

Impact of founder population, drift and selection on the genetic diversity of a recently translocated tree population

F Lefèvre, B Fady, D Fallour-Rubio¹, D Ghosn² and M Bariteau

INRA, Unité de Recherches Forestières Méditerranéennes, Avenue Antonio Vivaldi, 84000 Avignon, France

Recently established, temperate tree populations combine a high level of differentiation for adaptive traits, suggesting rapid genetic evolution, with a high level of genetic diversity within population, suggesting a limited impact of genetic drift and purifying selection. To study experimentally the evolutionary forces in a recently established population, we assessed the spatial and temporal patterns of genetic diversity within a disjunct population of *Cedrus atlantica* established 140 years ago in south-eastern France from a North African source. The population is expanding through natural regeneration. Three generations were sampled, including founder trees. We analysed 12 isozyme loci, three of which were previously found in tight association with selected genes, and quantitative traits. No bottleneck effect was detected in the founder generation, but a simple test of allelic association revealed an initial disequilibrium which

disappeared in the following generations. The impact of genetic drift during secondary evolution was limited, as suggested by the weak temporal differentiation. The genetic load was not reduced after 3 generations, and the quantitative variation for adaptive traits did not change either. Thus, initial genetic changes first proceed from a rapid re-organisation of the diversity through mating and recombination, whereas genetic erosion through drift and selection is delayed due to temporal and spatial stochasticity. Two life-history traits of trees contribute to slowing down the processes of genetic erosion: perenniality and large spatial scale. Thus, one would expect recently established tree populations to have a higher diversity than older ones, which seems in accordance with experimental surveys.

Heredity (2004) 93, 542–550. doi:10.1038/sj.hdy.6800549
Published online 18 August 2004

Keywords: population admixture; drift; selection; disequilibrium; *Cedrus*

Introduction

The high level of within-population diversity generally found in tree species (Hamrick *et al*, 1992) is a striking feature for two reasons. First, because in the temperate zone, the rapid post-glacial recolonisation involved long-distance founder events (Le Corre *et al*, 1997). However, these founder events have not resulted in genetic drift. By using a metapopulation model, Austerlitz *et al* (2000) showed that the long juvenile phase of trees could account for the high level of within-population diversity and the low differentiation for neutral genes. In this model, migrants cumulate over years before the new generation arises, that is, the long juvenile phase of trees increases the impact of migration. Only in very isolated situations, such as that described by Ledig (2000) for *Pinus coulteri*, where highly isolated populations are restricted to high elevations and separated by semiarid habitats which severely limit gene flow, do founder events lead to genetic drift.

The second process that is expected to reduce within-population diversity is purifying selection. Trees

usually show a marked differentiation among populations for adaptive traits: a typical example of clinal variation was observed on *Quercus* (Ducouso *et al*, 1996). From a large-scale study of 125 *Pinus contorta* provenances tested in a broad range of climatic conditions and based on the model of quantitative genetics for response to selection, Rehfeldt *et al* (2001) predicted that only one to 12 generations are sufficient for evolutionary adjustments after climatic changes occur, depending on the geographical zone. Indeed, a rapid genetic evolution for adaptive traits has been observed in translocated populations, as soon as the first seed generation, and this evolution was generally considered a result of selection (Espinell *et al*, 1995; Skråppa and Kohmann, 1997; Pinyopusarerk and Williams, 2000; Peterken, 2001). From a survey of experimental results, Reznick and Ghalambor (2001) concluded that demographic growth is a key factor preventing population decline that could result from directional selection.

Thus, tree populations show the capacity of rapid adaptive changes while maintaining a high level of genetic diversity within population. Mutation is not the main evolutionary force that shapes diversity at the time scale of a few generations of trees. Gene flow is particularly important in trees and interferes with selection in several ways (Lenormand, 2002). Genetic drift and selection are the other evolutionary forces controlling genetic diversity. In a recently established tree population, the result of combined evolutionary

Correspondence: F Lefèvre. E-mail: lefevre@avignon.inra.fr

¹Current address: Parc National des Pyrénées 59, route de Pau, 65000 Tarbes, France.

²Current address: Department of Environmental Management Mediterranean Agronomic Institute of Chania, PO Box 85, GR 73100 Chania, Crete, Greece.

Received 11 June 2003; accepted 1 June 2004; published online 18 August 2004

forces depends on the initial genetic diversity and the environmental conditions in the destination site. Several scenarios are possible.

When initial population establishment results in a bottleneck, genetic drift leads to a random change of allelic frequencies and a reduction in gene diversity in the colonising population compared to the source. Furthermore, we expect a reduction of diversity after inbreeding in the following generations if detrimental genes are purged due to selective sweep (reduction of diversity around the target genes of selection due to hitch-hiking). Indeed, most tree species are allogamous and their high level of diversity is also related to an important genetic load (Ledig, 1986). By contrast, for quantitative traits, a bottleneck results in reduced gene interaction effects and increased additivity (see Carson (1990) for review). If foundation results from population admixture, the allelic diversity is increased. Tree populations transplanted for afforestation programs provide an experimental support to study evolutionary processes. For example, Zheng and Ennos (1999) detected a higher diversity for *Pinus caribaea* populations introduced in China compared to the natural range of the species, due to population admixture, contrasting with increased inbreeding in the populations introduced in Australia, due to flowering defaults and consequently reduced population size. Hybridisation is another important factor of secondary evolution when there is an indigenous species related to the exotic, which is not discussed here.

