# Population differentiation of sessile oak at the altitudinal front of migration in the French Pyrenees

F. ALBERTO,*† J. NIORT,*† J. DERORY,*† O. LEPAIS,*† R. VITALIS,‡ D. GALOP§ and A. KREMER*†

*INRA, UMR 1202 Biodiversité Gènes et Communautés, Cestas, F-33610, France, †Université de Bordeaux, UMR1202 Biodiversité Gènes et Communautés, Talence, F-33400, France, ‡CNRS, UMR CBGP (INRA – IRD – CIRAD – Montpellier SupAgro), Campus International de Baillarguet, CS 30016, Montferrier-sur-Lez, F-34988, France, §CNRS, UMR 5602 Géographie de l'Environnement, Maison de la Recherche de l'Université du Mirail, 5 Allée A. Machado, Toulouse, F-31058, France

## Abstract

**To assess the effects of altitude on the level and structure of genetic diversity, a genetic survey was conducted in 12 populations of sessile oak (*Quercus petraea*) located between 130 and 1660 m in two parallel valleys on the northern side of the Pyrenees Mountains. Genetic diversity was monitored at 16 nuclear microsatellite loci and 5 chloroplast DNA (cpDNA) markers. The cpDNA survey suggested that extant populations in both valleys shared the same source populations from the plain. There was no visible trend of nuclear genetic diversity along altitude, even if indirect estimates of effective population sizes revealed a consistent reduction at higher altitudes. Population differentiation, although low, was mostly present among populations of the same valleys and reached similar levels than differentiation across the range of distribution of sessile oak. Contribution to the overall differentiation in the valleys was mostly due to the genetic divergence of the highest populations and the altitudinal variation of allelic frequencies at a few loci. Bayesian inference of migration between groups of populations showed that gene flow is preferentially unidirectional from lower altitudes in one valley to other groups of populations. Finally, we found evidence of clonal reproduction in high altitude populations. The introgression of *Quercus robur* and *Quercus pubescens* was also more frequent at the altitudinal margin suggesting that this mechanism may have contributed to the present migration and adaptation of *Q. petraea* and may also facilitate its future upslope shift in the context of climate change.**

*Keywords*: altitudinal gradient, chloroplast haplotypes, genetic diversity, hybridization, nuclear microsatellites, *Quercus*

*Received 28 October 2009; revision received 23 February 2010; accepted 3 March 2010*

## Introduction

Altitudinal gradients encompass contrasted environments and impose strong adaptive challenges for plants. At high altitude, environmental constraints as cold temperatures, frost and shorter growing season induce strong selection pressures that can affect reproduction and survival of populations (Premoli 2003). Because located at the ecological margins of the species,

Correspondence: Antoine Kremer, Fax: +33 557 12 28 81;
E-mail: kremer@pierroton.inra.fr

these peripheral populations could exhibit peculiar demographic and biological features as reduced population sizes or ability for clonal reproduction (Ettl & Peterson 2001; Premoli 2003). Isolation of populations can also be strongly increased by the topographical barriers and sharp variations of phenology between lower and higher altitudes. These physical and temporal constraints on gene flow can limit the recruitment and may consequently increase genetic drift. High altitude populations are not only exposed to extreme environmental conditions, but they also face genetic and demographic stochasticity, which may further challenge their survival

(Quiroga & Premoli 2007). As a result strong genetic differentiation between high and low altitude populations may be expected not only for adaptive traits but also for neutral markers (Premoli 2003; Quiroga & Premoli 2007).

Interest for high altitude populations has risen recently, as they may be expanding at even higher altitudes due to climate change (Jump *et al.* 2006; Aitken *et al.* 2008). Surveys of plant inventories have indicated that during the last century their distributions have shifted about 29 m per decade in altitude (Lenoir *et al.* 2008). Consequently, the populations that are currently located at the altitudinal tree species line constitute source populations for their altitudinal expansion. But the upslope shift of these populations as a response to changing climatic conditions will not only depend on their evolutionary potential (i.e. the genetic variability for adaptive traits) but will also depend on gene flow (Byars *et al.* 2007). Gene flow may be particularly important as alleles required for adaptation may actually be present at lower altitudes. In the context of climate change, evolutionary responses of populations at their range limits may therefore be dependent on alleles coming from warmer climates (Bridle & Vines 2007) contradicting the traditional view that gene swamping from central populations will actually limit species expansion (Kirkpatrick & Barton 1997). It is therefore of upmost importance to draw attention to the processes that shape the genetic diversity of populations located along altitudinal gradients (Engler *et al.* 2009). In Europe, as mountains are mostly present under southern latitudes, altitudinal migration is expected for a large spectrum of species whose southern distribution will be challenged by climatic changes (see for example Peñuelas *et al.* 2007). European white oaks are designated case studies in this respect. Temperate white oaks (*Quercus petraea*, *Quercus robur*) are widely distributed across Europe from Spain to the Ural Mountains and their southern extant limits extend right at the southern slopes of European Mountains (Pyrenees, Alps, Carpathians), where they cooccur with some of the Mediterranean white oaks (mainly *Quercus pubescens* and *Quercus pyrenaica*). Along their post glacial migration from the glacial refugial zones (Petit *et al.* 2002b), the two temperate oak species have also expanded within valleys of European mountains (Csaikl *et al.* 2002). Among the four temperate white oak species present along the northern side of the Pyrenees Mountains (*Q. petraea*, *Q. robur*, *Q. pubescens* and *Q. pyrenaica*), forest inventory data indicate that only *Q. petraea* and *Q. robur* extend at higher altitudes and can reach up to 1800 m. Relative abundance of *Q. petraea* increases above 1600 m, where *Q. robur* is only occasionally present. *Q. petraea* and *Q. robur* had already

reached these limits between 11 000 and 10 000 BP (Aubert 2001; Belet 2001). A recent study on hybridization dynamics among these four oak species showed that introgression is possible between all species but depends on species abundance (Lepais *et al.* 2009). A comparative study of various broadleaved species distributed in the Pyrenees Mountains has suggested that oaks may benefit more from climatic changes in terms of growing season length and are prone to faster altitudinal expansion than other species during the next decades (Vitasse *et al.* 2009a,b).

To understand the processes that may stimulate future altitudinal expansion of oaks we have concentrated our attention to a retrospective analysis of the processes that have shaped the current genetic variation along altitudinal gradients in the Pyrenees. We chose to focus on *Q. petraea* which reaches higher altitudes than the other white oaks and monitored genetic diversity along the altitudinal distribution of this species. The investigations were replicated within two altitudinal transects corresponding to two valleys. We used chloroplast DNA markers and nuclear microsatellites to retrace the historical migration patterns along the valleys and to characterize the extant genetic differentiation among populations. We wanted first to identify historical and contemporary processes that may have proceeded along the whole altitudinal gradient and contributed to the current genetic structure of populations. Second, we attempted to check whether high altitude populations located at the margin of the species' range exhibit peculiar traits that would facilitate their maintenance or future expansion as climate change is proceeding.

