

# Genetic diversity within and among *Pinus pinaster* populations: comparison between AFLP and microsatellite markers

STÉPHANIE MARIETTE\*, DAVID CHAGNÉ, CÉLINE LÉZIER, PATRICK PASTUSZKA, ANNIE RAFFIN, CHRISTOPHE PLOMION & ANTOINE KREMER

INRA, Equipe de Génétique et Amélioration des Arbres Forestiers, BP45, F-33610 Cestas, France

Twenty-three populations of *Pinus pinaster* (13 Aquitaine populations and 10 Corsican populations) were analysed at three microsatellite loci and 122 AFLP loci. The aims of the study were: (i) to compare levels of within-population and among-population diversity assessed with both kinds of markers; (ii) to compare Aquitaine and Corsican provenances of *P. pinaster*; and (iii) to know if both markers gave the same information for conservation purposes. Classical population genetics statistics were estimated and the ranking of populations obtained using different markers and/or parameters were compared by computing Spearman's rank correlations. Even though microsatellites showed a higher within-population diversity, they showed the same level of differentiation as AFLP markers. Moreover, both markers also showed a higher genetic diversity in the Aquitaine provenance and a higher differentiation among Corsican populations. AFLPs and microsatellites gave different population diversity rankings. Consequently, the results do not support the potential population identification within each provenance for conservation purposes.

**Keywords:** AFLP, genetic diversity, genetic differentiation, microsatellite, *Pinus pinaster*.

## Introduction

Different marker systems are currently available for monitoring genetic diversity in natural populations. Examples of markers with contrasting properties are AFLP markers (Vos *et al.*, 1995) and microsatellites (Tautz, 1989). First, AFLPs are considered to be dominant and biallelic, whereas microsatellites are codominant and usually reveal a high number of alleles. Secondly, the AFLP technique provides many markers randomly spread in the genome whereas, in a diversity study, the number of analysed microsatellites is usually low. Although mutation rates may be different for different markers and may be considered 'marker specific', other evolutionary factors such as migration or drift affect the whole genome regardless of the genomic regions that are monitored by different marker systems. The comparative analysis between different markers is therefore intended to reveal the balance between different evolutionary forces contributing to genetic diversity. In this paper, we compare the levels of diversity within and among *Pinus pinaster* Ait. popula-

tions estimated by both types of markers (AFLP and microsatellites).

The natural range of *P. pinaster* is restricted to the western part of the Mediterranean basin, including continental France (mainly south-west and south-east), Corsica, Spain, Portugal, continental Italy (Liguria and Tuscany), Italian islands (Sardinia and Pantelleria), Morocco and Tunisia. Previous genetic diversity studies have been performed using terpene, protein and allozyme markers (Baradat & Marpeau-Bezard, 1988; Baradat *et al.*, 1991; Bahrman *et al.*, 1992; Petit *et al.*, 1995; Bahrman *et al.*, 1997). The results of these studies indicated a subdivision of the species into three groups: Atlantic (including south-west France, Portugal and Galicia in Spain), peri-Mediterranean (including Andalusia in Spain, Corsica, Maures in south-east France and Liguria in Italy) and North African (Moroccan Atlas, Algeria and Tunisia). They also suggested that the species differentiated into regional "races" after the last glaciation, beginning in the south-west of the Iberian Peninsula. Salvador *et al.* (2000) studied the genetic variability with 18 isozyme loci and concluded that the southern Iberian range could be the putative recolonization centre of the species after the last glaciation.

\*Correspondence. E-mail: stephanie.mariette@pierroton.inra.fr

Chloroplast microsatellites were recently used in *P. pinaster* for studying genetic diversity and colonization processes of the species (Vendramin *et al.*, 1998; Ribeiro *et al.*, 2000). As a whole, few studies focused on the fine scale description of genetic variability of *P. pinaster* within regional 'races'.

Maritime pine is a widely afforested species in France. An intensive breeding programme is conducted with breeding populations originating from the Aquitaine and Corsican provenances. Natural resources of the species might be threatened in France. In south-west France, the artificial selection in the improvement programme may cause a reduction of the genetic variability. In the south-east, the threat may be due to the fires and to the bast scale *Matsucoccus feytaudi* Duc., which is a specific pest of *P. pinaster* (Burban *et al.*, 1999). After proliferating in south-east France, the pest recently appeared in Corsica (Jactel *et al.*, 1996) where it is spreading and producing extensive damage (Jactel *et al.*, 1998).

In this paper, 13 Aquitaine populations and 10 Corsican populations of *P. pinaster* were screened with two different types of molecular markers: three nuclear microsatellite loci and 122 AFLP loci. The first objective of the study was to compare the levels of within- and among-population diversity assessed by both types of markers. The second objective was to analyse and compare the genetic diversity of *P. pinaster* populations from the Aquitaine provenance and from the Corsican provenance. Finally, we intended to draw recommendations for conservation purposes based on the comparative analysis of genetic diversity with different markers.

## Materials and methods

### Plant material and DNA extraction

We sampled 13 populations in Aquitaine (south-west France). Populations were selected on the basis of historical records about their native origin [French National Forest Office (ONF), personal communication]. From palynological data, Paquereau (1964) indicates that *P. pinaster* appeared in Aquitaine about 8000 years ago and spread along the coast, where the species constituted some monospecific pine forests during the last 3000 years. The populations used in this study most likely originated from original stands after successive natural regenerations. Populations were sampled to represent the natural genetic variation of this species in south-west France.

The Corsican set of populations was selected among 30 previously collected populations (Durel & Bahrman, 1995). Ten of those populations, regularly distributed along the island, were chosen to cover the range of altitudes, from 10 m to 1105 m.

For the Aquitaine populations, needles were collected from about 30 young trees (3–5-years-old) randomly distributed over an area from 10 to 30 hectares for each population. In the Corsican populations, one seedling was sampled for each of about 30 open pollinated progenies per population. The mother trees of the open pollinated progenies were distributed on approximately 50 hectares. Genomic DNA was extracted from needles as described by Doyle & Doyle (1990).

### Microsatellite and AFLP markers

The 23 populations were screened with three nuclear microsatellites (*FRPP91*, *FRPP94* and *ITPH4516*) and with two AFLP primer–enzyme combinations. Microsatellite amplifications were performed as described by Mariette *et al.* (2001) and AFLP markers were obtained following the protocol by Costa *et al.* (2000). Two primer–enzyme combinations, *EcoRI* + *ACC/MseI* + *CCTG* and *EcoRI* + *ACG/MseI* + *CCCA*, coded *PEC 1* and *PEC 2*, respectively, were used. They were selected according to the number of amplified loci (respectively 110 and 81) and the number of polymorphic loci (respectively 29 and 28) detected in a mapped population (Costa *et al.*, 2000). Microsatellite and AFLP loci were scored with the *RFLP-SCAN* version 3.0 (Scanalytics) software.

