

Spatial genetic structure in continuous and fragmented populations of *Pinus pinaster* Aiton

A. I. DE-LUCAS,* S. C. GONZÁLEZ-MARTÍNEZ,† G. G. VENDRAMIN,‡ E. HIDALGO* and M. HEUERTZ§

*Departamento de Producción Vegetal y Recursos Forestales, E.T.S. Ingenierías Agrarias, Universidad de Valladolid, Avenida de Madrid 44, E-34004 Palencia, Spain, †Departamento de Sistemas y Recursos Forestales, CIFOR-INIA, Carretera de La Coruña km 7.5, E-28040 Madrid, Spain, ‡CNR Istituto di Genetica Vegetale, Sesto Fiorentino, Firenze, Italy, §Behavioural and Evolutionary Ecology, CP 160/12, Faculty of Science, Université Libre de Bruxelles, 50 Av. F.D. Roosevelt, B-1050 Bruxelles, Belgium

Abstract

Habitat fragmentation, i.e., the reduction of populations into small isolated remnants, is expected to increase spatial genetic structure (SGS) in plant populations through nonrandom mating, lower population densities and potential aggregation of reproductive individuals. We investigated the effects of population size reduction and genetic isolation on SGS in maritime pine (*Pinus pinaster* Aiton) using a combined experimental and simulation approach. Maritime pine is a wind-pollinated conifer which has a scattered distribution in the Iberian Peninsula as a result of forest fires and habitat fragmentation. Five highly polymorphic nuclear microsatellites were genotyped in a total of 394 individuals from two population pairs from the Iberian Peninsula, formed by one continuous and one fragmented population each. In agreement with predictions, SGS was significant and stronger in fragments ($Sp = 0.020$ and $Sp = 0.026$) than in continuous populations, where significant SGS was detected for one population only ($Sp = 0.010$). Simulations suggested that under fat-tailed dispersal, small population size is a stronger determinant of SGS than genetic isolation, while under normal dispersal, genetic isolation has a stronger effect. SGS was always stronger in real populations than in simulations, except if unrealistically narrow dispersal and/or high variance of reproductive success were modelled (even when accounting for potential overestimation of SGS in real populations as a result of short-distance sampling). This suggests that factors such as nonrandom mating or selection not considered in the simulations were additionally operating on SGS in Iberian maritime pine populations.

Keywords: forest fragmentation, genetic diversity, Iberian Peninsula, *Pinus*, spatial genetic structure

Received 10 December 2008; revision received 19 August 2009; accepted 25 August 2009

Introduction

Fine-scale spatial patterns of genetic variation in plant populations result from the complex interplay of local dispersal with genetic drift and other evolutionary processes such as natural selection (Slatkin 1973; Epperson 1990; Rousset 2004). The development of the Sp statistic for quantification of spatial genetic structure (SGS)

based on the decay of inter-individual kinship with spatial distance (Vekemans & Hardy 2004) has eased the research into the factors governing SGS. Sp , which allows to compare SGS between populations or species and between different marker types (e.g., Hardy *et al.* 2006), revealed that life history traits influencing pollen and seed dispersal distances are the strongest determinants of SGS (reviewed by Vekemans & Hardy 2004). SGS is on average stronger in herbs than in trees and stronger in selfing than in mixed-mating or outcrossing species (Vekemans & Hardy 2004). SGS is also stronger

Correspondence: Ana I. de Lucas, Fax: +34979108302; E-mail: ailucas@pvs.uva.es

in animal- than in wind-pollinated temperate trees, and stronger in tropical trees with narrow-range seed dispersal mediated by gravity, wind and/or scatter-hoarding animals than in those with far-ranging seed dispersal by monkeys, birds or bats (Dick *et al.* 2008).

Habitat fragmentation, i.e., the reduction of populations into small, isolated patches, poses genetic and ecological threats to plant populations (Young *et al.* 1996; Lowe *et al.* 2005; Aguilar *et al.* 2008; Kramer *et al.* 2008). Isolated remnant populations are predicted to suffer genetic erosion resulting from founder events caused by fragmentation, increased genetic drift because of small population size and inbreeding (Young *et al.* 1996; Aguilar *et al.* 2008). Mating system and rarity status influence this genetic erosion, with outcrossing species and recently rare species being most affected (Aguilar *et al.* 2008). Within species, the negative effects are more pronounced in smaller (Dunphy & Hamrick 2007) and in older fragments (for instance, in a metapopulation context, Young *et al.* 1996). The degree of fragment isolation also plays a decisive role, as fragments may remain connected by long-distance gene flow (Bittencourt & Sebbenn 2007; Kramer *et al.* 2008). Indeed, multiple cases of increased gene flow distances by pollen and/or seed have been observed in fragmented landscapes compared with intact populations, counteracting the adverse effects of fragmentation (El-Kassaby & Jaquish 1996; White *et al.* 2002; Dick *et al.* 2003; Bacles *et al.* 2006; but see Jump & Peñuelas 2006). Heterozygosity losses generally respond slowly to fragmentation, while higher progeny inbreeding and ecological consequences, such as lower reproductive output and lower fitness, are often observed within few generations (Lowe *et al.* 2005; Aguilar *et al.* 2008). The latter are the consequences of quantitative pollen limitation, increased selfing and mating between relatives (i.e., biparental inbreeding) and/or lower effective number of pollen donors in small, isolated populations (Robledo-Arnuncio *et al.* 2004; O'Connell *et al.* 2006; Mimura & Aitken 2007). While such mating system changes similarly affect anthropically fragmented (e.g., O'Connell *et al.* 2006) and naturally small or isolated populations (for instance at the range periphery, Mimura & Aitken 2007), both types of populations may evolve differently as a result of different effects on fitness of each process and the typically much shorter time frame of human-mediated fragmentation.

Habitat fragmentation is expected to increase SGS in plant populations because the interaction of local dispersal with increased mating between relatives, correlated mating and/or selfing will lead to higher levels of kinship at short distance (see also Wells & Young 2002). Moreover, the generally lower population density in fragments (Young & Merriam 1994; Nason & Hamrick

1997; Jump & Peñuelas 2006) leads to a reduced overlap of seed shadows, which increases the probability that nearby individuals are sibs and hence contributes to stronger SGS (Young & Merriam 1994; Hardy *et al.* 2006; reviewed by Vekemans & Hardy 2004). The potentially resulting spatial aggregation of reproductive individuals is further expected to increase SGS (Doligez *et al.* 1998). Increased SGS in fragments has indeed been observed in some species (Young & Merriam 1994; Van Rossum & Triest 2007; Yamagishi *et al.* 2007). However, in the tropical tree *Aucoumea klaineana*, weaker SGS has been reported in a fragmented population, reflecting the previously mentioned compensation effect of higher gene flow under fragmentation (Born *et al.* 2008). In slash pine (*Pinus elliottii* var. *densa*), weaker SGS was also observed in hurricane-disturbed vs. undisturbed fragments (Williams *et al.* 2007). It seems hence that the degree of fragment isolation and disturbance are fundamental factors controlling the extent of SGS through interaction with fragment population size and local population density.

Mediterranean maritime pine (*Pinus pinaster* Aiton) is a long-lived wind-pollinated conifer with a scattered distribution in the Iberian Peninsula, principally as a result of ecological disturbances (e.g. forest fires) and human-induced habitat fragmentation. Maritime pine is shade intolerant and often forms monospecific even-aged stands. Most Mediterranean populations of maritime pine are well-adapted to recurrent forest fires and trees from these populations have thick bark and a high percentage of serotinous cones, sometimes, over 20 years old. The aerial seed bank allows maritime pine (an obligatory seeder) to recolonize stands after forest fires, often competing with advantage against other forest trees. The putative glacial refugial areas and the different colonization pathways of this species in the Iberian Peninsula have been the object of numerous palynological, palaeo-climatological and genetic studies (Salvador *et al.* 2000; González-Martínez *et al.* 2001; Ribeiro *et al.* 2001; Derory *et al.* 2002; Bucci *et al.* 2007; de-Lucas *et al.* 2009). Fossil charcoal analyses have shown the presence of *Pinus pinaster* in western central Portugal since 33 000 BP (Figueiral 1995) and palynological data have demonstrated that the species survived through the last glacial maximum in eastern Iberia (31 000 BP) (Carrión *et al.* 2000). Pollen grains of species from the subgenus *Pinus* are very similar; however, a 40× microscopic magnification generally allowed distinction of *P. pinaster* pollen from that of other native pines of the Iberian Peninsula (*P. halepensis*, *P. nigra*, *P. pinea* and *P. sylvestris*, Carrión *et al.* 2000). Wood anatomical characters are similar in *P. pinaster* and in *P. pinea*, but sufficiently large and well-preserved samples allowed distinction of the species (Figueiral 1995).