## Materials and methods

### Study area and sampling

Twelve populations of *Q. petraea* were sampled along two parallel valleys (Ossau and Luz) on the northern side of the Pyrenees Mountains (Fig. 1a and b). These two valleys are approximately 30 km apart and extend South–North. Oaks are spread along the northern side of the Pyrenees and occur in continuous stands at lower altitudes and in scattered isolated stands at their higher limits. Four species are present with relative abundances decreasing from *Q. robur*, *Q. petraea*, *Q pubescens*, to *Q. pyrenaica*. *Q. pyrenaica* extends only in lowlands, while *Q. pubescens* spreads throughout the foothills below 800 m (http://www.ifn.fr). The altitudinal range of pedunculate oak (*Q. robur*) extends within the Pyrenees valleys up to 1600 m, with occasional stands present up to 1800 m. Similarly the current altitudinal distribution of sessile oak (*Q. petraea*) ranges from quite
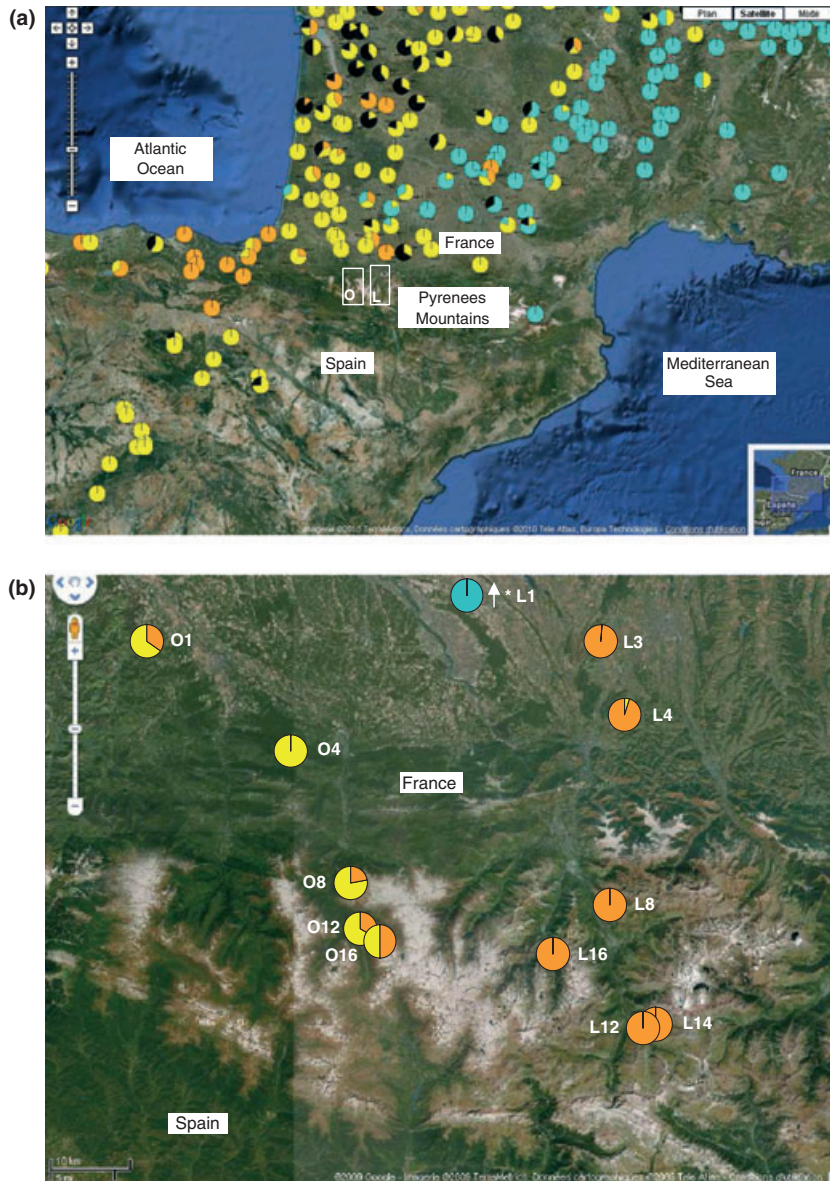
(a)



Fig. 1 Geographic maps of chloroplast haplotype frequencies (a) localization of the two studied Pyrenean valleys and distribution of chloroplast haplotypes along the Pyrenees foothills (adapted from Petit *et al.* (2002a) and (b) localization and chloroplast haplotype frequencies of the studied populations. Pie charts represent the haplotype frequencies in the populations sampled in Petit *et al.* (2002a). Blue: haplotype 7; Yellow: haplotype 10; Orange: haplotype 12; Black: any other haplotypes bulked together. The two white rectangles indicates the location of the two valleys in which populations were sampled (O: Ossau valley; L: Luz valley). * L1 is located approximately 50 km north of the position indicated on the map.

(b)



continuous forests in the plains with tall trees of almost 30 m to stands of smaller trees distributed in scattered populations at high altitudes (up to 1800 m), and is more frequent at higher altitudes than *Q. robur*. We explored the two valleys and used the data of the French National Inventory (http://www.ifn.fr) to sample the study populations. In each valley, we sampled natural populations distributed along an altitudinal gradient varying between 131 and 1630 m (Table 1 and Fig. 1) and over a distance of 80 km. At least 20 individuals were sampled in each population except in populations O8 and O16 where only 10 and 14 trees, respectively, were available. While the oak study populations at lower altitudes were sampled from widespread stands (a few hundreds of hectares), study populations at higher altitudes were sampled within scattered patches of a few hectares and in two cases the sampling was exhaustive (O8 and O16). The sampling was therefore representative of the overall distribution of oaks in the Pyrenees valleys. We collected either leaves in mid-summer 2004 or buds in winter 2005 on a total of 295 adult trees distributed among the 12 sampled populations. The taxonomic status of each tree was checked based on the leaf morphology (Kremer *et al.* 2002) assessed during the growing season on fully elongated leaves. DNA was extracted from buds and leaves following a protocol adapted from Doyle & Doyle (1990).

**Table 1** Description of the 12 natural populations these apostrophise sampled

| Valley | Site | Code | Sample size | Altitude (m) | Latitude | Longitude | Group |
|--------|------|------|-------------|--------------|----------|-----------|-------|
| Luz valley | Laveyron | L1 | 20 | 131 | 43°45′N | 00°13′W | LL |
| | Ibos | L3 | 25 | 387 | 43°15′N | 00°00′W | |
| | Adé | L4 | 40 | 427 | 43°08′N | 00°00′W | |
| | Chèze | L8 | 22 | 803 | 42°55′N | 00°02′W | |
| | Gèdre | L12 | 32 | 1235 | 42°47′N | 00°01′E | LH |
| | Gèdre haut | L14 | 23 | 1349 | 42°47′N | 00°02′E | |
| | Péguère | L16 | 25 | 1630 | 42°52′N | 00°07′W | |
| Ossau valley | Josbaig | O1 | 28 | 259 | 43°13′N | 00°44′W | OL |
| | Bager | O4 | 22 | 422 | 43°07′N | 00°32′W | |
| | Le Hourcq | O8 | 10 | 841 | 42°54′N | 00°26′W | |
| | Gabas | O12 | 34 | 1194 | 42°53′N | 00°25′W | OH |
| | Artouste | O16 | 14 | 1614 | 42°53′N | 00°24′W | |

Groups meanings:
LL: Luz valley low altitudes.
LH: Luz valley high altitudes.
OL: Ossau valley low altitudes.
OH: Ossau valley high altitudes.

## Chloroplast haplotypes genotyping

In this study we used a set of five chloroplast markers (μdt1, tf42, dt73b, dt13, dt74b) described by Deguilloux *et al.* (2003). These markers were initially developed in oaks to reconstruct the colonization pathway of *Quercus* sp. after the last glacial episode (Petit *et al.* 2002a) and allow identifying the eight major haplotypes present in France (Petit *et al.* 2002c). We used the same scoring as in Petit *et al.* (2002c) and Deguilloux *et al.* (2003).

## Nuclear microsatellites genotyping

A total of 16 nuclear microsatellite markers previously developed on oaks species (Dow *et al.* 1995; Steinkellner *et al.* 1997; Kampfer *et al.* 1998) were used in this study (QpZAG110, QrZAG11, QrZAG112, QrZAG39, QrZAG96, QrZAG20, QrZAG5, QrZAG65, QrZAG7, QrZAG87, QpZAG15, QpZAG46, QpZAG9, QpZAG1∕5, QpZAG36, MSQ13). These dinucleotide microsatellite markers are located on nine linkage groups of the *Quercus* map (Barreneche *et al.* 2004), and were considered as unlinked. They are anonymous markers, i.e. we do not know whether they are located in coding or noncoding regions of the genome. Among these 16 microsatellites, 10 markers were amplified in two multiplex kits following Lepais *et al.* (2006) and the other 6 markers were grouped in two distinct multiplex kits (QpZAG15 with QpZAG46, QpZAG9, and QpZAG1∕5, and QpZAG36 with MSQ13). Polymerase Chain Reaction (PCR) amplifications were all performed in the conditions described by Lepais *et al.* (2006). PCR product electrophoresis were performed with a Megabace1000 multicapillary sequencer (Amersham Biosciences Molecular Dynamic, Uppsala, Sweden) and the genotypes were scored and visually controlled using the Fragment Profiler software version 1.2 provided by the manufacturer.