### Statistical analyses

*Data analysis of microsatellite markers.* For each locus, the allelic richness ( $A$ ), the allelic frequencies, the observed heterozygosity ( $H_O$ ), the effective number of alleles ( $A_E = 1/(1 - H_E)$ ) and the fixation index ( $F_{IS}$ ) were calculated as described by Brown & Weir (1983). The analysis of diversity was also performed using Nei's (1987) single-locus estimators: the within-population gene diversity ( $H_i$  within population  $i$  and  $H_S$  mean value of  $H_i$  over several populations), the total gene diversity ( $H_T$ ) and the coefficient of gene differentiation among populations ( $G_{ST}$ ). These parameters were computed by using the *DIPLOIDE* program (Antoine Kremer, Equipe de Génétique et Amélioration des Arbres Forestiers, Cestas, France).

*Data analysis of AFLP markers.* To analyse AFLP markers, we assumed that each AFLP amplification product, regardless of its relative intensity, corresponded to a dominant allele at a unique locus. Polymorphic amplified loci were scored as '1' for the presence and '0' for the absence of a locus. Only loci with clearly amplified bands were used for data analysis.

The analysis of genetic diversity with AFLP fragments was done after estimating the frequencies of the

alleles responsible for the presence and absence of fragments. If  $p$  and  $q$  are the alleles responsible for the presence and absence of an AFLP fragment, and if  $Q$  is the frequency of trees that do not exhibit the fragment, then  $Q = q^2(1 - F_{IS}) + qF_{IS}$ , where  $F_{IS}$  is the fixation index. Assuming that the true value of  $F_{IS}$  is known, an asymptotically unbiased estimate of  $q$  is obtained by the use of a second order Taylor expression:

$$\hat{q} = \frac{F_{IS}}{2(1 - F_{IS})} + \frac{\Delta^{\frac{1}{2}}}{2(1 - F_{IS})} \left[ \frac{Q(1 - Q)}{N} \right] \quad (1)$$

with  $\Delta = F_{IS}^2 + 4(1 - F_{IS})Q$  and  $N$  the number of trees sampled per population.

The analyses were performed using the  $F_{IS}$  value that was estimated based on the average value from the three microsatellites. We estimated the diversity over all loci and we also restricted the analysis to loci that showed an observed frequency smaller than  $1 - (3/N)$ , where  $N$  is the sample size, as recommended by Lynch & Milligan (1994). These authors showed that the bias introduced on the estimation of  $q$  due to small sample size was substantial when the null allele was rare.  $H_S$ ,  $H_T$  and  $G_{ST}$  were computed by using the allelic frequencies as estimated by formula (1).

Analyses were performed by using the HAPDOM program (Antoine Kremer, Equipe de Génétique et Amélioration des Arbres Forestiers, Cestas, France).

For both types of marker, diversity parameters were estimated for each population and provenance (Corsica and Aquitaine). Mean values were estimated over all 23 populations, over all the Aquitaine populations and over all the Corsican populations.

Levels of differentiation ( $G_{ST}$ ) were calculated among all 23 populations, among the Aquitaine populations, among the Corsican populations and between the Aquitaine provenance and the Corsican one.

*Comparison of diversity and differentiation between the Corsican and Aquitaine provenance.* Resampling methods were used to estimate the differences between the two provenances [Aquitaine ( $A$ ) and Corsica ( $C$ )] for the parameters  $A_E$ ,  $H_O$ ,  $H_S$ ,  $H_T$ ,  $F_{IS}$  and  $G_{ST}$ . For example, in the case of  $H_S$ , bootstrapping was used, where all the individuals were resampled with replacement within each population of each provenance  $A$  and  $C$ , and  $H_S$  was computed in each provenance as well as the difference of  $H_S$  computed for each provenance ( $H_S^{Aq} - H_S^{Cor}$ ). Resampling was repeated 1000 times. The distribution of ( $H_S^{Aq} - H_S^{Cor}$ ) was then compared with the null hypothesis ( $H_S^{Aq} - H_S^{Cor} = 0$ ). Sampling variances were obtained by bootstrapping over individ-

uals in populations for  $A_E$ ,  $H_O$ ,  $H_S$ ,  $H_T$  and  $F_{IS}$  and by bootstrapping over populations for  $G_{ST}$ . The sample sizes in the Aquitaine and in the Corsican provenances were uneven, therefore the rarefaction method as described by Petit *et al.* (1998) was used for estimating the unbiased allelic richness at each locus, using 50, 100 and 200 as sample sizes for each provenance.

*Comparison of allele frequencies among populations within provenances.* The geographical pattern of the studied populations was analysed by using the principal component analysis (PCA) and Mantel tests. PCA was performed using the allele frequencies for each microsatellite and the frequencies of loci obtained in the genotypic AFLP analysis. In the case of microsatellites only allele frequencies of the five most frequent alleles were used. The ACP program (Philippe Baradat, Equipe de Génétique et Amélioration des Arbres Forestiers, Cestas, France) was used to compute the PCA analysis.

Mantel tests were performed for comparing the pairwise population geographical distance (in kilometres) with the pairwise population genetic distance matrix (Nei, 1978). The strength of the relationship between both matrices was measured using the Mantel matrix-correspondence test (Mantel, 1967; Sokal, 1979). The null hypothesis refers to the absence of association between the elements of the pairs of matrices. A normalized Z-test was performed in which the observed value after 999 permutations should be significantly larger than that expected by chance in order for an association to be accepted as valid. The program used to compute Mantel tests was TPGA version 1.3 (Mark P. Miller, Northern Arizona University, U.S.A.).

*Intra- and inter-marker comparisons.* Diversity assessed at different markers results from either 'marker specific' factors (such as mutation) or 'marker non-specific' factors (as for example drift or migration). Hence one would expect that values of diversity of a given population as assessed with different markers should be different. However, if two populations went through different evolutionary histories involving drift and migration, then they would differ for their level of diversity in a same manner for two different kinds of markers. In other words, one expects congruent rankings of the populations between markers. As a result, the comparative analysis of diversity between markers is not based on the comparison of levels of diversity within a population, but on the comparison of the ranking of different populations by correlation analysis.

