# Genetic diversity, population genetic structure and gene flow in the North African endemic and vulnerable *Quercus afares* Pomel. Revealed by microsatellites

Yacine Torche[1]*, Markus Mueller[2], Oliver Gailing[2]

[1]Biotechnology Department, Nature and Life Sciences Institute, Abdelhafid Boussouf University Center of Mila, Mila, Algeria.

[2]Faculty for Forest Sciences and Forest Ecology, Forest Genetics & Forest Tree Breeding, University of Göttingen, Büsgen weg 2, 37077 Göttingen, Germany.

*Corresponding author: torche.yacine@centre-univ-mila.dz / telephone number: +213667090748

## Abstract

Due to substantial population decline and habitat fragmentation, the North African Afares oak (*Quercus afares* Pomel.) is currently vulnerable and highly threatened. Like many endangered species, Afares is susceptible to biological and anthropogenic threats that can lead to genetic diversity loss and, ultimately, a heightened risk of extinction. Therefore, assessing genetic diversity is imperative for informing conservation strategies. In this study, genetic diversity and structure of the extant Algerian populations of *Q. afares* were analyzed. Genetic diversity, population structure, as well as gene flow and potential bottleneck effects were estimated using 16 polymorphic nuclear microsatellite markers. We found an unexpectedly high level of genetic diversity ($H_o$ = 0.64, $H_e$ = 0.60, $A_r$ = 7.16) across the studied localities for a threatened species with restricted distribution range. Moreover, there were high levels of gene flow based on pairwise genetic differentiation estimates and minimal genetic structuring, among the fragmented, outbreeding localities ($F_{is}$ = -0.11, $p$ > 0.05) of *Q. afares*. Additionally, no correlation was found between genetic and geographical distances ($R_{xy}$ = -0.29, $p$ = 0.5). These findings suggest that *Q. afares* stands represent a metapopulation. Despite the high genetic diversity of *Q. afares*, local populations in each small habitat patch remain at risk of extinction due to habitat loss and increasing anthropogenic threats. To minimize further loss of genetic diversity and habitat degradation, effective protection measures should be applied to *Q. afares* forests in Algeria. Conservation efforts should focus on habitat management and raising awareness among residents neighboring Afares oak forests.

Keywords: Quercus afares, genetic diversity, genetic structure, threatened species, conservation.

## Introduction

Forests play an important role in carbon sequestration as well as in soil and water management. They are considered complex ecosystems with high levels of biodiversity in terms of genetic resources, species, and habitats. Oak (*Quercus* spp.) forests hold significant ecological, economic, and cultural value. The *Quercus* genus comprises an estimated 400 to 600 species worldwide and is widely distributed as the dominant genus in the Mediterranean region of the Northern Hemisphere (Backs and Ashle 2021).

Among oak species, Afares oak (*Quercus afares* Pomel., section *Cerris*), a threatened deciduous oak endemic to North Africa, is Red listed as vulnerable by the IUCN (Jerome et al. 2020). This species is limited to restricted areas in northeastern Algeria and a single locality in western Tunisia. Afares oak has been hypothesized to originate from the hybridization of two distant parents, Cork oak (*Q. suber* L., section *Cerris*) and Zeen oak (*Q. canariensis* Willd., section *Quercus*) based on allozymes (Mir et al. 2006) and cpDNA (Sakka et al. 2015); the two species are sympatric with *Q. afares* over most of their geographical distribution, with contact zones between *Q. suber* (mixed stands) commonly found in low lands (200~700 m), while contact zones with *Q. canariensis* are more frequent at mid-altitudes (700~1200 m), above 1200 m, *Q. afares* constitutes monospecific stands (Mir et al. 2006). However, recent studies have questioned the hybrid origin of *Q. afares* (Simeone et al. 2018), stating that there is no evidence supporting a hybrid origin. Specifically, genome-wide SNP data confirmed that the two putative parent species *Quercus canariensis* (section *Quercus*) and *Quercus suber* (section *Cerris*) belong to different sections (Hipp et al. 2020) and there is no unambiguous evidence that hybridization occurs between members of different sections (Gailing and Curtu 2014). In addition, no evidence of a hybrid

DOI:10.2478/sg-2025-0010
edited by the Thünen Institute of Forest Genetics

origin of *Q. afares* was detected based on genome-wide SNP data (Denk et al. 2023).