Cedrus atlantica Manetti (cedar), which was introduced to France from North-Africa in the 1860s for afforestation of eroded land in the Mediterranean region (Cointat, 1996), provides such an opportunity to study evolutionary processes in trees. The species is now well established and regenerates naturally, the youngest seed-producing trees belong to the third- or fourth-generation after establishment. In the common garden experiments planted in France, seeds collected in the French artificial populations perform better in growth than seeds collected in the natural range (Bariteau and Ferrandes, 1992). Assuming that juvenile growth is partly under additive genetic control and does not only reflect maternal effects (Cornelius, 1993), the improved genetic value of transplanted populations could be interpreted as a response to selection in the local environment. Under this selection hypothesis, we expect the superiority of the introduced populations to be maintained with the selection pressure. Such a rapid and efficient response to selection requires that no major effect of genetic drift interferes with the process of selection. As an alternative hypothesis, we may consider that old populations were spatially structured in their native environment, leading to a certain amount of inbreeding due to mating among neighbour relatives (Epperson, 1992). Just by mixing the seeds collected on different trees and by sowing them at random, foresters could have broken the original neighbourhood structure and therefore decreased inbreeding in the first generation of seeds. Under this inbreeding hypothesis, which can be extended to heterosis as a result of population admixture, the superiority of seeds produced in the artificial population may decrease as the introduced population gets 'old' and structured. The selection and the inbreeding hypotheses lead to opposite expectations

regarding long-term sustainability of the introduced germplasm.

To infer drift and selection processes, we followed neutral and selected gene diversity across three consecutive generations, including surviving founder trees, in a forest which has undergone no silvicultural management since its founding in the mid 19th century. In previous work based on isozymes, we described the heterogeneity of segregation patterns and recombination rates among progenies, which suggested selection associated with some of the markers (Fallour *et al.*, 2001). Therefore, we assessed spatial and temporal variation of the genetic diversity by using three different measures that are generally poorly correlated (Reed and Frankham, 2001): neutral genes (some isozymes), single gene markers tightly associated to selection targets (or even directly selected genes, other isozymes), and quantitative traits measured in half-sib progenies. We addressed the following questions: did initial foundation result in a bottleneck or, conversely, in population admixture? What was the impact of genetic drift during secondary evolution? What was the evolution of selected genes and adaptive traits during secondary evolution? Finally, what expectations can be drawn for the genetic diversity of this population?

Material and methods

Material

The cedar forest under study lies on the ridge of a small mountain (Luberon, south east of France, 700 m max. elevation) subject to strong dry winds in the Mediterranean climate zone. Afforestation was used to stop soil erosion after centuries-long grazing: environmental conditions were particularly stressful for afforestation, several species were initially mixed but only *Cedrus* survived (Vicaire, 1864). According to the archives of the French National Forest Service, afforestation was performed by direct sowing of seeds collected in Algeria, but the exact provenance of origin remains uncertain (Cointat, 1996). It started in 1865–68 in Ménerbes and Lacoste districts (Figure 1), but only in 1893 in Bonnieux district (Fallour, 1998). No further introduction has been reported in this forest since that time. The forest expanded through natural regeneration in several waves, depending on the decrease of grazing intensity and forest fire regime. The forest is now continuous (Figure 1). Many seeds were initially planted but survival was irregular: in particular, an aerial photograph reveals that, in 1939, the Lacoste district (zone 1) consisted of an open landscape with scattered surviving founder trees (we counted a maximum of 800 trees for ca. 100 h), whereas survival rate was much higher in the Ménerbes district. No thinning or exploitation has occurred until very recently.

In all, 212 adult trees were sampled, representing the different size classes observed in the Luberon forest. They were distributed into three zones (z1, z2, z3), corresponding to the three introduction areas (Figure 1). Tree age was estimated by three wood cores at 60 cm height. A correction term was used to take into account undetected annual rings at 60 cm: from a destructive analysis, we estimated as 10 years the mean age of 60 cm tall seedlings sampled in the forest. Furthermore, the

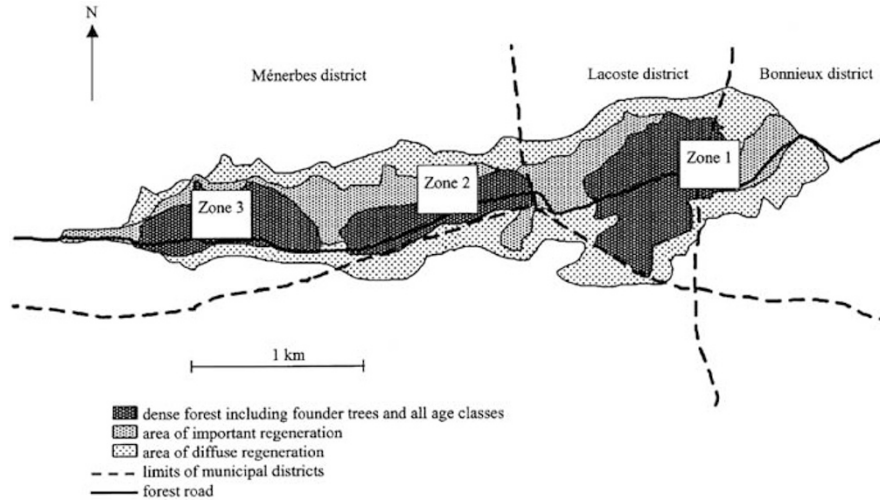


Figure 1 General map of Luberon Cedar forest. This forest expanded through natural regeneration after initial afforestation in three districts in the 1860s. Zones 1–3 indicate the sampling area for this study.

initial growth of the founder trees was very slow and counting the first annual rings was much more uncertain for them than for young trees. Since the archives give a precise information on the date of establishment, we used another correction term for the founder trees so that the oldest age estimate corresponds to the first establishment date, that is, 1863. Finally, we estimated tree age as follows:

$$\text{Age} = \text{number of annual rings} + 10$$

or

$$\text{Age} = \text{number of annual rings} + 15,$$

for the founder trees.

Three generations were then clearly identified (Figure 2):

- G0 correspond to the surviving founder trees.
- G1 correspond to the first regeneration step.
- G2 is potentially derived from G0 and G1 trees.

A zone \times generation group of progenies is hereafter referred to as a subpopulation.

Based on the information available in the archives of the National Forest Service, we consider that age variation observed within G0 does not correspond to a long period of introduction but rather reflects the error of our age estimation for the oldest trees. We did not find any tree corresponding to the afforestation of 1893 mentioned in Bonnieux district in our sampling area. To separate G1 from G2, we considered that significant flower production of *Cedrus* in this forest starts at the age of 40 (Toth, 1978), which is also in accordance with the starting date of G1.