For each type of marker (AFLP or microsatellites),  $A$ ,  $A_E$ ,  $H_O$  and  $H_S$  were estimated for each population (and provenance). Estimates of these diversity statistics were then compared between populations by computing

Spearman's rank coefficient correlation (Sokal & Rohlf, 1995). A significant correlation would indicate congruence between different factors and the impact of 'marker non-specific' factors on levels of diversity. Conversely, lack of correlation is likely to reveal that 'marker specific' factors are sufficiently different between populations to generate different levels of diversity.

## Results

### Within-population diversity

**Microsatellites.** Considering all the 23 populations, the three microsatellites showed a high level of polymorphism with an average allelic richness  $A$  of 27.33, ranging from 23 to 31 alleles.  $H_O$  was 0.645, ranging from 0.584 to 0.690, and  $H_T$  was 0.832, ranging from 0.742 to 0.903 (Table 1).

**AFLP markers.** For the *PEC 1* and the *PEC 2*, 56 and 66 amplified loci were analysed, respectively. The number of polymorphic loci in at least one population was 23 (41%) for *PEC 1* and 36 (55%) for *PEC 2*. Thus, on a whole, 122 loci were analysed and 48% of them were polymorphic.

Using the fixation index  $F_{IS}$  obtained from microsatellites to estimate allelic frequencies (eqn 1), the total diversity was 0.161 for *PEC 1*, 0.194 for *PEC 2* and

0.179 for all the studied loci. Using only the 36 polymorphic markers, as suggested by reducing the number of loci from Lynch & Milligan (1994), the levels of diversity were increased [0.321, 0.297 and 0.305 for *PEC 1*, *PEC 2* and both *PECs*, respectively (Table 2)].

### Comparison of diversity in Aquitaine and Corsica

At the single population level, Aquitaine populations showed higher within-population diversity levels than Corsican populations. This result was significant at the 5% level both for microsatellites (Table 1) and AFLP markers (Table 2). However, locus *FRPP94* showed a significantly higher effective number of alleles and within-population diversity for Corsican populations (Table 1). The fixation index within Corsican populations was significantly higher for locus *FRPP91* and over all loci (Table 1).

At the provenance level, the Aquitaine provenance showed significantly higher levels of diversity than the Corsican one, both for microsatellite and AFLP markers (Table 1 and Table 2). However, loci *FRPP94* and *PEC 1* showed significant higher diversity parameters for the Corsican provenance.

With 50, 100 and 200 as sample sizes for each provenance, the rarefaction method gave a higher number of alleles in Corsica for *FRPP91* (27.4 in Corsica and 21.5 in Aquitaine for a sample size of 200)

**Table 1** Genetic diversity statistics for microsatellite loci in Corsican and Aquitaine populations of *Pinus pinaster*

|                 |           | A                  | $H_O$ | $H_T$    |          |        |        |        |                     |  |
|-----------------|-----------|--------------------|-------|----------|----------|--------|--------|--------|---------------------|--|
| <i>FRPP91</i>   |           | 31                 | 0.661 | 0.849    |          |        |        |        |                     |  |
| <i>FRPP94</i>   |           | 23                 | 0.584 | 0.742    |          |        |        |        |                     |  |
| <i>ITPH4516</i> |           | 28                 | 0.690 | 0.903    |          |        |        |        |                     |  |
| Three loci      |           | 27.33              | 0.645 | 0.832    |          |        |        |        |                     |  |
|                 |           | $A_P$              | $A_T$ | $A_{EP}$ | $A_{ET}$ | $H_O$  | $H_S$  | $H_T$  | $F_{IS}$            |  |
| <i>FRPP91</i>   | Aquitaine | 11.31*             | 25    | 5.43*    | 5.79*    | 0.805* | 0.827* | 0.829* | 0.026*              |  |
|                 | Corsica   | 8.60               | 30    | 2.14     | 2.12     | 0.434  | 0.523  | 0.531  | 0.170               |  |
| <i>FRPP94</i>   | Aquitaine | 6.62 <sup>ns</sup> | 14    | 2.69*    | 2.82*    | 0.534* | 0.641* | 0.647* | 0.167 <sup>ns</sup> |  |
|                 | Corsica   | 7.60               | 20    | 4.25     | 4.86     | 0.661  | 0.772  | 0.796  | 0.144               |  |
| <i>ITPH4516</i> | Aquitaine | 12.69*             | 25    | 7.29*    | 8.22*    | 0.741* | 0.871* | 0.883* | 0.149*              |  |
|                 | Corsica   | 8.70               | 20    | 4.37     | 5.76     | 0.605  | 0.765  | 0.829  | 0.210               |  |
| Three loci      | Aquitaine | 10.21*             | 21.33 | 4.24*    | 4.62*    | 0.693* | 0.780* | 0.786* | 0.114*              |  |
|                 | Corsica   | 8.3                | 23.33 | 3.07     | 3.52     | 0.567  | 0.687  | 0.718  | 0.175               |  |

$A$ , allelic richness;  $A_P$ , allelic richness at the population level;  $A_T$ , allelic richness at the provenance level;  $A_{EP}$ , effective number of alleles at the population level;  $A_{ET}$ , effective number of alleles at the provenance level;  $H_O$ , observed heterozygosity;  $H_S$ , mean within-population diversity;  $H_T$ , total gene diversity;  $F_{IS}$ , fixation index within populations.

\* Significantly different at the 5% level between the Aquitaine and the Corsican provenance.

<sup>ns</sup> Values estimated for Aquitaine and Corsica are not significantly different at the 5% level.

**Table 2** Genetic diversity statistics for AFLP loci in Corsican and Aquitaine populations of *Pinus pinaster*

|                 |           | Total diversity†    |                     | Total diversity‡ |                     |
|-----------------|-----------|---------------------|---------------------|------------------|---------------------|
| <i>PEC 1</i>    |           | 0.161               |                     | 0.321            |                     |
| <i>PEC 2</i>    |           | 0.194               |                     | 0.297            |                     |
| Two <i>PECs</i> |           | 0.179               |                     | 0.305            |                     |
|                 |           | $H_S$ †             | $H_T$ †             | $H_S$ ‡          | $H_T$ ‡             |
| <i>PEC 1</i>    | Aquitaine | 0.140 <sup>ns</sup> | 0.147*              | 0.326*           | 0.374 <sup>ns</sup> |
|                 | Corsica   | 0.149               | 0.156               | 0.286            | 0.375               |
| <i>PEC 2</i>    | Aquitaine | 0.180*              | 0.183*              | 0.303*           | 0.353*              |
|                 | Corsica   | 0.168               | 0.178               | 0.240            | 0.318               |
| Two <i>PECs</i> | Aquitaine | 0.162 <sup>ns</sup> | 0.166 <sup>ns</sup> | 0.310*           | 0.361*              |
|                 | Corsica   | 0.159               | 0.168               | 0.255            | 0.340               |

$H_S$ , mean within-population diversity;  $H_T$ , gene diversity at the provenance level.