Afares oak is adapted to typically Mediterranean climates (Csa, Csb as classified by Köppen-Geiger climate zones) (Beck et al. 2018), characterized by mild, wet winters and dry, warm summers. Its ecological requirements align with those of Mediterranean sclerophyllous vegetation; it prefers humid, mountainous environments with acidic, non-limestone soils and exhibits moderate frost tolerance, similar to Zeen oak. Unlike *Q. suber* and *Q. canariensis*, *Q. afares* is severely constrained, with fragmented habitat across North Africa, in a recognized biodiversity hotspot, made of forests, mountains and coastal ecosystems critically threatened by human activity. Despite the protected status of *Q. afares* in Algeria (Journal Officiel de la Republique Algerienne, 2012), oak forests in North Africa generally remain vulnerable and face increasing conservation threats, including habitat loss, invasive species, climate shifts, and pest/pathogen damages (Kadi-Bennane et al. 2020).

In recent decades, deforestation, fragmentation of natural habitats and urbanization — along with illegal logging, uncontrolled resources extraction, fires and livestock overgrazing — threaten forest ecosystems, hinder their development, and prevent natural regeneration. Global warming also exacerbates these threats by increasing the intensity and the frequency of wildfires, droughts, and pest outbreaks, causing the degradation and disappearance of natural forests leading to serious biodiversity losses (Feng et al. 2024). Due to these environmental and anthropogenic pressures, genetic variation in natural habitats is subjected to decrease leading to impediments in the adaptation of tree species to local environmental changes. Additionally, habitat fragmentation of natural populations causes irreversible genetic diversity losses within tree populations, reducing gene flow, increasing levels of inbreeding, and raising the risk of genetic bottlenecks. This, in turn, increases the likelihood of local extinctions (Dobeš et al. 2017). Genetic diversity, one of the three pillars of biodiversity, is determined by gene flow, genetic drift, selection, mutation, and other factors in forest ecosystems (Neophytou et al. 2022). For tree species with a relatively long-life cycle, genetic diversity plays a critical role for the long-term survival and adaptability to changing environments by providing raw material for adaptation and evolution. Moreover, forest trees' responses to climate change are conditioned by the interplay of genetic variation, adaptation, gene flow, and interspecific hybridization (Kremer et al. 2012). Therefore, studying the genetic diversity and population structure of tree species is crucial for forest management and can help mitigate biodiversity loss, maintain genetic variation, and identify populations with high genetic variability, a prerequisite for effective conservation measures (de Vries et al. 2015). This is particularly relevant for species like *Q. afares*, with small population size, fragmented habitats and limited geographic ranges, which are likely subjected to reduced gene dispersal and increased genetic divergence, due to genetic drift and inbreeding and can be subjected to loss of genetic diversity and local extinction (Ortego et al. 2010). In this context, it is of great importance to assess the genetic diversity and structure of *Q. afares* to understand its current evolutionary and adaptive trends, and for the development and implication of conservation strategies.

The aim of this study was to investigate the genetic characteristics of 98 individuals from seven localities of *Quercus afares* that cover most of the species' distribution range, using microsatellites (SSRs) with the following objectives: (1) assess genetic diversity and structure, (2) analyze the correlation between genetic distance and geographical distance; and (3) examine potential bottleneck effects. To the best of our knowledge, this is the first study on the genetic diversity and structure of *Q. afares* in North Africa based on microsatellite markers. To date, this species has only been studied through morpho-ecological (Mhamdi et al. 2013; Kadi-Bennane et al. 2020) and hybridization surveys at allozymes and cpDNA (Mir et al. 2006; Sakka et al. 2015). The findings are expected to shed light on the possible genetic consequences of fragmentation and small population size, so as to provide a reference for effective conservation and germplasm resource management of *Q. afares* within the still poorly understood Mediterranean forest ecosystems.

