

Extensive spatial genetic structure revealed by AFLP but not SSR molecular markers in the wind-pollinated tree, *Fagus sylvatica*

ALISTAIR S. JUMP and JOSEP PEÑUELAS

Unitat d'Ecofisiologia CSIC-CEAB-CREAF, CREAF (Centre de Recerca Ecològica i Aplicacions Forestals), Universitat Autònoma de Barcelona, 08193, Bellaterra, Catalonia, Spain

Abstract

Studies of fine-scale spatial genetic structure (SGS) in wind-pollinated trees have shown that SGS is generally weak and extends over relatively short distances (less than 30–40 m) from individual trees. However, recent simulations have shown that detection of SGS is heavily dependent on both the choice of molecular markers and the strategy used to sample the studied population. Published studies may not always have used sufficient markers and/or individuals for the accurate estimation of SGS. To assess the extent of SGS within a population of the wind-pollinated tree *Fagus sylvatica*, we genotyped 200 trees at six microsatellite or simple sequence repeat (SSR) loci and 250 amplified fragment length polymorphisms (AFLP) and conducted spatial analyses of pairwise kinship coefficients. We re-sampled our data set over individuals and over loci to determine the effect of reducing the sample size and number of loci used for SGS estimation. We found that SGS estimated from AFLP markers extended nearly four times further than has been estimated before using other molecular markers in this species, indicating a persistent effect of restricted gene flow at small spatial scales. However, our SSR-based estimate was in agreement with other published studies. Spatial genetic structure in *F. sylvatica* and similar wind-pollinated trees may therefore be substantially larger than has been estimated previously. Although 100–150 AFLP loci and 150–200 individuals appear sufficient for adequately estimating SGS in our analysis, 150–200 individuals and six SSR loci may still be too few to provide a good estimation of SGS in this species.

Keywords: European beech, gene flow, genetic diversity, population structure, spatial autocorrelation

Received 3 August 2006; revision received 23 September 2006; accepted 14 October 2006

Introduction

Forest trees are generally highly genetically diverse and characterized by high levels of genetic diversity within populations and comparatively little genetic diversity between them (Loveless & Hamrick 1984; Merzeau *et al.* 1994; Leonardi & Menozzi 1996; Takahashi *et al.* 2000; Mariette *et al.* 2002). Within populations, genetic diversity is rarely distributed homogeneously; individuals often become less genetically similar as the distance between them increases. Such spatial genetic structure (SGS) may result from the interaction of factors such as population

history, limited gene flow and natural selection (Merzeau *et al.* 1994; Streiff *et al.* 1998). Adequate knowledge of SGS is important for the conservation and management of natural populations in order not to misrepresent species or population genetic diversity (Epperson 1989; Shapcott 1995). Accurate estimates of SGS can inform decisions on the targeting of scarce conservation resources at particular populations and the sampling area necessary to represent the genetic diversity within them (Escudero *et al.* 2003). If SGS is ignored, particular genotypes may be over or under-represented within samples and potentially important adaptive variation may therefore be missed (Epperson 1989).

In tree species with restricted seed and pollen dispersal, low gene flow is expected to result in significant genetic differentiation within continuous populations (Cavers *et al.* 2005; Hardy *et al.* 2006). However, significant SGS has been

Correspondence: Alistair S. Jump, Present address: Environment Department, The University of York, Heslington, York, YO10 5 DD, UK. Fax: 0044 1904432998; E-mail: aj523@york.ac.uk

detected even in high gene flow species such as outcrossing, wind-pollinated trees (Merzeau *et al.* 1994; Leonardi & Menozzi 1996; Streiff *et al.* 1998; Takahashi *et al.* 2000; Cottrell *et al.* 2003; Scalfi *et al.* 2005). In spatially structured populations, the likelihood of mating between individuals within the population is determined partly by the distance between them. The detection of SGS, however, may depend on the strategy used to sample the population. SGS may be underestimated if sampled individuals are too widely separated, or if comparisons are made between population subsamples where SGS within subsamples is not assessed (Merzeau *et al.* 1994; Shapcott 1995; Streiff *et al.* 1998; Escudero *et al.* 2003). The number of individuals sampled, and the number of loci used to estimate genetic distances between them, may also have a significant impact on whether the estimated SGS accurately reflects the actual SGS of the population (Cavers *et al.* 2005). Additionally, different classes of molecular markers may not always give congruent estimates of SGS within the same population (Gaudeul *et al.* 2004; Hardesty *et al.* 2005; Hardy *et al.* 2006).

High gene flow within populations and the longevity of individual trees may result in many forest tree species exhibiting low levels of SGS (Leonardi & Menozzi 1996; Streiff *et al.* 1998). In wind-pollinated forest tree species such as the common ash (*Fraxinus excelsior*, Heuertz *et al.* 2003) the oaks (*Quercus petraea* and *Quercus robur*, Bacilieri *et al.* 1994; Streiff *et al.* 1998) and the beeches (*Fagus sylvatica* and *Fagus crenata* Merzeau *et al.* 1994; Leonardi & Menozzi 1996; Takahashi *et al.* 2000; Asuka *et al.* 2004; Vornam *et al.* 2004; Scalfi *et al.* 2005), significant SGS is rarely detectable beyond approximately 30–40 m. This weak structuring within populations has been reported based on analysis of molecular marker data from allozymes (Bacilieri *et al.* 1994; Merzeau *et al.* 1994; Leonardi & Menozzi 1996; Streiff *et al.* 1998; Takahashi *et al.* 2000) and microsatellites (simple sequence repeats or SSRs) (Streiff *et al.* 1998; Heuertz *et al.* 2003; Asuka *et al.* 2004; Vornam *et al.* 2004; Scalfi *et al.* 2005). However, as Cavers *et al.* (2005) demonstrate by simulating SGS in a model tropical tree population, the accurate detection of SGS is heavily dependent on selecting appropriate sample sizes, both in terms of the number of individuals sampled and the number of loci used to analyse them.

Cavers *et al.* (2005) report that approximately 100 individuals and 5–10 SSR loci or 150 individuals and 100 amplified fragment length polymorphism (AFLP) loci will provide a close approximation (a correlation of $r \approx 0.9$) of real levels of SGS in tropical forest stands, given the relatively limited pollen and seed dispersal of many tropical trees. They state that, for high gene flow species such as wind-pollinated trees, the number of individuals should be increased. However, published studies of SGS in wind-pollinated species often do not achieve the sample sizes recommended by Cavers *et al.* (2005) even for species with relatively restricted gene flow. SGS in wind-pollinated trees

may therefore have been underestimated, particularly where small sample sizes are combined with few loci that show low levels of polymorphism.

To investigate this possibility, we genotyped 200 *F. sylvatica* trees with 6 SSR loci and 250 AFLP loci and conducted spatial analyses of pairwise kinship coefficients between individuals. We aimed to determine how observed patterns of SGS vary with the number of loci used for kinship coefficient calculation and the number of individuals included in the analysis. Our study compared two contrasting sampling strategies that can be used to assess spatial genetic structure using molecular markers: (i) selection of highly informative markers at a few loci (SSRs), and (ii) sampling numerous less informative markers randomly distributed throughout the genome (AFLPs). The number of loci compared for each marker system required comparable genotyping effort.

Materials and methods

Study species

Fagus sylvatica (European beech) is a monoecious, diploid, late-successional tree that dominates temperate forests over c. 17 million ha of Europe. It is highly outcrossing and largely self-incompatible with irregular synchronous flowering (masting) events. Reproduction does not begin until the species is 40–50 years old (Nilsson & Wastljung 1987; Comps *et al.* 2001). Average seed dispersal in this species is estimated to be less than 25 m (Gregorius & Kownatzki 2005), with long distance seed dispersal effected primarily by the European jay (*Garrulus glandarius*) (Nilsson 1985).