## Genetic diversity and population differentiation

The frequencies of each chloroplast haplotype were calculated in each population.

Gene diversity statistics (gene diversity $H_e$ (Nei 1987) and allelic richness A) were estimated for nuclear microsatellites using the program Fstat version 2.9.3.2 (Goudet 2001). Allelic richness ($A$) was calculated using the rarefaction method developed by El Mousadik & Petit (1996a). Rarefaction was done based on a minimum sample size of nine diploid individuals, which corresponded to the smallest number of individuals successfully genotyped for a given locus in a population. Pairwise differences of genetic diversity and allelic richness were tested by means of a paired *t*-test, following the method of Zhang & Allard (1986) with the statistical program R (R Development Core Team 2005). We performed Bonferroni corrections to account for multiple testing.

For both the chloroplast haplotypes and the nuclear microsatellites we measured among population differentiation as Wright's $F_{ST}$ according to Weir & Cockerham (1984) by performing a locus by locus analysis of molecular variance (AMOVA) using Arlequin version 3.01 (Excoffier *et al.* 2005). We calculated the overall $F_{ST}$ and the $F_{ST}$ between populations within each valley. A hierarchical AMOVA was used to partition the genetic

variation at three levels, among all populations ($F_{ST}$), among populations within valleys ($F_{SV}$) and among valleys ($F_{VT}$).

## Selection imprints

We used the software package BayeScan version 1.0 (Foll & Gaggiotti 2008) to test for the neutrality of the microsatellite markers. BayeScan extends Beaumont & Balding's (2004) Bayesian method that implements the multinomial-Dirichlet likelihood, which arises in a wide range of neutral population genetics models. The rationale is to discriminate between the effects on $F_{ST}$ values that are specific to each population and to each locus. A locus is suspected to be under selection if the locus-specific effect is significantly different from zero. The method developed by Foll & Gaggiotti (2008) uses a hierarchical Bayesian approach to estimate the posterior probabilities of two alternative models, one including the effects of selection and one excluding it. The results are expressed as a Bayes factor which indicates for each locus the ratio of posterior probability of the selection model against the neutral model and can be translated as different levels of evidence of selection according to Jeffreys' interpretation (1998). We ran BayeScan with 10 pilot runs with a burn-in of 500 000 followed by 50 000 iterations each, a sample size of 50 000 and a thinning interval of 20.

For each outlier locus, we examined the overall $F_{ST}$ and the $F_{ST}$ within each valley and we inspected the altitudinal variations of allelic frequencies along the two altitudinal gradients in order to check for potential clinal patterns of variation in both valleys. The loci departing significantly from neutral expectation were removed from further analyses of genetic structure and gene flow (see below).

## Genetic structure of populations

We used the Bayesian clustering method implemented in Structure version 2.3.1 (Pritchard *et al.* 2000) to determine the genetic structure of the sampled populations. We used the new clustering method implemented in this version which is not only based on the individual multilocus genotypes but also takes into account the sampling locations (Hubisz *et al.* 2009). The LocPrior model considers that the prior distribution of cluster assignments can vary among populations. This approach is recommended by the authors when the genetic data are not very informative to help the detection of population structure. A parameter $r$ indicates the extent to which the sampling locations are informative (small values <1 indicate that locations are informative). We conducted five independent runs for each

value of $K$ (the number of putative clusters) ranging from 1 to 12 (the number of populations sampled). We used the admixture model with sampling locations as prior information (Hubisz *et al.* 2009) and assumed correlated allele frequencies among populations (Falush *et al.* 2003). Each run consisted in 500 000 burn-in steps followed by $10^6$ iterations. To determinate the most likely number of clusters we used the method developed by Evanno *et al.* (2005), based on an *ad hoc* statistic, $\Delta K$, which depends on the rate of change in the log probability of the data between successive values of $K$ and we also followed the recommendations of the software documentation (Pritchard *et al.* 2009). Once the most likely $K$ value was determined, we chose the run with the higher posterior probability and lower variance for interpreting results.

## Gene flow and population sizes

We used the Bayesian approach implemented in Migrate version 3.0.3 (Beerli & Felsenstein 2001; Beerli 2006) to assess the direction and amount of gene flow among populations. Based on the genetic data this software estimates the scaled effective population size ($\theta = 4N\mu$) for each population and the scaled immigration rates ($M = m/\mu$) between each pair of populations. The program Migrate calculates the posterior probability distribution of these parameters, using a Metropolis–Hastings algorithm to explore all possible genealogies. A coalescent simulation explores the likelihood space for $\theta$ and $M$. Within each valley we grouped populations in two groups according to altitude (Table 1) which limited the number of parameters to estimate. We performed Migrate using the Brownian motion model with starting values of $\theta$ and $M$ generated from the $F_{ST}$ estimates, exponential window prior for both parameters (min = 0, mean = 5, max = 50 and $\Delta$ = 5 for $\theta$; min = 0, mean = 5, max = 50 and $\Delta$ = 5 for $M$), slice sampling (Neal 2003) for the proposal distribution, and mutation rates calculated from data for each locus. We replicated two long chains of $10^6$ genealogies recorded every 20 steps after a burn-in period of 20 000. Then we used the resulting $\theta$ and $M$ estimates as starting values for another run in order to test the reliability of the analysis. We repeated this procedure until the starting parameters and the resulting estimates were congruent. If necessary we adjusted the $\Delta$ values of the exponential window priors to get a better sample of the parameter landscapes.

## Introgression

We further used the software Structure for species assignment and introgression analysis since morphological

variation is multimodal but overlapping across trees of different species (Kremer *et al.* 2002) and introgressed trees could not be excluded. Introgression analysis was previously performed by Lepais *et al.* (2009) using a larger sample of 2107 oak trees representing the four oak species present in the study area (*Q. petraea*, *Q. robur*, *Q. pubescens*, *Q. pyrenaica*) and including the sampled trees and was conducted using a subset of 10 microsatellites (for details see Lepais *et al.* 2009). Bayesian clustering of the trees was achieved by using Structure version 2.1 (Pritchard *et al.* 2000; Falush *et al.* 2003) and results were analyzed for the most likely number of clusters ($K = 4$). Each species was represented by one cluster and the admixture coefficient $Q$, corresponding to the assignment probability of each tree to each cluster (Pritchard *et al.* 2000), was used to infer its taxonomic status. Following Lepais *et al.* (2009), individuals were classified as pure species for $Q > 0.90$ (or $Q < 0.90$ for one cluster but $Q < 0.10$ for the three others) and as hybrids for $Q < 0.90$.

## Results

In the two higher altitude populations (L16 and O16), a few trees shared the same multilocus genotype for the 16 loci and were assumed to be clonal copies. Two multilocus genotypes were found in two exact copies in the L16 population and four copies of a single genotype were found in the O16 population. In the L16 population trees sharing the same genotype were separated by one meter, but in O16 population they were located in a sloping avalanche corridor and separated by 5 m each. In addition one tree exhibited a tribanded electrophoregram at 10 loci in L3 population, which suggested a natural triploid genotype (Dzialuk *et al.* 2007). These seven trees were excluded for the genetic analyses, leading to a final data set of 288 genotypes.