Seeds were collected from 186 trees. Allocation of sample trees among age classes and forest zones is described in Table 1. The youngest trees (G2) in zone 3 were not sampled because they rarely produced any cones.

Three other samples were included in the isozyme analysis as reference populations: a natural population from Algeria (Tala-Guilef), and two French populations

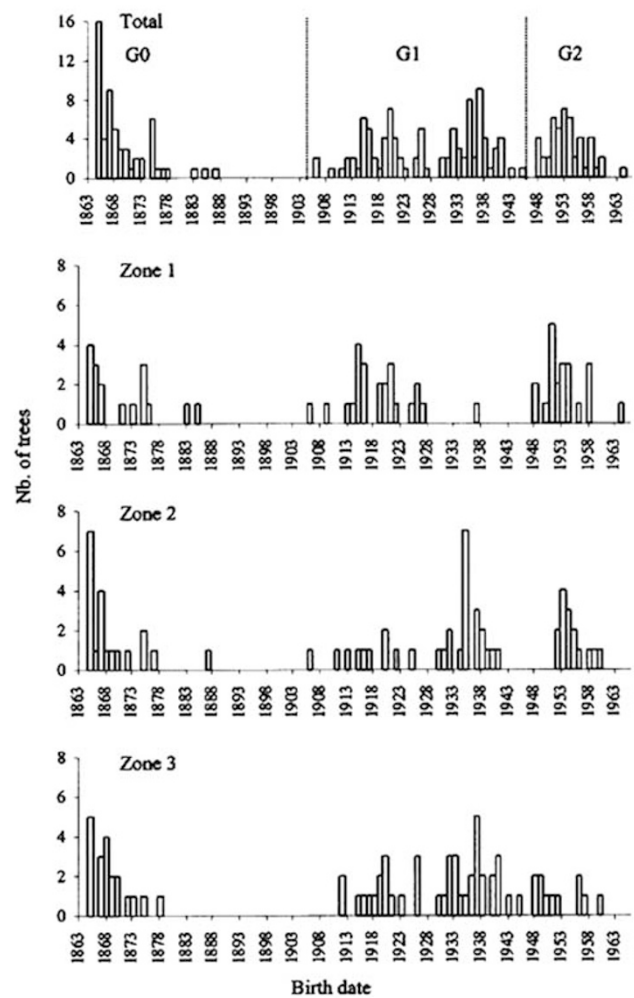


Figure 2 Distribution of estimated ages and limits of the three generations: founder trees (G0), first generation issued from founder trees (G1) and second generation (G2), starting 40 years later with possible contribution of G1 trees to the gamete pool. The G2 trees from zone 3 were not included in the genetics study.

Table 1 Sample size for isozyme analysis (respectively progeny test) within the Luberon forest

	G0	G1	G2	Unknown age	Total
Z1	16 (10)	24 (10)	21 (10)	1 (0)	62 (30)
Z2	16 (10)	30 (20)	14 (10)	2 (0)	62 (40)
Z3	20 (10)	33 (20)	0 (0)	9 (0)	62 (30)
Total	52 (30)	87 (50)	35 (20)	12 (0)	186 (100)

(Ventoux and Saumon), which were established at the same period as the Luberon forest, possibly from the same gene pool (at least by the same forest engineer; Cointat, 1996). Seeds from Tala-Guilef were kindly provided by Ms F Krouchi (Univ. Tizi-Ouzou), and commercial seed lots were used for the French populations.

Isozyme markers

For the Luberon sample, we used the segregation data of 11–29 megametophytes per progeny (Fallour *et al.*, 2001) to determine the genotype of each seed tree; a tree was declared as heterozygote for a locus when segregation occurred whatever the relative frequency of each allele. Otherwise, it was declared as homozygote. Out of the 12 polymorphic loci, three showed constant segregation distortion with a deficit of the rarest alleles in all the progenies from heterozygous seed trees (*Idh*, *Acp-c*, *Got-a*): tight linkage with selected genes or even direct selection was thus suspected for these loci (Fallour *et al.*, 2001), and we refer to them as selected markers hereafter. The nine other loci that showed Mendelian segregation in all progenies (*Acp-a*, *Lap-a*, *Lap-b*), or segregation distortion in some progenies but not always in the same direction (*Got-b*, *Mdh-c*, *Mnr-a*, *Mnr-c*, *Pgi-b*, *Skdh*), were considered as neutral markers. Two markers were linked, *Acp-c* and *Got-b*; the distance between these markers was estimated as 17 cM but the recombination rate varied from 0.05 to 0.28 among seed trees, suggesting the existence of selected genes in the genomic region (Fallour *et al.*, 2001). For the reference populations, diploid genotypes were determined on germinated embryos with the same isozyme markers, following the protocols described by Fallour *et al.* (2001).

Monolocus diversity parameters such as mean number of alleles per locus, observed (H_o) and expected heterozygosity (H_e , unbiased estimate following Nei, 1978), and F_{is} values were computed using the GENETIX software (Belkhir *et al.*, 1998). Sampling error and confidence intervals were based on a bootstrap procedure (1000 iterations) using default parameter settings. For differentiation among subpopulations, the GENEPOP software (Raymond and Rousset, 1995) was used to perform Fisher's exact test of differentiation on allelic contingency tables (1000 iterations) and F_{st} parameters were estimated with the GENETIX software following Weir and Cockerham (1984). For a multilocus approach, combined two-locus genotypic disequilibria were estimated within each subpopulation following Weir (1990), and tested by bootstrap on monolocus genotypes (1000 iterations) using GENEPOP (Raymond and Rousset, 1995). Bonferroni's correction was used to take into account the high number of tests.

Under several mutation models, a recent bottleneck in a population under mutation-drift equilibrium reduces the number of alleles faster than it affects their frequencies, leading to an excess of diversity (H_e) compared to its expectation with the same number of alleles under the hypothesis of constant population size (Cornuet and Luikart (1996). We used the BOTTLENECK software (Piry *et al.*, 1998) to perform this test although isozymes do not strictly follow such mutation models.