† The analysis is genotypic.

‡ The analysis is genotypic and restricted to loci whose observed frequency is less than  $1 - (3/N)$  with  $N$  the population sample size, following Lynch & Milligan (1994).

\* Significantly different at the 5% level between the Aquitaine and the Corsican provenance.

<sup>ns</sup> Values estimated for Aquitaine and Corsica are not significantly different at the 5% level.

and for *FRPP94* (18.3 in Corsica and 11.7 in Aquitaine). In contrast, *ITPH4516* showed more alleles in the Aquitaine provenance (22.4 in Aquitaine and 18.8 in Corsica). The Corsican provenance showed a higher mean allelic richness ( $A = 23.33$ ) than the Aquitaine provenance [ $A = 21.33$  (Table 1)].

Considering the distribution of allelic frequencies of microsatellites, the provenances did not share the same common alleles for *FRPP91* and for *ITPH4516* (Fig. 1). For *FRPP91*, one allele showed divergent frequencies in Corsican and Aquitaine provenances (0.680 in Corsica and 0.004 in Aquitaine). For all loci, each provenance exhibited private alleles: for *FRPP91*, one rare allele was specific to Aquitaine and six were specific to Corsica; for *FRPP94*, three alleles were found in Aquitaine against nine in Corsica and for *ITPH4516*, eight alleles were found in Aquitaine vs. three in Corsica (Fig. 1). Corsica showed a higher number of rare alleles (18 against 12 in Aquitaine).

AFLP markers also showed private alleles. Considering both *PECs*, seven polymorphic loci exhibited bands in Aquitaine which were absent from Corsica. No specific allele was found in Corsica.

#### Subdivision of genetic diversity among populations and between provenances

The average differentiation among all populations was 11.1% for the microsatellite markers (Table 3), varying from 6% for *FRPP94* to 18.2% for *FRPP91*. For AFLP markers the level of differentiation was 10.2% for the

genotypic analysis, when all 122 loci were considered, and reduced to 6.1%, following Lynch & Milligan (1994). The differentiation levels estimated with the two primer–enzyme combinations were found to be significantly different. The *PEC 2* showed a significantly higher differentiation than the *PEC 1* (data not shown).

The genetic differentiation assessed with microsatellite markers was significantly higher than the differentiation assessed with AFLP markers only when we used the Lynch & Milligan (1994) method of analysis.

The differentiation among Corsican populations was higher than the differentiation among Aquitaine populations with both microsatellite and AFLP markers (Table 3). The difference was significant for microsatellites and for the analysis of AFLP when the 122 loci were considered.

The differentiation between provenances (9.2% for microsatellites and 4.9% for AFLP markers) represented more than half of the total differentiation (Table 3).

#### Comparison of levels of diversity assessed with different measures ( $A$ and $H_i$ )

Allelic richness and expected heterozygosity estimated with microsatellite loci within each population exhibited similar rankings of populations within each provenance. The rankings of populations given by  $A$  calculated over all microsatellites were positively and significantly (5% level) correlated with rankings given by  $H_i$ , both for Aquitaine populations ( $r_S = 0.765$ ,  $P = 0.018$ ) and for Corsican populations ( $r_S = 0.710$ ,  $P = 0.035$ ). In the

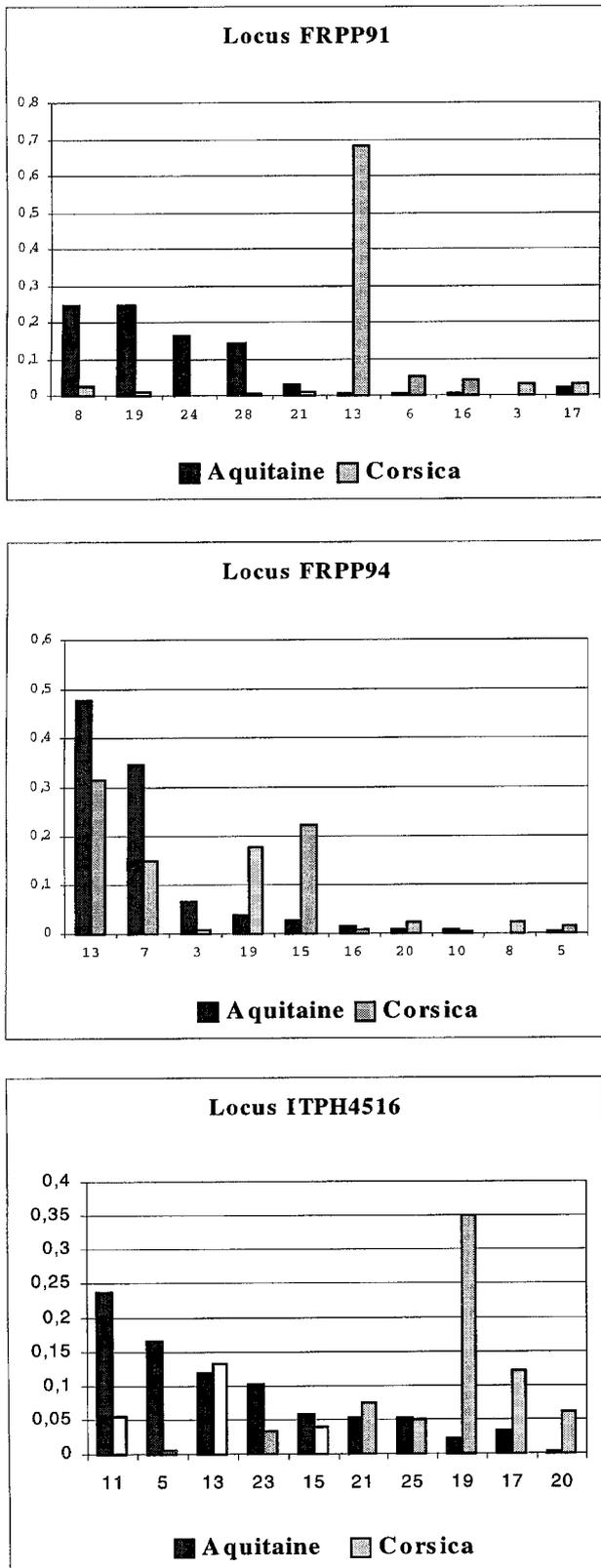


Fig. 1 Microsatellite allelic frequencies for Corsican and Aquitaine populations of *Pinus pinaster* used in the paper.