Refuge locations in eastern and southern Iberia and north and northwest-ward colonization pathways were confirmed with molecular markers (Salvador *et al.* 2000; Bucci *et al.* 2007). The previous studies on mating system and SGS in maritime pine have shown extensive pollen and seed dispersal (González-Martínez *et al.* 2006; de-Lucas *et al.* 2008) and weak fine-scale genetic structure, slightly stronger in natural regeneration than in mature trees (González-Martínez *et al.* 2002).

In this work, we used nuclear microsatellite markers to characterize SGS in two population pairs of *P. pinaster* containing each one continuous and one fragmented stand. To check whether the detected SGS patterns may generally apply, population pairs were located in very different areas, corresponding to a glacial refuge and to a presumably recently recolonized region. As population size and degree of isolation are fundamental factors to explain SGS under fragmentation (assuming constant local density, see above), we investigated their effect on SGS formation under different pollen and seed dispersal kernels in simulated populations mirroring the studied *Pinus pinaster* stands.

Materials and methods

Study area and sampling

Two pairs of native *Pinus pinaster* populations were sampled in the Iberian Peninsula, containing one continuous and one fragmented population each (Fig. 1). Two plots, Olba and Quatretonda, were installed in eastern Spain (Mediterranean coast), near the hypothesized nuclei of post-glacial expansion of the species (Carrión *et al.* 2000; Salvador *et al.* 2000). Olba is located within a relatively large (~18 500 ha) continuous stand, which, to our knowledge, has not suffered any disturbances in historical time. Quatretonda is a small population (~50 ha) near the Mediterranean Sea which has been affected by recurrent forest fires throughout centuries and is threatened by postfire interspecific competition with *Pinus halepensis*, a sympatric Mediterranean pine. The other two plots, Cuéllar and Fuentelapeña, were installed in the Castilian Plateau (central inland Spain), a region that is thought to have been colonized by *Pinus pinaster* well after the last

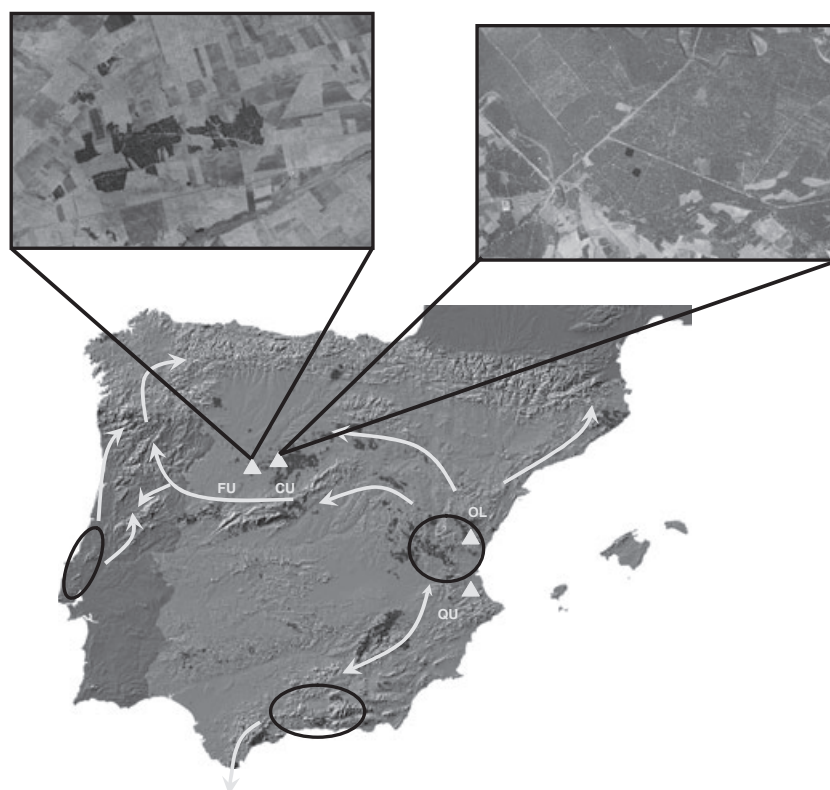


Fig. 1 Study plot location map, showing also putative glacial refugia (encircled), postglacial colonization routes and distribution of present-day native maritime pine populations (shaded) in the Iberian Peninsula (after Salvador *et al.* 2000; Ribeiro *et al.* 2001; Burbán & Petit 2003; Bucci *et al.* 2007). Sampling populations: CU: Cuéllar (continuous); FU: Fuentelapeña (fragmented); OL: Olba (continuous) and QU: Quatretonda (fragmented). Insets are aerial photographs (provided by the Spanish National Geographical Institute) of the fragmented population of Fuentelapeña (left) and the continuous population of Cuéllar (right).

glacial maximum (Salvador *et al.* 2000; Bucci *et al.* 2007; but see de-Lucas *et al.* 2009). Cuéllar is located within a large (~120 000 ha) maritime pine forest that is often mixed with stone pines (*Pinus pinea*). Fuentelapeña (~100 ha) is located at the westernmost edge of the Castilian Plateau, under similar ecological conditions as Cuéllar's. However, this part of the species' range is characterized by fragmentation (see Fig. 1) and a marked decline of maritime pine stands caused by a lack of natural regeneration, competition by an abundant regeneration of *Pinus pinea* and cultivation of former forest land (see photo report in Supplemental Material).

In each stand, 67–85 mature trees were sampled within a 70 × 70 m plot placed in areas with similar mature-tree (80–90 years-old) density (173 trees/hectare; see plot characteristics and photographic material in Supplemental Material). Mixed areas with *Pinus pinea* or *Pinus halepensis* were avoided. To check for within-population variability in SGS, two 70 × 70 m plots were sampled in Cuéllar, but as no significant differences in SGS were found (S_p of 0.010 and 0.005 respectively) only pooled results are presented. Mature trees from each plot were mapped using a total station (Topcon GTS-229). Needles were collected, lyophilized and stored at –80 °C.

DNA isolation and molecular markers

Total genomic DNA was extracted from lyophilized needles following Doyle & Doyle's (1990) protocol modified for small volumes. Six polymorphic nuclear microsatellites markers (number of alleles: 4–20) were analysed. Amplification conditions are given elsewhere (*frpp94*, Mariette *et al.* 2001; *itph4516* and *frpp91*, González-Martínez *et al.* 2002; *ssrPt_ctg4363*, *RPtest11* and *ssrPt_ctg275*, Chagné *et al.* 2004). Microsatellite fragments were separated in an ABI-PRISM 310 genetic analyzer and genotype scores were obtained by comparison with an internal size standard (GeneScan ROX-500) using GeneScan version 2.0.1 software (Applied Biosystems, Foster City, CA, USA). To detect scoring errors resulting from the presence of null alleles (Brookfield 1996), stuttering or large allele dropout, microsatellite data were tested using Micro-Checker software (Van Oosterhout *et al.* 2004). One locus (*frpp91*) consistently showed a high null allele frequency and was removed from further analyses.

Data analyses

Genetic diversity and differentiation. For each population, allelic richness and number of private alleles were computed following a rarefaction method that compensates

uneven population sample sizes, as implemented in the HP-Rare software (Kalinowski 2005). Expected heterozygosity (H_e) was calculated following Nei (1978) and genetic differentiation (F_{ST}) following Weir & Cockerham (1984). Inbreeding coefficients (F) were computed as kinship coefficients between genes within individuals and tested against the null hypothesis of Hardy-Weinberg equilibrium by permutations using SPAGeDi version 1.2. (Hardy & Vekemans 2002). The population-scaled mutation parameter, $\theta = 4N_e\mu$ (where N_e is the effective diploid population size and μ is the mutation rate) was computed following a moment estimator based on summary statistics, $\hat{\theta}_H$ (Kimmel *et al.* 1998; see also RoyChoudhury & Stephens 2007). To reduce the effect of narrow-range sampling, diversity and F estimates were based on data sets where one individual was removed from all pairs with higher than average kinship located in the distance classes with significant average kinship coefficients, $\hat{F}_k < 10$ m in Olba, < 20 m in Fuentelapeña and Quatretonda, see below).