### Chloroplast DNA diversity and differentiation

A total of 280 chloroplast haplotypes could be reconstructed, out of 288, due to 8 incomplete genotypes. Three chloroplast haplotypes were identified over the 12 populations (haplotypes 7, 10 and 12 according to Petit *et al.* 2002a; see also Fig. 1b). According to these authors who carried out a continental wide survey of cpDNA variation, haplotypes 10 and 12 belong to the western lineage (lineage B), and haplotype 7 to an eastern lineage (lineage A). These haplotypes were unevenly distributed among the two valleys and populations (Table 2). Five out of seven populations of the Luz valley were fixed for haplotype 12 and one (L1) for haplotype 7. The remaining population of the Luz val-

**Table 2** Frequencies of the chloroplast haplotypes and genetic diversity ($H_e$) and allelic richness ($A$) of the nuclear microsatellites

| diversity | Chloroplast haplotypes frequencies | | | Nuclear microsatellites | |
| | 7 | 10 | 12 | $H_e$ | $A$ |
|---|---|---|---|---|---|
| L1 | 1.00 | 0.00 | 0.00 | 0.832 | 7.847 |
| L3 | 0.00 | 0.00 | 1.00 | 0.820 | 7.823 |
| L4 | 0.00 | 0.05 | 0.95 | 0.819 | 7.917 |
| L8 | 0.00 | 0.00 | 1.00 | 0.837 | 8.231 |
| L12 | 0.00 | 0.00 | 1.00 | 0.836 | 7.929 |
| L14 | 0.00 | 0.00 | 1.00 | 0.829 | 7.653 |
| L16 | 0.00 | 0.00 | 1.00 | 0.836 (0.833) | 8.300 (8.133) |
| O1 | 0.00 | 0.65 | 0.35 | 0.819 | 7.896 |
| O4 | 0.00 | 1.00 | 0.00 | 0.765 | 6.766 |
| O8 | 0.00 | 0.78 | 0.22 | 0.823 | 7.136 |
| O12 | 0.00 | 0.67 | 0.33 | 0.812 | 7.379 |
| O16 | 0.00 | 0.50 | 0.50 | 0.844 (0.821) | 8.089 (7.211) |
| Overall | 0.06 | 0.26 | 0.68 | 0.822 | 8.400 |

Allelic richness ($A$) was computed using the rarefaction method (El Mousadik & Petit 1996a) based on nine individuals. Values between parentheses correspond to the genetic diversity ($H_e$) and allelic richness ($A$) of populations L16 and O16 calculated when clonal copies are included in the sample size.

ley (L4) was almost fixed for haplotype 12 (38 out of 40 trees carried haplotype 12). In the Ossau valley, one population (O4) was fixed for haplotype 10 and the remaining populations were all of mixed composition with haplotypes 10 and 12 distributed roughly as 70–30% in each stand. Hence while the populations of the Luz valley were almost all fixed for two haplotypes and thus highly differentiated the level of differentiation between populations in the Ossau valley was extremely low (Table 3).

**Table 3** Genetic differentiation statistics

| | $F_{SV}$ | $F_{VT}$ | $F_{ST}$ Luz | $F_{ST}$ Ossau |
|---|---|---|---|---|
| Nuclear microsatellites | 0.022*** | 0.003** | 0.018*** | 0.028*** |
| Chloroplast haplotypes | 0.509*** | 0.592** | 0.901*** | 0.106* |

$F_{SV}$: differentiation among populations within valleys.
$F_{VT}$: differentiation between the Luz and Ossau valley.
$F_{ST}$ Ossau: differentiation among populations within the Ossau valley.
$F_{ST}$ Luz: differentiation among populations within the Luz valley.
Differentiation for nuclear microsatellites and chloroplast haplotypes was estimated by AMOVA using Arlequin.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

**Table 4** Genetic statistics and neutrality test performed for the 16 nuclear microsatellites

| Locus | $A$ | $H_e$ | Overall $F_{ST}$ | $F_{ST}$ Luz | $F_{ST}$ Ossau | BF |
|---|---|---|---|---|---|---|
| QpZAG110 | 8.005 | 0.820 | 0.035 | 0.029 | 0.053 | 0.2 |
| QrZAG11 | 6.617 | 0.757 | 0.018 | 0.021 | 0.013 | 0.5 |
| QrZAG112 | 3.388 | 0.372 | 0.015 | 0.002 | 0.029 | 0.5 |
| QrZAG39 | 9.357 | 0.844 | 0.047 | 0.039 | 0.038 | 270.7 |
| QrZAG96 | 8.990 | 0.858 | 0.019 | 0.020 | 0.019 | 0.2 |
| QrZAG20 | 9.181 | 0.878 | 0.016 | 0.008 | 0.026 | 0.3 |
| QrZAG5 | 10.186 | 0.900 | 0.024 | 0.016 | 0.040 | 0.2 |
| QrZAG65 | 12.428 | 0.937 | 0.022 | 0.022 | 0.022 | 0.2 |
| QrZAG7 | 10.535 | 0.901 | 0.028 | 0.024 | 0.030 | 0.2 |
| QrZAG87 | 7.862 | 0.843 | 0.016 | 0.004 | 0.011 | 0.3 |
| QpZAG15 | 7.578 | 0.793 | 0.039 | 0.042 | 0.043 | 7.7 |
| QpZAG46 | 8.094 | 0.859 | 0.015 | 0.012 | 0.010 | 0.3 |
| QpZAG9 | 7.667 | 0.846 | 0.025 | 0.017 | 0.046 | 0.3 |
| QpZAG1/5 | 7.622 | 0.842 | 0.020 | 0.019 | 0.022 | 0.3 |
| QpZAG36 | 9.553 | 0.888 | 0.015 | 0.009 | 0.023 | 0.9 |
| MSQ13 | 7.330 | 0.821 | 0.013 | 0.004 | 0.024 | 1.3 |
| All | 8.400 | 0.823 | 0.023 | 0.018 | 0.028 | |

Genetic diversity ($H_e$) and allelic richness ($A$) were calculated with Fstat, genetic differentiation ($F_{ST}$) was calculated with Arlequin and the Bayes factor BF was obtained with BayeScan.

### Nuclear genetic diversity and allelic richness

Oak populations exhibited high levels of gene diversity across all loci (Table 4). Fifteen of the 16 microsatellites presented high levels of genetic diversity (from 0.757 to 0.937) and allelic richness (from 6.62 to 12.43). Only locus QrZAG112 showed significantly lower levels of genetic diversity (0.372) and allelic richness (3.388). After Bonferroni corrections, no pairwise comparisons of the average level of genetic diversity or the allelic richness were significant. However, both parameters were of similar magnitude across populations except for the O4 population, in which we found the lowest values as compared to the other populations (Table 2). No difference of genetic diversity was noticeable between the two valleys.

### Nuclear genetic differentiation among populations

The partitioning of genetic variance into different components (between and within valleys) using AMOVA, indicated that differentiation resides rather between populations within valleys than between the two valleys. While the differentiation among all 12 populations amounted to 0.023, $F_{VT}$ value between the two valleys amounted only to 0.003 ($p < 0.01$; Table 3). Differentiation was however higher within Ossau valley than within Luz valley ($F_{ST} = 0.028$ and 0.018, respectively). Larger differentiation within Ossau valley was mainly

caused by the divergence of O4 from the other populations as indicated by pairwise $F_{ST}$ values (data not shown). The BayeScan analysis revealed that two loci, QrZAG39 and QpZAG15, presented patterns of genetic differentiation among populations that were not compatible with neutral expectations (Table 4). For QrZAG39, the model including selection effects was clearly favoured (BF = 270.7) corresponding to a decisive evidence of selection according to Jeffrey's interpretation (Jeffreys 1998), whereas QpZAG15 presented a substantial evidence of selection (BF = 7.7). The Bayes factors of the remaining 14 markers were lower than 1.3 which did not witness any selection signature.

Genetic differentiation within valleys varied across loci (Table 4) with three loci, QrZAG110, QrZAG39 and QpZAG15, contributing to most of the genetic differentiation in both valleys. For loci QrZAG39 and QpZAG15 we detected congruent variation of allelic frequencies with altitude. Allele 109 at locus QrZAG39 decreased with increasing altitude and allele 113 at locus QpZAG15 decreased with decreasing altitude (Fig. 2). Although the trend of variation was the same in both valleys, the linear regressions were significant in the Luz valley (slope = $-1.35 \times 10^{-4}$, $r = -0.87$ and $p < 0.01$ for QrZAG39, and slope = $1.70 \times 10^{-4}$, $r = 0.91$ and $p < 0.01$ for QpZAG15), but not in the Ossau valley.