Sites and sampling

This work was conducted in the Montseny Mountains of Catalonia (northeastern Spain), which lie 50 km north-northwest of Barcelona and 20 km inland from the Mediterranean Sea. *Fagus sylvatica* reaches the southern edge of its distribution in Europe in the Montseny Mountains, occurring in the temperate zone above approximately 1000 m above sea level (a.s.l.) and forming c. 2830 ha of near-continuous forest along the Turó de l'Home (1712 m a.s.l.) — Les Agudes (1706 m a.s.l.) ridge. The *F. sylvatica* forest is naturally occurring uneven-aged high forest where this species is monodominant with a relatively even distribution of individual trees. The forest has been managed at low intensity by the selective removal of large trees coupled with natural regeneration from seed. However, the impact of forest management on the upper and lower limits of the forest has been low (Peñuelas & Boada 2003).

We sampled 210 young trees (< 80 years old) in forest with no evidence of any recent disturbance. Trees were

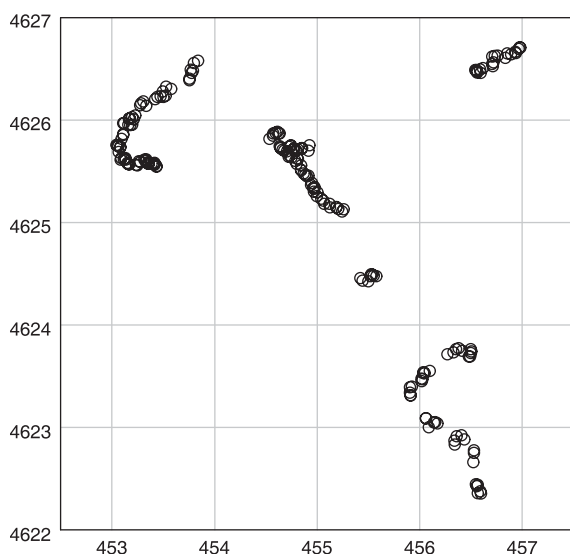


Fig. 1 The distribution of *Fagus sylvatica* individuals sampled in the Montseny Mountains, Catalonia, northeastern Spain, for the analysis of spatial genetic structure. Gridlines are numbered with Universal Transverse Mercator (UTM) coordinates in kilometres.

sampled from the largest expanse of continuous forest during spring 2004 and the location of each tree was mapped. Samples were collected at random along a series of transects from trees grouped at the upper and lower altitudinal limits of *F. sylvatica* on Turó de l'Home and from a more central region of the forest. The distance between samples varied from a minimum of 2.2 m to a maximum of 5 km (Fig. 1). From each tree, we collected newly unfolding leaves or leaf buds and dried these immediately in fine-grain silica gel.

DNA extraction

Dried leaf tissue (~0.5 cm²) was ground in liquid nitrogen for 30 s at 30 revolutions per second using a mixer mill (Tissue Lyser, QIAGEN Inc.) and two glass beads. Genomic DNA was extracted from ground tissue using a DNeasy Plant Mini Kit (QIAGEN) and quantified using a NANODROP ND-1000 spectrophotometer running software version 3.0.1 (NanoDrop Technologies) following the manufacturer's instructions.

SSR analysis

Individuals were genotyped at six highly variable microsatellite loci originally isolated in *F. sylvatica* (FS1-03, FS1-15, FS3-04, FS4-46) (Pastorelli *et al.* 2003), *Fagus crenata* (FCM5) (Tanaka *et al.* 1999) and *Castanea sativa* (CsCAT14) (Marinoni *et al.* 2003). Polymerase chain reaction (PCR) was performed with 5 ng template DNA in 15 µL buffer consisting of 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 0.001% gelatine,

200 µM each dATP, dCTP, dGTP, dTTP, 1% polyvinyl pyrrolidone-40, 0.125% BSA(V), 5 pmol each forward and reverse primer, 0.25 U (approximate) DNA polymerase (IRTA, isolated according to a standard protocol). Forward primers carried a fluorescent label (Applied Biosystems – ABI). MgCl₂ concentrations and cycling conditions were as reported in the original publications. Each set of 96 PCRs included a negative (water) and positive (known genotype) control. Alleles belonging to the six different loci were segregated on an ABI 3100 Avant genetic analyser according to the manufacturer's instructions. Fragment sizes were determined with reference to a LIZ-500 size standard (ABI) for each sample injection using GENESCAN software version 3.7.

AFLP analysis

The AFLP protocol followed is detailed in Jump *et al.* (2006). After digestion of genomic DNA with *EcoRI* and *Tru9I* (an isoschizomer of *MseI*), adaptors were ligated on both ends of the DNA fragments. Two sequential PCR amplifications were performed. The first PCR using *Eco-A* and *Mse-C* primers and the second PCR using one of five *Eco-ANN*/*Mse-CNN* primer combinations (where N represents any nucleotide). *Eco-ANN* primers were radiolabelled using $\gamma^{32}\text{P}$. Primer combinations used were (i) *Eco-AAG*/*Mse-CAG* (ii) *Eco-ACT*/*Mse-CTC* (iii) *Eco-ACA*/*Mse-CAG* (iv) *Eco-ACA*/*Mse-CAC* (v) *Eco-AAG*/*Mse-CTC*. Oligonucleotide sequences for PCR adapters and primers are as published by Vos *et al.* (1995).

Electrophoresis of radiolabelled PCR products was performed on 6% polyacrylamide gels. These were dried and exposed to X-ray film for 3–5 days at –80 °C. Films were developed following a standard protocol and genotypes scored by hand as a binary matrix of band presence/absence. Each set of 60 reactions included two positive (known genotype) and two negative (H₂O) controls carried from restriction digest through to the final selective AFLP PCR.

Statistical analysis

Eight individuals failed to amplify in either the SSR analysis or the AFLP analysis. These eight individuals were excluded from the final data set, together with two additional randomly selected individuals, to give a total of 200 individuals that amplified for both marker types and were included in the final analysis. The spatial distribution of these included individuals is shown in Fig. 1.

A single AFLP locus found to violate assumptions of marker neutrality by Jump *et al.* (2006) was excluded from the data set. No SSR loci were found to deviate from neutrality (data not shown). Of the total 254 polymorphic AFLP loci, a further three randomly selected loci were excluded from the analysis to give a final data set of 250

AFLP loci and 6 SSR loci per individual. In the following analyses, we compared five AFLP primer combinations with six SSR loci to provide diversity and SGS estimates based on a similar genotyping effort for both marker types while ensuring that minimum comparisons (one AFLP primer combination vs. two SSR loci) still involved multilocus genotypes.

Population genetic diversity

SSR allele frequencies, observed and expected heterozygosities and Shannon's diversity index were calculated using the software POPGENE version 1.32 (Yeh & Boyle 1997). Calculation of allelic richness, F_{IS} (using the estimator of Weir & Cockerham (1984)) and deviation from Hardy–Weinberg Equilibrium (HWE) was performed in FSTAT version 2.9.3.2 (Goudet 2001).

For the estimation of genetic diversity from AFLP data, we calculated allele frequencies based on the assumption of HWE within populations. Although the population samples we describe here deviated significantly from HWE when analysed using SSR markers, previous work using the same SSR markers on seven nearby quadrat subsamples of the same forest provided no evidence of significant deviation from HWE (Jump & Peñuelas 2006). The potential cause of deviation from HWE in this study is discussed below. We estimated the frequency of the dominant allele at each AFLP locus for the population as $1 - \text{the square root of the frequency of band absence at that locus}$ (Krauss 2000). Genetic diversity was calculated from AFLP data as mean expected heterozygosity and mean Shannon diversity index over all loci using POPGENE.

