HIGH LEVELS OF GENETIC DIVERSITY IN A LONG-TERM EUROPEAN GLACIAL REFUGIUM OF *PINUS SYLVESTRIS* L.[1]

J. J. Robledo-Arnuncio[1,2], R. Alía[2,3] & L. Gil[1]

[1)] Unidad de Anatomía, Fisiología y Genética, ETSIM, Ciudad Universitaria s/n, E-28040 Madrid, Spain
[2)] Unidad de Genética Forestal, CIFOR-INIA. P.O. Box 8111, E-28080 Madrid, Spain;
e-mail: alia@inia.es, Fax:+34913572293
[3)] Corresponding author

ABSTRACT

Genetically valuable regions of a species on the basis of previous broad-range genetic surveys must be considered for further studies. This can be especially true in areas with diverse ecological conditions. Southernmost European populations of Scots pine (*Pinus sylvestris* L.) from the Iberian Peninsula are considered as Tertiary relicts that have persisted throughout Quaternary glaciations in this area, which is regarded as an independent evolutionary unit. Fragmentation and a higher interpopulation genetic variation than that found in northern regions are characteristics of Scots pine Spanish populations. Six chloroplast microsatellite (cpSSR) loci were examined in 324 individuals from 13 Iberian populations in the Northern Meseta region and different approaches for determining population genetic divergence were performed to obtain information for conservation purposes. Very high levels of haplotypic diversity were found in relation to those detected in other pine species from the Iberian Peninsula, providing a new insight into the evolutionary history of different taxa in southern refugia. Most of the genetic diversity is within populations, with very little but significant variation among populations (R_{ST}=0.024). Most of variation is due to a few highly differentiated populations, which should be priority candidates for conservation strategies. Big genetic differences were found between closed populations and there was not a clear geographic pattern in the distribution of genetic variation.

Key words: *Pinus sylvestris*, chloroplast microsatellites, gene conservation, glacial refugia, Iberian Peninsula

INTRODUCTION

The assessment of genetic diversity within and among populations is central to single-species conservation strategies, since genetic variation is a fundamental component of biodiversity. Until more comprehensive information becomes available on the variation of quantitative adaptive traits and genes underlying quantitative traits, neutral molecular markers remain as useful tools for rapid decision making on conservation priorities (MCKAY & LATTA 2002). Range-wide surveys of neutral genetic diversity are now available for many tree species and are providing valuable information on broad geographic patterns of genetic variation distribution and on evolutionary history. Nevertheless, diagnosing distinct populations and selecting priority candidates for conservation may require a more intensive sampling strategy, since evolutionary processes may be taking place on a small scale (CRANDALL *et al.* 2000; GRAM & SORK 2001). Thus, particularly interesting areas, as revealed by broad-scale genetic surveys, should be the target of subsequent studies performed with suitable high resolution markers. Chloroplast microsatellites (cpSSRs) have demonstrated high levels of intraspecific variability and represent useful markers for population genetic analysis on small scales (PROVAN *et al.* 2001).

Biochemical, isozyme and mitochondrial DNA markers together with paleobotanical information suggest that southernmost European marginal populations of Scots pine (*Pinus sylvestris* L.) from the Iberian Peninsula are Tertiary relicts and belong to a different evolutionary unit from those in northern European locations (LEBRETON *et al.* 1990; PRUS-GŁOWACKI & STEPHAN 1994; SINCLAIR *et al.* 1999; SORANZO *et al.* 2000; WILLIS *et al.* 1998). These studies also show higher levels of inter-population differentiation between the Spanish populations of the species than the levels in northern regions. Scots pine occurs in isolated mountainous

[1] This paper has been presented at the IUFRO Symposium on Population and Evolutionary Genetics of Forest Trees held in Stará Lesná, Slovakia, on August 25–29, 2002.

areas in the Iberian Peninsula, following the characteristic fragmented distribution pattern of southern localities across its range, which is the most widespread among Eurasian pines. Physical isolation, together with a complex geography and diverse climatic and edaphic conditions could be promoting the genetic divergence of Scots pine populations in this region even on small scales, as shown by the differences found between closed locations in adaptive traits (ALÍA *et al.* 2001). Therefore, populations of *Pinus sylvestris* L. in this long-term glacial refugium may represent a precious portion of the global genetic resources of the species. Existing data suggest that substantial genetic variation between closed localities can be found and call for intensive sampling to avoid overlooking valuable populations.

The present work focuses on the Scots pine populations from the Iberian Peninsula that are scattered over a perimetric range around the Northern Meseta (Fig. 1). Populations of the species in this area were found genetically different from other Spanish populations by means of isozyme analysis (Prus-Glowacki *et al.* in this symposium). *Pinus sylvestris* occurs on the different mountain chains that enclose the Meseta Miocenic sedimentary plateau, mostly between 1000 and 1800 m and usually defining timberline. Two small relict localities from glacial times (FRANCO *et al.* 2001), Cuéllar and Coca, remain isolated at low altitude in the dry Meseta plains, within the present distribution range of other Mediterranean pine species, namely *Pinus pinaster* Ait. and *Pinus pinea* L. The main goal of this study was to increase the available molecular information on Scots pine populations in the Northern Meseta region and provide a detailed picture of neutral genetic variation distribution for genetic conservation purposes. To achieve this aim a cpSSR survey across thirteen sampling sites was performed.