### Genetic structure

The logarithm of the probability of the data [LnP(D)] as a function of K reached a peak for K = 4 (mean values: LnP(D) = −18169.4, Var[LnP(D)] = 934.3) and then decreased, but we found a higher ΔK value for K = 2 than for K = 4 using Evanno's criterion (Evanno *et al.* 2005). According to the software documentation (Pritchard *et al.* 2009), we choose to interpret the results for K = 4 because the two additional clusters exhibited strong assignments with two existing populations (L12
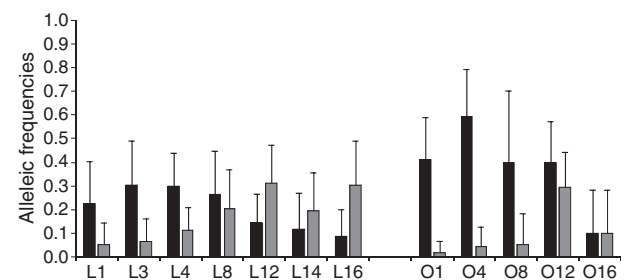


**Fig. 2** Allelic frequencies of allele 109 (locus QrZAG39) and allele 113 (locus QpZAG15) along the two altitudinal gradients. Plots in black: allele 109 of microsatellite QrZAG39. Plots in grey: allele 113 of microsatellite QpZAG15. Bars represent the 95% confidence intervals.
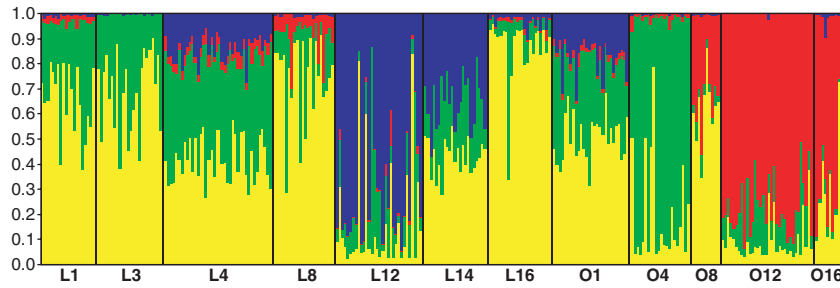
**Fig. 3** Structure clustering results obtained for $K = 4$ with the LocPrior model. Each individual is represented by a thin bar corresponding to the sum of the assignment probabilities to the $K$ clusters. Populations are separated by black bars and identified at the bottom.

for the blue cluster and O4 for the green cluster; Fig. 3). For the most likely run the $r$ value was 0.67, indicating that the sample locations were informative and helped greatly to find the population structure. The genetic structure detected followed the altitudinal gradient with many populations presenting high levels of admixture and few populations almost completely assigned to one cluster (Fig. 3). Low altitude populations were mainly admixed between the yellow and the green cluster. The green cluster was considerably represented in populations located until 400 m with an average assignment probability amounting to 78.5% in population O4 and lower values in the other populations (between 29.9% in O1 and 40.0% in L4). High assignment probabilities to the yellow cluster were found in populations L1, L3, L4, L8, O1 and O8 (between 42.1% in L4 and 74.7% in L8), whereas trees of population L16 were almost totally assigned to this cluster (87.0% on average). The red cluster characterized high altitude populations in Ossau valley where assignment probabilities increased along the altitudinal

gradient from 32.6% in population O8 to 77.5% and 68.5% in populations O12 and O16, respectively. Finally the blue cluster was mainly represented in two high altitude populations of Luz valley (67.2% and 35.2% for L12 and L14 populations, respectively).

*Gene flow and population size*

After three runs of Migrate the results were congruent and we choose the latter run for interpretations. We observed a clear trend between groups of populations located in the two valleys (Table 5). At a same altitudinal level, populations located in Luz valley presented higher $\theta$ values compared to the Ossau valley. In both valleys the scaled effective population sizes decreased noticeably with increasing altitude. The LL group showed thus the highest $\theta$ value. Moreover, almost all significant $M$ values were found between this group and the other groups of populations. The scaled immigration rates $M$ were higher from the LL group to the OL group than to the two high altitudes groups

**Table 5** Results from the Migrate analysis on groups of populations (a) scaled effective population sizes ($\theta$) and (b) scaled immigration rates ($M$)

(a)

| $\theta_{LLc}$ | $\theta_{LH}$ | $\theta_{OL}$ | $\theta_{OH}$ |
|---|---|---|---|
| 14.7 (11.6–17.9) | 9.7 (7.7–11.8) | 7.0 (5.1–9.1) | 2.9 (1.4–4.5) |

(b)

| | LL | LH | OL | OH |
|---|---|---|---|---|
| LL | — | 1.6 (0.4–2.7) | 1.0 (0.1–1.9) | 0.5 (0.0–1.4) |
| LH | 2.3 (1.0–3.5) | — | 1.2 (0.2–2.2) | 0.0 (0.0–0.7) |
| OL | 2.6 (1.4–3.9) | 0.3 (0.0–0.7) | — | 0.0 (0.0–0.7) |
| OH | 2.4 (1.0–3.7) | 0.0 (0.0–0.7) | 0.0 (0.0–0.7) | — |

$\theta = 4N_e\mu$ with $N_e$: effective population size and $\mu$: mutation rate per site per generation.

$M = m/\mu$ with $m$: immigration rate.

Source groups of populations for migration rates are in columns and sink groups of populations are in rows. Values between parentheses correspond to the 95% confidence interval.

**Table 6** Number (and percentage) of pure species and hybrid oaks in the sampled populations

|  | Q. petraea | Q. robur | Q. pubescens | Q. pyrenaica | $H_{PetRob}$ | $H_{PetPub}$ | $H_{PetPyr}$ | $H_{RobPub}$ | $H_{RobPyr}$ | $H_{PubPyr}$ | % Hyb |
|---|---|---|---|---|---|---|---|---|---|---|---|
| L1 | 16 | | | | 2 | 1 | | 1 | | | 20.0% |
| L3 | 16 | | | | 4 | | 3 | | 1 | | 33.3% |
| L4 | 38 | | | | 1 | | 1 | | | | 5.0% |
| L8 | 17 | | 2 | | 1 | | 1 | 1 | | | 13.6% |
| L12 | 24 | | 2 | | 3 | 2 | 1 | | | | 18.8% |
| L14 | 16 | | | | 2 | 2 | 3 | | | | 30.4% |
| L16 | 12 | | | | 8 | 1 | 1 | | | 1 | 47.8% |
| O1 | 23 | | | | 4 | 1 | | | | | 17.9% |
| O4 | 21 | | | | | | 1 | | | | 4.5% |
| O8 | 7 | 1 | | | 1 | | 1 | | | | 20.0% |
| O12 | 27 | | | | 3 | 2 | 2 | | | | 20.6% |
| O16 | 6 | | | | 2 | 1 | | 1 | | | 40.0% |
| Total | 223 | 1 | 4 | 0 | 31 | 10 | 14 | 3 | 1 | 1 | 20.8% |
| % | 77.4% | 0.3% | 1.4% | 0.0% | 10.8% | 3.5% | 4.9% | 1.0% | 0.3% | 0.3% | |

Assignment to pure species or hybrid status is based on the admixture coefficient Q (Lepais *et al.* 2009).
$H_{PetRob:}$ hybrid between *Q. petraea* and *Q. robur*.

($M_{LL->OL} = 2.6$ against $M_{LL->LH} = 2.3$ and $M_{LL->OH} = 2.4$) whereas the reciprocal migration rates were lower and preferentially within valleys ($M_{LH->LL} = 1.6$ against $M_{OL->LL} = 1.0$). Finally a low migration rate was also found from low altitudes in Ossau valley to high altitudes in Luz valley ($M_{OL->LH} = 1.2$).