Aquitaine populations, the correlation was also significantly positive for each microsatellite.

For Corsican populations, pairwise comparison of diversity statistics between loci did not always indicate congruent results. For *FRPP91*, the ranking given by  $H_O$  was positively and significantly correlated with the ranking given by  $H_O$  on *FRPP94*. The ranking given by  $A$  on *FRPP91* was also positively and significantly correlated with the ranking given by  $H_O$  on *FRPP94* and by  $H_O$  on *ITPH4516*. The same analysis was performed for the Aquitaine provenance and no correlation was ever found.

#### Comparison of levels of diversity assessed with different markers

For each type of analysis, the rankings of populations given by *PEC 1* and *PEC 2* were not significantly correlated for both Aquitaine and Corsican populations (Table 4). Rankings given by *PEC 1* and *PEC 2* were checked separately but were never significantly correlated (Table 5); they were positively and significantly correlated, as expected, with rankings given by both *PECs*, in general.

When comparing  $H_i$  values given by microsatellites and AFLP markers, no positive and significant correlation was found (Table 4).

#### Geographical pattern

The two first axes of the principal component analysis separated Aquitaine populations from Corsican ones (Fig. 2).

Mantel tests performed on Aquitaine populations did not detect any geographical pattern with AFLP markers or with microsatellites. A significant correlation ( $r=0.360$ ,  $P=0.018$ ) between the genetic distances matrix and the geographical distances matrix was found for Corsican populations when microsatellites were used.

#### Discussion

Maritime pine is a highly polymorphic species as shown by the values of diversity measures and comparisons with other species. For example, in this study, the average value of  $H_O$  for microsatellites was 0.645 and the average value of  $A$  was 27.33. These values are similar to the values found in other pine species (Smith & Devey, 1994; Echt *et al.*, 1996). The ultimate aim of this study was to derive conservation decisions based on a comparative survey of genetic diversity conducted with different markers. The comparative analysis was intended to reveal the differential impact of 'marker

**Table 3** Levels of differentiation (with standard deviations) among Aquitaine and Corsican populations of *Pinus pinaster*

| $G_{ST}^T$            | $G_{ST}^{Aq \text{ vs. Cor}}$ | $G_{ST}^{Aq}$ | $G_{ST}^{Cor}$ |               |
|-----------------------|-------------------------------|---------------|----------------|---------------|
| Three microsatellites | 0.111 (0.012)                 | 0.092 (-)     | 0.009 (0.004)  | 0.044 (0.008) |
| AFLP (two PECs)†      | 0.102 (0.004)                 | 0.066 (-)     | 0.032 (0.005)  | 0.054 (0.008) |
| AFLP (two PECs)‡      | 0.061 (0.003)                 | 0.049 (-)     | 0.017 (0.004)  | 0.027 (0.005) |

$G_{ST}^T$  Nei's  $G_{ST}$  among all 23 populations;  $G_{ST}^{Aq \text{ vs. Cor}}$ , Nei's  $G_{ST}$  between Aquitaine provenance and Corsican provenance;  $G_{ST}^{Aq}$ , Nei's  $G_{ST}$  among 13 Aquitaine populations;  $G_{ST}^{Cor}$ , Nei's  $G_{ST}$  among 10 Corsican populations.

† The analysis is genotypic.

‡ The analysis is genotypic and restricted to loci whose observed frequency is less than  $1 - (3/N)$  with  $N$  the population sample size, following Lynch & Milligan (1994).

**Table 4** Spearman's rank correlation analysis among diversity values obtained with different markers in *Pinus pinaster*

|                         | $H_i^\dagger$ (2 PECs) | $H_i^\dagger$ (PEC 1) | $H_i^\dagger$ (PEC 2) | $H_i^\ddagger$ (2 PECs) | $H_i^\ddagger$ (PEC 1) | $H_i$ (microsatellites) |
|-------------------------|------------------------|-----------------------|-----------------------|-------------------------|------------------------|-------------------------|
| Aquitaine               |                        |                       |                       |                         |                        |                         |
| $H_i^\dagger$ (2 PECs)  |                        |                       |                       |                         |                        | -0.220 (0.441)          |
| $H_i^\dagger$ (PEC 1)   | 0.830* (0.004)         |                       |                       |                         |                        | 0.016 (0.962)           |
| $H_i^\dagger$ (PEC 2)   | 0.703* (0.015)         | 0.258 (0.376)         |                       |                         |                        | -0.445 (0.121)          |
| $H_i^\ddagger$ (2 PECs) | 0.955* (0.001)         | 0.839* (0.004)        | 0.569* (0.050)        |                         |                        | -0.272 (0.341)          |
| $H_i^\ddagger$ (PEC 1)  | 0.582* (0.045)         | 0.764* (0.080)        | -0.011 (0.962)        | 0.713* (0.014)          |                        | -0.060 (0.827)          |
| $H_i^\ddagger$ (PEC 2)  | 0.874* (0.003)         | 0.588* (0.043)        | 0.874* (0.003)        | 0.831* (0.004)          | 0.319 (0.274)          | -0.462 (0.108)          |
| Corsica                 |                        |                       |                       |                         |                        |                         |
| $H_i^\dagger$ (2 PECs)  |                        |                       |                       |                         |                        | -0.370 (0.260)          |
| $H_i^\dagger$ (PEC 1)   | 0.442 (0.191)          |                       |                       |                         |                        | -0.127 (0.689)          |
| $H_i^\dagger$ (PEC 2)   | 0.845* (0.012)         | 0.146 (0.675)         |                       |                         |                        | -0.517 (0.117)          |
| $H_i^\ddagger$ (2 PECs) | 0.733* (0.030)         | 0.212 (0.537)         | 0.760* (0.024)        |                         |                        | -0.430 (0.191)          |
| $H_i^\ddagger$ (PEC 1)  | 0.018 (0.971)          | 0.539 (0.110)         | -0.243 (0.455)        | 0.248 (0.467)           |                        | -0.248 (0.445)          |
| $H_i^\ddagger$ (PEC 2)  | 0.733* (0.029)         | 0.067 (0.856)         | 0.863* (0.010)        | 0.867* (0.010)          | -0.212 (0.513)         | 0.248 (0.445)           |

Each Spearman's rank correlation is followed by the associated probability.

$H_i$ : within-population diversity.

† The analysis of AFLP markers is genotypic.

‡ The analysis of AFLP markers is genotypic and restricted to loci whose observed frequency is less than  $1 - (3/N)$  with  $N$  the population sample size, following Lynch & Milligan (1994).