## Material and Methods

### *Plant material*

A total of 98 georeferenced samples from 7 forest stands in northeastern Algeria were collected (Fig 1; Table 1). These sampling sites cover most the species' distribution range, with the exception of one locality in west Tunisia and another close to the Tunisian border. The sites are located in the four Wilayas (Governorates) of Skikda, Jijel, Béjaïa and Tizi Ouzou, only one site was located in the Taza National Park, representing the largest *Q. afares* forest. The distance between pairs of sampled trees was over 200 m.

Table 1
Sampling information of the study localities, number of individuals (N).

| Forest stands | ID | N | Geographical Location | Altitude (m) | Latitude | Longitude |
|---|---|---|---|---|---|---|
| Akfadou | BA | 15 | Béjaïa | 450 ~ 800 | 36.6922 | 4.7892 |
| Lac Noir | BLN | 9 | Béjaïa | 1000 ~ 1400 | 36.6885 | 4.6460 |
| Yakouren | BY | 15 | Tizi Ouzou | 600 ~ 1200 | 36.7284 | 4.4953 |
| Djimla | JJM | 15 | Jijel | 900 ~ 1300 | 36.5643 | 5.9118 |
| Selma Ben Ziada | JS | 17 | Jijel (Taza National Park) | 750 ~1250 | 36.6812 | 5.6419 |
| Tamesguida | JT | 14 | Jijel | 900 ~ 1600 | 36.5567 | 5.8127 |
| Siouane | SI | 13 | Skikda | 600 ~ 1160 | 36.9694 | 6.4210 |