To check whether significant heterozygosity deficits (see Results) were caused by Wahlund effects, we investigated population substructure using a Bayesian method with MCMC estimation implemented in STRUCTURE version 2.2 (Pritchard *et al.* 2000). In each population, an admixture model with correlated allele frequencies between clusters was run ten times for a given number of inferred clusters, K , from $K = 1$ to 8 using burn-in lengths of 5×10^4 and run lengths of 5×10^5 Markov Chain Monte Carlo steps. The optimum number of clusters, K , was determined following guidelines from the authors (Pritchard & Wen 2004).

Fine-scale spatial genetic structure and dispersal estimates. Fine-scale spatial genetic structure (SGS) was analysed by linear regression of pairwise kinship coefficients (or coancestry estimators, \hat{F}_{ij}) on the logarithm of inter-tree distance using SPAGeDi version 1.2 software (Hardy & Vekemans 2002). Kinship coefficients \hat{F}_{ij} were computed between individuals i and j as a correlation between allelic states, as proposed by J. Nason in Loiselle *et al.* (1995). Significance of the SGS pattern was tested in each population by comparing the observed regression slope \hat{b}_F with its expected distribution obtained from 10 000 permutations of rows and columns of the inter-tree distance matrix. The SGS pattern was visualized by plotting average kinship coefficients \hat{F}_k against distance for $k = 7$ distance classes (0–10 m, 10–20 m, 20–30 m, 30–40 m, 40–50 m, 50–60 m, >60 m). Approximate standard errors for \hat{F}_k were estimated by jackknifing (i.e. deleting information from one locus at a time). SGS was quantified using the S_p statistic, defined by the ratio $-\hat{b}_F / (1 - \hat{F}_1)$, where \hat{F}_1 is the mean \hat{F}_{ij} in the first distance class (0–10 m) that includes all pairs of

neighbours (Vekemans & Hardy 2004). The Sp statistic was also used to estimate overall historical dispersal distances as Wright's neighbourhood size Nb ($Nb = 1/Sp$). This inference is valid if SGS is caused solely by isolation by distance, i.e. the interaction of restricted dispersal and genetic drift, and has reached dispersal-drift equilibrium (Vekemans & Hardy 2004). Nb is defined as $Nb = 4\pi D_e \sigma_g^2$, where D_e is the effective population density and σ_g^2 the axial variance of gene dispersal, equal to half the mean squared axial parent-offspring distance. Estimates of σ_g were also obtained, assuming ratios of effective to census density, D_e/D , of 0.1 and 0.5 (Vekemans & Hardy 2004). While the slope of kinship-distance plots (i.e., Sp) is informative on total gene flow, its shape, and notably its curvature at short distance can provide insights into the relative contributions of seed and pollen dispersal to total gene flow (Heuertz *et al.* 2003). Shapes of dispersal curves were estimated by fitting polynomial functions of the third degree (i.e., $y = a + bx + cx^2 + dx^3$) of the logarithm of distance to the standardized residuals ($\hat{F}_{ij} - F_{ij(\text{exp})}$) plots.

Simulated SGS scenarios. To determine the conditions necessary for reaching observed levels of SGS in maritime pine populations with different levels of fragmentation, multiple SGS scenarios were simulated using C++ code kindly provided by Dr. Olivier Hardy (Université Libre de Bruxelles, Belgium) and used in Heuertz *et al.* (2003). Expected Sp values from simulations were then compared with those found in real populations. Also relevant for our case study was the number of generations it took to reach SGS equilibrium (immigration – drift equilibrium) in simulations. Simulated populations maintained a constant density of 173 trees/ha (equal to maximal density of the real populations). Trees were placed on a square grid with an inter-tree distance of 7.6 m and initial genotypes of each individual (20 loci and 20 alleles per locus) were randomly assigned considering even allele frequencies within loci. Then, populations were left to mate according to predefined seed and pollen dispersal functions, giving each individual a 50% chance of surviving and a 50% chance of being replaced by an offspring per generation (paralleling even-aged mature stands with a juvenile cohort, a realistic scenario for the species).

We considered a combination of (i) population sizes: 225 (15×15), 900 (30×30), 4900 (70×70), 22 500 (150×150) and 90 000 (300×300) individuals and (ii) immigration rates: 0% (total isolation), 5% seed and 20% pollen (low immigration), 20% seed and 40% pollen (moderate immigration) and 40% seed and 85% pollen (high immigration, approximately corresponding to observations in large, continuous populations; González-

Martínez *et al.* 2006). The 20 demographic scenarios were simulated using pollen and seed dispersal kernels that represented direct estimates in native maritime pine populations (González-Martínez *et al.* 2006; de-Lucas *et al.* 2008) or indirect (historical) dispersal estimates from the current study. The modelled kernels corresponding to direct estimates were power exponential functions (dispersal scenario DE) with scale parameter $\alpha = 6.09$ and shape parameter $\beta = 1$ for seed (González-Martínez *et al.* 2006), and $\alpha = 0.0028$ and $\beta = 0.2229$ for pollen (de-Lucas *et al.* 2008), corresponding to average dispersal distances, δ , of 12.19 m for seed and 111.9 m for pollen, and to an axial standard deviation of gene dispersal $\sigma_g = 125.67$ m. A shape parameter $\beta < 1$ (used here for pollen) indicates that the dispersal kernel is fat-tailed, i.e. there are more long-range dispersal events than for the exponential distribution ($\beta = 1$) and much more than for the normal distribution (Austerlitz *et al.* 2004). The kernels corresponding to indirect dispersal estimates (three scenarios) were normal (IN1 and IN2) or power exponential functions (IE), defined from the axial standard deviation, σ_g , of average gene dispersal of the three plots with significant SGS (Fuentelapeña, Olba and Quatretonda, see Results). For computation of σ_g , we assumed an effective density of $D_e = D$ (173 trees/ha) yielding very narrow-ranged dispersal, $\sigma_g = 16.5$ m (IN1), or an effective density of $D_e = 0.1D$ yielding $\sigma_g = 52.23$ m (IN2 and IE). An effective density equalling census density (IN1) can be considered very high; realistic D_e/D ratios are comprised between 0.1 (IN2 and IE) and 0.5 (Vekemans & Hardy 2004). The seed and pollen components in these three scenarios were derived from total gene dispersal σ_g following Crawford (1984), $\sigma_g^2 = \sigma_s^2 + \sigma_p^2/2$, assuming a nonsignificant shape component of SGS (using the ratio $\sigma_s^2/0.5\sigma_p^2 = 0.055$, determined from simulations in Heuertz *et al.* 2003). The shape parameters β of the seed and pollen dispersal kernels modelled for the IE scenario were the same as above. Formulas in Austerlitz *et al.* (2004) were used to relate parameters and statistical properties of dispersal kernels.

Finally, as simulated scenarios generally failed to explain observed levels of SGS unless extreme narrow-range dispersal (IN1) and almost complete isolation were assumed (see Results), we developed simulations considering a high range of reproductive success variance (from 5 to 50), as estimated by the variance of the total number of offspring per individual. These scenarios were only tested for the smallest plot (225 individuals) and no immigration case as a result of computer-time constraints and because of the fact that, even in this extreme case, the variance in reproductive success did not influence SGS to a large extent under most dispersal scenarios (see Results).