MATERIALS AND METHODS

Plant material, DNA extraction and sizing of polymerase chain reaction (PCR) products

13 sampling sites were selected to cover the full range of the species on the Northern Meseta, both over surrounding mountain chains and the two relict populations (Cuéllar and Coca) located in the Meseta plains (Fig. 1, Table 1). Needles from 25 randomly selected trees from each population were collected. Needle tissue from each tree was ground in liquid nitrogen and then total genomic DNA was extracted following the protocol of DELLAPORTA *et al.* (1983). Primers designed by VENDRAMIN *et al.* (1996) were used to amplify six microsatellites regions of the chloroplast genome (Pt15169, Pt26081, Pt30204, Pt36480, Pt71936, Pt87268; codes

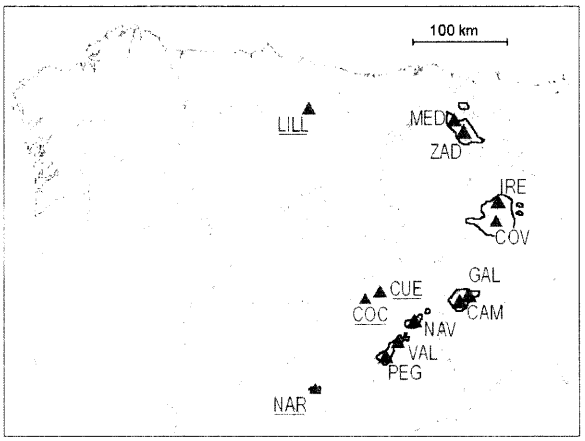

Figure 1. Map of the Northern Meseta region (central and northwestern Spain) showing the distribution of *Pinus sylvestris* and the locations of sampled populations for this study. The codes for populations are in Table 1. Small isolated populations are underlined.

Table 1. Label, name, altitude and location for te thirteen *Pinus sylvestris* L. populations analyzed.

Label	Population	Province	Altitude (m)	Latitude	Longitude
Lill	Puebla de Lillo	León	1550	43° 04′ N	5° 15′ W
Zad	San Zadornil	Burgos	1000	42° 50′ N	3° 11′ W
Med	Medina de Pomar	Burgos	860	42° 57′ N	3° 16′ W
Cov	Covaleda	Soria	1550	41° 56′ N	2° 48′ W
Ire	Hoyos del Iregua	La Rioja	1350	42° 05′ N	2° 39′ W
Gal	Galve de Sorbe	Guadalajara	1400	41° 15′ N	3° 07′ W
Cam	Campisábalos	Guadalajara	1400	41° 13′ N	3° 12′ W
Nav	Navafría	Segovia	1700	41° 00′ N	3° 50′ W
Val	Valsaín	Segovia	1500	40° 49′ N	4° 01′ W
Peg	Peguerinos	Ávila	1500	40° 39′ N	4° 12′ W
Nar	Navarredonda	Ávila	1550	40° 21′ N	5° 07′ W
Cue	Cuéllar	Segovia	800	41° 16′ N	4° 13′ W
Coc	Coca	Segovia	790	41° 12′ N	4° 30′ W

as in VENDRAMIN *et al.* 1996). PCR was done in a total volume of 10 mml containing 2.5 mM $MgCl_2$, 1 x reaction buffer (16 mM $(NH_4)_2SO_4$, 67 mM Tris-HCl; Ecogen), 275 mmM dNTPs, 1.5 pmol IRD-800 labelled forward primer, 1.5 pmol reverse primer, 0.16 U *Taq* polymerase (Ecogene) and 10 ng of genomic DNA. Reactions were done on a Perkin Elmer model 9700 thermal cycler using the following profile: initial denaturation at 95 °C for 5 min, followed by 15 cycles of 1 min at 94 °C, 45 s at 55 °C and 1 min at 72 °C, and a final extension step of 8 min at 72 °C. Amplification products were resolved on 6 %, 25-cm-long, 0.25-mm-thick, denaturing polyacrylamide gels containing 7 M urea and 1 × TBE buffer. Gels were run at 45 W constant power for approximately 1 h using a Li-Cor 4200 Series automatic sequencer. Sizing of the amplified fragments was carried out by Gene ImagIR ver. 3.56 software (Scanalytics) using external standards.

Data analysis

Different haplotypes were defined as a unique combination of size fragments across the six chloroplast microsatellites. Effective number of haplotypes was calculated as $N_e = 1/(\Sigma p_i^2)$, where p_i is the population frequency of the i-th haplotype. Unbiased haplotype diversity, H_e, was computed for each populations (NEI 1987), as well as the mean genetic distance among individual haplotypes following a stepwise mutation model approach, D^2_{sh}, as described in VENDRAMIN *et al.* (1998).

The amount of genetic variation among populations was estimated using an analysis of molecular variance (AMOVA, EXCOFFIER *et al.* 1992), defining distances among haplotype pairs with measure and computing the significance of obtained R_{ST} values by permuting haplotypes among populations 1,000 times. An among-population Nei's genetic distance (NEI *et al.* 1987) matrix based on haplotype frequencies was used to build a dendrogram of population relationships via the unweighted pair group clustering method (UPGMA). Clustering of the populations was also carried out by a Principal Component Analysis (PCA) performed on a coancestry coefficient distance matrix [$D = -\ln(1-F_{ST})$; REYNOLDS *et al.* 1983]. Assuming equal size stationary populations in equilibrium, Reynolds ′ distance is expected to be approximately proportional to divergence time. Programs used to conduct these analyses were Arlequin ver. 2.000 (SCHNEIDER *et al.* 2000), Microsat 1.5 (Eric Minch, Stanford University, USA), Neighbor and Drawtree modules of the Phylip package ver. 3.5c (FELSENSTEIN 1993) and Simca-P ver. 3.01 (Umetrics AB, Sweden).