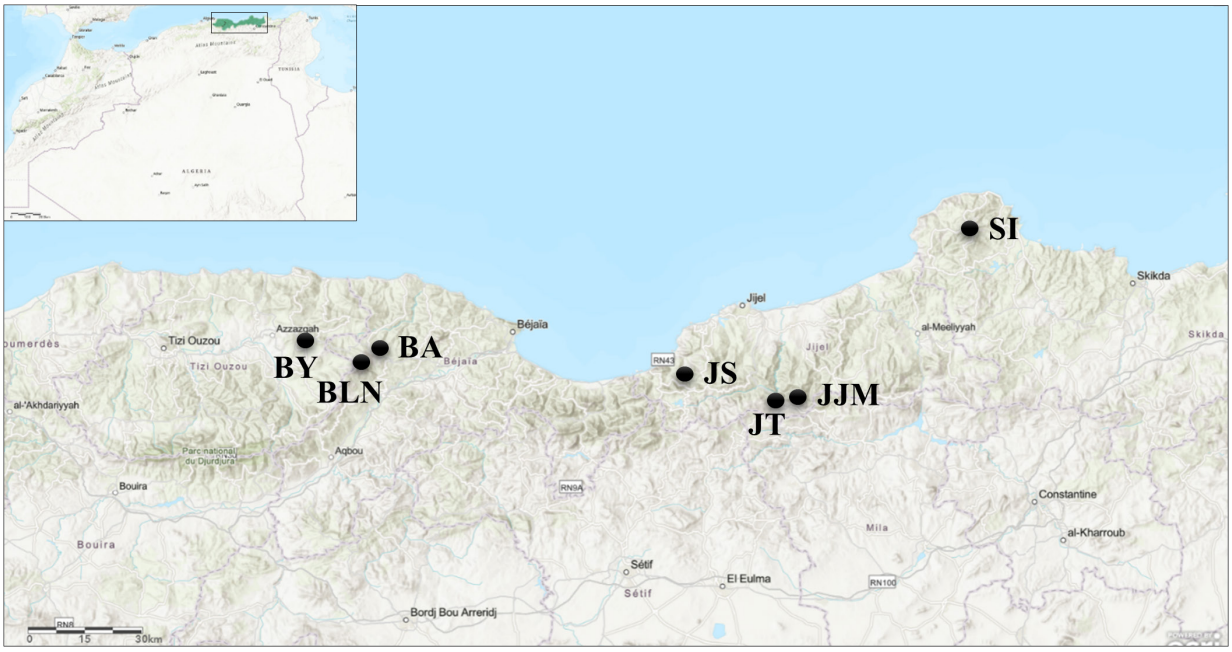


Fig 1.
Geographical location of the sampled localities (black dots) in Northeast Algeria. Map created in ArcGIS Online (Esri, California, USA).

### *DNA extraction and SSR analysis*

Total genomic DNA was extracted from young leaves and buds (approx. 1 cm2 piece of leaf or 2 to 3 buds) using the DNeasy 96 Plant Kit from Qiagen (Hilden, Germany) according to the manufacturer's instructions.

For initial testing, 25 nuclear microsatellite markers were amplified with six multiplexes (Supplementary Table S1) in a 13 µl PCR mix using 1 µl DNA (ca. 0.6 ng/ µl) as template following the steps described by Burger and Gailing, (2022) (Supplementary file 1).

All PCR reactions were conducted using a Biometra Thermal Cycler (MJ Research PTC 200, Analytik Jena, Germany) using the following touchdown program: 15 min initial denaturation at 95 °C, followed by 10 touchdown cycles at 94 °C for 1 min, 1 min at 60 °C (decreasing 1 °C each cycle) and 1 min at 72 °C. Subsequently, 25 cycles at 94 °C for 1 min, annealing at 50 °C for 1 min and elongation at 72 °C for 1 min. The amplification process concluded with a final elongation step at 72 °C for 20 min. Amplification products were resolved on an ABI 3130xl Genetic Analyzer (Applied Biosystems, Foster City, USA) using the GeneScan™ Rox-500 and Liz-500 (only for multiplex 3) size markers. For the fragment length analysis, PCR multiplexes 5 and 6 were run together. Scoring of alleles was conducted using GeneMapper® version 4.1 (Applied Biosystems, Foster City, USA).

## Data analysis

### Genetic diversity indices

Genotyping errors due to the presence of null alleles can affect the estimation of population differentiation. Null allele frequencies were estimated using the Windows®-based software MicroChecker version 2.2.3 (Van Oosterhout et al. 2004).

Intra- and interpopulation genetic diversity parameters were calculated. Observed ($N_a$) and effective ($N_e$) number of alleles, the observed heterozygosity ($H_o$) and expected heterozygosity ($H_e$), and the Shannon diversity index ($I$) were all calculated at each locus, over all loci, and for each locality using the software GenAlEx version 6.51b2 (Peakall and Smouse 2012; Smouse et al. 2017). The inbreeding coefficient Fis for individual markers and across all markers as well as the fixation index Fst were computed using Arlequin 3.5 (Excoffier et al., 2010). The statistical significance was tested with a nonparametric approach described in Excoffier et al. (1992) with 1000 permutations.

The estimation of the mean number of alleles per locus as a measure of allelic richness ($A_r$) can be affected when the sample sizes of the populations are either low on average or highly unbalanced among populations. For this reason, $A_r$ and the private $P_{Ar}$ were calculated by the rarefaction (to the lowest sample size) method using HP-Rare 1.1 software (Kalinowski 2005). This approach uses the frequencies of alleles at the locus to estimate the expected number of alleles and/or private alleles in a subsample of $N$ individuals selected at random from a sample of $N$ individuals in each locality.