Spatial genetic structure

Analyses of SGS were conducted based on the calculation of the kinship coefficient between individuals (F_{ij}) which summarizes the genetic co-ancestry between individuals i and j and can be defined as $F_{ij} = (Q_{ij} - Q_m) / 1 - Q_m$ (where Q_{ij} is the probability of identity in state for random genes from i and j , and Q_m is the average probability of identity in state for genes coming from random individuals from the sample). Kinship coefficients were calculated from AFLP data according to the dominant marker estimator described by Hardy (2003) and from SSR data according to the codominant marker estimator reported by Loiselle *et al.* (1995) (equations 15 and 16, respectively, in Hardy 2003). For estimation of kinship coefficients from AFLP data, the inbreeding coefficient (F_{IS}) was set to 0.088, as calculated from SSR data in this study, however, Hardy (2003) notes that estimation of kinship using this method is relatively robust to errors in F_{IS} estimation.

It was clear from an initial investigation of SGS in this population that the mean estimate of F_{ij} over pairs of

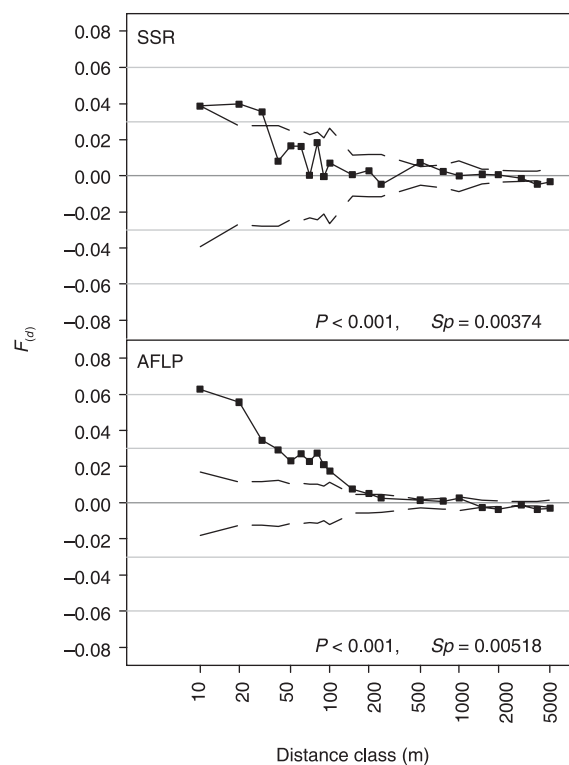


Fig. 2 Analysis of spatial genetic structure in *Fagus sylvatica* forest based on SSR and AFLP data over the full distance range contained within our data set. Solid lines indicate the mean kinship coefficient per distance class [$F_{(d)}$] and dashed lines the limits of its 95% confidence interval. Individuals are significantly more similar than would be expected by chance when $F_{(d)}$ lies above its 95% confidence limit. P values are given for the regression slope of individual pairwise kinship coefficients on the corresponding logarithmic geographical distance between individuals (see text). Sp represents the rate of decrease of pairwise kinship with distance.

individuals within a given distance interval [$F_{(d)}$] does not decrease linearly with distance or log distance but shows a rapid reduction in slope beyond approximately 200 m from each tree (Fig. 2). Consequently, analyses were restricted to a maximum distance of 200 m between trees to avoid parameter estimates being biased by this nonlinearity. $F_{(d)}$ was computed based on 20 equal intervals rising by 10 m from 0 to 200 m and plotted against distance to visualize spatial autocorrelation patterns. Ninety-five per cent confidence intervals for $F_{(d)}$ were calculated based on 10 000 random permutations of individuals among geographical locations. To test the hypothesis that there was significant spatial structure, the observed regression slope of F_{ij} on $\ln(r_{ij})$, b , was compared with those obtained after 10 000 random permutations of individuals among locations and statistical significance determined using a Mantel test [where $\ln(r_{ij})$ represents the natural logarithm of the physical distance between individuals i and j]. This procedure has the advantage that all of the information is contained in

a single test statistic and the results are not dependent on arbitrarily set distance intervals (Vekemans & Hardy 2004). Spatial genetic structure was also quantified by the S_p statistic, which represents the rate of decrease of pairwise kinship with distance (Vekemans & Hardy 2004). S_p is calculated as $-b/[1 - F_{(1)}]$ [where $F_{(1)}$ is the mean F_{ij} between individuals in the first distance class]. Following Hardy *et al.* (2006), we present the standard error (SE) of b (calculated by jackknifing over loci) as an estimate of the variability of S_p . This estimate does not take into account the SE of $F_{(1)}$, since its impact on the SE of S_p is very small in comparison with that of b (OJ Hardy, personal communication). All analyses of SGS were conducted using the program SPAGEdi version 1.2d (Hardy & Vekemans 2002).

We subsampled the complete data sets ($N = 200$ individuals with 6 SSR loci and 250 AFLP loci) over individuals and over loci to assess the effect of varying sample sizes on estimates of spatial genetic structure. To assess the effect of reducing the number of individuals, N was reduced from $N = 200$ by the random exclusion of individuals to give data sets of $N = 150$ and $N = 100$. To assess the effect of reducing the number of loci, spatial analyses were conducted based on 2, 3, 4, 5 and 6 SSR loci arbitrarily ordered as follows: (i) FS1-03 (ii) FS4-46 (iii) FS1-15 (iv) FCM5 (v) CsCAT14 (vi) FS3-04. AFLP loci were numbered from 1 to 250 according to primer combination and product size within each primer combination. Spatial analyses of AFLP data were based on 50, 100, 150, 200 and 250 polymorphic loci selected in order from the full data set. Distance matrices were also calculated based on $N = 200$ for individual SSR loci and AFLP primer combinations to quantify how closely SGS for individual loci/primer combinations corresponded to overall estimates of SGS. The Pearson product-moment correlation coefficient (r) was calculated to compare $F_{(d)}$ calculated from reduced numbers of individuals and loci with $F_{(d)}$ based on the full data set for each marker type. For all such comparisons, r is calculated based on a sample size of $N = 20$ distance classes.

Results

Marker summary information

A total of 75 alleles were recorded over all SSR loci. Alleles corresponded to individual loci as follows: FS1-03: 24 alleles; FS1-15: 17; FS3-04: 4; FS4-46: 12; FCM5: 15; CsCAT14: 3. Observed and expected mean heterozygosities per locus ranged from 0.345 and 0.412 (CsCAT14) to 0.850 and 0.875 (FS1-15).

A total of 254 easily scored polymorphic AFLP markers were recorded over 209 individuals using five primer combinations. The number of polymorphic markers per individual primer combination was as follows: (i) *Eco*-

AAG/*Mse*-CAG, 50 markers; (ii) *Eco*-ACT/*Mse*-CTC, 44; (iii) *Eco*-ACA/*Mse*-CAG, 53; (iv) *Eco*-ACA/*Mse*-CAC, 50; (v) *Eco*-AAG/*Mse*-CTC, 57.

Population genetic diversity

Per-locus tests of Hardy–Weinberg equilibrium (HWE) indicated that three SSR loci showed a significant departure from HWE ($F_{IS} \neq 0$, $\alpha = 0.05$ after Bonferroni correction, 6000 randomizations); all involved heterozygote deficit. When tested over all SSR loci, a significant heterozygote deficit remained ($F_{IS} = 0.088$, $\alpha = 0.05$). Genetic diversity of the forest was high, with an allelic richness of 12.5 (averaged over loci) from SSR data. SSR-based estimates of genetic diversity were higher than those based on AFLP markers: Shannon diversity index, SSR: 1.593; AFLP: 0.380, expected heterozygosity SSR: 0.703; AFLP: 0.244.

Spatial genetic structure

When all 200 individuals were included in the analysis, significant SGS was detected with both 6 SSR and 250 AFLP markers (Fig. 3a). However, whereas SSR markers revealed only very weak structure for distance classes greater than 10 m and less than 30 m, AFLP markers revealed significant SGS extending nearly four times further — up to 110 m from each tree. For AFLP data, the mean kinship coefficient per distance class, $F_{(d)}$, is described by a near monotonic decrease with increasing distance, a pattern that is not repeated for SSR data (Fig. 3a). We found no consistent evidence for differences in SGS between the sample regions of the forest (data not shown).