Each combination of parameters was simulated 100 times, letting SGS build up during 1000 generations, and the spatial genetic structure of the simulated populations was quantified using the *Sp* statistic, calculating the slope of the linear regression between kinship and distance as:

$$b_r = \frac{\sum_k n_k (x_k - \bar{x})(\bar{F}_k - \bar{F})}{\sum_k n_k (x_k - \bar{x})^2} \quad (\text{eqn 1})$$

where n_k is the number of pairs of individuals included in distance class k ; x_k is the average of \ln (distance between individuals) for distance class k ; \bar{x} is the average of x_k weighted by the number of pairs of individuals belonging to each class k ; \bar{F}_k is the average of Loiselle's kinship coefficient (F_{ij}) in class k (over 100 replicates for each generation) and \bar{F} is the average of \bar{F}_k , weighted by the number of pairs of individuals belonging to each class k . *Sp* statistics were computed for the entire pairwise distance range of the simulated scenarios using fourteen distance classes. They were further computed for distance classes between 0 and 100 m, to evaluate the effect of short-distance sampling, a common case in experimental SGS studies, on detectability of fine-scale spatial genetic structure patterns.

Results

Genetic diversity and differentiation

Genetic variation, as estimated by expected heterozygosity, was similar in all populations ($H_e \sim 0.75$) irrespectively of their location (putative glacial refugium or colonization edge areas) and status of fragmentation. However, there was a slight tendency of loss of alleles in the presumably recently colonized Castilian Plateau (see A , A^S , and A^P values of Table 1) compared with refugial populations. This pattern was also reflected in the population-scaled mutation parameter, $\hat{\theta}_H$ (average

$\hat{\theta}_H = 11.11$ in the Castilian Plateau and $\hat{\theta}_H = 14.45$ in eastern Spain), and in inbreeding coefficients. Inbreeding coefficients were moderate (0.028–0.117), except for the fragmented Fuentelapeña population whose F value (0.181) was higher than those of other populations at four out of five loci.

Genetic differentiation among populations was slightly higher between putative glacial refugium (pairwise $F_{ST} = 0.060$) than between colonization edge populations (pairwise $F_{ST} = 0.028$). STRUCTURE did not detect substructure in any population and the highest posterior probability of the data was observed for $K = 1$ cluster in each population, suggesting that each population harboured a single gene pool (see Supplemental Material) and the absence of Wahlund effects.

Fine-scale spatial genetic structure and dispersal estimates

Significant SGS, i.e., a negative regression slope of kinship coefficients on the logarithm of spatial distance, was detected in the two fragmented populations (Quatretonda and Fuentelapeña), and also, but weaker, in the continuous population located in the glacial refugium range (Olba) (Table 2, Fig. 2a). Kinship coefficients in the two fragmented populations were significant for the two first distance classes (20 m; $P < 0.05$), whereas in the Olba continuous population, inter-individual kinship was only significant at shorter distances (10 m; $P < 0.05$). Replicate plots installed within Cuéllar population did not show any difference in terms of SGS.

In the fragmented *P. pinaster* populations, *Sp* values ($Sp = 0.0196$ in Quatretonda and $Sp = 0.0264$ in Fuentelapeña) were higher than in the continuous populations ($Sp = 0.0104$ in Olba and $Sp = 0.0065$ in Cuéllar, the latter not significantly different from 0). They were also higher than *Sp* values generally observed in wind-dispersed temperate trees (mean $Sp = 0.012$, $n = 5$ species, Vekemans & Hardy 2004). Wright's neighbourhood

Table 1 Genetic diversity and differentiation in four maritime pine populations from Spain; N : sample size; A : number of alleles per locus; A^S and A^P : Allelic richness and number of private alleles per locus following rarefaction (Kalinowski 2005); H_e : Expected heterozygosity (Nei 1978); F : Inbreeding coefficient (Hardy & Vekemans 2002); $\hat{\theta}_H$: Population-scaled mutation parameter (Kimmel *et al.* 1998); F_{ST} : Genetic differentiation (Weir & Cockerham 1984)

Region	Pop	Type	N	A	A^S	A^P	H_e	$\hat{\theta}_H$	F	Pairwise F_{ST}		
										CU	FU	OL
Castilian Plateau	CU	C	151	10.2	8.9	0.6	0.758	12.00	0.117	—	—	—
	FU	F	59	8.0	7.9	0.4	0.734	10.22	0.181	0.028	—	—
Eastern Spain	OL	C	70	10.0	9.7	0.8	0.754	12.82	0.065	0.020	0.023	—
	QU	F	67	10.8	10.4	2.3	0.743	16.08	0.028	0.041	0.090	0.060
Average				9.8	9.2	1.0	0.747	12.78	0.098	0.044		

CU, Cuéllar; FU, Fuentelapeña; OL, Olba; QU, Quatretonda; C, Continuous; F, Fragmented.

Table 2 Fine-scale population genetic structure in four maritime pine populations from Spain; *N*: Sample size; *b*-log (\hat{b}_F): Slope of the regression of kinship with $\ln(\text{dist})$; \hat{F}_1 and \hat{F}_2 : mean \hat{F}_{ij} between individuals belonging to the first (0–10 m) and the second (10–20 m) distance classes respectively; *Sp*: statistic defined by the ratio $-\hat{b}_F/(1-\hat{F}_1)$, where \hat{b}_F is the regression slope of the autocorrelogram and \hat{F}_1 is the mean \hat{F}_{ij} between individuals belonging to the first distance class that includes all pairs of neighbours (Vekemans & Hardy 2004); *Nb*: Wright’s neighbourhood size; σ_g : estimate of gene flow from *Sp*, assuming different ratios of effective (D_e) to census density (*D*); *, average values over two 70 × 70 m plots

Region	Pop	Type	<i>N</i>	<i>b</i> -log (\hat{b}_F)	<i>P</i> -value (<i>b</i> -log)	\hat{F}_1	\hat{F}_2	<i>Sp</i>	<i>Nb</i>	σ_g (m) ($D_e/D = 0.5$)	σ_g (m) ($D_e/D = 0.1$)
Castilian Plateau	CU	C	76*	-0.0065*	ns	0.0115*	0.0040*	0.0065	—		
	FU	F	78	-0.0254	0.000	0.0370	0.0260	0.0264	37.86	18.7	41.7
Eastern Spain	OL	C	80	-0.0101	0.003	0.0233	0.0033	0.0104	96.54	29.7	66.5
	QU	F	85	-0.0190	0.000	0.0319	0.0165	0.0196	51.03	21.7	48.4

CU, Cuéllar; FU, Fuentelapeña; OL, Olba; QU, Quatretonda; C, Continuous; F, Fragmented.

*Average between the two replicates.