### Tests of Hardy–Weinberg equilibrium and linkage disequilibrium.

Deviations from Hardy-Weinberg Equilibrium (HWE) and Linkage disequilibrium (LD) were estimated using GENEPOP version 4.8.3 (Rousset 2008) with deviations from expected proportions evaluated with a Bonferroni corrected $p$-value threshold of 0.05 (Rice 1989). HWE was estimated for the 16 maintained loci (see results section) with the Monte Carlo based exact test for a null hypothesis of random union of gametes, and p-values obtained from a global test across samples using Fisher's method (dememorization 10.000; batches 20; iterations 10.000). LD was calculated for each pair of loci in the seven localities with the default Monte Carlo based G-test for the null hypothesis of independence among genotypes at different loci, and the log likelihood ratio as test statistic (dememorization 10.000; batches 100; iterations 5000).

### Population structure analysis

To characterize the patterns of genetic structure of the *Q. afares* localities we used complementary approaches as follows.

Principal coordinates analysis (PCoA) based on a dissimilarity matrix of Nei's genetic distance (Nei 1978) was performed using the software GenAlEx 6.51b2 (Peakall and Smouse 2012; Smouse et al. 2017).

Population structure was assessed using a Bayesian approach implemented in the software STRUCTURE version 2.3.4 (Hubisz et al. 2009). The method attempts to identify possible populations based on the microsatellite dataset. The analysis was performed using the admixture model for its consideration of the ancestry of admixed individuals, and correlated allele frequencies were selected because of their potential to improve clustering in closely related populations. The possible range of clusters ($K$) tested was from 1 to 9 (the putative number of populations plus 2) with 10 runs per each $K$. A burn-in period of 50,000, Markov chain Monte Carlo (MCMC) replicates of 105, the default setting 'admixed model without LOCPRIOR' was set to identify population structure solely based on genetic information. However, to detect subtle population structures we additionally ran the analysis with the LOCPRIOR option. Furthermore, to determine the "best $K$" from the logarithmic results and the $\Delta K$ method (Evanno et al. 2005), the online software CLUMPAK (Kopelman et al. 2015) was used. In addition, for post-processing and graphical representation of the results of the model-based population, the CLUMPAK software (Cluster Markov Packager Across $K$) was applied to STRUCTURE analysis (Kopelman et al. 2015). Subsequently, the observed ($H_o$) and expected ($H_e$) heterozygosity, allelic richness ($A_r$) and private allele richness ($P_{Ar}$) were measured based on STRUCTURE results.

Using Arlequin 3.5 software, the hierarchical genetic structure was investigated to assess the presence of a differential genetic structure across the total sample range and among putative geographical groupings of samples with hierarchical AMOVA (Excoffier et al. 1992), as well as the pairwise differentiation Fst between populations (Weir and Cockerham 1984) and gene flow estimates ($N_m$) using the equation $N_m = 0.25 (1 - F_{st})/F_{st}$ (Excoffier et al. 1992; Schneider et al. 2000). Since the high mutation rate of microsatellites can lead to underestimates of genetic differentiation by means of *F-statistics* (Balloux et al. 2000), we also computed the coefficient of genetic differentiation $G''_{st}$, which is considered more accurate in many contexts than $G'_{st}$, especially when heterozygosity is high and sample sizes are small (Meirmans and Hedrick 2011). The relationship between genetic (Linearized $F_{st}$ ($\text{lin}F_{st} = F_{st}/(1-F_{st})$; (Slatkin 1995)) and geographical distance was calculated via Mantel test (GenAlEx 6.51b2; 9999 permutations) to test for isolation-by-distance (IBD).

Heterozygosity excess caused by recent population bottlenecks was tested using Wilcoxon signed-rank test under the Stepwise Mutation Model (SMM) and Two-Phase Model (TPM) in BOTTLENECK v.1.2.02 software (Luikart and Cornuet 1998; Piry et al. 1999). The variance for TPM was set to 12, and the proportion of SMM in TPM was 95 %, with 1000 iterations (Piry et al. 1999). Mode-shift tests were also used to detect genetic bottlenecks. The remaining relevant parameters were set as the software default values.