### Introgression

The assignment of the sampled trees to the four oak species showed that the sampled populations are mainly composed by pure *Quercus petraea* trees at 77.4% (Table 6). However four trees were assigned to the *Q. pubescens* cluster and one tree to the *Q. robur* cluster and hybrid trees were present in all populations. Among the 60 trees considered as hybrids, more than half (31) were *Q. petraea* x *Q. robur* hybrids, 14 were *Q. petraea* x *Q. pyrenaica* hybrids and 10 were *Q. petraea* x *Q. pubescens* hybrids. Five trees were considered as hybrids between other species than *Q. petraea* but showing also sizeable assignment probabilities to the *Q. petraea* cluster (>0.20). Along the altitudinal gradient we observed a significant increase in the proportion of hybrids with increasing altitude (regression coefficient significant at $p < 0.05$), ranging 17.9% to 20% at low altitude to 47.8% to 40% at high altitude in the Luz and Ossau valley, respectively (Table 6).

### Discussion

#### Oak colonization dynamics of Pyrenean valleys

Palinological historical reconstruction indicates that deciduous oaks have colonized the northern side of the

Pyrenees between 13 000 and 12 000 BP stemming from the Spanish glacial refugial areas (Reille & Andrieu 1995). Migration pathways from Spain skirted the Pyrenees on the western edge in the Basque region and oaks spread on the northern foothills from west to east (Petit *et al.* 2002b). An eastern colonization route coming from the Mediterranean basin is also suggested from phylogeographical investigations (Petit *et al.* 2002c). There is indeed a sharp geographical split in chloroplast lineages (between the eastern lineage A and the western lineage B) indicating that both movements met at mid distance from the Atlantic ocean and the Mediterranean sea at the northern side of the Pyrenees (Fig. 1a). We found signatures of these records in our data set since all sampled populations comprise two chloroplast haplotypes that belong to the western lineage (haplotypes 10 and 12), and only one plain population is fixed for a haplotype belonging to the Eastern lineage (haplotype 7). Pollen records at higher altitude indicate that deciduous oaks reached their upper limit (between 1700 and 2000 m) in less than 2000 years (Aubert 2001; Belet 2001) since the earliest records mentioned oak presence between 11 000 and 10 000 BP. We suspect that both valleys were colonized by source populations located in the plain and our data suggest that the source populations were composed of different haplotypes (haplotypes 10 and 12 for the Ossau valley, and haplotype 12 for the Luz valley). Interestingly haplotype 10 is the most frequent haplotype occurring in extant oak populations located in the plain along the Pyrenees Mountains (Petit *et al.* 2002b), and the persistence of mixed populations (haplotypes 10 and 12) in the Ossau valley suggests past or recent contributions from the Luz valley that is entirely fixed for haplotype 12. Persistence of

extant mixed populations can be obtained in only two different scenarios. Either the founder populations were already mixed and large enough to prevent the fixation of one haplotype by drift over the 200 to 500 generations since foundation, or recurrent seed flow was maintained from the source population of the neighbouring Luz valley that is fixed for one haplotype. The latter scenario is less probable as average seed-dispersal distances were reported to be limited to less than 1 km in oak species (Valbuena-Carabaña et al. 2005).

*Preferential unidirectional pollen flow*

Colonization along altitudinal gradients has generated population differentiation within each valley (Table 3) that has been tempered by pollen flow since establishment of the populations. Because of the strong asymmetry between pollen and seed flow in oak (El Mousadik & Petit 1996b), we suspect that most of the gene flow occurring since stand establishment is due to pollen movement. Even if the $F_{ST}$ estimates are low, they are close to those observed over the entire natural range (Mariette et al. 2002). The population clustering analysis using Structure further showed that the overall differentiation was mainly due to distinct altitudinal trends of variation occurring in each valley. Whereas populations located below 800 m are mostly admixed (with each individual being assigned to several putative clusters), populations located at higher altitudes are not (with each individual being predominantly assigned to a single cluster). In the Ossau valley, populations from the highest elevations (O12, O16 and to a lesser extent O8) appear as a distinct genetic cluster differentiated from lower altitude populations. Similarly, in the Luz valley populations located from 800 m upward exhibited differences from lower populations. Populations L12 on one hand and L16 and L8 on the other hand are each assigned to a single cluster, whereas population L14 is admixed between these two clusters. Divergence of high altitude populations is observed even if the markers presumably under altitudinal selection are removed from the analysis which indicates that the genetic differentiation is caused by demographic events. Genetic differentiation of high altitude populations is found in several studies focusing on the genetic diversity of tree species along an altitudinal gradient (reviewed by Ohsawa & Ide 2008). Our survey suggests that disruption of gene flow along altitude or introgression (see below) may cause the observed divergence. Limitations to pollen flow can be created by physical barriers such as the mountain topography disrupting wind streams (Fournier et al. 2006; Herrera & Bazaga 2008). Differences in flowering phenology between populations located at extreme altitudes can also create temporal restrictions

to gene flow (Schuster et al. 1989; Premoli 2003). Temperature is the main environmental factor controlling the vegetative and reproductive phenology of most tree species (Chuine 2000; Vitasse et al. 2009b) and populations along altitudinal gradients are exposed to highly contrasted temperature regimes. Leaf unfolding phenology of the same oak populations that we sampled was monitored by Vitasse et al. (2009b). Flushing dates which are correlated to flowering dates, varied by 50 days between the earliest populations located at lower altitudes to the latest at the highest altitudes. Even if a large variation of flushing dates persisted also among trees within populations (Vitasse et al. 2009b), such delays between populations separated by only a few tens of kilometres can limit pollen flow. Furthermore, the Migrate analysis suggested that gene flow occurred preferentially from lower to higher altitudes. This directional gene flow may be the direct consequences of altitudinal variation in phenological dates and demographic differences in population sizes. While populations at higher altitudes flush later, they also exhibit lower variations of flushing dates among trees (F. Alberto, INRA—Université de Bordeaux, Bordeaux; unpublished). Hence, because of the partial overlap of flushing dates between altitudes and the larger between tree variation at lower than higher altitudes, pollen migration rates are expected to be higher from low to high altitude than the reverse.

*Imprints of natural selection*

Among the 16 nuclear microsatellites, two loci (QrZAG39, QpZAG15) exhibited divergence patterns which deviated from neutral expectations. The congruent allelic variations displayed by these loci in the two valleys illustrate directional selection acting along an environmental gradient. Previous studies have already reported clinal variations of allozyme frequencies with respect to altitude (Mitton et al. 1980; Belletti & Lanteri 1996; Kara et al. 1997). More recently, Jump et al. (2006) found a sharp differentiation of AFLP markers in European beech populations located at different altitudes, which departed significantly from neutral expectation. Interestingly, QpZAG15 is located on linkage group 9 of the Quercus genetic map, where the strongest QTL for bud burst is located (Derory et al. 2010). As genetic differentiation for date of bud burst has also been observed on the same populations in common garden experiments (Vitasse et al. 2009c), one may conclude that allelic frequencies have been shifted by hitchhiking effects between the QTL of budburst and QpZAG15, although strong linkage is needed between the microsatellite and QTL to prevent recombination in natural populations. QrZAG39 is located close to one extremity

of linkage group 5, with no known QTL position in its vicinity, but hitchhiking effects with another adaptive trait shaped by altitudinal variation cannot be excluded. An alternative interpretation to the clinal variation of allelic frequencies is selective introgression. As hybridization rates increase with altitude, allele frequencies at loci involved in species divergence may vary along an altitudinal gradient. Although located at less than 20 centimorgans (cM) from genomic hotspots of species divergence, the reported interspecific $F_{ST}$ values of the two loci were in the range of neutral expectations, suggesting that the loci are not affected by introgression (Scotti-Saintagne *et al.* 2004). We therefore suspect that the congruent variations of allelic frequencies are caused by diversifying natural selection occurring along the altitudinal gradients in the two valleys.