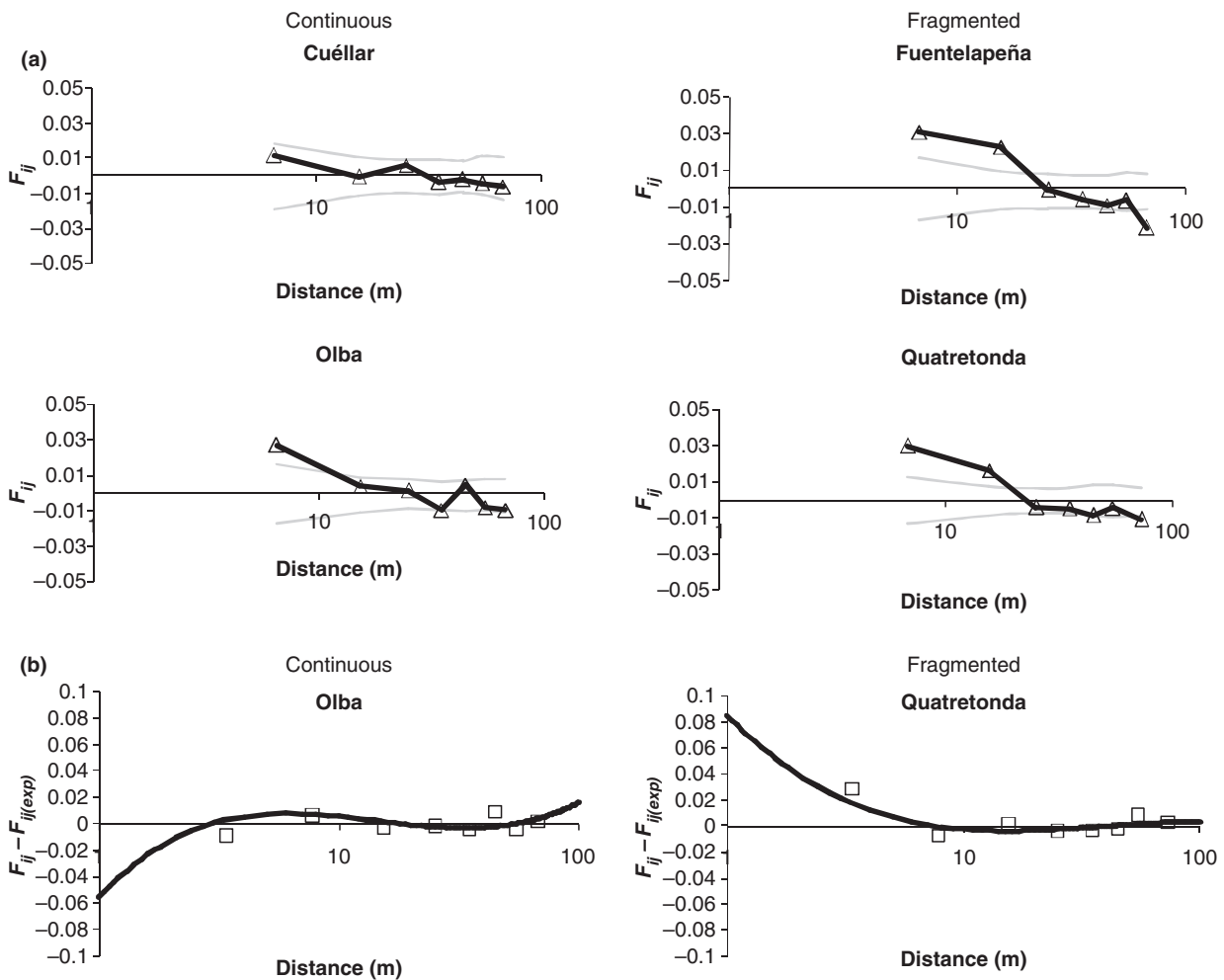


Fig. 2 Fine-scale spatial genetic structure (SGS) in continuous and fragmented maritime pine populations: (a) autocorrelograms and (b) polynomial regressions on residuals. SGS was stronger in fragmented than continuous populations (a). The polynomial regressions are represented only for the two populations in which one or more terms of the polynomial regression were statistically significant (b).

sizes inferred from SGS were 38 and 51 in fragmented populations and 97 in the Olba continuous population, and σ_g ranged from ca. 20 to 45 m in fragmented and from 30 to 67 m in continuous populations (Table 2).

The shape of kinship-distance plots informed little about the relative contribution of seed vs. pollen dispersal to global gene flow. The third degree polynomial functions fitted to residuals plots had marginally signifi-

cant terms ($P < 0.1$) only in Quatretonda and Olba (Fig. 2b). In the fragmented Quatretonda, a concave curve at short distance indicated a restricted short-scale component of gene flow, commonly attributed to restricted seed dispersal (pollen disperses at least about six times further than seeds, Barbará *et al.* 2008). In Olba, the polynomial curve was convex, indicating no particular restriction of seed dispersal.

Simulated SGS scenarios

Numerical simulations of SGS scenarios were characterized by a rapid build-up and stabilization of SGS after only a few generations (10–20) and high stochasticity. Sp values for the simulated scenarios are graphically represented in Fig. 3. For power exponential dispersal kernels (DE, IE), small population size had a much stronger effect than isolation on SGS at equilibrium. SGS was about 10 to 15 times stronger in the smallest than in the largest populations, regardless of immigration rate, while the isolated population scenarios had only about twice the SGS of the high immigration scenarios. Under normal dispersal kernels (IN1, IN2), isolation had a stronger impact on SGS and this impact was more pronounced in large populations (10 to 60-fold increase of Sp under isolation compared with high immigration) than in small ones (3 to 12-fold increase). For identical dispersal variances, SGS was weaker under power exponential (IE) than under normal dispersal kernels (IN2).

Restricting the distance range for regression analysis to short distances (0–100 m) instead of using the total available distance range resulted in a systematic bias towards overestimating SGS. For moderate-size populations (ca. 900 individuals), the overestimation of Sp was on the order of two-fold, whereas for large populations, it could reach one order of magnitude. Most simulated scenarios were characterized by very weak SGS, with Sp usually well under 0.005. Sp values as those observed in the study populations ($Sp > 0.01$) were only found with unrealistically narrow dispersal parameters, especially concerning pollen dispersal: normal dispersal kernels with $\sigma_s = 3.8$ m and $\sigma_p = 22.4$ m (scenario IN1) or $\sigma_s = 5.3$ m and $\sigma_p = 32.2$ m under complete isolation and small population sizes (results not shown). Even hypothesizing that we substantially overestimated Sp in the real populations (i.e., real populations were large), no simulated scenario could explain the observation. Modelling a high variance of reproductive success resulted in $Sp > 0.01$ for the normal dispersal scenarios, but power exponential scenarios (more realistic for pollen dispersal) still failed to explain the observed Sp values (Fig. 4).

On the basis of the comparison of simulated scenarios and fine-scale genetic structure in real maritime pine

populations, we can hence conclude that simulated scenarios with realistic dispersal functions and variance of reproductive success cannot explain the level of structure observed, suggesting that other factors not considered here can substantially affect SGS formation too.

Discussion

Genetic diversity and differentiation

Pinus pinaster populations from the Castilian Plateau, with a presumably more recent origin, were very similar to those in one putative glacial refugium in terms of expected heterozygosity and only marginally different for allelic richness, a parameter that is more sensitive to founder events (Nei *et al.* 1975). The only evidence of a more stable demography and older origin of eastern Iberian populations came from the distribution of private and rare alleles (see also Petit *et al.* 2003). Eastern Iberian populations contained a higher number of private alleles, especially Quatretonda ($A^P = 2.3$), in agreement with results from chloroplast microsatellite markers assayed in wide-range studies of the species (i.e., private haplotypes mostly distributed in the east of Spain, Bucci *et al.* 2007). This observation provides support to the view that rear edge populations could preserve high genetic distinctiveness and contain valuable ecotypes (Hampe & Petit 2005 and references therein).

In the fragmented Fuentelapeña population, high inbreeding ($F = 0.181$) was observed. Null alleles were ruled out as a possible cause as the locus with suspected null alleles had been removed. A Wahlund effect was also unlikely, as population substructure was not detected. Therefore, high F in Fuentelapeña is probably related to strong family structure as a result of recent fragmentation. Quantitative pollen limitation may occur in forest fragments (reviewed by Kramer *et al.* 2008) and in naturally small and isolated populations, increasing the frequency of selfing, mating between relatives and correlated mating (Robledo-Arnuncio *et al.* 2004; Mimura & Aitken 2007). In Fuentelapeña, genes from neighbouring trees ($\hat{F}_1 = 0.037$) were substantially less related than genes within individuals, suggesting either that inbreeding was largely due to selfing or to nonspatial biparental inbreeding (assortative mating between relatives).

Fine-scale spatial genetic structure

Large continuous populations of maritime pine had either weak (Olba) or no (Cuéllar) spatial genetic structure, irrespectively of their past demographic history. This is in agreement with results obtained in another large continuous maritime pine population, Coca, in central Spain (only ~50 km from Cuéllar), where,