When correlograms were constructed based on smaller sample sizes ($N = 150$, Fig. 3b; $N = 100$, Fig. 3c), it is clear that in each case, greater SGS was detected with AFLP rather than SSR markers. In all cases, SGS became weaker as sample size was reduced. When calculated based on AFLP markers, the minimum SGS estimate (SGS_{min}) falls from 110 m when $N = 200$, to 90 m when $N = 150$ and 20 m when $N = 100$. (SGS_{min} is defined here as the point at which $F_{(d)}$ first becomes statistically indistinguishable from 0 at $P = 0.05$). The maximum SGS estimate (SGS_{max}) is identical to SGS_{min} when calculated based on 200 or 150 individuals. However, when calculated based on 100 individuals, SGS_{max} is 120 m, 100 m greater than the minimum estimate. (SGS_{max} is defined here as the greatest distance at which $F_{(d)} > 0$ at $P = 0.05$ before $F_{(d)}$ crosses the x -axis). When calculated based on SSR markers, SGS_{min} falls from 30 m when $N = 200$ and $N = 150$ to 20 m when $N = 100$. In each case, no significant SGS is detected for the first distance class of 0–10 m and SGS_{min} and SGS_{max} are equivalent. Average sample size per distance class fell from 71.5 (± 3.0) when $N = 200$ to 38.9 (± 1.7) when $N = 150$ and 17.4 (± 1.1) when $N = 100$ (means are followed by standard error in

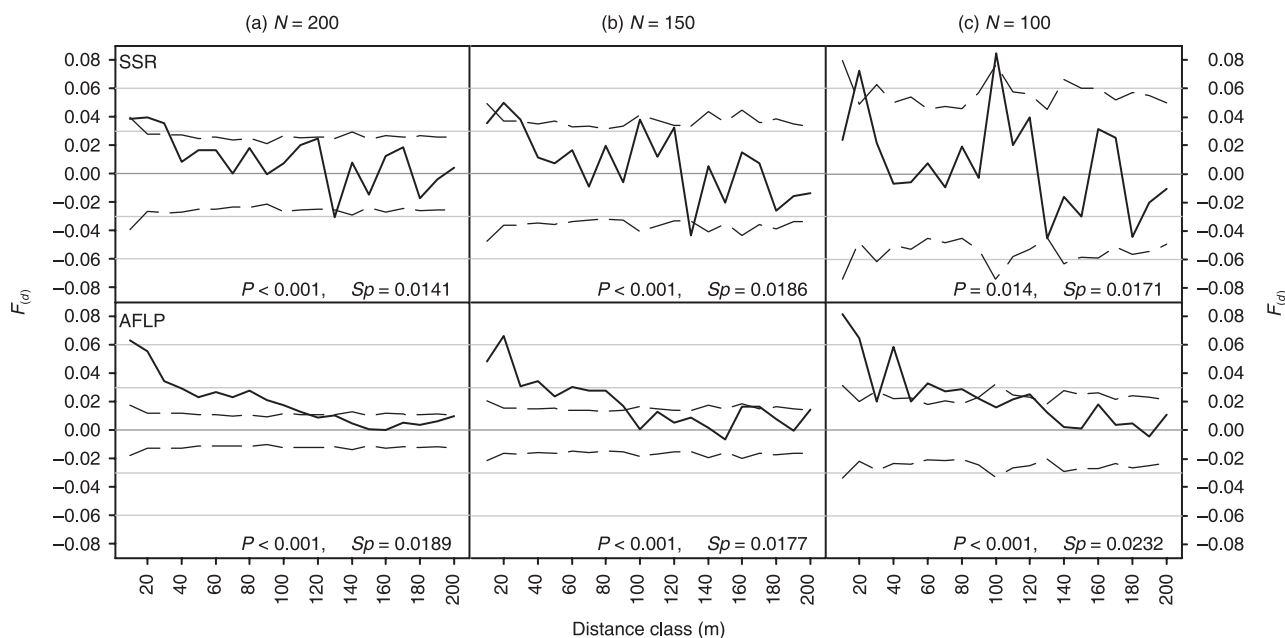


Fig. 3 Analysis of spatial genetic structure in *Fagus sylvatica* forest from the mean kinship coefficient per distance class [$F_{(d)}$] calculated from SSR and AFLP data over 10-m distance classes up to a maximum distance of 200 m. The effect of decreasing sample size (number of individuals, N) is shown by decreasing N from left to right. Explanation of lines, P values and Sp is as for Fig. 2.

parentheses). Correlograms based on $N = 150$ closely resemble the correlogram based on $N = 200$ for both marker types: $r = 0.91$, SSR; $r = 0.89$, AFLP. However, reducing sample sizes to $N = 100$ has a greater impact on the SGS estimates based on SSR markers: $r = 0.73$, SSR; $r = 0.90$, AFLP.

When comparing correlograms for SSR and AFLP data in Fig. 3, it is clear that the confidence interval for $F_{(d)}$ is always greater when $F_{(d)}$ is derived from SSR data and when $F_{(d)}$ is calculated from fewer individuals. Figure 4 shows the effect of decreasing the number of loci used for the calculation of the kinship coefficient for both marker types on overall estimates of $F_{(d)}$ and its confidence interval. The confidence interval for $F_{(d)}$ based on SSR data is always wider than that calculated from AFLP data in this study, even when 50 AFLP loci are compared against six SSR loci. For both marker types, increasing the number of markers decreases the confidence interval of $F_{(d)}$. This effect is most noticeable for AFLP data, where the confidence interval decreases consistently with every increase in the number of marker loci used for genetic distance calculation, although the decrease becomes progressively smaller as more AFLP loci are included. For SSR markers, there is no overall decrease in the width of the confidence interval when more than four loci are used for genetic distance calculation (Fig. 4).

Altering the number of loci used also results in a marked variation of $F_{(d)}$ itself. For AFLP data, if 150 loci are used for the calculation of genetic distance, the correlogram closely

resembles ($r = 0.97$) that based on the AFLP data set as a whole (250 loci, Fig. 4). If a correlation of 0.9 is taken as the critical correlation of $F_{(d)}$ values based on subsets of the data with $F_{(d)}$ values based on the full data set (following Cavers *et al.* 2005), then 100 AFLP loci give a good approximation of the final correlogram ($r = 0.93$, Fig. 5). However, whereas for SSR markers, increasing the number of loci does cause the correlogram to become more similar to that based on all six loci, each increase in the number of loci used for its calculation still results in a marked change of $F_{(d)}$ (Figs 4 and 5). Five SSR loci are needed for the calculation of $F_{(d)}$ values that approximate those based on the SSR data set as a whole ($r = 0.90$, Fig. 5).

The full data set of $N = 200$ with 6 SSR loci/250 AFLP loci gives our best estimates of the Sp statistic in this population. Values of Sp calculated from our full data set over a maximum of 5000 m, $Sp_{5000\text{ m}}$ (SSR: 0.00374 ± 0.00111 (SE); AFLP: 0.00518 ± 0.00055) are considerably lower than the values of this statistic calculated over a maximum distance of 200 m, $Sp_{200\text{ m}}$ (SSR: 0.01412 ± 0.00316 ; AFLP: 0.01888 ± 0.00217). When calculated over a maximum distance of 200 m and compared over the analyses based on reduced numbers of individuals or loci, AFLP-based estimates of Sp are higher than all but one estimate of Sp based on SSRs markers (Figs 3 and 4) with a mean Sp over all analyses of 0.01382 for SSR and 0.02071 for AFLP markers.