Haplotype richness (number of haplotypes) of each population and total haplotype richness were standardized using the rarefaction method described by EL MOUSADIK and PETIT (1996). The contribution of each population to the total haplotypic richness (CR_T) was computed following the approach in PETIT *et al.* (1998) for allelic richness, distinguishing two components, one related to its level of diversity (CR_S) and the other to its divergence from the other populations (CR_D). Additionally, a measure of genetic divergence was obtained for each population by recalculating R_{ST} index after removing that specific population from the original data set and comparing the new estimation ($R_{ST,i}$) to that obtained when all populations are included. It is expected that lower R_{ST} values than the original estimate will only be obtained after the removal of a population genetically different from the remaining ones, and therefore this approach has been used to determine genetic uniqueness of tree populations for conservation purposes (DAVIDSON & EL-KASSABY 1997). Finally, an exact test of population differentiation (RAYMOND & ROUSSET 1995) based on haplotypic frequencies was carried out among population pairs. This procedure tests the hypothesis of random distribution of haplotypes among populations using a Markov chain probability estimation on a $r \times k$ contingency table, being r the number of haplotypes and k the number of populations. The number of significantly different populations from one single population, n_s, can be considered as a measure of its genetic divergence. The test was computed with Arlequin ver. 2.000 (SCHNEIDER *et al.* 2000), performing 1,000 dememorisation and 10,000 Markov chain steps and setting a significance level of 0.05.

RESULTS

The six cpSSR regions analysed were all polymorphic, giving a total of 29 size variants (3–7 per locus). The different variants combined in 139 different haplotypes among the 322 individuals surveyed from 13 populations. Many of these haplotypes (55 %) were detected only once, 19 % were detected just in two individuals and the rest were found in 3–15 trees. A high proportion (61 %) of the haplotypes found in the study were unique to a particular population, but most of them were observed just in one or two individuals from a single population, which seems to be a result of their low frequency and the limited sample size and would

suggest that they cannot be considered as population markers. None of the detected haplotypes was common to all populations, and the most frequent was found in eight out of the thirteen populations with an overall frequency of 4.6 %. A description of the variants at each cpSSR and the definition of haplotypes and their frequencies can be obtained from the corresponding author upon request.

A very high intrapopulation haplotypic diversity (mean H_e = 0.977) was observed, being the Guadarrama Chain the most diverse geographical area (Navafría, Valsaín and Peguerinos populations; Fig. 1). The highest value of H_e was found in Valsaín and Navafría (0.993) and the lowest (0.943) in Cuéllar relict population (Table 2). The mean number of different haplotypes per population was 19 out of 25 sampled trees (22 trees in Hoyos del Iregua). A two-fold difference was observed between the highest effective number of haplotypes (21.5 of Valsaín and Navafría) and the lowest observed value (10.6 in Cuéllar), being the average across the thirteen populations 16.8 effective haplotypes. A rough positive correlation can be observed between H_e and the within-population mean pairwise distance D^2_{sh} (Table 2). The population with the lowest haplotype diversity, Cuéllar, also showed one of the lowest within-population mean pairwise distances (mean D^2_{sh} = 3.746). However, the isolated population in Navarredonda de Gredos, with the maximum D^2_{sh} displays an haplotypic diversity below average values (Table 2).

A very low (R_{ST} = 0.024) but significant (p = 0.011) amount of the haplotype variation is attributable to between population differentiation. The UPGMA dendrogram based on Nei's genetic distance calculated on haplotype frequencies reveals two principal groups of populations (Fig. 2), one mainly including northern locations of the species and the other clustering populations of the species at the southern mountainous limit of the Meseta. However, the geographical pattern of the dendrogram is not clear, since Navarredonda and Campisábalos cluster with the group of northern populations and Covaleda appears within the southern group. In addition, some geographically closed populations show an important genetic distance, like Galve de Sorbe and Campisábalos or like Cuéllar and Coca, whereas others distant populations cluster together, like Covaleda and Valsaín. The elongated branch lengths of Lillo, San Zadornil, Cuéllar and Navarredonda populations point out their genetic distinctiveness (Fig. 2).

Table 2. Sample size and measures of genetic diversity and divergence for each Scots pine population based on six cpSSRs.

Population	n	N	N_e	H_e	SD	D^2_{sh}	CR_T	CR_S	CR_D	$R_{ST,i}$	n_s
Lill	25	18	15.244	0.973	(0.018)	4.527	0.030	–0.363	0.393	0.028	7
Zad	25	20	16.892	0.980	(0.017)	4.658	1.243	0.105	1.138	0.023	8
Med	25	18	14.535	0.970	(0.019)	4.383	–0.685	–0.366	–0.319	0.029	5
Cov	25	19	16.892	0.980	(0.016)	3.786	0.667	–0.126	0.793	0.028	5
Ire	22	19	17.267	0.987	(0.017)	4.494	0.522	–0.088	0.610	0.025	5
Gal	25	17	12.755	0.960	(0.023)	3.158	–0.905	–0.603	–0.302	0.023	4
Cam	25	20	18.939	0.987	(0.015)	4.472	–0.555	0.342	–0.897	0.018[a]	1
Nav	25	23	21.552	0.993	(0.013)	4.509	0.759	0.811	–0.052	0.025	2
Val	25	23	21.552	0.993	(0.013)	4.844	1.267	0.811	0.456	0.028	3
Peg	25	21	18.939	0.987	(0.015)	4.539	0.080	0.342	–0.262	0.023	2
Nar	25	20	15.244	0.973	(0.022)	5.158	1.501	0.102	1.399	0.019	9
Cue	25	15	10.593	0.943	(0.026)	3.746	0.335	–1.078	1.413	0.016[a]	12
Coc	25	20	17.857	0.983	(0.015)	4.279	0.421	0.108	0.313	0.029	5
Average	–	19	16.789	0.977		4.35	0.36	0	0.36	0.024	5.2
All populations R_{ST}										0.024	