## Results

From the twenty-five initial nuclear microsatellites used for the study, nine were excluded from further analysis. The SSRs FS_C2791 and VIT023 were monomorphic, while the eight remaining markers (3D15; 1P10; FS_C8183; FS_C 2361; PIE215;

PIE223; FIR13) were non-informative due to non-specific amplification. The sixteen remaining markers were highly polymorphic and then used for further analysis.

## Genetic diversity

The number of alleles detected for each locus varied between 2.00 (PIE239) and 10.86 (PIE040) (Table 2), while the expected heterozygosity $H_o$ ranged from 0.16 (Qr0332) to 0.87 (FS_C2660). Locus PIE040 expressed the highest $H_e$ (0.82), while the lowest (0.14) value was recorded for locus Qr0332. Allelic richness values varied between 2.00 (PIE239) and 8.89 (PIE040). Fis showed positive and significant values for FIR043 (0.15). The analysis conducted with MicroChecker software indicated the absence of null alleles (Table S2). The estimated relative genetic differentiation was found to be very low, with 0.016 and 0.044 for $F_{st}$ and $G''_{st}$ respectively, however, $F_{st}$ values for FIR35 (0.043) and PIE125 (0.082) were significant and considerably higher than the average. When the seven localities are considered as a metapopulation following the STRUCTURE results, genetic parameters increase compared to the mean diversity observed within individual localities (e.g., $N_a$ = 7.25, $A_r$ = 7.16, Mean T, Table 2).

Exact tests of Hardy-Weinberg Equilibrium (HWE) showed that all localities and all loci (Fisher's method) were in HW equilibrium after Bonferroni correction except locus PIE239 (Table 2 and 3). Tests of linkage disequilibrium (LD) for all pairs of loci showed no LD therefore, all microsatellite loci were included in the analysis.

Table 2

Sixteen microsatellite loci analyzed: number of alleles ($N_a$); number of effective alleles ($N_e$); Shannon's index ($I$); observed heterozygosity ($H_o$); expected heterozygosity ($H_e$); Allelic richness ($A_r$); within-population inbreeding coefficient ($F_{is}$) and genetic differentiation coefficient ($F_{st}$). ($P_{HWE}$) $p$-value for Hardy–Weinberg equilibrium exact test, adjusted with the Bonferroni correction, $p$ = 0.00044. (Significance of inbreeding coefficient $F_{is}$ and fixation index $F_{st}$ were tested using a nonparametric approach described in Excoffier et al. (1992) with 1000 permutations: *$p < 0.05$. Mean T = Total diversity of the metapopulation after STRUCTURE results (see below). $P_{HWE}$ significance values of exact test for Hardy–Weinberg deviations (values in bold are significant).