Values of $F_{(d)}$ based on individual SSR loci show only a low to moderate correlation with those based on the entire SSR data set (mean $r = 0.47$ range: 0.35–0.65) and only half

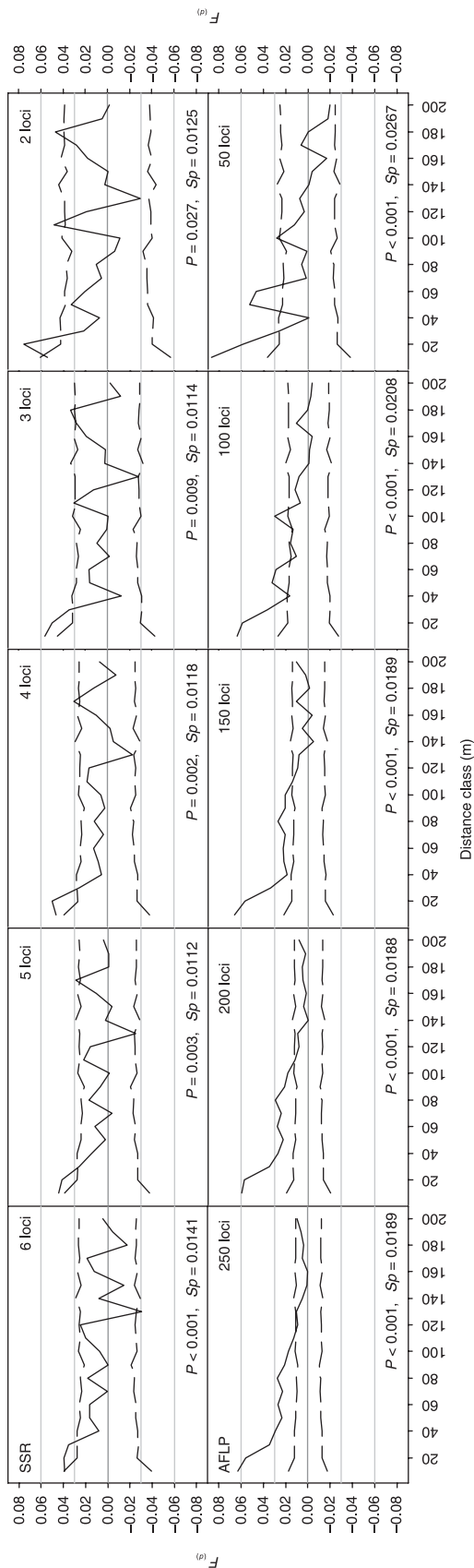


Fig. 4 The effect of decreasing the number of SSR and AFLP loci used for the calculation of individual pairwise kinship coefficients on the estimation of spatial genetic structure in *Fagus sylvatica* forest. Spatial analysis of the mean kinship coefficient, $F_{(d)}$, was based on 200 individuals and 10-m distance classes. Explanation of lines, P values and Sp is as for Fig. 2.

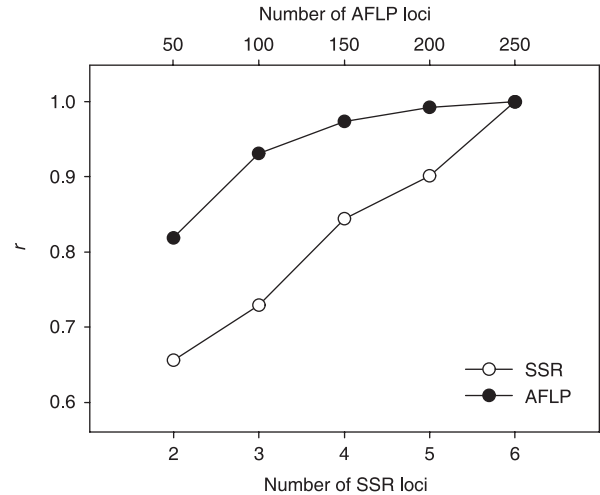


Fig. 5 Correlation of $F_{(d)}$ values calculated from the full data set of 200 individuals and 6 SSR/250 AFLP loci with $F_{(d)}$ values calculated from reduced numbers of SSR and AFLP loci. All correlations are significant at or below $P = 0.001$.

Table 1 Correlation of $F_{(d)}$ values for individual SSR loci/AFLP primer combinations with values from all SSR/AFLP data

Comparison	A/L	r	P
SSR loci vs. total SSR data			
1) FS1-03	24	0.59	0.006
2) FS4-46	12	0.37	0.108
3) FS1-15	17	0.52	0.019
4) FCM5	15	0.37	0.110
5) CsCAT14	3	0.35	0.129
6) FS3-04	4	0.65	0.002
AFLP primer combinations vs. total AFLP data			
1) <i>Eco</i> -AAG/ <i>Mse</i> -CAG	50	0.79	< 0.001
2) <i>Eco</i> -ACT/ <i>Mse</i> -CTC	53	0.82	< 0.001
3) <i>Eco</i> -ACA/ <i>Mse</i> -CAG	44	0.76	< 0.001
4) <i>Eco</i> -ACA/ <i>Mse</i> -CAC	50	0.69	0.001
5) <i>Eco</i> -AAG/ <i>Mse</i> -CTC	53	0.93	< 0.001

A/L, number of alleles per SSR locus or loci per AFLP primer combination used to estimate SGS. Significant correlations ($P < 0.05$) are presented in bold text.

of individual locus vs. full data set correlations are statistically significant at $P = 0.05$. In contrast, all individual AFLP primer combinations result in values of $F_{(d)}$ that resemble values based on the entire AFLP data set at least moderately well (mean $r = 0.80$, range: 0.69–0.93) with all correlations significant at $P = 0.001$ or below (Table 1). SGS estimated from individual SSR loci shows a low correlation with SGS estimated from other loci, whereas the correlation between individual AFLP primer sets is much higher (SSR: mean $r = 0.07$, range: –0.29–0.35; AFLP: mean $r = 0.55$, range: 0.35–0.74).

Discussion

It is apparent from our results that genetic diversity is not distributed homogeneously over the studied forest. The spatial genetic structure we detect is likely to be a consequence of a persistent effect of limited gene flow through restricted seed dispersal, despite the potential for disruption of SGS by long-distance gene flow by wind-dispersed pollen in this species (Bacilieri *et al.* 1994; Streiff *et al.* 1998; Heuertz *et al.* 2003; Asuka *et al.* 2004).

Our results show that, based on AFLP data, significant spatial genetic structure in *Fagus sylvatica* extends up to 110 m from each tree, three to four times further than has previously been estimated for this species (Merzeau *et al.* 1994; Leonardi & Menozzi 1996; Vornam *et al.* 2004; Scalfi *et al.* 2005). In contrast, SGS detected from microsatellite data is very weak, extending from 10 to 30 m based on the same sample size ($N = 200$, Fig. 3a). Although Vekemans & Hardy (2004) caution against comparing the extent of SGS between studies based on correlograms since these are dependent on the sampling scheme used and arbitrarily set distance intervals, our SSR-based SGS estimate is consistent with estimates from previous studies. The confidence interval for AFLP data is much narrower than for SSRs, resulting in greater power to detect SGS from our AFLP data. However, this does not entirely explain the weaker spatial structure we detected using SSR rather than AFLP markers. When calculated from SSRs, $F_{(d)}$ drops to near 0 by 70 m [$F_{(d)} = 0.0001$, $P = 0.96$], compared to 150 m for AFLP data ($F_d = 0.0005$, $P = 0.84$, Fig. 3a), likewise S_p is lower when calculated from SSR rather than AFLPs from the full data set (Figs 2 and 3a) and lower in all but one case when reduced data sets are compared (Figs 3b, 3c and 4).