n = sample size. N = number of haplotypes. N_e = effective number of haplotypes. H_e = unbiased haplotypic diversity with standard error between brackets (NEI 1987). D^2_{sh} = average distance between individuals within populations (VENDRAMIN *et al.* 1998). CR_T = population contribution (percentage) to total haplotypic richness (following PETIT *et al.* 1998). CR_S = population contribution (percentage) to total haplotypic richness due to its own diversity (following PETIT *et al.* 1998). CR_D = population contribution (percentage) to total haplotypic richness due to its divergence (following PETIT *et al.* 1998). $R_{ST,i}$ = value of R_{ST} when the i-th population is removed from the calculation. n_s = number of significant different populations based on pairwise exact test of population differentiation (RAYMOND & ROUSSET 1995). [a] = value not significantly different from cero (p = 0.005; 1000 permutations).

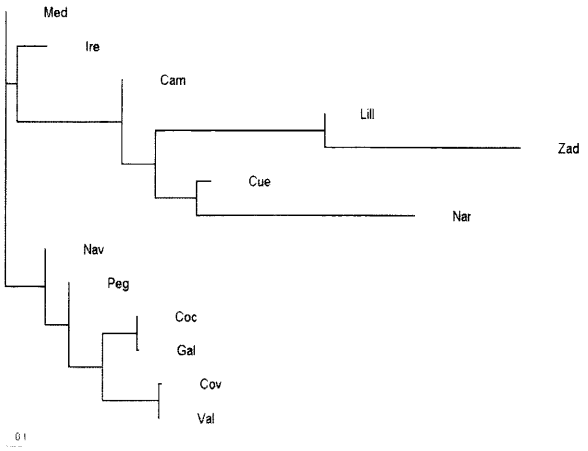

Figure 2. Dendrogram of populations generated using UPGMA method based on the matrix of Nei's genetic distances (NEI 1987) calculated on haplotype frequencies.

In terms of all the three parameters of genetic divergence calculated for each population, namely (i) its contribution to the total haplotypic diversity due to its own divergence, CR_D, (ii) its contribution to the total portion of genetic variation among populations, $R_{ST,i}$ and (iii) the number of populations from which it is significantly different, n_s, it comes out that Cuéllar, Navarredonda, and to a lesser extent San Zadornil are the most differentiated populations (Table 2, Fig. 3). The contribution of each population to diversity has a low absolute value (1.4 % maximum), but PETIT *et al.* (1998) point out that comparisons across populations remain appropriate even when an increasing number of populations had substantially decreased absolute contributions. The average contribution to total haplotypic richness is positive, which may happen when a large differentiation for this variable exists in the species (PETIT *et al.* 1998).

The exact test of population differentiation based on haplotype frequencies show that Cuéllar small marginal population is the only one significantly different from all the other populations in the study. This population is also the one with the highest contribution to total haplotypic richness due to its genetic divergence (CR_D= 1.413 %). Additionally, R_{ST} estimation is not significantly different from cero (p = 0.05; 1000 permutations) when Cuéllar population is removed from the analysis. Considering the contribution to the total haplotypic diversity due to the diversity level of the populations, it is shown again that the Guadarrama Chain (Valsaín, Navafría and Peguerinos populations) is the Northern Meseta region harbouring the highest levels of genetic variation (Table 2, Fig. 3). PCA based on the pairwise coancestry coefficients groups populations without following a geographical pattern, showing a main compact cluster of populations from different regions and four outliers, namely Cuéllar, Navarredonda, Campisábalos and Peguerinos (Fig. 4). The percentage of total variation explained by the two principal components was 76 %. Again, it is demonstrated the genetic distinctiveness of Cuéllar and Navarredonda populations, but PCA analysis also points at the genetic divergence of Campisábalos and Peguerinos, which strikingly are the two populations less differentiated in terms of the exact test of RAYMOND & ROUSSET (1995) based on haplotype frequencies, being significantly different just from one and two populations respectively (Fig. 4, Table 2). Besides this, these two populations have negative contributions to total haplotype richness due to their genetic divergence (CR_D; Table 2). On the contrary, the decrease of *RST* when any of these populations are not considered in the analysis would support the results obtained from PCA on coancestry coefficients (Table 2).