\* Significant positive correlation at the 5% level.

specific' factors vs. 'marker non-specific' factors on diversity. In addition, as conservation should be based on genetic diversity at the whole genome level, the comparative analysis will provide insights on the genomic heterogeneity for diversity and differentiation. Therefore, we will first discuss congruencies and discrepancies between markers, and then consider their consequences on conservation strategies.

**Congruent results of differentiation obtained with microsatellites and AFLPs**

As for diversity, some evolutionary factors such as migration may have an impact on differentiation regardless of the marker, whereas others (such as

mutation) may be considered as 'marker specific'. For most markers, the contribution of mutation to differentiation has generally been ignored because mutation rates of traditional isozyme markers are negligible when compared to migration rates. This may not be the case when comparing microsatellites and AFLP. Because microsatellite markers show higher mutation rates than other molecular markers (Weber & Wong, 1993), they are expected to exhibit lower levels of differentiation. Furthermore, as reported by Estoup *et al.* (1995), homoplasy due to the high mutation rate may increase the underestimation of differentiation. On the other hand, the dominant property of AFLP markers is also responsible for a bias in the estimation of differentiation. Jenczewski *et al.* (1999) showed, both

**Table 5** Comparison of diversity in Corsican and Aquitaine populations of *Pinus pinaster*

| Aquitaine               | Micro-satellites | Aquitaine               | AFLP<br>122 loci | Aquitaine               | AFLP<br>(Lynch & Milligan) |
|-------------------------|------------------|-------------------------|------------------|-------------------------|----------------------------|
| Usagère de Biscarosse   | 0.798 A          | Lège                    | 0.172 A          | Lit-et-Mixe             | 0.331 A                    |
| Hourtin                 | 0.776 A          | Lit-et-Mixe             | 0.171 A          | Lège                    | 0.331 A                    |
| St-Julien-en-Born       | 0.775 A          | St-Julien-en-Born       | 0.171 A B        | St-Julien-en-Born       | 0.330 A B                  |
| Pointe de Grave         | 0.771 A          | Vielle-St-Girons        | 0.168 A B        | Vielle-St-Girons        | 0.328 A B                  |
| Lège                    | 0.765 A          | Ste-Eulalie-en-Born     | 0.167 A B C      | Ste-Eulalie-en-Born     | 0.318 A B                  |
| Mimizan                 | 0.763 A          | Pointe de Grave         | 0.163 A B C D    | Carcans                 | 0.312 A B                  |
| Carcans                 | 0.762 A          | Carcans                 | 0.163 A B C D    | Mimizan                 | 0.312 A B                  |
| Lit-et-Mixe             | 0.759 A          | Mimizan                 | 0.162 A B C D    | Pointe de Grave         | 0.311 A B                  |
| Lacanau                 | 0.758 A          | Hourtin                 | 0.161 A B C D    | Domaniale de Biscarosse | 0.306 A B C                |
| Domaniale de Biscarosse | 0.752 A          | Boulevard des Allemands | 0.158 B C D      | Hourtin                 | 0.299 A B C                |
| Boulevard des Allemands | 0.751 A          | Domaniale de Biscarosse | 0.154 C D E      | Usagère de Biscarosse   | 0.295 B C                  |
| Ste-Eulalie-en-Born     | 0.750 A          | A Usagère de Biscarosse | 0.151 D E        | Boulevard des Allemands | 0.293 C                    |
| Vielle-St-Girons        | 0.744 A          | Lacanau                 | 0.144 E          | Lacanau                 | 0.269 C                    |

| Corsica    | Microsatellites | Corsica    | AFLP 122 loci | Corsica    | AFLP<br>(Lynch & Milligan) |
|------------|-----------------|------------|---------------|------------|----------------------------|
| Vero       | 0.751 A         | Pineto     | 0.168 A       | Ominanda   | 0.288 A                    |
| Tova       | 0.714 A B       | Ominanda   | 0.167 A       | Pineto     | 0.274 A B                  |
| Cagna      | 0.709 A B C     | Cagna      | 0.165 A       | Ventilegne | 0.267 A B C                |
| Pinia      | 0.679 A B C     | Ventilegne | 0.164 A       | Restonica  | 0.264 A B C                |
| Pineto     | 0.679 A B C     | Bonifatu   | 0.159 A B     | Cagna      | 0.260 A B C                |
| Aullene    | 0.662 B C       | Aullene    | 0.158 A B     | Aullene    | 0.250 B C D                |
| Ventilegne | 0.636 B C       | Vero       | 0.158 A B     | Tova       | 0.245 C D                  |
| Restonica  | 0.630 B C       | Restonica  | 0.154 B       | Vero       | 0.244 C D                  |
| Bonifatu   | 0.624 C         | Tova       | 0.153 B       | Bonifatu   | 0.241 C D                  |
| Ominanda   | 0.599 C         | Pinia      | 0.150 B       | Pinia      | 0.220 D                    |

Populations having the same letter do not show a significant difference in their level of diversity.

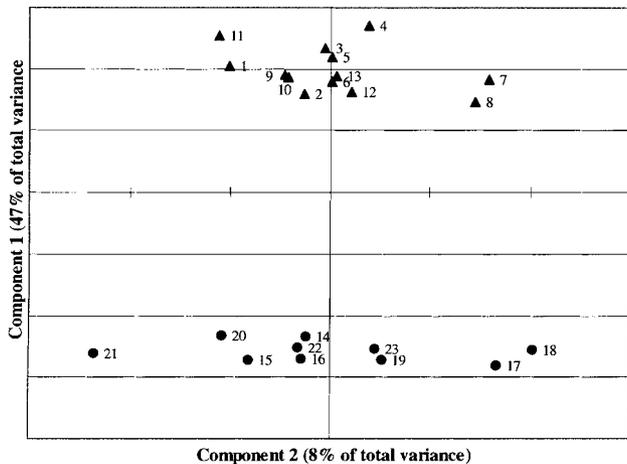


Fig. 2 Principal component analysis on *Pinus pinaster* populations: ● Aquitaine populations; ▲ Corsican populations.