## Maintenance of genetic diversity associated to a decrease of population size

Gene diversity and allelic richness reached similar levels across the altitudinal gradient and between the two valleys except for one population located in low altitude (O4). The small population size and the isolation of this stand on a north-exposed mountainside may have exposed O4 to drift. In their review on genetic diversity of plants along altitudinal gradients Ohsawa & Ide (2008) reported maintenance of diversity in 29% of the studies. As an example, Truong *et al.* (2007) found similar levels of heterozygosity in *Betula pubescens* populations located between 450 and 1200 m, which was explained by the maintenance of extensive pollen and seed flow of this species along the altitudinal gradient. On the opposite, in studies reporting steadily decrease of genetic diversity, limitations of gene flow, small population size or founder effects at higher altitudes were advocated (Premoli *et al.* 2003; Quiroga & Premoli 2007). In our study, despite the disruption of gene flow suggested by genetic differentiation, genetic diversity was maintained along the altitudinal gradients. However, we found a trend towards decreasing population size with increasing altitude in both valleys and population size was higher in Luz than in Ossau valley. In both valleys, the scaled population size parameter ($\theta = 4N_e\mu$) was always lower at higher altitude. The apparent paradox between maintenance of diversity and decrease of population size can be explained by the relationship linking diversity and population size ($H_e = \theta/(\theta + 1)$), at equilibrium between drift and mutation and under the Wright Fisher model (Hartl & Clark 1989). In the upper ranges of $N_e$ ($N_e > \sim 5000$ for $\mu = 5 \times 10^{-4}$) population size may further increase while diversity saturates (i.e. variations of $N_e$ will have little impact on $H_e$). Assuming an average

mutation rate of $5 \times 10^{-4}$ for microsatellites (Selkoe & Toonen 2006), our estimates of $N_e$ vary between 1450 and 7350, corresponding to $H_e$ varying between 0.744 and 0.936 which lie in the range where $H_e$ is not linearly related to $N_e$ (Table 2).

## Peculiar reproductive characteristics at higher altitudes

The survey of genetic diversity across a wide range of altitudinal variation showed that populations at the higher margin of distribution share peculiar reproductive characteristics. First we found evidence of vegetative propagation. Clonal copies of trees sharing the same multilocus genotypes were separated by more than 5 m in O16 population, and at smaller distances in the L16 population. Oaks are known to produce stumps sprouts when they are logged or hedged but clonal copies may be separated at most by a few meters if repeated over generations. Re-rooting of shrubby trees (Copini *et al.* 2006), root suckers (Mayes *et al.* 1998; Valbuena-Carabaña *et al.* 2008) or repeated stump sprouting caused by snow coverage and avalanches may cause vegetative reproduction of *Q. petraea* at high altitudes.

Finally this study showed that the high altitudinal *Q. petraea* populations in both valleys comprised a significant proportion of hybrids between *Q. petraea* and *Q. robur* or *Q. pubescens*, although species status was checked from leaf morphology (Kremer *et al.* 2002). The more frequent occurrence of hybrids at the margins of the distribution fits to the scenario of species dispersion facilitated by unidirectional gene flow between a recipient local species (*Q. petraea*) and the migrant species (*Q. robur* or *Q. pubescens*) (Petit *et al.* 2004). Our study indicated that hybridization occurred between one local parental species (*Q petraea* acting as female) and one distant species (acting as male parent *Q. robur* or *Q. pubescens*). As *Q. robur* and *Q. pubescens* are present at lower altitudes, our observations suggested that pollen flow occurs from lower to higher altitude. Interestingly, this is the reverse situation to the one observed at the latitudinal margin of the species distribution, where *Q. robur* is the local species and *Q. petraea* is the migrant species. Our results therefore suggest that the mechanism of dispersion facilitated by hybridization is bidirectional depending on the spatial and demographic distribution of the parental species. Additionally, as suggested by Premoli (2003) for *Nothofagus pumilio* populations, hybridization may also contribute to the maintenance of diversity in high altitude populations. By providing new genetic variation and creating new combinations of alleles, hybridization increases genetic diversity. Furthermore, this gain of genetic diversity can improve the ecological tolerance of a species and

facilitate its adaptation to extreme environments (Lewontin & Birch 1966; Rieseberg *et al.* 2003). Under climate warming, introgression from species inhabiting lower altitudes may rather enhance then constrain adaptation. Hybridization may therefore play a major role in the adaptation of *Q. petraea* along the altitudinal gradients in the Pyrenees Mountains and these oak populations represent an interesting case for the study of natural hybrid zones (Lexer *et al.* 2004).

In conclusion our results show that *Q. petraea* populations maintain high level of genetic diversity at higher altitudes despite a reduction in population size and weak differentiation along altitudinal transects. Due to the high overall level of population sizes, genetic diversity is not affected by the altitudinal variation. Pollen flow is preferential from lower to higher altitudes and contributes to the maintenance of diversity. Hybrids are more frequent at higher altitudes and the direction of introgression is the opposite of the latitudinal limit of the distribution suggesting that this mechanism depends of the context of the local and migrant species. Hybridization may play an important role in the adaptation of *Q. petraea* in these peripheral populations. Consequently, by bringing new genetic variability, pollen flow and hybridization may both facilitate the migration of oak at higher altitudes as a response to climate change.

## Acknowledgements

## References

Aitken SN, Yeaman S, Holliday JA, Wang TL, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, **1**, 95–111.

Aubert S (2001) *Limite supérieure de la forêt, climat et anthropisation: dynamique tardiglaciaire et holocène de la végétation dans la vallée du Marcadau (Hautes-Pyrénées, France)*. Thèse de doctorat, Université Toulouse III, pp. 365.

Barreneche T, Casasoli M, Russell K *et al.* (2004) Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats (SSRs). *Theoretical and Applied Genetics*, **108**, 558–566.

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.

Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*, **22**, 341–345.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences, USA*, **98**, 4563–4568.

Belet JM (2001) *Variations tardiglaciaires et Holocène de la limite supérieure de la forêt dans les Pyrénées centrales: exemple du vallon d'Estibère (Hautes Pyrénées, France)*. Thèse de doctorat. Université Toulouse III, pp. 241.

Belletti P, Lanteri S (1996) Allozyme variation among European beech (*Fagus sylvatica* L.) stands in Piedmont, north-western Italy. *Silvae Genetica*, **45**, 33–37.

Bridle JR, Vines TH (2007) Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology & Evolution*, **22**, 140–147.

Byars SG, Papst W, Hoffmann AA (2007) Local adaptation and cogradient selection in the alpine plant, *Poa hiemata*, along a narrow altitudinal gradient. *Evolution*, **61**, 2925–2941.

Chuine I (2000) A unified model for budburst of trees. *Journal of Theoretical Biology*, **207**, 337–347.

Copini P, Buiteveld J, Sass-Klaassen U, Ouden JD (2006) Oak clusters on the Veluwe. *Vakblad Natuur Bos Landschap*, **3**, 24–27.

Csaikl UM, Burg K, Fineschi S, König AO, Matyas G, Petit RJ (2002) Chloroplast DNA variation of white oaks in the alpine region. *Forest Ecology and Management*, **156**, 131–145.

Deguilloux MF, Pemonge MH, Bertel L, Kremer A, Petit RJ (2003) Checking the geographical origin of oak wood: molecular and statistical tools. *Molecular Ecology*, **12**, 1629–1636.

Derory J, Scotti-Saintagne C, Bertocchi E *et al.* (2010) Contrasting relationships between the diversity of candidate genes and variation of bud burst in natural and segregating populations of European oaks. *Heredity*, **104**, 438–448.

Dow BD, Ashley MV, Howe HF (1995) Characterization of highly variable (GA/CT)$_n$ microsatellites in the bur oak, *Quercus macrocarpa*. *Theoretical and Applied Genetics*, **91**, 137–141.

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus*, **12**, 13–15.

Dzialuk A, Chybicki I, Welc M, Śliwińska E, Burczyk J (2007) Presence of triploids among oak species. *Annals of Botany*, **99**, 959–964.

El Mousadik A, Petit RJ (1996a) High level of genetic differentiation for allelic richness among populations of the

argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics*, **92**, 832–839.

El Mousadik A, Petit RJ (1996b) Chloroplast DNA phylogeography of the argan tree of Morocco. *Molecular Ecology*, **5**, 547–555.

Engler R, Randin CF, Vittoz P *et al.* (2009) Predicting future distributions of mountain plants under climate change: does dispersal capacity matter? *Ecography*, **32**, 34–45.