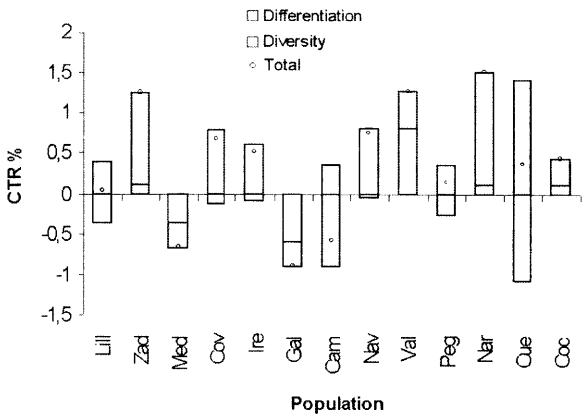

Figure 3. Contribution to total haplotypic richness (CTR%) of each population of Scots pine, subdivided into a diversity and a differentiation component following PETIT *et al.* (1998).

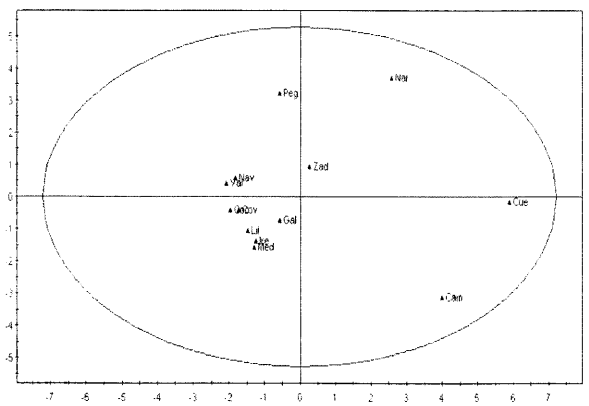

Figure 4. Principal component analysis of the 13 populations of Scots pine based on the matrix of coancestry coefficients (REYNOLDS *et al.* 1983). The proportion of the variance explained is 76%.

It must be emphasized that both R_{ST} and coancestry coefficients are calculated assuming a cpSSR stepwise mutation model to define interhaplotypic distance (D^2_{sh}), whereas the exact test of population differentiation and the contributions to the total haplotype richness are based solely on haplotype frequencies and do not take size differences among cpSSR variants into account. Indeed, when the number of different cpSSR size variants is the metrics considered for interhaplotypic distance calculation (infinite allele model), the portion of total variation attributable to among population differentiation drops to a half (data not shown).

DISCUSSION

High diversity values (0.943–0.993) based on haplotype frequencies were observed in thirteen native Scots pine populations from the Spanish Northern Meseta region using six polymorphic chloroplast microsatellites. This substantial genetic diversity is much higher than that found on Spanish populations of the species using isozymes (0.267–0.342; PRUS-GŁOWACKI *et al.* in this symposium) or mitochondrial DNA markers (0.067–0.515; SORANZO *et al.* 2000), which confirms previously reported differences between hypervariable cpSSR and other genetic markers (PROVAN *et al.* 2001).

More interestingly, levels of variation of other Mediterranean pine species, also in the Iberian Peninsula and at the same cpSSRs (or at a greater number including the six used in this study), were always lower than for *Pinus sylvestris* L. (Table 3). The difference is especially striking between Scots pine and the most termophile taxa (*Pinus pinea* L. and *Pinus halepensis* Mill.), meanwhile Maritime pine (*Pinus pinaster* Ait.) shows intermediate diversity levels. Present geographical distribution of all the four species follows a similarly fragmented pattern in Spain, with many marginal and isolated populations. Historical processes rather than present distribution and patterns of ongoing gene flow must be considered to interpret present genetic structure and diversity levels of Mediterranean plant species (THOMPSON 1999). Climatic oscillations throughout the Quaternary may have played a major role in shaping the observed genetic variation of plants, especially when this had been surveyed using highly variable molecular markers (COMES & KADEREIT 1998). Since the abovementioned pine species are adapted to very different ecological conditions, also very different demographic responses to climatic shifts can be expected. Scots pine is out of the four species the one better adapted to low temperatures, followed by Maritime pine. Indeed, we know that *P. sylvestris* occurred in wider and more continuous areas at low elevation in the Spanish Northern Meseta during the colder early Holocene (FRANCO *et al.* 2001). The low level of interpopulation differentiation at the cpSSR (R_{ST} = 0.024) in this region is consistent with the hypothesis of a recent fragmentation (SAVOLAINEN & KUITTINEN 2000). Uphill altitudinal migration would have provided on the other hand an effective means to endure warm stages avoiding long latitudinal movements, as shown by present day distribution. We can hypothesize that this short-distance demographic dynamics, together with the mountainous orographic profile of the Iberian Peninsula, would have allowed the

Table 3. Population genetic diversity and differentiation parameters of four different *Pine* species from the Iberian Peninsula based on cpSSRs. Sample size was 24–25 individuals per population for all the species.

Species	cpSSR [1]	Populations [2]	N [3]	N_e [4]	H_e [5]	D^2_{sh} [6]
P. pinea [7]	9	10	3.42	2.02	0.46	0.09
P. halepensis [8]	9	15	5.00	2.52	0.56	0.55
P. pinaster [9]	6	7	14.00	11.82	0.93	4.23
P. sylvestris [10]	6	13	19.46	16.79	0.98	4.35

[1] Number of cpSSR loci analyzed.
[2] Number of populations analyzed.
[3] Average number of haplotypes per population.
[4] Average effective number of haplotypes per population.
[5] Average unbiased haplotype diversity across populations.
[6] Average across populations of the mean distance of individuals within populations.
[7] GÓMEZ *et al.* (2000).
[8] GÓMEZ (1998).
[9] VENDRAMIN (pers. com.).
[10] This study.