Effect of sample size used to estimate SGS

Reduction of sample sizes below $N = 200$ reduces the distance over which SGS is detected and causes unpredictable variation in the value of S_p (Fig. 3). However, whereas estimates of SGS for $N = 150$ (Fig. 3b) for both marker types still closely resemble the overall estimate ($r \approx 0.9$), the reduction of sample sizes to $N = 100$ (Fig. 3c) has a much greater effect on the SSR based estimate ($r = 0.73$) than that based on AFLPs ($r = 0.90$). The impact of reducing sample sizes on SGS estimates is visible in Fig. 3, where the plots of $F_{(d)}$ against distance class become more variable as sample size decreases, both in terms of larger confidence intervals for each distance class and greater stochasticity of $F_{(d)}$ across distance classes. This effect is particularly pronounced for SSR data. This effect of reduced sample size is apparent in other studies that have compared SGS in populations of varying size (Takahashi *et al.* 2000; Cottrell *et al.* 2003; Asuka *et al.* 2004). Mean sample sizes per distance

class for $N = 100$ are low however (17.4 pairwise comparisons), and below the minimum recommended sample size of 30 pairwise comparisons per distance class for statistically robust estimation of SGS (Escudero *et al.* 2003; Cavers *et al.* 2005). Low sample size will reduce power to detect SGS in the smallest subsample in this test in our study and in other studies where low sample sizes are used. However, even with such low sample sizes per distance class, the AFLP correlogram for $N = 100$ still closely resembles that from the full data set ($r = 0.90$).

Effect of the number of loci used to estimate SGS

Reducing the number of loci used for calculation of F_{ij} between individuals has a similar effect to the reduction in sample size (Figs 3–5). Correlograms show greater departure from the near monotonic decrease in $F_{(d)}$ with distance of the full data set AFLP-based correlogram, and the $F_{(d)}$ confidence interval becomes wider, thereby decreasing the power to detect SGS, as the number of loci is decreased. The change in S_p with decreasing numbers of loci is not predictable (Fig. 4), although the estimate based on AFLP data does not alter until its calculation is based on less than 150 loci.

If we assume that the effort to genotype individuals with one AFLP primer combination (here equivalent to approximately 50 AFLP loci) is comparable to that required for a single SSR locus, then the reduction of loci used for the SGS estimate has a greater effect on the SSR data. When the number of loci used to estimate SGS is reduced to 100 AFLP/3 SSR loci then the relationship between $F_{(d)}$ from the reduced data set and $F_{(d)}$ calculated from the full data set is $r = 0.73$ for SSRs but still $r = 0.93$ for AFLP data. From Fig. 5 it is apparent that using more than 100–150 AFLP loci makes little improvement to the overall estimate of SGS in *F. sylvatica*, whereas each additional SSR locus markedly improves the SSR-based estimate. Cavers *et al.* (2005) demonstrate that when sufficient markers are used to accurately estimate the real SGS of the population, the curves shown in Fig. 5 should reach a plateau, with each additional marker (or primer combination) yielding little increase in the accuracy of the SGS estimate. The contrasting curves shown in Fig. 5 suggest that whereas the number of AFLP markers used in the full data set is in excess of that needed for an accurate estimate of SGS, the inclusion of additional SSR markers would be likely to lead to further improvement. Consequently, we conclude from our data that whereas 100–150 AFLP loci accurately estimate SGS, six SSR loci are still too few to give a good estimate of SGS in this species. Unfortunately, as in Fig. 5 we can only compare reduced numbers of markers to our full data set, we cannot estimate how many SSR markers should be used to provide an estimate of SGS comparable to that obtained from 150 AFLP loci.

SGS estimated from AFLP vs. SSR data

The 110 m extent of SGS in *F. sylvatica* that we detected using AFLP markers is substantially higher than that reported in most published studies of this and similar species and with the SSR data we present here. AFLP-based estimates of SGS have been published for the wind-pollinated species *Quercus dentata* (> 50 m < 100 m) and *Quercus crispula* (no significant SGS) (Ishida & Kimura 2003). However, as the authors acknowledge, their study was hampered by very small sample sizes ($N = 54$ and $N = 37$, respectively), which are well below that suggested by Cavers *et al.* (2005) and the work we present here. The estimates of SGS in *Q. dentata* and *Q. crispula* should therefore be treated with caution. Previous studies of SGS in *F. sylvatica* have estimated its maximum extent to be generally less than 40 m (Merzeau *et al.* 1994; Leonardi & Menozzi 1996; Vornam *et al.* 2004; Scalfi *et al.* 2005). Our estimate from SSR data agrees with these studies and with comparable patterns reported for similar species, such as *Fagus crenata*, *Fraxinus excelsior*, *Quercus petraea* and *Quercus robur* (Bacilieri *et al.* 1994; Streiff *et al.* 1998; Takahashi *et al.* 2000; Heuertz *et al.* 2003; Asuka *et al.* 2004). Studies comparing differentiation between populations using AFLP and SSR markers cite two potential causes of the detection of higher population differentiation with AFLPs than with SSRs that are relevant to this study: SSR-based estimates of population differentiation may be lower due to the potential for maternal inheritance of some AFLP markers due to possible localization in the chloroplast or mitochondrial genome, and the greater likelihood that some AFLPs will be linked to adaptive traits (Maguire *et al.* 2002; Mariette *et al.* 2002; Gaudeul *et al.* 2004).

The size of the nuclear genome of *F. sylvatica* is estimated to be 544×10^6 bp (Gallois *et al.* 1999). No direct estimates of the size of the chloroplast or mitochondrial genome are available in this species. However, if we take upper estimates across plant species for the size of the genome of these organelles as 0.2×10^3 bp (Martin & Herrmann 1998) and 2.4×10^3 bp (Ohta *et al.* 1998), respectively, then less than two of the 250 AFLP loci we analysed here are likely to be located in maternally inherited genomes, based on probability alone.

We excluded one AFLP marker that was found to be potentially non-neutral in an earlier study in this species (Jump *et al.* 2006). Low levels of linkage disequilibrium are expected for *F. sylvatica* (Jump *et al.* 2006) and our 250 markers give an average of only one marker per 2.18×10^6 bp of the genome. The probability of additional AFLP markers from our data set being associated with spatially selected adaptive traits in this population is therefore also very low. Potential localization of AFLP markers in maternally inherited genomes in *F. sylvatica*, or association of markers with adaptive traits is therefore unlikely to be

responsible for the strong SGS we detect in this species using AFLPs but not using SSRs.

Patterns of variation at a single SSR locus are often poorly predicted by other SSR loci. Consequently, six SSR loci are likely to be too few to provide an exact view of the whole genome (Maguire *et al.* 2002; Mariette *et al.* 2002; Cottrell *et al.* 2003; Gaudeul *et al.* 2004). In our data, SGS estimated from individual SSR loci shows a low correlation with that estimated from other SSR loci (mean $r = 0.07$) and with SGS estimated from all SSR data (mean $r = 0.47$, Table 1). In contrast, this correlation is higher between individual AFLP primer combinations (mean $r = 0.55$) and between individual primer combinations and the overall AFLP-based SGS estimate (mean $r = 0.80$, Table 1). In combination with the typically higher S_p values from AFLP data rather than SSR data, these low SSR-based correlations suggest that the greater SGS that we report from AFLP-based estimates may result from the much greater coverage of the *F. sylvatica* genome with AFLPs. Therefore six SSR loci may be insufficient to accurately estimate SGS in this species (Cottrell *et al.* 2003; Gaudeul *et al.* 2004; Fig. 5).

Estimates of SGS in similar wind-pollinated species

If six SSR loci are not sufficient to accurately estimate SGS in *F. sylvatica* from 150 to 200 individuals in our study, it is possible that SGS in this and similar species has previously been underestimated. Our SSR-based estimate of SGS in this species is in agreement with other published work (Bacilieri *et al.* 1994; Merzeau *et al.* 1994; Leonardi & Menozzi 1996; Streiff *et al.* 1998; Takahashi *et al.* 2000; Heuertz *et al.* 2003; Asuka *et al.* 2004; Vornam *et al.* 2004; Scalfi *et al.* 2005). These studies report estimates of SGS based on similar or smaller sample sizes than those used here, generally between three and six SSR loci and 4 to 11 allozyme loci with 30–220 individuals per population. Larger sample sizes (number of individuals) are included only in the studies of Asuka *et al.* (2004), who genotyped a population of $N = 506$ over three SSR loci and Takahashi *et al.* (2000) who include one population of $N = 486$ genotyped with nine allozyme loci.