To test the hypothesis of population admixture at introduction, we used the Bayesian approach developed by Dawson and Belkhir (2001): their PARTITION software provides a statistical test of the null hypothesis $k=1$ vs an alternative $k>1$, where k is the number of source populations assumed to be at Hardy-Weinberg and linkage equilibrium. For this analysis, we only considered eight unlinked loci out of the 12: *Acp-c* was dropped because of linkage with *Got-b*, whereas *Lap-a*, *Lap-b* and *Mnr-b* were dropped because of missing data.

As usually observed with isozymes, most of the polymorphism consisted of 'rare alleles', resulting in small sample size within each class of individuals, and a low power of statistical tests based on monolocus parameters. However, there may be a biologically sound information at a global level even when individual tests are not statistically significant, for example, when all rare allele frequencies at different loci vary in the same direction. Here we performed a basic global test of intra- and interlocus equilibrium using the eight unlinked loci previously mentioned. For each locus, we made a distinction between 'common' and 'rare' alleles based on their frequency in each subpopulation. At equilibrium, and under the null hypothesis of panmixia, we expect no particular associations of the same class of alleles among themselves, neither within locus nor among loci. Rare alleles were scored '1' within each locus, common alleles were scored '0'. For each genotype, we computed the number of rare alleles for these eight loci. This number potentially ranged from 0 to 16. For each subpopulation, we compared the distribution of the number of rare alleles per genotype to its expectation after bootstrapping alleles within each subpopulation (1000 iterations). Computations were made using S-PLUS (1996).

To study the evolution of the genetic load associated with the three selected markers, we compared the distribution of segregation ratio among generations by using a nonparametric Kruskal-Wallis test.

Progeny test

The nursery test involved 100 progenies representing the different zones and generations (Table 1). The seeds collected in 1995 were dehydrated and stored in vacuum-sealed bags at +4°C until spring 1997. In March 1997, the seeds were germinated in the State Experimental Nursery of Aix-en-Provence (France). Emerging seedlings were individually transplanted in 400 cm³ containers with a planting medium (2/3 peat and 1/3 pine bark) usually used for *Cedrus* reforestation material. Five blocks were consecutively installed this way during the 15 days of seedling emergence.

The experimental trial consisted of five complete blocks with no within-block replication, each progeny being represented by a single unit plot of five seedlings

per block, that is, 25 seedlings per progeny in total. The five blocks were uniformly treated until August 21st when two different treatments were applied:

- in the stressed treatment (S, blocks 1–3), irrigation was completely stopped in order to generate drought conditions;
- in the normal treatment (N, blocks 4 and 5), seedlings were regularly watered during the whole growing season.

Total plant height was measured in both treatments at the end of the growing season. In order to check the mortality in the stressed treatment, plants were rehydrated from November 15th until December 9th when mortality was scored.

Within each block, the 100 progenies were distributed in a regular grid of 5×20 unit plots. The five blocks were contiguous, resulting in a global regular grid of 25×20 unit plots. Due to material constraints, the design was only partially randomised within each block and subpopulations were partially confounded with row structure within a block. Therefore, rather than using a split-plot model to analyse the data, we used the iterated Papadakis procedure (Pichot, 1993) to adjust the raw data to the heterogeneity of environmental conditions. This method does not require any *a priori* assumption on the distribution of spatial heterogeneity (Pichot, 1993). It also allowed us to include the spatial information at a broader scale than just the block. The method consists of a succession of ANOVAs, as follows: at iteration n , the ANOVA model combines the genotype and a covariate corresponding to the mean residual value of the eight nearest neighbours from the previous $n-1$ step. Iterations proceed until convergence. This procedure provided the data set adjusted for spatial heterogeneity (including block effect) that we used for further analysis. For height, the raw data were the individual measures. For survival, the raw data were the survival rates per unit plot arcsine transformed for homoscedasticity. Narrow sense heritabilities were estimated as

$h^2 = 4\sigma_{\text{fam}}^2 / (\sigma_{\text{fam}}^2 + \sigma_e^2)$, where σ_{fam}^2 and σ_e^2 represent the between- and within-family components of variances, respectively, using restricted maximum likelihood estimates. The genotypic variance (measured on family means) was compared among the different generations using a Bartlett test. All computations were made with S-PLUS (1996). Routines for Papadakis procedure are available on request.

Results

Impact of founding: bottleneck or admixture?

The overall isozyme diversity in the Luberon forest ($H_e = 0.209$) was similar to that of native populations of *C. atlantica*, and significantly higher than that of other introduced populations in France (Table 2). The overall fixation index was lower in the Luberon forest ($F_{\text{is}} = -0.043$) than in the Algerian population under study ($F_{\text{is}} = +0.064$), although not significant (Table 2). Globally, when the three generations were pooled together, we observed no spatial structure among the three sampled zones.

For the founder population (generation G0), there was no significant differentiation (F_{st}) among zones, and no significant deficit in heterozygotes within each zone that would have suggested a Wahlund effect (Table 3, all F_{is} values were nonsignificant). We observed a slightly lower diversity in zone 3 compared to the other zones at this generation.

The test of population admixture performed in the founder gene pool following Dawson and Belkhir (2001) was not significant. Similarly, at multilocus level, no single two-locus genotypic disequilibrium was significant at G0 after Bonferroni's correction for multiple tests. However, the distribution of the number of rare alleles per individual in generation G0 differed significantly from its expectation under the hypothesis of random association of alleles (Figure 3); there was an excess of individuals with no rare alleles and a deficit of individuals with one rare allele, that is, rare alleles were

Table 2 Genetic diversity of isozyme loci in various *Cedrus* populations

	Sample size	H_e	F_{is}	Reference
<i>C. atlantica</i> artificial populations in France				
Luberon (all together)	186	0.209 (0.013)	-0.043 NS	This study (data source: Fallour <i>et al</i> , 2001)
Ventoux	32	0.125 (0.019)	+0.023 NS	This study
Saumon	58	0.158 (0.018)	-0.027 NS	This study
Marcelly	>75	0.124		Scaltsoyiannes (1999) ^a
<i>C. atlantica</i> populations in the natural range				
Tala-Guilef (Algeria)	130	0.186 (0.012)	+0.064 NS	This study
Djurdjura (Algeria)	>75	0.202		Scaltsoyiannes (1999)
Babors (Algeria)	>75	0.228		Scaltsoyiannes (1999)
Ouarsenis (Algeria)	>75	0.229		Scaltsoyiannes (1999)
Aures (Algeria)	>75	0.155		Scaltsoyiannes (1999)
Ich_N'Timghilt (Morocco)	>75	0.158		Scaltsoyiannes (1999)
<i>C. libani</i> (mean of 10 populations)		0.232		Scaltsoyiannes (1999)
<i>C. brevifolia</i> (one population)		0.316		Scaltsoyiannes (1999)
<i>C. deodara</i> (mean of four populations)		0.136		Scaltsoyiannes (1999)

H_e : expected heterozygosity, unbiased estimate (sampling standard deviation); F_{is} : deficit of heterozygotes (NS: not significantly different from 0).