with simulation and experimental results, that  $F_{ST}$  is biased upward when using RAPD phenotypes to estimate allelic frequencies and Krutovskii *et al.* (1999) showed that RAPDs can reasonably estimate population differentiation if sample sizes are larger than about 30 individuals. In our case, sample sizes were around 30 individuals for each population and some loci were fixed for the presence of amplified loci so an overestimation of differentiation for AFLPs cannot be excluded. Despite these expected discrepancies, our results lead to similar conclusions for genetic differentiation assessed with microsatellite and AFLP markers. Although 'absolute values' of differentiation may slightly differ between markers, they lead to similar conclusions (Table 3): (i) when all populations were considered, a high differentiation among populations was found both with microsatellites and AFLP markers; (ii) furthermore, both markers indicate that a predominant part of the total differentiation is due to the subdivision between the Aquitaine and Corsican provenances; and (iii) lastly the two markers reveal that there is more differentiation among the Corsican populations than among the Aquitaine populations. The congruent results obtained with different markers most likely indicate that marker differences (mutation rates, homoplasy, estimation biases) had only a minor impact on differentiation. Conversely, migration and drift may still be the prevailing factors involved. This is illustrated by the contrasting difference between the Aquitaine and Corsican forest conditions. The Aquitaine populations were sampled within a broad continuous forest occupying a flat area of about 1 million hectares. In contrast, the Corsican populations came from an heterogeneous discontinuous forest distributed in mountainous regions. Consequently there is likely to

be more gene flow among stands in Aquitaine than in Corsica.

Interestingly our results (Table 3) could be compared with previous reports based on different markers, but also on different sampling of populations.  $G_{ST}$  values estimated on six populations chosen throughout the range of the species (Petit *et al.*, 1995) were 0.170, 0.161 and 0.139, respectively, for protein, isozyme and terpene loci. Since the sampling area was larger, differentiation was also higher. However this study also indicated congruent results of differentiation across three different marker systems.

#### *Congruent results of diversity obtained with microsatellites and AFLPs*

Among the different evolutionary factors that have an impact on genetic diversity, mutation can be considered as 'marker specific'. However there are other factors such as genetic drift and migration that can be considered as 'non-marker specific'. For example, a reduction in population size in a given population will decrease the level of diversity for any marker (AFLP and microsatellites) regardless of its mutation rates. As a result, one would expect that populations that had different evolutionary histories would also differ in their levels of diversity, following a similar trend for both markers. Most of the diversity measures estimated with different markers showed a significantly lower level in Corsican populations than in Aquitaine populations (Table 1 and Table 2). For microsatellites, the Corsican populations showed a lower allelic richness than the Aquitaine ones but the alleles tended to be different in each population. For AFLPs, there are more pronounced differences between the two provenances for the analysis restricted to the polymorphic markers (according to the rules of Lynch & Milligan, 1994) than for the analysis conducted on the whole set of data.

The congruent results of diversity levels between the two markers suggest that population differences in diversity are more related to 'marker non-specific' than 'marker specific' causes. Among the former, population history may be advocated to interpret these results. The frequent occurrence of fire over large areas, the more severe climatic conditions in Corsica and the mountainous topography may have induced repetitive sequences of restriction and expansion of maritime pine forest, leading to progressive genetic erosion. Furthermore genetic erosion can not be prevented by gene flow from the continent, as Corsica is isolated by a few hundred kilometres from France or Italy. An alternative hypothesis, and not exclusive with the previous one, may be that the number of founder populations that became established on the Corsican island was limited.

### Discrepancy between AFLPs and microsatellites for measuring within-population diversity

If the estimates of levels of diversity with different markers were congruent for the comparison between the Corsican and Aquitaine provenances, they do not lead to the same conclusions for the comparison of populations within each of the two provenances. There is no significant correlation among the diversity estimates calculated for the AFLPs and microsatellites (Table 4). Considering AFLP markers alone, a striking result was that the two primer-enzyme combinations provided different rankings of populations, particularly in Corsica. Furthermore, rankings given by two different microsatellite loci were rarely significantly and positively correlated. When marker systems were compared, neither  $A$  nor  $H_i$ , estimated by microsatellites, gave similar rankings to  $H_i$  estimated by AFLP markers. The discrepancy found between markers may be due to either biological and/or technical causes. 'Marker non-specific' factors (drift, migration, population history) did not differ sufficiently between different populations to generate different levels of diversity. This would result in the lack of differentiation for population diversity. Hence populations within the Corsica or Aquitaine provenance would be homogeneous for their level of diversity due, for example, to intensive and equal gene flow, and the absence of drift within each provenance. This hypothesis is supported by the statistical multipopulation comparison of diversity (Table 5), showing that, except with microsatellites in Corsican populations, no diversity difference could be detected among populations. The technical interpretation is related to the sampling used. The difference in levels of diversity may be so small among populations that our sampling strategy did not allow the detection of any significant differences. All differences among diversity estimates are therefore so small that the ranking observed is only generated by the noise of the estimation.

### Conservation implications

The comparative analysis of genetic diversity with different markers has provided interesting insights not only into the organization and subdivision of diversity but also into the genetic mechanisms that contributed to the extant structure of diversity. All markers indicated that the Corsican and Aquitaine provenance are strongly differentiated. This is suggested by the high  $G_{ST}$  values between the two provenances and by the principal component analysis (Fig. 2). It is also illustrated by the important differences in allele frequencies for the microsatellite locus *FRPP91* (Fig. 1). Conservation

actions are therefore recommended to save diversity for each of the two provenances. Our results indicate however, that sampling of candidate populations or trees for conservation should be achieved differently in the two provenances. The Aquitaine provenance is characterized by a low differentiation for allele frequencies (extremely low  $G_{ST}$  values) and a low differentiation for within-population diversity, and these results are again consistent across markers. As a result *ex situ* or *in situ* conservation could be based on a random sample of trees throughout the Aquitaine area. On the other hand the Corsica provenance is characterized by a higher population differentiation for allele frequencies ( $G_{ST}$  values) and a slightly higher differentiation for within-population diversity. Consequently, it is expected that different populations may contain different alleles, as suggested by the distribution of allelic richness among populations. Sampling guidelines for conservation would therefore recommend concentrating on a few populations that would sample the total diversity. At this stage, our results indicate that a more intense monitoring of diversity should be conducted in the Corsican provenance in order to identify the candidate populations for conservation. To conclude on the conservation issue, we would also recommend complementing the present study with a parallel investigation on adaptive traits (drought, pest and insect resistance) and other economic traits (growth and form) before a definite strategy is adopted.

### Acknowledgements

We are very grateful for the technical help provided by Cécile Cabrero and Audrey Lartigue. We thank Dominique Charon and the French National Forest Office (ONF) for the help in collecting material in Aquitaine. This work was supported by grants from France (Ministère de l'Agriculture et de la Pêche-DERF no. 61.21.04/98 and Bureau des Ressources Génétiques) and the European Union (INCO, ERBIC-08CT-970200). The authors thank two anonymous reviewers and Don Stratton for their comments on the manuscript.