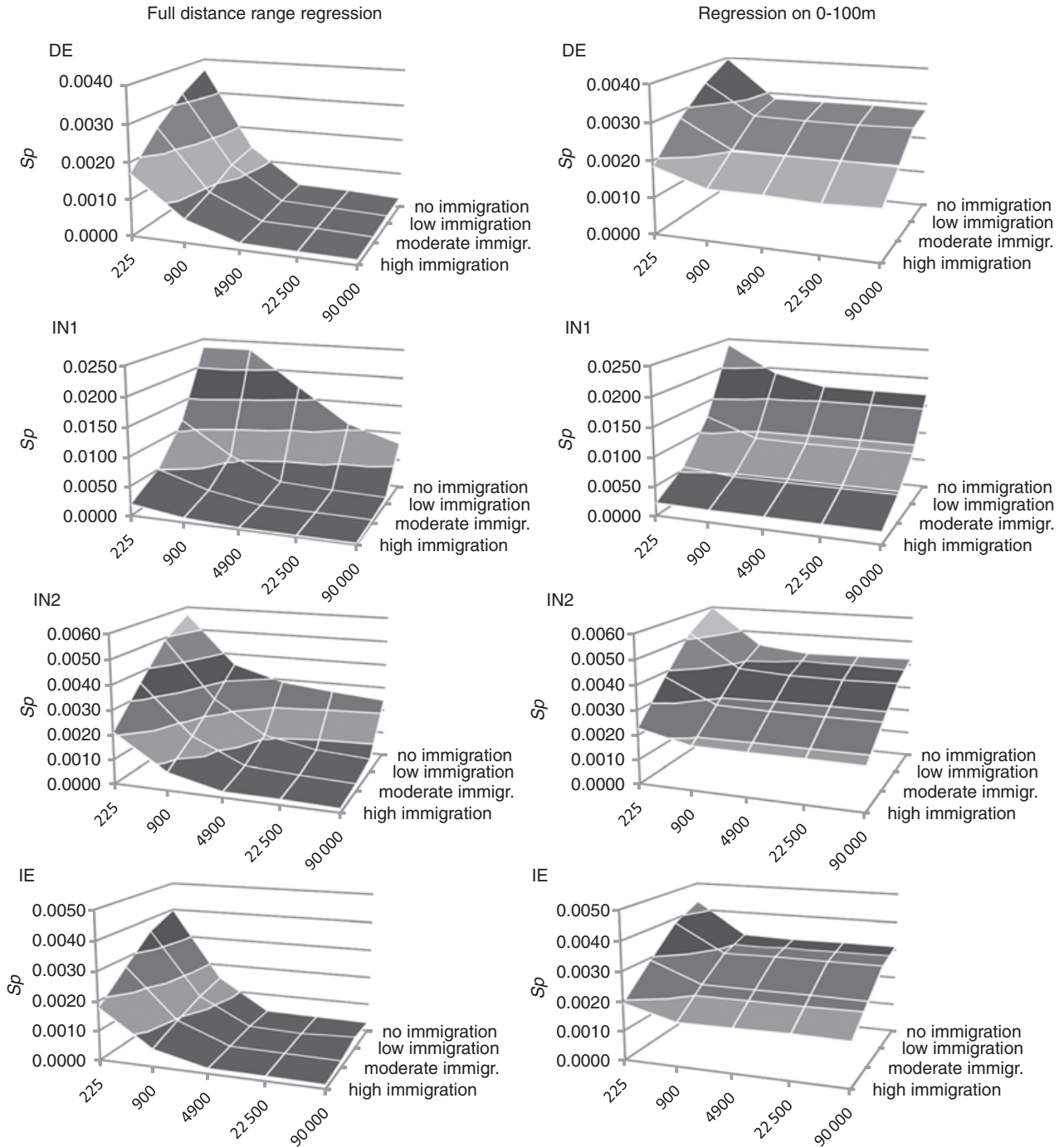


Fig. 3 S_p values obtained in four simulated dispersal scenarios as a function of population size (horizontal axis) and level of immigration (depth axis). DE, power exponential kernels inferred from direct estimates; IN1 and IN2, normal kernels inferred from indirect (historical) estimates; IE, power exponential kernels inferred from indirect estimates. Scenarios IN2 and IE have the same axial variance of dispersal distances. See Materials and methods for details.

despite a very skewed distribution of female effective reproductive success (González-Martínez *et al.* 2006), SGS was weak at the seedling stage and could not even be detected in mature trees (González-Martínez *et al.* 2002). By contrast, and in agreement with theoretical predictions for small, isolated populations, relatively

strong SGS was found in the two fragmented populations Quatretonda and Fuentelapeña. This stronger SGS may result from an impact of fragmentation (via lowering the effective density) on the mating system, i.e. a higher incidence of selfing, correlated mating and/or biparental inbreeding (Robledo-Arnuncio *et al.* 2004;

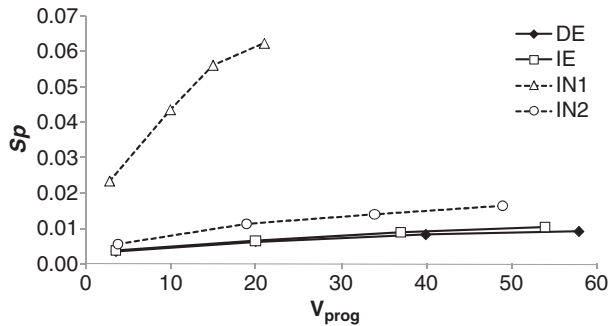


Fig. 4 Sp values as a function of the variance of reproductive success expressed as the variance of the lifetime total number of offspring per individual, V_{prog} , obtained in four simulated dispersal scenarios for a small, isolated population (15×15 individuals). Codes as in Fig. 3.

Mimura & Aitken 2007). In Fuentelapeña, our data suggested indeed substantial selfing (see above). Lower census density (Vekemans & Hardy 2004) or spatial aggregation of individuals (Doligez *et al.* 1998) were unlikely causes for higher SGS under fragmentation in our study as current densities were similar in continuous and fragmented populations, although we ignore details about historical demographic fluctuations. From its interpretation under drift-dispersal equilibrium, $Sp = 1/(4\pi D_e \sigma_g^2)$, it can be seen that restricted dispersal can also cause stronger SGS. In outcrossing and wind-pollinated species, pollen dispersal is generally extensive (Streiff *et al.* 1999; Bacles & Ennos 2008; de-Lucas *et al.* 2008), so that restricted seed dispersal becomes the main determinant of within-population SGS. In forest fragments, effective seed dispersal may be reduced compared with continuous populations, because seed dispersal beyond the fragment boundaries will usually not contribute to the next generation unless it reaches another fragment. Robledo-Arnuncio and Rousset (pers. comm.) illustrated this in simulations, finding a reduction of gene flow with increased spatio-temporal correlation of local tree density (e.g., in local forest fragments vs. nonforested landscape over time). In Quatretonda, the polynomial regression of residuals on the logarithm of spatial distance was concave at short distance, consistent with such restricted seed dispersal (about six times lower than pollen dispersal, Barbará *et al.* 2008). Overall, the stronger SGS observed in *P. pinaster* fragments suggested that compensation of the fragmentation impact on SGS through increased gene flow did not occur in our study populations, unlike in the tropical *Aucoumea klaineana* (Born *et al.* 2008). Furthermore, SGS patterns did not vary according to past demographic history in maritime pine, in contrast to Sitka spruce, where peripheral populations displayed stronger SGS than core populations (Gapare & Aitken 2005).

Dispersal estimates

When assuming an effective density of 0.1–0.5 times the census density, the axial standard deviation of gene flow σ_g estimated from SGS (20–66 m) was two to five-fold lower than direct estimates from continuous forest ($\sigma_g = 125$ m, González-Martínez *et al.* 2006; de-Lucas *et al.* 2008). Both estimates are not incompatible, however, as dispersal is a highly variable stochastic process (Nathan *et al.* 2000; Muller-Landau *et al.* 2008) and our direct estimates were based on only a single season each (see Smouse & Sork 2004). Our recent investigations in the populations of Cuéllar and Coca showed that seed dispersal occurred mostly during summer storms and that the shape of the best-fitting seed dispersal kernel and average seed dispersal distance strongly varied over years (L. Juez, S. C. González-Martínez *et al.*, unpublished). Moreover, because of the stochasticity of dispersal, fitting dispersal kernels can be mathematically challenging, especially for pollen flow (Robledo-Arnuncio *et al.* 2006). Our indirect estimates were performed over a too short distance range, as the correct range to reliably estimate σ_g with nuSSRs is ca. $\sigma_g-20\sigma_g$ (Rousset 2000; Vekemans & Hardy 2004). The comparison of our SGS patterns with simulations, however, suggested that our probable overestimation of Sp did not exceed a factor of two and that therefore, we underestimated historical gene dispersal σ_g by less than ca. 40%. It is also arguable that the fragmented study populations were not at drift-dispersal equilibrium, which would invalidate the dispersal estimates (Vekemans & Hardy 2004). However, in such case, high positive and negative kinship coefficients without significant SGS would be expected (Wells & Young 2002). The time of population fragmentation is not known for both fragmented populations, so that we can be confident that it is not a recent process (i.e., older than a couple of hundred years), adding support to the existence of drift-dispersal equilibrium.