^aBased on isozyme analysis of 75–100 megagametophytes per population (haploid data).

Table 3 Genetic diversity within each subpopulation of the Luberon forest

		Nine neutral loci	Three 'selected' loci	12 loci together
<i>Whole stand</i>				
A	G0	2.667 (0.106)	2.333 (0.119)	2.583 (0.084)
	G1	2.556 (0.043)	2.667 (0.121)	2.583 (0.044)
	G2	2.444 (0.091)	2.667 (0.188)	2.500 (0.085)
H_e	G0	0.203 (0.013)	0.195 (0.022)	0.201 (0.011)
	G1	0.212 (0.011)	0.208 (0.018)	0.211 (0.010)
	G2	0.219 (0.017)	0.209 (0.032)	0.216 (0.015)
H_o	G0	0.212	0.199	0.208
	G1	0.216	0.226	0.218
	G2	0.230	0.216	0.226
<i>Zone 1</i>				
A	G0	2.222 (0.121)	2.000 (0.191)	2.167 (0.100)
	G1	2.000 (0.089)	2.667 (0.220)	2.167 (0.084)
	G2	2.222 (0.098)	2.667 (0.200)	2.333 (0.089)
H_e	G0	0.181 (0.021)	0.194 (0.037)	0.184 (0.019)
	G1	0.199 (0.017)	0.248 (0.039)	0.211 (0.016)
	G2	0.212 (0.020)	0.264 (0.041)	0.225 (0.019)
H_o	G0	0.191	0.200	0.193
	G1	0.216	0.245	0.223
	G2	0.210	0.287	0.229
<i>Zone 2</i>				
A	G0	2.333 (0.122)	2.000 (0.186)	2.250 (0.102)
	G1	2.333 (0.080)	2.000 (0.084)	2.250 (0.064)
	G2	2.250 (0.102)	2.000 (0.254)	2.182 (0.098)
H_e	G0	0.238 (0.025)	0.191 (0.036)	0.227 (0.021)
	G1	0.236 (0.020)	0.211 (0.029)	0.230 (0.016)
	G2	0.261 (0.028)	0.097 (0.040)	0.216 (0.023)
H_o	G0	0.235	0.197	0.225
	G1	0.248	0.221	0.241
	G2	0.303	0.100	0.248
<i>Zone 3</i>				
A	G0	1.889 (0.077)	2.000 (0.085)	1.917 (0.058)
	G1	2.444 (0.090)	2.000 (0.162)	2.333 (0.079)
H_e	G0	0.171 (0.019)	0.205 (0.036)	0.179 (0.016)
	G1	0.195 (0.017)	0.172 (0.022)	0.189 (0.014)
H_o	G0	0.191	0.200	0.193
	G1	0.185	0.215	0.192

G0, G1, G2: three generations of trees; A: mean nb of alleles per locus (sampling standard deviation); H_e : expected heterozygosity, unbiased estimate (sampling standard deviation); H_o : observed heterozygosity.

preferentially grouped by two or more in a few individuals.

Secondary evolution of the nine neutral markers

Isozyme diversity did not decrease from the oldest to the youngest generation, but rather increased both globally

and within each zone, in particular in zone 3 (Table 3). Only one temporal difference was significant: between G0 and G1 in zone 1 ($F_{st}=0.030$, $P=0.05$). No bottleneck effect was detected at any generation, neither in any zone nor globally. A slight spatial differentiation appeared at generation G1 between zones 1 and 3 ($F_{st}=0.017$, $P=0.034$). There was no differentiation between zone 2 and zones 1 and 3 at any generation.

The observed heterozygosity, H_o , globally increased from the founder G0 to the G2 generation (from 0.208 to 0.226). It also increased within zones 1 and 2 and was stable within zone 3 (Table 3). Thus, the younger trees were more heterozygous than the older ones. However, all F_{is} values remained nonsignificant.

The pattern of multilocus associations revealed a clear redistribution of rare alleles among individuals over generations (Figure 3). In particular, the number of individuals with 0 or 1 rare allele reached their expected values as soon as generation G1.

Secondary evolution of the three selected markers

The rare alleles in constant segregation distortion were not lost in the youngest generation, and the diversity at these loci did not decrease, neither globally nor within each zone (Table 3). We did not detect any difference among generations in the distribution of segregation ratio for *Idh* and *Acp-c* loci (29 and 76 heterozygous seed trees, resp.), that is, the segregation distortion was the same in the younger age classes as in the founder trees. This could not be tested for *Got-a* because there were only three heterozygous seed trees at this locus.

Diversity and evolution of quantitative traits

Total height in normal conditions showed a high genetic variation among families ($h^2=0.43$, $P=0.00$). The various subpopulations differed for this trait, as revealed by the significant zone \times generation interaction, but there was no general difference, neither among zones nor among age classes (Table 4). The genotypic variance did not differ among zones ($P=0.65$), nor among generations ($P=0.53$).

In stressed conditions, we observed a surprisingly high level of genetic variation for survival, ranging from 1 to 99% depending on the family ($h^2=0.15$, $P=0.02$), but no difference among zones or generations as for the previous trait (Table 5). Similarly, the genotypic variance did not differ among zones ($P=0.54$), nor among generations ($P=0.89$).