maintenance of large and rather stable populations of Scots pine, which could explain its high levels of genetic diversity (HEWITT 2000). On the other hand, *P. halepensis* and *P. pinea* are clearly termophile species. For this reason, they may have undergone important range contractions during the harshest glacial periods, which would have been followed by long expansions back from refugia during subsequent climatic warmings (MORGANTE *et al.* 1998). These processes may lead to losses of genetic diversity in tree species due to colonization effects or to small refuge sizes (AUSTERLITZ *et al.* 2000; COMPS *et al.* 2001). Diversity measures based on cpSSRs as high as the ones found in this study have just been reported for populations of the same species from Scotland (PROVAN *et al.* 1998), which are supposed to have been partly derived from a closed refugium in Scottish lands or southwest Ireland (SINCLAIR *et al.* 1998). It seems that Scots pine biology has ensured the maintenance of its neutral genetic diversity throughout glacial periods and across its range despite having undergone probably different demographic processes in different areas.

On the other hand, the hypervariability of the markers used in this study must be considered carefully when interpreting results. We have seen that 55 % of the haplotypes are found in a single individual and that 61 % were unique to a single population. This means a large proportion of haplotypes with very low population frequencies (<0.05). The probability of not detecting one of this haplotypes in a random sample of 25 trees is closed to 30 %. Thus, distances and comparisons between populations based on haplotype frequencies must be interpreted with caution, especially since size homoplasy of the cpSSRs could be an additional factor confounding the estimates (JARNE & LAGODA 1996). HEDRICK (1999) has demonstrated that the level of differentiation between groups is greatly influenced by highly variable loci, since G_{ST} estimations can be very small even when populations have no alleles in common and since differentiation cannot exceed the average level of within population homozygosity. However, our estimation of R_{ST} yields a low but significant value (0.024; $p = 0.01$) that seems consistent with the estimation for six populations in the same area based on nuclear lowly polymorphic isozyme markers ($G_{ST} = 0.025$; PRUS-GŁOWACKI *et al.* in this symposium). This variation is mainly due to the genetic divergence of Cuéllar, Navarredonda and Campisábalos populations, as shown by $R_{ST,i}$ parameters (Table 2). Indeed, when these three populations are simultaneously removed from the analysis, the R_{ST} value is zero (data not shown). Distribution of among-population genetic variation does not show a clear geographic pattern, as shown by UPGMA clustering (Fig. 2). Anyway, this grouping method is based on genetic distances between populations calculated on haplotypic frequencies. For the reasons mentioned above, this observation should be regarded cautiously until a more intensive sampling effort is performed. A phylogeographic approach on the same haplotype data should provide useful complementary information on this issue (ROBLEDO *et al.* in prep.).

The different methods used in this study for single-population genetic differentiation analysis have yielded consistent results (Table 2, Figs. 3 and 4). Cuéllar and Navarredonda are the most distinct populations based on molecular cpSSR information. Both are in geographically marginal locations. Navarredonda is isolated between Gredos and La Paramera mountain Chains, more than 50 km apart from the closest population. Cuéllar is, together with Coca, the only Scots pine population growing in the Northern Meseta Plains. Its census size reaches 8000 trees, growing on Quaternary sandy soils under a dry climate atypical for the species (annual rainfall 430 mm) thanks to a continuous phreatic supply of water. Most of the trees occur in a narrow strip (5–50 meters) 15 kilometers along river Cega, exactly where the underground water rises. Haplotype diversity in Cuéllar is the minimum across all populations in this study, whereas populations in the Guadarrama Chain, 30 km to the southeast, show the highest levels of genetic diversity in the Northern Meseta (Fig. 1, Table 2). According to hypothesis inferred from the macrofossil and pollen record, this unique location can be explained as a relict community from glacial times (FRANCO *et al.* 2001), during which Scots pine would have grown in wider woodlands on the plains of the Northern Meseta. Progressive retreat of *Pinus sylvestris* L. from the plains towards higher altitudes must have followed Holocene climatic warming, as ecological conditions in the plateau became unsuitable for the species. Therefore, geographical isolation of Cuéllar population in the microhabitat where it still survives probably started in the early Holocene, 8,000–10,000 years BP. Although overall genetic diversity of Scots pine populations in the Northern Meseta is high, it seems that demographic processes during the last 10000 years may have lead to a severe depletion of genetic variation in some particular situations (two-fold difference in the effective number of haplotypes between Cuéllar and Navafría or Valsaín). Since a long-term effective size similar to present day census size does not seem small enough to have produced such effect (ADAMS & BURCZYK 2000) there are two plausible complementary explanations for this

diversity loss. One is that Cuéllar population had undergone one or several population bottlenecks in the past following Scots pine range fragmentation in the early Holocene. The other possibility is that its long-term effective size had been much smaller than its census size due to its singular linear spatial structure, which would make mating among relatives much more likely than in a non-linear shaped population (LOVELESS & HAMRICK 1984). The genetic distinctiveness of this population, as revealed by all the methods used in this study, together with its marginal ecological situation, small size and unusual structure make it a priority candidate for conservation efforts.