A study conducted on *Q. petraea* and *Q. robur* by Cottrell *et al.* (2003) supports the possibility that SGS in some wind-pollinated species may have been underestimated. Cottrell *et al.* (2003) report extensive SGS from large samples of these species genotyped at six SSR loci (*Q. petraea*, $N = 214$, SGS < 80 m; *Q. robur*, $N = 388$, SGS < 160 m; six SSR loci), although they report much lower SGS for samples of $N = 58$. Cottrell and colleagues sampled *Q. robur* from open woodland intermixed with other species, rather than from dense monodominant stands typical of *Q. petraea*. *Quercus robur* may therefore be less comparable with *F. sylvatica* in our study. Cottrell *et al.* (2003) suggest that the extensive SGS recorded for *Q. robur* may be due in part to forest management practices that might have involved

planting of related material throughout the forest they studied. However, forest managers maintain that only natural regeneration from seed has occurred in the forest we studied at Montseny (M. Boada, personal communication). The strong spatial genetic structure that we report for *F. sylvatica* based on AFLP data is therefore unlikely to be of anthropogenic origin, particularly as strong SGS is not detected from the same samples using SSR markers.

Comparison of the values of the S_p statistic that we present here with published values for other species is problematic. If the relationship between $F_{(d)}$ and $\ln(r)$ departs strongly from linearity then the value of the S_p statistic becomes dependent on the distance range implicit within the sample (Vekemans & Hardy 2004). It is clear from Fig. 2 that this condition applies to our data. Although the values of $S_{p_{5000\text{ m}}}$ (SSR: 0.00374; AFLP: 0.00518, Fig. 2) are comparable with values reported by Hardy *et al.* (2006) for tree species with similar pollination and dispersal characteristics ($S_p = 0.00196\text{--}0.01076$), our values for $S_{p_{200\text{ m}}}$ (SSR: 0.01412; AFLP: 0.01888, Fig. 3) are nearly four times larger than for $S_{p_{5000\text{ m}}}$. Consequently, we use the values of the S_p statistic that we present only for comparison between tests within this study.

Genetic diversity, inbreeding and sampling scale

Levels of genetic diversity that we report for the *F. sylvatica* forest at Montseny are in broad agreement with estimates published in studies of this and similar species (Streiff *et al.* 1998; Mariette *et al.* 2002; Cottrell *et al.* 2003; Vornam *et al.* 2004). The lower level of diversity we report with AFLP markers is related to their di-allelic character, which limits heterozygosity to a maximum value of 0.5 (Maguire *et al.* 2002; Mariette *et al.* 2002; Gaudeul *et al.* 2004). We detected low but significant levels of inbreeding at three individual loci and over all loci ($F_{IS} = 0.088$). In a previous study using the same SSR loci, Jump & Peñuelas (2006) report a value of F_{IS} of 0.062 that is not statistically distinguishable from zero based on *F. sylvatica* quadrat subsamples taken within the same area of forest. The inbreeding that we detect here may therefore result from a Wahlund effect caused by the pooling of samples taken from a large area of forest within which significant substructuring occurs as we demonstrate.

Jump & Peñuelas (2006) reported extensive gene flow over a larger area of the same *F. sylvatica* forest, in apparent contrast to the work we present here. However, rather than highlighting a true disparity, the findings of these two studies demonstrate why sampling strategy for investigating genetic structure must be appropriate to the hypotheses being tested (Merzeau *et al.* 1994; Shapcott 1995; Streiff *et al.* 1998; Escudero *et al.* 2003; Cavers *et al.* 2005). Jump & Peñuelas (2006) investigated large-scale patterns of gene flow over the forest by the assessment of isolation-by-

distance patterns and population differentiation between quadrat subsamples. As individual subsamples were equal to or larger than the maximum extent of SGS we report here, the small-scale genetic structure that we report in the present study was not detected in this previous work. The work we present here shows that SGS does occur, but it is limited to distances much smaller than those covered by the forest population as a whole.

The utility of AFLP markers for estimating SGS

Historically, most studies of SGS have been based on analysis of allozyme data, with only more recent studies based on analysis of SSRs. Allozymes are less sensitive for the detection of SGS than SSR markers, potentially as a consequence of their typically lower polymorphism (Asuka *et al.* 2004; Hardy *et al.* 2006). In a recent review of 61 papers assessing SGS in 67 species, Escudero *et al.* (2003) report that 87% of the papers used allozymes. Although 5% of the studies reported by Escudero *et al.* (2003) were based on random amplified polymorphic DNA (RAPD) data, none were based on analysis of AFLPs and few AFLP-based papers have since been published.

Although individual AFLP loci provide relatively little information, our work shows that a large number of them distributed throughout the genome may provide a better estimate of SGS than the small number of more informative SSRs or allozymes commonly used for SGS estimation. Despite a comparable genotyping effort for AFLPs and SSRs in our study, we report a much higher estimate of SGS in *F. sylvatica* using AFLPs than that detected using SSRs or allozymes in this and other studies. Similarly, Hardy *et al.* (2006) report stronger SGS for AFLP vs. allozyme-based estimates of SGS in neotropical trees where data from both markers were available. However, although Hardesty *et al.* (2005) detected weak SGS (generally < 40 m) using SSR markers in the animal dispersed neotropical tree *Simarouba amara*, they detected no SGS using AFLP markers on the same individuals ($N = 100$, 5 SSR loci/155 AFLP loci). Consequently, although our work suggests AFLPs may be highly effective for the estimation of SGS, additional estimates of SGS from AFLP markers are necessary, particularly in wind-pollinated forest trees, in order to better understand SGS in these species and to determine the generality of our results.

Conclusions

We detected significant spatial structure in *Fagus sylvatica* extending up to 110 m from each tree based on AFLP data. This degree of spatial structure is nearly four times further than has been reported previously for this species and others that are ecologically similar. When we repeated the same analysis based on SSR data, spatial structure extended

only 30 m, in agreement with other published studies. The larger spatial structure we detected using AFLP markers is therefore not a peculiar characteristic of the population studied, but a consequence of the marker data used for the analysis. Subsampling our data over loci and individuals suggests that whereas our data set includes sufficient AFLP loci (> 150) to effectively saturate the SGS estimate, the total number of SSR loci used (six) may still be too few to provide an accurate estimate of SGS in this species. As previous studies of SGS using SSR and allozyme markers have generally been based on similarly low numbers of loci to those used for SSRs here, we conclude that spatial genetic structure in wind-pollinated trees may have been underestimated previously.

Acknowledgements

We are extremely grateful to J. Garcia and P. Arus (IRTA, Cabriels, Barcelona) for access to their genetic analyser and to J.A. Martínez-Izquierdo for providing laboratory facilities. We thank J.M. Hunt for assistance in the field and E. Ramallo, L. Rico and J. Quilis for help in the laboratory. A.S.J. received a Marie Curie EIF fellowship (Contract MC-MEIF-CT-2003-501475, FOREST RISE) while conducting this research. Additional support was provided by the European Union (Contract 506675, ALARM), the Catalan Government (grant SGR2005-00312), the Spanish Government (grants REN2003-04871 and CGL2004-01402/BOS) and Fundación BBVA. We are grateful to Jake Gratten for discussion and comments on this work and to three anonymous reviewers for helpful comments on a previous version of this manuscript.