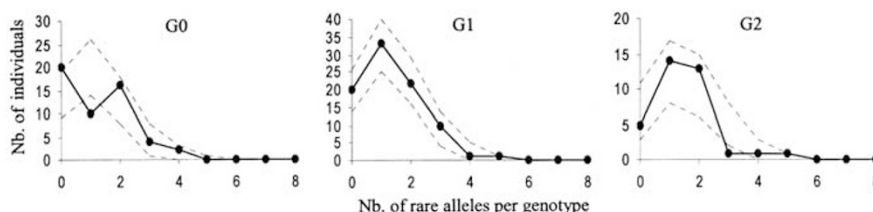


Figure 3 Distribution of the number of rare alleles per genotype over eight unlinked loci (potentially 0–16) for each generation. Dotted lines represent the 95% confidence intervals of expected values under the null hypothesis of random association within each generation (trees from the same generation were pooled among zones).

Table 4 Analysis of variance for total height in normal conditions

	<i>Df</i>	<i>Mean square</i>	<i>F</i>	<i>P-value</i>
Genotype (total)	99	10 111.51	7.55	0.000
Zone	2	28 781.48	0.69	0.565
Generation	2	19 699.27	0.47	0.662
Zone × Generation	3	41 481.19	4.79	0.004
Genotype in (Zone × Generation)	92	8661.79	6.46	0.000
Residuals	774	1339.62		

Data were previously adjusted to environmental effects using Papadakis' method (see text). Adjusted mean squares are provided, main effects are tested relative to their interaction.

Table 5 Analysis of variance for survival (arcsine transformed) in stressed conditions

	<i>Df</i>	<i>Mean square</i>	<i>F</i>	<i>P-value</i>
Genotype (total)	99	0.2599	1.44	0.020
Zone	2	0.4122	0.75	0.544
Generation	2	0.3768	0.69	0.568
Zone × Generation	3	0.5485	2.22	0.091
Genotype in (Zone × Generation)	92	0.2471	1.37	0.041
Residuals	160	0.1802		

Data were previously adjusted to environmental effects using Papadakis' method (see text). Adjusted mean squares are provided, main effects are tested relative to their interaction.

Discussion

A high level of diversity was found in the Luberon cedar forest, similar to that of the native populations. No bottleneck was detected in the G0 founder gene pool at monolocus level, and no differentiation was observed among the founder gene pools of the three sampled zones. This suggests that many seed trees had been originally collected in Algeria and that the same seed lot was used for afforesting the three districts. No departure from equilibrium could be detected using either monolocus parameters (Wahlund effect on F_{is}), two-locus disequilibrium, or the test of general equilibrium developed by Dawson and Belkhir (2001). We suggest that this is due to the low power of these tests because of many rare alleles. We developed a simple test of multilocus associations of these rare alleles, and detected an initial genetic disequilibrium in the founder gene pool, replaced by a random association of alleles in the following generations. This strongly suggests population admixture at founding.

During secondary evolution after establishment, we neither observed a reduction of genetic diversity on adaptive nor on neutral genes. A single temporal difference was detected in zone 1 at the first regeneration phase. This indicates that genetic drift did not occur severely in the Luberon except in zone 1, where initial survival was also the lowest according to aerial photographs. The stable frequency of the distorted alleles associated to deleterious genes and the stability of segregation distortion over the three generations suggest that no purge of the genetic load has occurred in this population. Furthermore, the maintenance of distorted alleles in the natural populations (all heterozygous founder trees showed the same distortion with a deficit

of the same allele) indicates that purging the genetic load was not efficient in a larger time scale either. Byers and Waller (1999) did not find clear evidence of a purge of the genetic load in small populations, and attributed this to the effect of genetic drift which reduces the impact of selection. This cannot be the only explanation in our case, because genetic drift was not detected for the neutral markers. Neither can the limited impact of selection on the genetic load be explained by demographic changes after founding because selection is even more effective in expanding populations (Otto and Whitlock, 1997). Therefore, we conclude that the accumulation of individual genetic contributions over years participates in the maintenance of the genetic diversity. Annual stochasticity of pollen and seed production, combined with annual variation of climatic conditions, simultaneously reduces the impact of fertility variation (Dodd and Silvertown, 2000) and selection in two ways: (i) temporal stochasticity increases the chance of reproductive success of any individual over years; (ii) the variation of cumulative fertility over years may be lower than the variation of any single year (Krouchi *et al*, 2004).

Transplantation is a major perturbation that enhances the effect of selection on adaptive traits (Endler, 1986), and we would expect a rapid evolution of quantitative traits that are not purely neutral. Indeed, the studied population was originally established in stressful conditions that influenced survival rate in the founder population. In the nursery test, however, we observed a significant additive genetic variation for survival and growth traits under stress, but no change over generations, neither in means nor in variance. The maintenance of genetic variance in introduced tree populations was also observed by Kaufman and Smouse (2001). Excluding the effect of mutation at this time scale, the maintenance of genetic variance is classically explained by the infinitesimal model, which considers that many genes of small effect are controlling the traits under study (see Barton and Keightley (2002) for a review). Here, we observed that individual markers associated with selectively controlled genes did not change in frequency either, although they do not correspond to the infinitesimal model. Therefore, we conclude that two mechanisms contribute to reduce the impact of selection in this tree population: temporal stochasticity, as previously invoked, and spatial heterogeneity of environmental conditions. Such an interpretation seems likely in the Luberon forest. A pedological study (Fallour, 1998) has shown that soil conditions were highly homogeneous on a broad scale over the whole forest but could drastically vary at local scale due to cracks in the substratum, which is a key factor locally determining water availability to trees.