### References

- BAHRMAN, N., BARADAT, P. AND PETIT, R. J. 1992. Structuration de la variabilité génétique du Pin maritime dans l'ensemble de son aire naturelle. *Colloque international en hommage à Jean Pernès*. Publications du Bureau des Ressources Génétiques, pp. 352–368.
- BAHRMAN, N., PLOMION, C., PETIT, R. J. AND KREMER, A. 1997. Contribution of two-dimensional electrophoresis of proteins to maritime pine genetics. *Ann. Sci. For.*, **54**, 225–236.

- BARADAT, P. AND MARPEAU-BEZARD, A. 1988. *Le pin maritime, Pinus pinaster* Ait. *Biologie et génétique des terpènes pour la connaissance et l'amélioration de l'espèce*. PhD Thesis, University of Bordeaux I.
- BARADAT, P., MARPEAU, A. AND WALTER, J. 1991. Terpene markers. In: Müller-Starck, G. and Ziehe, M. (eds) *Genetic Variation in European Populations of Forest Trees*, pp. 40–66. Sauerländer's Verlag, Frankfurt-am-Main.
- BROWN, A. H. D. AND WEIR, B. S. 1983. Measuring genetic variability in plant populations. In: Tanksley, S. D. and Orton, T. J. (eds) *Isozymes in Plant Genetics and Breeding Part A*, pp. 219–239. Elsevier Science Publishers, Amsterdam.
- BURBAN, C., PETIT, R., CARCREFF, E. AND JACTEL, H. 1999. Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. *Mol. Ecol.*, **10**, 1593–1602.
- COSTA, P., POT, D., DUBOS, C., FRIGERIO, J. M. ET AL. 2000. A genetic map of Maritime pine based on AFLP, RAPD and protein markers. *Theor. Appl. Genet.*, **100**, 39–48.
- DOYLE, J. J. AND DOYLE, J. L. 1990. Isolation of plant DNA from fresh tissue. *Focus*, **12**, 13–15.
- DUREL, C. E. AND BAHRMAN, N. 1995. Analyse de la diversité génétique des peuplements de Pin maritime de Corse. Gestion et exploitation de la ressource génétique. *Rev. For. Fr.*, **47**, 509–522.
- ECHT, C. S., MAY-MARQUARDT, P., HSEIH, M. AND ZAHORCHAK, R. 1996. Characterization of microsatellite markers in eastern white pine. *Genome*, **29**, 1102–1108.
- ESTOUP, A., TAILLIEZ, C., CORNUET, J.-M. AND SOLIGNAC, M. 1995. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Mol. Biol. Evol.*, **12**, 1074–1084.
- JACTEL, H., MÉNASSIEU, P. AND BURBAN, C. 1996. Découverte en Corse de *Matsucoccus feytaudi* Duc. (Homoptera: Margarodidae), cochenille du Pin maritime. *Ann. Sci. For.*, **53**, 145–152.
- JACTEL, H., MÉNASSIEU, P., CERIA, A., BURBAN, C. ET AL. 1998. Une pullulation de la cochenille *Matsucoccus feytaudi* provoque un début de dépérissement du Pin maritime en Corse. *Rev. For. Fr.*, **50**, 33–45.
- JENCZEWSKI, E., PROSPERI, J.-M. AND RONFORT, J. 1999. Differentiation between natural and cultivated populations of *Medicago sativa* (Leguminosae) from Spain: analysis with random amplified polymorphic DNA (RAPD) markers and comparison to allozymes. *Mol. Ecol.*, **8**, 1317–1330.
- KRUTOVSKII, K. V., EROFEEVA, S. Y., AAGAARD, J. E. AND STRAUSS, S. H. 1999. Simulation of effects of dominance on estimates of population genetic diversity and differentiation. *J. Hered.*, **60**, 499–502.
- LYNCH, M. AND MILLIGAN, B. G. 1994. Analysis of population genetic structure with RAPD markers. *Mol. Ecol.*, **3**, 91–99.
- MANTEL, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.*, **27**, 209–220.
- MARIETTE, S., CHAGNÉ, D., DECROOQ, S., VENDRAMIN, G. G. ET AL. 2001. Microsatellite markers for *Pinus pinaster* Ait. *Ann. Sci. For.*, **58**, 203–206.
- NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.
- NEI, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- PAQUEREAU, M. M. 1964. Flores et climats post-glaciaires en Gironde. *Actes Société Linnéenne Bordeaux*, **101**, 1–156.
- PETIT, R. J., BAHRMAN, N. AND BARADAT, P. H. 1995. Comparison of genetic differentiation in maritime pine (*Pinus pinaster* Ait.) estimated using isozyme, total protein and terpenic loci. *Heredity*, **75**, 382–389.
- PETIT, R. J., EL MOUSADIK, A. AND PONS, O. 1998. Identifying population for conservation on the basis of genetic markers. *Cons. Biol.*, **12**, 844–855.
- RIBEIRO, M. M., PLOMION, C., PETIT, R. J., VENDRAMIN, G. G. ET AL. 2001. Variation of chloroplast single-sequence repeats in Portuguese maritime pine (*Pinus pinaster* Ait.). *Theor. Appl. Genet.*, **102**, 97–103.
- SALVADOR, L., ALA, R., AGÚNDEZ, D. AND GIL, L. 2000. Genetic variation and migration pathways of maritime pine (*Pinus pinaster* Ait.) in the Iberian peninsula. *Theor. Appl. Genet.*, **100**, 89–95.
- SMITH, D. N. AND DEVEY, M. E. 1994. Occurrence and inheritance of microsatellites in *Pinus radiata*. *Genome*, **37**, 977–983.
- SOKAL, R. 1979. Testing statistical significance of geographic variation patterns. *Syst. Zool.*, **28**, 227–232.
- SOKAL, R. AND ROHLF, F. J. 1995. *Biometry*. Freeman, New York.
- TAUTZ, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucl. Acids Res.*, **17**, 6463–6471.
- VENDRAMIN, G. G., ANZIDEI, M., MADAGHIELE, A. AND BUCCI, G. 1998. Distribution of genetic diversity in *Pinus pinaster* Ait. as revealed by chloroplast microsatellites. *Theor. Appl. Genet.*, **97**, 456–463.
- VOS, P., HOGERS, R., BLEEKER, M., REIJANS, M. ET AL. 1995. AFLP: a new technique for DNA fingerprinting. *Nucl. Acids Res.*, **23**, 4407–4414.
- WEBER, J. L. AND WONG, C. 1993. Mutation of human short tandem repeat. *Hum. Mol. Genet.*, **2**, 1123–1128.