The significant differences found in the diversity and divergence parameters between some sampling location pairs on a small scale are remarkable. This is the case of San Zadornil and Medina Pomar populations, located fifteen kilometers apart in the same ecological and geographical area. San Zadornil contributes importantly (CR_D = 1.138 %) to total haplotye richness due to its genetic divergence from the other populations, whereas the contribution of Medina Pomar is negative. Differences between Galve de Sorbe and Campisábalos populations are more striking, both at diversity (N_e) and genetic divergence parameters (CR_D, R_{ST}, i, n_s; Table 2). These two populations grow only eight kilometers apart at the same mountainous region, but there is a significant difference between their respective habitats, since Galve de Sorbe grows on acid substrates while Campisábalos does on limestone. Although cpSSRs are neutral markers and should not reflect variation in adaptive traits (MCKAY & LATTA 2002), it could be hypothesized that a marked difference in edaphic factors between the two locations, interacting with interspecific competence and climatic factors, could have contributed to diverse long-term demographic histories affecting differently neutral variation. Other explanation could be that there had been a recent admixture of populations in this area and that gene flow has not homogenized haplotype frequencies yet. SORANZO *et al.* (2000) found that Galve de Sorbe population was fixed for a mitotype that was predominant in eastern Spain, whereas the others analyzed populations from the Northern Meseta (Lillo, San Zadornil, Valsaín and Navarredonda de Gredos) showed a higher frequency for the other mitotype found in the study, predominant in western locations. This could be supporting the hypothesis of the admixture of populations genetically divergent in the area. Anyway, both ecological and molecular data suggest that this two Scots pine populations should be considered separately from a conservation point of view. One last remarkable population pair is that made up of Coca and Cuéllar. Coca, the second Scots pine population isolated in the Northern Meseta plains, shows much higher levels of diversity than Cuéllar, just thirty kilometers apart and under similar ecological conditions. None of the calculated parameters points at the genetic divergence of this population, that consists of only thirty six even-aged trees scattered over a *Pinus pinaster* woodland. Anyway, since the origin of this population is unclear, although a document from the 16th century mentions the occurrence of a Scots pine stand at that time in the same area (ÁLVAREZ & ALLUÉ 1998), it would be risky to make inferences about its evolutionary history until more information become available. This study has not produced illuminating results on the origin of Coca, but it has shown its important genetic distance from other closed population of undoubted natural origin.

Sampling genetic diversity on an adequate scale is central to both conservation strategies and evolutionary studies (CRANDALL *et al.* 2000). It has been shown that cpSSRs have high resolution for detecting neutral molecular variation on small scales and that, in some particular situations, an intensive site sampling can detect significant genetic variation between nearby locations within a continuous population of a species. In these cases, misleading information could be obtained from a genetic survey that had been performed on a larger scale.

ACKNOWLEDGMENTS

We wish to thank D. Agúndez, M. Sevilla, A. Álvarez, M.R. Chambel, A. Piñera and F. del Caño for field and technical assistance. Special thanks to A. Gómez for her technical assistance and valuable suggestions and comments on this article. The first author was supported by a predoctoral scholarship from the UPM (Universidad Politécnica de Madrid).

REFERENCES

ADAMS, W.T. & BURCZYK, J. 2000: Magnitude and implications of gene flow in gene conservation reserves. *In*: Forest Conservation Genetics. Principles and Practice (eds. Young, A., Boshier, D. and Boyle, T.). pp. 215–226. CSIRO Publishing, Collingwood.

ALÍA, R., MORO-SERRANO, J. & NOTIVOL, E. 2001: Genetic variability of Scots pine (*Pinus sylvestris*) provenances in Spain: growth traits and survival. *Silva Fenn.* 35(1): 27–38.

ÁLVAREZ, J.C. & ALLUÉ, M. 1997: Aspectos forestales en las Ordenanzas de la Comunidad de Villa y Tierra de Coca de 1.583. Actas del I Congreso Forestal

Hispano-LusoII Congreso Forestal Español. pp. 383–388. Pamplona.

AUSTERLITZ, F., MARIETTE, S., MACHON, N., GOUYON, P. & GODELLE, B. 2000: Effects of colonization processes on genetic diversity: differences between annual plants and tree species. *Genetics* **154**: 1309 –1321.

COMES, H.P. & KADEREIT, J.W. 1998: The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Sci.* **3**(11): 432–438.

COMPS, B., GÖMÖRY, D., LETOUZEY, J., THIÉBAUT, B. & PETIT, R.J. 2001: Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics* **157**: 389 –397.

CRANDALL, K.A., BININDA-EMONDS, O.R.P., MACE, G.M. & WAYNE, R.K. 2000: Considering evolutionary processes in conservation biology. *Trends Ecol. Evol.* **15**(7): 290 –295.

DAVIDSON, R. & EL-KASSABY, Y.A. 1997: Genetic diversity and gene conservation of pacific silver fir (*Abies amabilis*) on Vancouver Island, British Columbia. *Forest Genetics* **4**(2): 85–98.

DELLAPORTA, S.L., WOOD, J. & HICKS, J.B. 1983: A plant DNA minipreparation: Versión II. *Plant Mol. Biol. Rep.* **1**: 19–21.

EL MOUSADIK, A. & PETIT, R.J. 1996: High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic of Morocco. *Theor. Appl. Genet.* **92**: 832–839.

EXCOFFIER, L., SMOUSE, P. & QUATTRO, J.M. 1992: Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.

FELSENSTEIN, J. 1993: PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.

FRANCO, F., GARCÍA, M., MALDONADO, J., MORLA, C. & SAINZ, H. 2001: The Holocene history of *Pinus* forests in the Spanish Northern Meseta. *The Holocene* **11**(3): 343–358.

