graph between $F_{st}/(1 - F_{st})$ and the logarithm of the distance is shown in Figure 2. Most outliers in the graph included the ‘Sierra de Oria’ marginal population. The close genetic similarity between two groups of populations, ‘Sierra de Segura-Alcaraz’ and ‘Sierra Bermeja/Ronda’ is also remarkable. These populations, while lying several kilometers apart, are connected by the Guadalquivir valley.

Average pairwise gene interchange estimated using Migrate software was low ($N_m = 1.6029$). Some populations included in different provenance regions, but located close to each other (e.g., MA2 and MA4), were practically isolated ($N_m < 1$; Table 2). Moreover, altitudes as low as 1,000 m.a.s.l. (at present, maritime pine can grow at 1,800 m.a.s.l.) were enough to effectively isolate the populations AL (Sierra de Oria) and J1 (Cazorla), even when these two populations are separated by less than 50 km. The ‘Sierra de Oria’ population had a higher number of immigrants than emigrants, as shown by a likelihood ratio.
test performed with the three closest populations: MU, J1 and MA3 ($p < 0.000$; see also Figure 3). Asymmetric gene flow is expected at the edge of the distribution range of a species due to expansion/retreat processes.

**Table 2.** Pairwise gene flow (Nm) among six *P. pinaster* populations (one from each provenance region) in southeastern Spain

<table>
<thead>
<tr>
<th>POP (x)</th>
<th>Nm (x receiving population)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>J1,x MA3,x MA2,x MA4,x MU,x AL,x Average</td>
</tr>
<tr>
<td>J1</td>
<td>1.6482 0.8896 0.7716 0.8380 1.5401 1.1375</td>
</tr>
<tr>
<td>MA3</td>
<td>0.9082 1.1647 1.2346 0.5501 1.0674 0.9850</td>
</tr>
<tr>
<td>MA2</td>
<td>2.7209 3.7466 0.9551 0.2753 0.8005 1.6997</td>
</tr>
<tr>
<td>MA4</td>
<td>0.3352 0.2188 0.8776 0.6887 1.5348 0.7310</td>
</tr>
<tr>
<td>MU</td>
<td>1.9738 2.8169 3.8128 1.3741 2.1699 2.4295</td>
</tr>
<tr>
<td>AL</td>
<td>1.2184 2.8912 2.5738 1.6951 4.7951 2.6347</td>
</tr>
<tr>
<td>Average</td>
<td>1.4313 2.2643 1.8637 1.2061 1.4294 1.4225</td>
</tr>
</tbody>
</table>

**Figure 3.** Pairwise gene interchange (Nm) between AL (Sierra de Oria) marginal population and, (i) J1 and MA3, two close populations but separated from AL by different mountain ranges and (ii) MU, a population with a direct gene flow pathway.
Discussion

Temperate forest trees have contrasting spatial patterns of genetic structure at the natural range scale. Some show clinal variation of allele frequencies (e.g., *Quercus petraea*, Zanetto & Kremer 1995; *Pinus halepensis*, Agúndez *et al*. 1999; *Pinus sylvestris*, Prus-Glowacki & Stephan 1994; *Picea abies*, Bucci & Vendramin 2000), but no gene diversity structure has been found in others (e.g., *Castanea sativa*, Fineschi *et al*. 2000; *Prunus spinosa*, Mohanty *et al*. 2000; *Sorbus torminalis*, Demesure *et al*. 2000). Within a given species, it is also common to find genetic structure differences at a regional scale. Maritime pine, in particular, presented a geographical pattern within Mediterranean populations (Salvador *et al*. 2000), but no structure in the Atlantic populations (Portugal, Ribeiro *et al*. 2001; south-western France, Mariette *et al*. 2001). The transfer of seeds and the high gene flow among regions have probably erased the putative original spatial structure in Portugal (Ribeiro 2001).