Simulated SGS scenarios

Numerical simulations have shown that SGS builds up very fast, in just a few generations (Sokal & Wartenberg 1983). Small effective population size had a much stronger effect than isolation on SGS under power exponential dispersal kernels. This was expected, as pollen dispersal was modelled with a very fat-tailed kernel, i.e. a high proportion of long-distance dispersal events, which lead to a situation close to panmictic pollen dispersal in the model populations. Under such circumstances, immigration with similar allele frequencies as in the model population is not expected to produce a substantial decrease of SGS (a decrease of Sp by about

half was observed). By contrast, isolation had a much stronger effect than population size on SGS under normal dispersal kernels, reflecting the rarity of long-distance dispersal for this family of dispersal kernels.

SGS in simulations was generally low, suggesting that the fairly high S_p values observed in the maritime pine populations could not fully be accounted for by population size or degree of isolation, or even by the upward bias resulting from SGS estimation at relatively short distances (ca. 100 m). S_p around 0.02 as in fragmented populations was only obtained when modelling unrealistically narrow pollen and seed dispersal (IN1) in a small, completely isolated population (Fig. 3 and supplemental material). S_p around 0.01 as in the continuous Olba population could only be reached with realistic dispersal kernels if a small isolated population with a very high variance of reproductive success was modelled (Fig. 4). This suggests that in the real populations, SGS was increased by a combination of factors, probably ecological and/or demographical, that were not considered in our simulations. The study populations may in fact have a substantially smaller effective population size than simulated (observed $N_b \sim 38\text{--}96$ vs. $N_b \sim 270$ recorded in the small isolated DE scenario). A low number of reproducers may be the cause of low effective population size in fragmented populations. Indeed, in a strongly fragmented *Pinus sylvestris* population, Robledo-Arnuncio *et al.* (2004) showed that only 8% of possible mating pairs were observed, leading to an important decrease in effective population size. Assortative mating, for example, for tree size or phenology (e.g., Gérard *et al.* 2006; González-Martínez *et al.* 2006) may also lead to low numbers of reproducers. Finally, selection is a strong factor which contributes to shaping genetic patterns, especially in highly heterogeneous Mediterranean environments (Thompson 2005). Selection at the seedling stage occurring as a result of maternal effects (e.g. Bladé & Vallejo 2008) or environmental heterogeneity (e.g. Calvo *et al.* 2008) may substantially reduce the effective population size, contributing to relatively strong SGS in the studied *Pinus pinaster* populations.

Acknowledgements

The authors wish to thank O.J. Hardy for providing C++ code and valuable suggestions. The authors also extend their gratitude to C. Ordóñez and J.C. Santana, for field and laboratory assistance respectively and to P.C. Grant for the English grammar review. A.I. de Lucas was supported by a PhD scholarship from the 'Universidad de Valladolid' and S. C. González-Martínez by a 'Ramón y Cajal' fellowship RC02-2941. The work was funded by CICYT AGL2007-65795-C02-01, REPROFOR AGL2005-07440-C02-01/FOR and VaMPiro CGL2008-05289-C02-02/BOS projects. M. Heuertz was a postdoctoral researcher of the National Fund for Scientific Research of Belgium (FRS-FNRS)

and acknowledges an FNRS-funded scientific visit to CIFOR-INIA.

References

- Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Lobo J (2008) Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Molecular Ecology*, **17**, 5177–5188.
- Austerlitz F, Dick CW, Dutech C *et al.* (2004) Using genetic markers to estimate the pollen dispersal curve. *Molecular Ecology*, **13**, 937–954.
- Bacles CFE, Ennos RA (2008) Paternity analysis of pollen-mediated gene flow for *Fraxinus excelsior* L. in a chronically fragmented landscape. *Heredity*, **101**, 368–380.
- Bacles CFE, Lowe AJ, Ennos RA (2006) Effective seed dispersal across a fragmented landscape. *Science*, **311**, 628.
- Barbará T, Lexer C, Martinelli G *et al.* (2008) Within-population spatial genetic structure in four naturally fragmented plant species of a neotropical 'inselberg' radiation, *Alcantarea imperialis*, *A. geniculata*, *A. glaziouana*, and *A. regina* (Bromeliaceae). *Heredity*, **101**, 285–296.
- Bittencourt JVM, Sebbenn AM (2007) Patterns of pollen and seed dispersal in a small, fragmented population of the wind-pollinated tree *Araucaria angustifolia* in southern Brazil. *Heredity*, **99**, 580–591.
- Bladé C, Vallejo VR (2008) Seed mass effects on performance of *Pinus halepensis* Mill. seedlings sown after fire. *Forest Ecology and Management*, **255**, 2362–2372.
- Born C, Hardy OJ, Chevallier M-H *et al.* (2008) Small-scale spatial genetic structure in the Central African rainforest tree species *Aucoumea klaineana*: a stepwise approach to infer the impact of limited gene dispersal, population history and habitat fragmentation. *Molecular Ecology*, **17**, 2041–2050.
- Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, **5**, 453–455.
- Bucci G, González-Martínez SC, Le Provost G *et al.* (2007) Range-wide phylogeography and genetic zones in *Pinus pinaster* Ait. revealed by chloroplast microsatellite markers. *Molecular Ecology*, **16**, 2137–2153.
- Burban C, Petit RJ (2003) Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. *Molecular Ecology*, **12**, 1487–1495.
- Calvo L, Santalla S, Valbuena L *et al.* (2008) Post-fire natural regeneration of a *Pinus pinaster* forest in NW Spain. *Plant Ecology*, **197**, 81–90.
- Carrión JS, Navarro C, Navarro J, Munuera M (2000) The distribution of cluster pine (*Pinus pinaster*) in Spain as derived from palaeoecological data: relationships with phytosociological classification. *The Holocene*, **10**, 243–252.
- Chagné D, Chaumeil P, Ramboer A *et al.* (2004) Cross-species transferability and mapping of genomic and cDNA SSRs in pines. *Theoretical and Applied Genetics*, **109**, 1204–1214.
- Crawford TJ (1984) The estimation of neighbourhood parameters for plant populations. *Heredity*, **53**, 273–283.
- Derory J, Mariette S, González-Martínez SC *et al.* (2002) What can nuclear microsatellites tell us about maritime pine genetic resources conservation and provenance certification strategies? *Annals of Forest Science*, **59**, 699–708.