GÓMEZ A., 1998: Análisis de la variabilidad genética de *Pinus halepensis* en España mediante el uso de marcadores de ADN: RAPDs y Cp-microsatélites. PhD Thesis Doctoral. Polytechnical University of Madrid, Madrid.

GÓMEZ, A., AGUIRIANO, E., BUENO, M.A. & ALÍA, R. 2000: Microsatélites del cloroplasto en *Pinus pinea* L. *In* I Simposio del Pino Piñonero, Actas Tomo II, pp. 39–46. Junta de Castilla y León, Valladolid, Spain.

GRAM, W.K., SORK, V.L. 2001: Association between environmental and genetic heterogeneity in forest tree populations. *Ecology* **87**(7): 2012 –2021.

HEDRICK, P.W. 1999: Highly variable loci and their interpretation in evolution and conservation. *Evolution* **53**(2): 313–318.

HEWITT, G. 2000: The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907–913.

JARNE, P. & LAGODA, P.J.L. 1996: Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* **11**: 424–429.

LEBRETON, P., LARACINE-PITTET, C., BAYET, C. & LAURANSON, J. 1990 : Variabilite polyphenolique et systematique du pin sylvestre *Pinus sylvestris* L. *Ann. Sci. Forest.* **47**: 117–130.

LOVELESS, M.D. & HAMRICK, J.L. 1984: Ecological determinants of genetic structure in plant populations. *Ann. Rev. Ecol. Syst.* **15**: 65–95.

MCKAY, J.K. & LATTA, G. 2002: Adaptive population divergence: markers, QTL and traits. *Trends Ecol. Evol.* **17**(6): 285–291.

MORGANTE, M., FELICE, N. & VENDRAMIN, G.G. 1998: Analysis of hypervariable chloroplast microsatellites in *Pinus halepensis* reveals a dramatic genetic bottleneck. *In*: Molecular Tools for screening biodiversity (eds. Karp, A., Isaac, P.G. and Ingram, D.S.). pp. 407–412. Chapman & Hall, London.

NEI, M. 1987: Molecular Evolutionary Genetics. Columbia University Press, New York.

PETIT, R.J., EL MOUSADIK, A. & PONS, O. 1998: Identifying populations for conservation on the basis of genetic markers. *Conservation Biology* **12**(4): 844–855.

PROVAN, J., POWELL, W. & HOLLINGSWORTH, P. 2001: Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends Ecol. Evol.* **16**(3): 142–148.

PROVAN, J., SORANZO, N., WILSON, N.J., MCNICOL, J.W., FORREST, G.I., COTTRELL, J. & POWELL, W. 1998: Gene-pool variation in Caledonian and European Scots pine (*Pinus sylvestris* L.) revealed by chloroplast simple-sequence repeats. *Proc. R. Soc. Lond. B* **265**: 1697–1705.

PRUS-GLOWACKI, W. & STEPHAN, B.R. 1994: Genetic variation of *Pinus sylvestris* from Spain in relation to other European populations. *Silvae Genet.* **43**(1): 7–14.

RAYMOND, M. & ROUSSET, F. 1995: An exact test for population differentiation. *Evolution* **49**: 1280–1283.

REYNOLDS, J., WEIR, B.S. & COCKERHAM, C.C. 1983: Estimation for the coancestry coefficient: basis for a short-term genetic distance. *Genetics* **105**: 767–779.

SAVOLAINEN, O. & KUITTINEN, H. 2000: Small population processes. *In*: Forest Conservation Genetics. Principles and Practice (eds. Young, A., Boshier, D. and Boyle, T.). pp. 91–100. CSIRO Publishing, Collingwood.

SCHNEIDER, S., ROESSLI, D. & EXCOFFIER, L. 2000: Arlequin ver. 2,000: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

SINCLAIR, W.T., MORMAN, J.D. & ENNOS, R.A. 1998: Multiple origins for Scots pine (*Pinus sylvestris* L.) in Scotland: evidence from mitochondrial DNA variation. *Heredity* **80**: 233–240.

SINCLAIR, W.T., MORMAN, J.D. & ENNOS, R.A. 1999: The postglacial history of Scots pine (*Pinus sylvestris* L.) in western Europe: evidence from mitochondrial DNA variation. *Mol. Ecol.* **8**: 83–88.

SORANZO, N., ALIA, R., PROVAN, J. & POWELL, W. 2000: Patterns of variation at a mitochondrial sequence-tagged-site locus provides new insights into the ostgla-

cial history of European *Pinus sylvestris* populations. *Mol. Ecol.* **9**: 1205–1211.

THOMPSON, J.D. 1999: Population differentiation in Mediterranean plants: insights into colonization history and the evolution and conservation of endemic species. *Heredity* **82**: 229–236.

VENDRAMIN, G., LELLI, L., ROSSI, P. & MORGANTE, M. 1996: A set of primers for the amplification of 20 chloroplast microsatellites in *Pinaceae*. *Mol. Ecol.* **5**: 595–598.

VENDRAMIN, G.G., ANZIDEI, M., MADAGHIELE, A., & BUCCI, G. 1998: Distribution of genetic diversity in *Pinus pinaster* Ait. as revealed by chloroplast microsatellites. *Theor. Appl. Genet.* **97**: 456–463.

WILLIS, K.J., BENNETT, K.D. & BIRKS, H.J. 1998: The late Quaternary dynamics of pines in Europe. *In*: Ecology and biogeography of *Pinus* (ed. Richardson, D.M.). pp. 107–121. Cambridge Univ. Press, New York.