Ettl GJ, Peterson DL (2001) Genetic variation of subalpine fir (*Abies lasiocarpa* (Hook.) Nutt.) in the Olympic Mountains, WA, USA. *Silvae Genetica*, **50**, 145–153.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977–993.

Fournier N, Rigling A, Dobbertin M, Gugerli F (2006) Random amplified polymorphic DNA (RAPD) patterns show weak genetic differentiation between low- and high-elevation types of Scots pine (*Pinus sylvestris* L.) in dry continental valleys in the Alps. *Annals of Forest Science*, **63**, 431–439.

Goudet J (2001). *FSTAT, A Program to Estimate and Test Gene Diversities and Fixation Indices (Version 2.9.3)*. Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

Hartl DL, Clark AG (1989) *Principles of Population Genetics*. Sinauer Associates, Sunderland, Massachusetts.

Herrera CM, Bazaga P (2008) Adding a third dimension to the edge of a species' range: altitude and genetic structuring in mountainous landscapes. *Heredity*, **100**, 275–285.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Jeffreys H (1998) *Theory of Probability*. Oxford University Press, Oxford.

Jump AS, Hunt JM, Martinez-Izquierdo JA, Peñuelas J (2006) Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. *Molecular Ecology*, **15**, 3469–3480.

Kampfer S, Lexer C, GloÈssl J, Steinkellner H (1998) Characterization of (GA)$_n$ microsatellite loci from *Quercus robur*. *Hereditas*, **129**, 183–186.

Kara N, Korol L, Isik K, Schiller G (1997) Genetic diversity in *Pinus brutia* Ten.: altitudinal variation. *Silvae Genetica*, **46**, 155–161.

Kirkpatrick M, Barton NH (1997) Evolution of a species' range. *American Naturalist*, **150**, 1–23.

Kremer A, Dupouey JL, Deans JD *et al.* (2002) Leaf morphological differentiation between *Quercus robur* and *Quercus petraea* is stable across western European mixed oak stands. *Annals of Forest Science*, **59**, 777–787.

Lenoir J, Gégout JC, Marquet PA, de Ruffray P, Brisses H (2008) A significant upward shift in plant species optimum elevation during the 20th century. *Science*, **320**, 1768–1771.

Lepais O, Leger V, Gerber S (2006) Short note: High throughput microsatellite genotyping in oak species. *Silvae Genetica*, **55**, 238–240.

Lepais O, Petit RJ, Guichoux E *et al.* (2009) Species relative abundance and direction of introgression in oaks. *Molecular Ecology*, **18**, 2228–2242.

Lewontin RC, Birch LC (1966) Hybridization as a source of variation for adaptation to new environments. *Evolution*, **20**, 315–336.

Lexer C, Heinze B, Alia R, Rieseberg LH (2004) Hybrid zones as a tool for identifying adaptive genetic variation in outbreeding forest trees: lessons from wild annual sunflowers (Helianthus spp.). *Forest Ecology and Management*, **197**, 49–64.

Mariette S, Cottrell J, Csaikl UM *et al.* (2002) Comparison of levels of genetic diversity detected with AFLP and microsatellite markers within and among mixed *Q. petraea* (Matt.) Liebl. and *Q. robur* L. stands. *Silvae Genetica*, **51**, 72–79.

Mayes SG, McGinley MA, Werth CR (1998) Clonal population structure and genetic variation in sand-shinnery oak, *Quercus havardii* (Fagaceae). *American Journal of Botany*, **85**, 1609–1617.

Mitton JB, Sturgeon KB, Davis ML (1980) Genetic differentiation in ponderosa pine along a steep elevational transect. *Silvae Genetica*, **29**, 100–103.

Neal RM (2003) Slice sampling. *Annals of Statistics*, **31**, 705–741.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Ohsawa T, Ide Y (2008) Global patterns of genetic variation in plant species along vertical and horizontal gradients on mountains. *Global Ecology and Biogeography*, **17**, 152–163.

Peñuelas J, Ogaya R, Boada M, Jump AS (2007) Migration, invasion and decline: changes in recruitment and forest structure in a warming-linked shift of European beech forest in Catalonia (NE Spain). *Ecography*, **30**, 829–837.

Petit RJ, Csaikl UM, Bordacs S *et al.* (2002a) Chloroplast DNA variation in European white oaks—phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management*, **156**, 5–26.

Petit RJ, Brewer S, Bordacs S *et al.* (2002b) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, **156**, 49–74.

Petit RJ, Latouche-Halle C, Pemonge MH, Kremer A (2002c) Chloroplast DNA variation of oaks in France and the influence of forest fragmentation on genetic diversity. *Forest Ecology and Management*, **156**, 115–129.

Petit RJ, Bodenes C, Ducousso A, Roussel G, Kremer A (2004) Hybridization as a mechanism of invasion in oaks. *New Phytologist*, **161**, 151–164.

Premoli AC (2003) Isozyme polymorphisms provide evidence of clinal variation with elevation in *Nothofagus pumilio*. *Journal of Heredity*, **94**, 218–226.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Pritchard JK, Wen W, Falush D (2009) *Documentation for Structure Software: Version 2.3*. Available from http://pritch.bsd.uchicago.edu/structure.html.

Quiroga MP, Premoli AC (2007) Genetic patterns in *Podocarpus parlatorei* reveal the long-term persistence of cold-tolerant elements in the southern Yungas. *Journal of Biogeography*, **34**, 447–455.

R Development Core Team (2005) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available from URL: http://www.R-project.org.

Reille M, Andrieu V (1995) The late Pleistocene and Holocene in the Lourdes Basin, Western Pyrenees, France: new pollen analytical and chronological data. *Vegetation History and Archaeobotany*, **4**, 1–21.

Rieseberg LH, Raymond O, Rosenthal DM *et al.* (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science*, **301**, 1211–1216.

Schuster WS, Alles DL, Mitton JB (1989) Gene flow in limber pine—evidence from pollination phenology and genetic differentiation along an elevational transect. *American Journal of Botany*, **76**, 1395–1403.

Scotti-Saintagne C, Mariette S, Porth I *et al.* (2004) Genome scanning for interspecific differentiation between two closely related oak species [*Quercus robur* L. and *Q. petraea* (Matt.) Liebl.]. *Genetics*, **168**, 1615–1626.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, **9**, 615–629.

Steinkellner H, Fluch S, Turetschek E *et al.* (1997) Identification and characterization of (GA/CT)(*n*)-microsatellite loci from *Quercus petraea*. *Plant Molecular Biology*, **33**, 1093–1096.

Truong C, Palme AE, Felber F (2007) Recent invasion of the mountain birch *Betula pubescens* ssp. tortuosa above the treeline due to climate change: genetic and ecological study in northern Sweden. *Journal of Evolutionary Biology*, **20**, 369–380.

Valbuena-Carabaña M, Gonzalez-Martinez SC, Sork VL *et al.* (2005) Gene flow and hybridisation in a mixed oak forest (*Quercus pyrenaica* Willd. and *Quercus petraea* (Matt.) Liebl.) in Central Spain. *Heredity*, **95**, 457–465.

Valbuena-Carabaña M, Gonzalez-Martinez SC, Gil L (2008) Coppice forests and genetic diversity: a case study in *Quercus pyrenaica* Willd. from Central Spain. *Forest Ecology and Management*, **254**, 225–232.

Vitasse Y, Porté AJ, Kremer A, Michalet R, Delzon S (2009a) Response of canopy duration to temperature changes in four temperate tree species: relative contributions of spring and autumn leaf phenology. *Oecologia*, **161**, 187–198.

Vitasse Y, Delzon S, Dufrene E *et al.* (2009b) Leaf phenology sensitivity to temperature in European trees: do within-species populations exhibit similar responses? *Agricultural and Forest Meteorology*, **149**, 735–744.

Vitasse Y, Delzon S, Bresson CC, Michalet R, Kremer A (2009c) Altitudinal differentiation in growth and phenology among populations of temperate-zone tree species growing in a common garden. *Canadian Journal of Forest Research*, **39**, 1259–1269.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Zhang Q, Allard RW (1986) Sampling variance of the genetic diversity index. *Journal of Heredity*, **77**, 54–55.