References

- Asuka Y, Tomaru N, Nisimura N, Tsumura Y, Yamamoto S (2004) Heterogeneous genetic structure in a *Fagus crenata* population in an old-growth beech forest revealed by microsatellite markers. *Molecular Ecology*, **13**, 1241–1250.
- Bacilieri R, Labbe T, Kremer A (1994) Intraspecific genetic structure in a mixed population of *Quercus petraea* (Matt) Liebl and *Quercus robur* L. *Heredity*, **73**, 130–141.
- Cavers S, Degen B, Caron H *et al.* (2005) Optimal sampling strategy for estimation of spatial genetic structure in tree populations. *Heredity*, **95**, 281–289.
- Comps B, Gómory D, Letouzey J, Thiébaud B, Petit RJ (2001) Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics*, **157**, 389–397.
- Cottrell JE, Munro RC, Tabbener HE *et al.* (2003) Comparison of fine-scale genetic structure using nuclear microsatellites within two British oakwoods differing in population history. *Forest Ecology and Management*, **176**, 287–303.
- Epperson BK (1989) Spatial patterns of genetic variation within plant populations. In: *Plant Population Genetics, Breeding and Genetic Resources* (eds Brown AHD, Clegg MT, Kahler AL, Weir BS), pp. 229–253. Sinauer Associates, Sunderland, Massachusetts.
- Escudero A, Iriondo JM, Torres ME (2003) Spatial analysis of genetic diversity as a tool for plant conservation. *Biological Conservation*, **113**, 351–365.
- Gallois A, Burrus M, Brown M (1999) Evaluation of the nuclear DNA content and GC percent in four varieties of *Fagus sylvatica* L. *Annals of Forest Science*, **56**, 615–618.
- Gaudeul M, Till-Bottraud I, Barjon F, Manel S (2004) Genetic diversity and differentiation in *Eryngium alpinum* L. (Apiaceae): comparison of AFLP and microsatellite markers. *Heredity*, **92**, 508–518.
- Goudet J (2001) *FSTAT: A Program to Estimate and Test Gene Diversities and Fixation Indices*, version 2.9.3. Available from www.unil.ch/izea/software/fstat.html.
- Gregorius H-R, Kownatzki D (2005) Spatiogenetic characteristics of beech stands with different degrees of autochthony. *BMC Ecology*, **5**, DOI: 10.1186/1472-6785-5-8 available from www.biomedcentral.com/1472-6785/5/8.
- Hardesty BD, Dick DW, Kremer A, Hubbell S, Bermingham E (2005) Spatial genetic structure of *Simarouba amara* Aubl. (Simaroubaceae), a dioecious, animal-dispersed Neotropical tree, on Barro Colorado Island, Panama. *Heredity*, **95**, 290–297.
- Hardy OJ (2003) Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant markers. *Molecular Ecology*, **12**, 1577–1588.
- Hardy OJ, Vekemans X (2002) SPAGEBI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618.
- 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 JF, Palada M, Hardy OJ (2003) Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. *Molecular Ecology*, **12**, 2483–2495.
- Ishida TA, Kimura MT (2003) Assessment of within-population genetic structure in *Quercus crispula* and *Q. dentata* by amplified fragment length polymorphism analysis. *Ecological Research*, **18**, 619–623.
- Jump AS, Peñuelas J (2006) Genetic effects of chronic habitat fragmentation in a wind-pollinated tree. *Proceedings of the National Academy of Sciences, USA*, **103**, 8096–8100.
- Jump AS, Hunt JM, Martínez-Izquierdo JA, Peñuelas J (2006) Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. *Molecular Ecology*, **15**, 3469–3480.
- Krauss SL (2000) Accurate gene diversity estimates from amplified fragment length polymorphism (AFLP) markers. *Molecular Ecology*, **9**, 1241–1245.
- Leonardi S, Menozzi P (1996) Spatial structure of genetic variability in natural stands of *Fagus sylvatica* L. (beech) in Italy. *Heredity*, **77**, 359–368.
- 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.
- Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics*, **15**, 65–95.
- Maguire TL, Peakall R, Saenger P (2002) Comparative analysis of genetic diversity in the mangrove species *Avicennia marina* (Forsk.) Vierh. (Avicenniaceae) detected by AFLPs and SSRs. *Theoretical and Applied Genetics*, **104**, 388–398.
- Mariette S, Cottrell J, Csaikl UM *et al.* (2002) Comparison of levels of genetic diversity detected with AFLP and microsatellite markers within and among mixed *Q. petraea* (MATT.) LIEBL. and *Q. robur* L. stands. *Silvae Genetica*, **51**, 72–79.

- Marinoni D, Akkak A, Bounous G, Edwards KJ, Botta R (2003) Development and characterization of microsatellite markers in *Castanea sativa* (Mill.). *Molecular Breeding*, **11**, 127–136.
- Martin W, Herrmann RG (1998) Gene transfer from organelles to the nucleus: how much, what happens, and why? *Plant Physiology*, **118**, 9–17.
- Merzeau D, Comps B, Thiebaut B, Cuguen J, Letouzey J (1994) Genetic structure of natural stands of *Fagus sylvatica* L. (beech). *Heredity*, **72**, 269–277.
- Nilsson SG (1985) Ecological and evolutionary interactions between reproduction of beech *Fagus sylvatica* and seed-eating animals. *Oikos*, **44**, 157–164.
- Nilsson SG, Wastljung U (1987) Seed predation and cross-pollination in mast-seeding beech (*Fagus sylvatica*) patches. *Ecology*, **68**, 260–265.
- Ohta N, Sato N, Kuroiwa T (1998) Structure and organization of the mitochondrial genome of the unicellular red algae *Cyanidioschyzon merolae* deduced from the complete nucleotide sequence. *Nucleic Acids Research*, **26**, 5190–5198.
- Pastorelli R, Smulders MJM, Van't Westende WPC *et al.* (2003) Characterization of microsatellite markers in *Fagus sylvatica* L. and *Fagus orientalis* Lipsky. *Molecular Ecology Notes*, **3**, 76–78.
- Peñuelas J, Boada M (2003) A global change-induced biome shift in the Montseny Mountains (NE Spain). *Global Change Biology*, **9**, 131–140.
- Scalfi M, Piovani P, Piotti A, Leonardi S, Menozzi P (2005) Effects of habitat fragmentation on genetic structure of beech populations in central Italy. In: *The role of biotechnology for the characterisation and conservation of crop, forestry, animal and fishery genetic resources*, FAO electronic forum on biotechnology in food and agriculture, Turin, Italy. <http://www.fao.org/biotech/torino05.htm>.
- Shapcott A (1995) The spatial genetic structure in natural populations of the Australian temperate rainforest tree *Atherosperma moschatum* (Labill.) (Monimiaceae). *Heredity*, **74**, 28–38.
- Streiff R, Labbe T, Bacilieri R *et al.* (1998) Within-population genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. *Molecular Ecology*, **7**, 317–328.
- Takahashi M, Mukouda M, Kono K (2000) Differences in genetic structure between two Japanese beech (*Fagus crenata* Blume) stands. *Heredity*, **84**, 103–115.
- Tanaka K, Tsumura Y, Nakamura T (1999) Development and polymorphism of microsatellite markers for *Fagus crenata* and the closely related species, *F. japonica*. *Theoretical and Applied Genetics*, **99**, 11–15.
- Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, **13**, 921–935.
- Vornam B, Decarli N, Gailing O (2004) Spatial distribution of genetic variation in a natural beech stand (*Fagus sylvatica* L.) based on microsatellite markers. *Conservation Genetics*, **5**, 561–570.
- Vos P, Hogers R, Bleeker M *et al.* (1995) AFLP — a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.
- Weir B, Cockerham C (1984) Estimating *F*-Statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- Yeh FC, Boyle TJB (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany*, **129**, 157.

The authors share an interest in characterising genetic and ecological patterns and processes over a range of spatial and temporal scales. Their collaborative research is currently focussed on investigating the impacts of environmental change on genetic diversity and population genetic structure in *Fagus sylvatica*.