The high mountain ranges, which separate populations of maritime pine in southern Spain, have probably played an important role in shaping the present distribution of gene variation. Southern maritime pine populations showed a significant spatial structure, but only when the length of the shortest pathway at low altitude (< 800 m.a.s.l.) was considered (this study). Fine estimations of $N_m$ using matrix migration models for pairs of populations showed that pairwise gene interchange was very low, even at short distances, when mountains higher than 1,000 m.a.s.l. separated populations (e.g., AL and J1). The ability of mountains to serve as effective historical barriers to gene flow is well documented. Hewitt (2001) pointed out the importance of barriers such as the Pyrenees or the Alps in shaping the actual pattern of genetic diversity in several European animal and plant species. At a regional scale, mountains (i) promote more stable population dynamics due to displacements in altitude in response to climatic changes (Comes & Kadereit 1998), and (ii) increase the differentiation between populations as a consequence of the isolation of populations, even during the mildest periods of the glaciations. In maritime pine, a global $F_{st}$ value of 0.085 was found in the Iberian Peninsula. This value is 0.042 when the populations from southern Spain were removed from the analysis (data recalculated from González-Martínez *et al*. 2001).
The oscillating climates of the last glacial stage must have profoundly influenced the altitudinal location of plant species in southern Spain. Horizontal transfer of chloroplast DNA types among species of the *Armeria* complex (Plumbaginaceae) in Sierra Nevada showed a scenario that must have involved populations ascending or descending mountains (Gutiérrez-Larena *et al.* 2002). Our results indicate that the expansion during interglacials in maritime pine must have taken place by low altitude pathways, through territories presently covered by meso- and thermo-Mediterranean woods such as *Quercus faginea*, *Q. ilex rotundifolia*, *Q. suber*, *Q. coccifera*, *Pistacia lentiscus*, *Juniperus phoenicea*, *J. oxycedrus* and *Olea europaea*. An interesting case study is provided by the recolonization of oaks (*Quercus robur*, *Q. petraea* and *Q. pubescens*) across the Swiss Alps. Mátyás & Sperisen (2001) suggested that oak species coming from a glacial refugium in Italy were able to cross the Alps in the area of the Brenner pass (1,371 m.a.s.l.). For maritime pine, populations along the Guadalquivir Valley show high genetic similarity (‘Sierra de Segura-Alcaraz’ and ‘Sierra Bermeja/Ronda’ provenances). This valley may have acted as a corridor for different gene pools of the species during the mildest periods of the most recent glaciation.

Overall, this picture agrees with the palaeoecological information from on-going studies. The pollen records of Cañada de la Cruz, Siles, and Villaverde, situated across elevational and latitudinal gradients have been correlated to produce a picture of Upper Pleistocene and Holocene environmental history in the Segura region (Carrión 2002). In particular, the Siles lake palaeoecological record (2º 30′, 38º 24′ N, 1320 m.a.s.l.) shows that *P. pinaster*, together with a number of temperate and Mediterranean, mesothermophilous trees and shrubs, persisted in these mountains during the last glacial times.

Apart from the relevance of this persistence, which had also been shown in the southern Iberian System (Carrión & van Geel 1999) although not as prominently, it is worth stressing that maritime pine has been extremely sensitive to climatic changes, with rapid (century-scale) altitudinal patterns of displacement, not only during the arrival of late-glacial amelioration, but even throughout the Holocene (Carrión *et al.* 2001b). Furthermore, the Siles record provides support to the view that *Pinus pinaster* could survive in southern European mountains at relatively elevated locations during the last glacial stage. This
hypothesis was put forward by Bennett et al. (1991), who contended that tree survival would have been especially important in those mountain ranges that, like the Balkans, allowed rapid altitudinal displacements of tree populations in response to climatic pulses. As in the Balkans, the Segura mountains probably permitted latitudinal movements of tree populations owing to their more or less north-south orientation. Interestingly, *P. pinaster* is absent from the Holocene Villaverde pollen record (Carrión et al. 2001a), which suggests a more recent distribution in the northern platforms of Sierra de Alcaraz, in contrast with the Segura range, resulting from a recent expansion and/or introduction by humans.

Some practical implications can be drawn from the spatial structure analysis in southern Spain. First, the lack of genetic structure within ‘Sierra de Segura-Alcaraz’ provenance makes seed collection from multiple stands unnecessary in this region, and thus seed collection concentrated in few stands is recommended. Second, plantations with plant material from ‘Sierra de Segura-Alcaraz’ are recommended only locally and in populations connected by the Guadalquivir valley but not in the south-easternmost range of the species, where historical isolation has produced highly differentiated populations. Third, southern Spain is a focal area for *in situ* conservation of maritime pine genetic resources. Stands selected for conservation purposes should cover a wide range of locations, irrespectively of the geographical distance between them. The location of genetic reserves has to be considered carefully and genetic, ecological and demographic information should be combined to define conservation priorities.

**Acknowledgements**

The study was funded by the Cooperation project DGCN–INIA CC00-0035 and the INIA project SC97-118. Thanks to Diana Barba for her technical assistance in the lab and to Fernando del Caño and Alfonso Piñero for the field collections. Laura Salvador from the University of Huelva gently provided data from AB, MA1, GR and AL populations. Thanks are extended to Maria M. Ribeiro who discussed a draft of the manuscript and to Patricia C. Grant who revised the English.
References