Although French cedar landraces generally perform better than native populations in common garden experiments (Bariteau and Ferrandes, 1992), we could not detect any impact of selection on gene diversity across the three generations under study. The genetic load was maintained, and seed trees with 'poor' breeding values in terms of survival and growth in controlled conditions were maintained. More information comes from the evolution of the organisation of gene diversity, in particular heterozygosity. Heterozygosity is generally high in forest tree populations. It increases with the age of the trees and F_{is} estimates are generally

highly negative, that is, old trees are more heterozygous than expected (Farris and Mitton, 1984; Yazdani *et al*, 1985; Plessas and Strauss, 1986). This pattern of heterozygosity is considered to be the result of selection (Ledig, 1986). A counterexample was recently found in *Fraxinus* populations. In this case, Morand *et al* (2002) attributed the deficit of heterozygotes to assortative mating and/or the presence of null alleles in their microsatellites. Our naturalised population showed a particular trend: no positive F_{is} was observed in the sample of old founder trees despite population admixture; furthermore, observed heterozygosity increased during the secondary evolution phase and younger trees were slightly more heterozygous than older ones. Such pattern is expected when a Wahlund effect at foundation is balanced by a general trend of increased heterozygosity with age (we only have access to the old survivors of the founder population). Since we observed no decrease in gene diversity but increased heterozygosity in the younger generations, we conclude that the inbreeding hypothesis is a more likely explanation than selection to account for the superiority of this population as a seed source compared to natural populations. The inbreeding hypothesis is also supported by the overall higher heterozygosity (lower F_{is}) observed in the Luberon forest compared to an Algerian population. The genetic value of this germplasm may still evolve in one or another direction when the new generations come to pollen and seed production.

As expected, the initial allelic associations inherited from the structure of diversity among population sources were rapidly erased, as early as the second generation. By contrast, the effect of drift and selection on the genetic diversity was not yet visible. Gene flow is not the main force that balanced drift and selection in this population: no secondary introduction occurred in this area, and the nearest cedar forest, Ventoux, is located 40 km away (probably from the same seed source). We hypothesise that, besides the long juvenile phase mentioned by Austerlitz *et al* (2000), two other life history traits also contribute to the high within-population diversity found in trees: perenniality and large spatial scale. First, perenniality increases the effective population size when contributions to the regeneration are cumulated over years, due to temporal stochasticity. Second, the large spatial scale of tree populations limits the effect of selection due to spatial heterogeneity of environmental conditions at this scale. In other words, in a recently established tree population founded from various sources, new genotypic combinations emerge more rapidly than genetic erosion occurs through genetic drift and selection. Under this hypothesis, the rapid genetic changes observed in translocated tree populations may result from the creation of new genotypes, based on the diversity of founding sources, and not only from the erosion of pre-existing diversity. Therefore, when secondary gene flow is limited, we would expect a higher level of genetic diversity within recently established tree populations than in older ones. This mechanism could partly explain the spatial pattern of gene diversity observed in beech (Comps *et al*, 2001): gene diversity decreases with population age from the most recently colonised regions to the initial refugia, whereas allelic richness, which is more rapidly affected by founder events (Cornuet and Luikart, 1996), shows the opposite

trend. Referring to the review by Hamrick *et al* (1992), population age since establishment could partly explain the higher diversity found within populations of boreal-temperate tree species ($Hep = 0.204$) compared with tropical ones ($Hep = 0.125$), although these groups of species do not differ in their overall genetic diversity at species level ($Hes = 0.206$ and 0.191 , resp.). Hamrick *et al* (1992) also observed a significantly higher within-population diversity in long-lived woody perennials compared to short-lived woody perennials, but this may not directly be attributed to the effect of perenniality previously mentioned because the two groups also differ in their overall diversity at the species level. Finally, Hamrick *et al* (1992) did not find a significant difference among successional status for within-population diversity, which could be due to the fact that both recently established and 'old' populations can be found in each species.

The Luberon cedar forest is still a 'young' and diverse population, but genetic erosion through drift and selection may progressively become more effective: management should avoid accelerating this process. We observed a lower genetic diversity in the other French *Cedrus* forests (considering that commercial seed lots are supposed to be collected from more than 30 trees in each stand, we assume they provide a good sample of the populations). This difference should be further investigated but it is noticeable that, in these forests, the oldest trees were rapidly exploited after the first regeneration waves due to their limited economic value and bad shape (branchiness). This has reduced their long-term contribution across generations and potentially allowed a higher impact of genetic drift and natural selection on the diversity. In the Luberon forest, the old founder trees were preserved for their patrimonial value; they still contribute to the maintenance of the high level of genetic diversity in this stand.

Acknowledgements

This work was part of the project 'Adaptation and selection of Mediterranean *Pinus* and *Cedrus* for sustainable afforestation of marginal lands' FAIR CT95-0097 financed by the Commission of the European Communities DGVI. We thank D Hibbs and C Pichot for critically reviewing a previous version of the manuscript, and two anonymous reviewers for their helpful comments and suggestions. We thank A Aliberti, G Bettachini, A Gai-Checa, B Jouaud, F Rei and J Thévenet for seed collection and processing, I Hochu and T Nieuwjaer for laboratory assistance, and P Brahic and co-workers who took care of the nursery test in the State Experimental Nursery of Aix-en-Provence.

References

- Austerlitz F, Mariette S, Machon N, Gouyon PH, Godelle B (2000). Effects of colonization processes on genetic diversity: differences between annual plants and tree species. *Genetics* **154**: 1309–1321.
- Bariteau M, Ferrandes P (1992). Les Cèdres. In: Gallais A, Bannerot H (eds) *Amélioration des Espèces Végétales Cultivées*, INRA Editions: Paris, France. pp 732–743.
- Barton NH, Keightley PD (2002). Understanding quantitative genetic variation. *Nat Rev Genet* **3**: 11–21.

- Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F (1998) GENETIX, software for population genetic data analysis, Laboratoire Génome et Populations, CNRS UPR 9060, Université de Montpellier II, Montpellier, France.
- Byers DL, Waller DM (1999). Do plant populations purge their genetic load? *Effects of population size and mating history on inbreeding depression*. *Annu Rev Ecol Syst* **30**: 497–513.
- Carson HL (1990). Increased genetic variance after a population bottleneck. *Trends Ecol Evol* **5**: 228–230.
- Cointat M (1996). Le roman du cèdre. *Rev For Fr* **48**: 503–526.
- Comps B, Gömöry D, Letouzey J, Thiébaud B, Petit RJ (2001). Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics* **157**: 389–397.
- Cornelius J (1993). Heritabilities and additive genetic coefficients of variation in forest trees. *Can J Forest Res* **24**: 372–379.
- Cornuet JM, Luikart G (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**: 2001–2014.
- Dawson KJ, Belkhir K (2001). A Bayesian approach to the identification of panmictic populations and the assignment of individuals. *Genet Res* **78**: 59–77.
- Dodd ME, Silvertown J (2000). Size-specific fecundity and the influence of lifetime size variation upon effective population size in *Abies balsamea*. *Heredity* **85**: 604–609.
- Ducouso A, Guyon JP, Kremer A (1996). Latitudinal and altitudinal variation of bud burst in western populations of sessile oak (*Quercus petraea* (Matt.) Liebl.) *Ann Sci For* **53**: 775–782.
- Endler JA (1986). Natural Selection in the Wild. Monographs in Population Biology N°21, May RM (ed), Princeton University Press: Princeton, New Jersey, USA.
- Epperson BK (1992). Spatial structure of genetic variation within populations of forest trees. *New Forests* **6**: 257–278.
- Espinel S, Aragonés A, Ritter E (1995). Performance of different provenances and of the local population of the Monterey pine (*Pinus radiata* D. Don.) in northern Spain. *Ann Sci For* **52**: 515–519.
- Fallour D (1998). Evolution et structuration spatiale de la diversité du cèdre de l'Atlas sur le Petit Luberon: approches écologique, dendroécologique et génétique. PhD Thesis, University of Aix-Marseille III, France.
- Fallour D, Fady B, Lefèvre F (2001). Evidence of variation in segregation patterns within a *Cedrus* population. *J Hered* **92**: 260–266.
- Farris MA, Mitton JB (1984). Population density, outcrossing rate, and heterozygote superiority in Ponderosa pine. *Evolution* **38**: 151–1154.
- Hamrick JL, Godt MJW, Sherman-Broyles SL (1992). Factors influencing levels of genetic diversity in woody plant species. *New Forests* **6**: 95–124.
- Kaufman SR, Smouse PE (2001). Comparing indigenous and introduced populations of *Melaleuca quinquenervia* (Cav.) Blake: response of seedlings to water and pH levels. *Oecologia* **127**: 487–494.
- Krouchi F, Derridj A, Lefèvre F (2004). Year and tree effect on reproductive organisation of *Cedrus atlantica* in a natural forest. *Forest Ecol Manage* **197**: 181–189.
- Le Corre V, Machon N, Petit RJ, Kremer A (1997). Colonization with long-distance seed dispersal and genetic structure of maternally inherited genes in forest trees: a simulation study. *Genet Res* **69**: 117–125.
- Ledig FT (1986). Heterozygosity, heterosis, and fitness in outbreeding plants. In: Soulé ME (ed) *Conservation Biology: The Science of Scarcity and Diversity*, Sinauer Associates, Inc. Publishers: MA, Sunderland. pp 77–104.
- Ledig FT (2000). Founder effects and the genetic structure of Coulter pine. *J Hered* **91**: 308–315.
- Lenormand T (2002). Gene flow and the limits to natural selection. *Trends Ecol Evol* **17**: 183–189.
- Morand ME, Brachet S, Rossignol P, Dufour J, Frascaria-Lacoste N (2002). A generalized heterozygote deficiency assessed with microsatellites in French common ash populations. *Mol Ecol* **11**: 377–385.
- Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583–590.
- Otto SP, Whitlock MC (1997). The probability of fixation in populations of changing size. *Genetics* **146**: 723–733.
- Peterken GF (2001). Ecological effects of introduced tree species in Britain. *Forest Ecol Manage* **141**: 31–42.
- Pichot C (1993). Analyse de dispositifs par approches itératives prenant en compte les performances des plus proches voisins. *Agronomie* **13**: 109–119.
- Pinyopusarek K, Williams ER (2000). Range-wide provenance variation in growth and morphological characteristics of *Casuarina equisetifolia* grown in Northern Australia. *Forest Ecol Manage* **134**: 219–232.
- Piry S, Luikart G, Cornuet JM (1998). *BOTTLENECK. A program for detecting recent effective population size reductions from allelic frequency data*, INRA-URLB, Laboratoire de Modélisation et Biologie Evolutive, Montpellier, France.
- Plessas ME, Strauss SH (1986). Allozyme differentiation among populations, stands, and cohorts in Monterey pine. *Can J Forest Res* **16**: 1155–1164.
- Raymond M, Rousset F (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Hered* **86**: 248–249.
- Reed DH, Frankham R (2001). How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* **55**: 1095–1103.
- Rehfeldt GE, Wykoff WR, Ying CC (2001). Physiological plasticity, evolution, and impacts of a changing climate on *Pinus contorta*. *Climat Change* **50**: 355–376.
- Reznick DN, Ghalambor CK (2001). The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* **112/113**: 183–198.
- Skrøppa T, Kohmann K (1997). Adaptation to local conditions after one generation in Norway spruce. *Forest Genet* **4**: 171–177.
- S-PLUS (1996). Version 3.4. MathSoft, Inc.: Seattle, Washington.
- Scaltsioyannes A (1999). Allozyme differentiation and phylogeny of cedar species. *Silvae Genet* **48**: 61–68.
- Toth J (1978). Contribution à l'étude de la fructification et de la régénération du cèdre de l'Atlas (*Cedrus atlantica* M.) dans le sud de la France. PhD Thesis, Faculté des Sciences et Techniques Aix-Marseille III, France.
- Vicaire H (1864). Reboisement des montagnes. Compte-rendu des travaux de 1863. Ministère des Finances, Administration des Forêts, Paris, Imprimerie Impériale.
- Weir BS (1990). *Genetic Data Analysis. Methods for Discrete Population Genetic Data*, Sinauer Associates, Inc. Publishers: Sunderland, MA.
- Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.
- Yazdani R, Muona O, Rudin D, Szmidi AE (1985). Genetic structure of a *Pinus sylvestris* L. seed-tree stand and naturally regenerated understory. *Forest Sci* **31**: 430436.
- Zheng YQ, Ennos RA (1999). Genetic variability and structure of natural and domesticated populations of Caribbean pine (*Pinus caribaea* Morelet). *Theor Appl Genet* **98**: 765–771.