- Dick CW, Etchelecu G, Austerlitz F (2003) Pollen dispersal of tropical trees (*Dinizia excelsa*: Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rainforest. *Molecular Ecology*, **12**, 753–764.
- Dick CW, Hardy OJ, Jones FA, Petit RJ (2008) Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Tropical Plant Biology*, **1**, 20–33.
- Doligez A, Baril C, Joly HI (1998) Fine-scale spatial genetic structure with nonuniform distribution of individuals. *Genetics*, **148**, 905–919.
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *BRL Focus*, **12**, 13–15.
- Dunphy BL, Hamrick JL (2007) Estimation of gene flow into fragmented populations of *Bursera simarouba* (Burseraceae) in the dry-forest life zone of Puerto Rico. *American Journal of Botany*, **94**, 1786–1794.
- El-Kassaby YA, Jaquish B (1996) Population density and mating pattern in western larch. *The Journal of Heredity*, **87**, 438–443.
- Epperson B (1990) Spatial autocorrelation of genotypes under directional selection. *Genetics*, **124**, 757–771.
- Figueiral I (1995) Charcoal analysis and the history of *Pinus pinaster* (cluster pine) in Portugal. *Review of Palaeobotany and Palynology*, **89**, 441–454.
- Gapare WJ, Aitken SN (2005) Strong spatial genetic structure in peripheral but not core populations of Sitka spruce [*Picea sitchensis* (Bong.) Carr.]. *Molecular Ecology*, **14**, 2659–2667.
- Gérard P, Klein E, Austerlitz F *et al.* (2006) Assortative mating and differential male mating success in an ash hybrid zone population. *BMC Evolutionary Biology*, **6**, 96.
- González-Martínez SC, Salvador L, Agúndez D, Alía R, Gil L (2001) Geographical variation of gene diversity of *Pinus pinaster* Ait. in the Iberian Peninsula. In: *Genetic Response of Forest System to Changing Environmental Condition* (eds Muller-Stark G, Schubert R), pp. 161–171. Kluwer Academic Publisher, Dordrecht.
- González-Martínez SC, Gerber S, Cervera MT *et al.* (2002) Seed gene flow and fine-scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers. *Theoretical and Applied Genetics*, **104**, 1290–1297.
- González-Martínez SC, Burczyk J, Nathan R *et al.* (2006) Effective gene dispersal and female reproductive success in Mediterranean maritime pine (*Pinus pinaster* Aiton). *Molecular Ecology*, **15**, 4577–4588.
- Hampe A, Petit J (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, **8**, 461–467.
- Hardy OJ, Vekemans X (2002) SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620. <http://www.ebe.ulb.ac.be/ebe/Software.html> [accessed 21 September 2009].
- Hardy OJ, Maggia L, Bandou E *et al.* (2006) Fine-scale genetic structure and gene dispersal inferences in 10 neotropical tree species. *Molecular Ecology*, **15**, 559–571.
- Heuertz M, Vekemans X, Hausman F, Paladas M, Hardy OJ (2003) Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. *Molecular Ecology*, **12**, 2483–2495.
- Jump AS, Peñuelas J (2006) Genetic effects of chronic habitat fragmentation in a wind-pollinated tree. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 8096–8100.
- Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187–189. <http://www.montana.edu/kalinowski/Software/HPRare.htm> [accessed 21 September 2009].
- Kimmel M, Chakraborty R, King JP *et al.* (1998) Signatures of population expansion in microsatellite repeat data. *Genetics*, **148**, 1921–1930.
- Kramer AT, Ison JL, Ashley MV, Howe HF (2008) The paradox of forest fragmentation genetics. *Conservation Biology*, **22**, 878–885.
- Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, **82**, 1420–1425.
- Lowe AJ, Boshier D, Ward M, Bacles CFE, Navarro C (2005) Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for neotropical trees. *Heredity*, **95**, 255–273.
- de-Lucas AI, Robledo-Arnuncio JJ, Hidalgo E, González-Martínez SC (2008) Mating system and pollen gene flow in Mediterranean maritime pine. *Heredity*, **100**, 390–399.
- de-Lucas AI, González-Martínez SC, Hidalgo E, Bravo F, Heuertz M (2009) Admixture, one-source colonization or long-term persistence of maritime pine in the Castilian Plateau? Insights from nuclear microsatellite markers. *Investigación Agraria: Sistemas y Recursos Forestales*, **18**, 3–12.
- Mariette S, Chagné D, Decroocq S *et al.* (2001) Microsatellite markers for *Pinus pinaster* Ait. *Annals of Forest Science*, **58**, 203–206.
- Mimura M, Aitken SN (2007) Increase selfing and decrease effective pollen donor number in peripheral relative to central populations of *Picea sitchensis* (Pinaceae). *American Journal of Botany*, **94**, 991–998.
- Muller-Landau HC, Wright SJ, Calderón O, Condit R, Hubbell SP (2008) Interspecific variation in primary seed dispersal in a tropical forest. *Journal of Ecology*, **96**, 653–667.
- Nason JD, Hamrick JL (1997) Reproductive and genetic consequences of forest fragmentation: two case studies of neotropical canopy trees. *Journal of Heredity*, **88**, 264–276.
- Nathan R, Safriel UN, Noy-Meir I, Schiller G (2000) Spatiotemporal variation in seed dispersal and recruitment near and far from *Pinus halepensis* trees. *Ecology*, **81**, 2156–2169.
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.
- Nei M, Maruyana T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1–10.
- O'Connell LM, Mosseler A, Rajora OP (2006) Impacts of forest fragmentation on the mating system and genetic diversity of white spruce (*Picea glauca*) at the landscape level. *Heredity*, **97**, 418–426.
- Petit RJ, Aguinalde I, de Beaulieu JL *et al.* (2003) Glacial refuges: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563–1565.
- Pritchard JK, Wen W (2004) Documentation for STRUCTURE software version 2. <http://pritch.bsd.uchicago.edu> [accessed 21 September 2009].

- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Ribeiro MM, Plomion C, Petit R, Vendramin GG, Szmidt AE (2001) Variation in chloroplast single-sequence repeats in Portuguese maritime pine (*Pinus pinaster* Ait.). *Theoretical and Applied Genetics*, **102**, 97–103.
- Robledo-Arnuncio JJ, Alía R, Gil L (2004) Increased selfing and correlated paternity in a small population of a predominantly outcrossing conifer, *Pinus sylvestris*. *Molecular Ecology*, **13**, 2567–2577.
- Robledo-Arnuncio JJ, Austerlitz F, Smouse PE (2006) A new indirect method of estimating the pollen dispersal curve, independently of effective density. *Genetics*, **173**, 1033–1045.
- Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology*, **13**, 58–62.
- Rousset F (2004) *Genetic Structure and Selection in Subdivided Populations*. Princeton University Press, Princeton, NJ.
- RoyChoudhury A, Stephens M (2007) Fast and accurate estimation of the population-scaled mutation rate, θ , from microsatellite genotype data. *Genetics*, **176**, 1363–1366.
- Salvador L, Alía R, Agúndez D, Gil L (2000) Genetic variation and migration pathways of maritime pine (*Pinus pinaster* Ait.) in the Iberian peninsula. *Theoretical and Applied Genetics*, **100**, 89–95.
- Slatkin M (1973) Gene flow and selection in a cline. *Genetics*, **75**, 733–756.
- Smouse PE, Sork VL (2004) Measuring pollen flow in forest trees: an exposition of alternative approaches. *Forest Ecology and Management*, **197**, 21–38.
- Sokal RR, Wartenberg DE (1983) A test of spatial autocorrelation using an isolation-by-distance model. *Genetics*, **105**, 219–237.
- Streff R, Ducouso A, Lexer C *et al.* (1999) Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. *Molecular Ecology*, **8**, 831–841.
- Thompson JD (2005) *Plant Evolution in the Mediterranean*. Oxford University Press, Oxford.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538. <http://www.microchecker.hull.ac.uk/> [accessed 21 September 2009]
- Van Rossum F, Triest L (2007) Fine-scale spatial genetic structure of the distylous *Primula veris* in fragmented habitats. *Plant Biology*, **9**, 374–382.
- Vekemans X, Hardy J (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, **13**, 921–935.
- Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- Wells G, Young AG (2002) Effects of seed dispersal on spatial genetic structure in populations of *Rutidosis leptorrhychoides* with different levels of correlated paternity. *Genetical Research*, **79**, 219–226.
- White GM, Boshier DH, Powell W (2002) Increased pollen flow counteracts fragmentation in a tropical dry forest: an example from *Swietenia humilis* Zuccarini. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 2038–2042.
- Williams DA, Wang Y, Borchetta M, Gaines MS (2007) Genetic diversity and spatial structure of a keystone species in fragmented pine rockland habitat. *Biological Conservation*, **138**, 256–268.
- Yamagishi H, Tomimatsu H, Ohara M (2007) Fine-scale spatial genetic structure within continuous and fragmented populations of *Trillium camschatcense*. *Journal of Heredity*, **98**, 367–372.
- Young AG, Merriam HG (1994) Effects of forest fragmentation on the spatial genetic structure of *Acer saccharum* Marsh (Sugar Maple) populations. *Heredity*, **72**, 201–208.
- Young AG, Boyle T, Brown AHD (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology and Evolution*, **11**, 413–418.

This work is part of A.I.d.-L.'s PhD on gene flow, demography and genetic diversity in maritime pine. S.C.G.-M. is working in population genetics and genomics of forest trees and conservation genetics of Mediterranean plants. G.G.V. is interested in conservation genetics, genetic structure studies and genomics of forest trees. E.H. is professor at the University of Valladolid, her research focuses on genetics of grapevine and forest trees. M.H. is interested in empirical and simulation studies of mating systems, gene flow and genetic structure in plants.

Supporting information

Additional supporting information may be found in the online version of this article.

Data S1. Photo report.

Data S2. Population locations and their characteristics.

Data S3. Bar graphs of percentages of ancestry in each of $K = 2$ clusters inferred with the STRUCTURE software (Pritchard *et al.* 2000).

Data S4. *Sp* values obtained in simulated dispersal scenarios DE, IN1, IN2, IE (see Materials & methods for description of the simulated dispersal kernels).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.