| Locus | $N_a$ | $N_e$ | $I$ | $H_o$ | $H_e$ | $A_r$ | $F_{is}$ | $F_{st}$ | $P_{HWE}$ | $G''_{st}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| **3A05** | 6.71 | 3.56 | 1.48 | 0.76 | 0.69 | 5.70 | -0.08 | 0.026 | 0.690 | 0.055 |
| **2P24** | 3.29 | 2.16 | 0.90 | 0.59 | 0.53 | 3.08 | -0.07 | 0.002 | 0.710 | 0.00029 |
| **FIR28** | 4.71 | 2.15 | 1.01 | 0.58 | 0.51 | 4.13 | -0.10 | 0.031 | 0.564 | 0.063 |
| **FIR043** | 3.71 | 1.89 | 0.84 | 0.40 | 0.45 | 3.37 | 0.15* | 0.025 | 0.061 | 0.051 |
| **FIR35** | 7.71 | 5.84 | 1.84 | 0.84 | 0.81 | 6.97 | 0.01 | 0.043* | 0.031 | 0.284 |
| **GOT040** | 6.14 | 4.55 | 1.63 | 0.77 | 0.78 | 5.73 | 0.05 | 0.009 | 0.260 | 0.041 |
| **GA0C11** | 4.57 | 2.69 | 1.19 | 0.60 | 0.62 | 4.32 | 0.06 | -0.001 | 0.294 | -0.005 |
| **PIE040** | 10.86 | 5.77 | 2.03 | 0.83 | 0.82 | 8.89 | 0.03 | -0.001 | 0.324 | -0.007 |
| **VIT107** | 5.86 | 3.48 | 1.42 | 0.92 | 0.71 | 5.16 | -0.25 | 0.004 | 0.153 | 0.035 |
| **PIE125** | 2.29 | 1.76 | 0.63 | 0.47 | 0.42 | 2.18 | -0.06 | 0.082* | 0.949 | 0.160 |
| **Qr0332** | 2.14 | 1.17 | 0.28 | 0.16 | 0.14 | 1.92 | -0.06 | -0.012 | 0.999 | -0.015 |
| **FS_C2660** | 4.29 | 2.94 | 1.20 | 0.87 | 0.66 | 3.92 | -0.29 | 0.0179 | 0.073 | 0.062 |
| **Qr0057** | 3.71 | 3.06 | 1.17 | 0.85 | 0.67 | 3.50 | -0.22 | -0.002 | 0.006 | 0.001 |
| **Qr1423** | 2.29 | 1.52 | 0.53 | 0.38 | 0.33 | 2.19 | -0.12 | 0.019 | 0.999 | 0.040 |
| **PIE227** | 3.29 | 2.49 | 1.00 | 0.55 | 0.59 | 3.23 | 0.11 | 0.029 | 0.805 | 0.084 |
| **PIE239** | 2.00 | 1.91 | 0.67 | 0.80 | 0.48 | 2.00 | -0.64 | -0.005 | <0.001 | -0.014 |
| **Mean** | 4.60 | 2.93 | 1.11 | 0.65 | 0.57 | 4.14 | -0.09 | 0.016 | 0.46 | 0.044 |
| **Mean T** | 7.25 | 3.21 | 1.24 | 0.64 | 0.60 | 7.15 | | | | |

Table 3
Genetic diversity parameters for the seven *Q. afares* localities analyzed: mean number of different alleles ($N_a$), mean number of effective alleles ($N_e$), Shannon's index ($I$), observed heterozygosity ($H_o$), expected heterozygosity ($H_e$), unbiased expected heterozygosity ($UH_e$), allelic richness ($A_r$), and private allelic richness ($PA_r$), inbreeding coefficient ($F_{is}$). ($P_{HWE}$) $p$-value Hardy–Weinberg equilibrium exact test, adjusted with the Bonferroni correction, $p = 0.00044$.

Significance of inbreeding coefficient $F_{is}$ was tested using a nonparametric approach described in Excoffier et al. (1992) with 1000 permutations. Mean T = Total diversity of the metapopulation after STRUCTURE results (see below). The full localities names are given in Table 1

| Pop. ID | $N_a$ | $N_e$ | $I$ | $H_o$ | $H_e$ | $UH_e$ | $A_r$ | $PA_r$ | $F_{is}$ | $P_{HWE}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| **BA** | 4.44 | 2.79 | 1.09 | 0.66 | 0.57 | 0.59 | 3.95 | 0.16 | -0.13 | 0.394 |
| **BLN** | 4.31 | 2.83 | 1.11 | 0.72 | 0.58 | 0.62 | 4.31 | 0.10 | -0.19 | 0.737 |
| **BY** | 4.81 | 3.01 | 1.13 | 0.64 | 0.58 | 0.60 | 4.20 | 0.16 | -0.08 | 0.500 |
| **JJM** | 4.75 | 2.95 | 1.11 | 0.64 | 0.57 | 0.59 | 4.11 | 0.23 | -0.08 | 0.220 |
| **JS** | 4.63 | 2.92 | 1.09 | 0.60 | 0.55 | 0.57 | 4.06 | 0.09 | -0.13 | 0.053 |
| **JT** | 4.75 | 3.13 | 1.16 | 0.69 | 0.59 | 0.62 | 4.29 | 0.33 | -0.17 | 0.370 |
| **SI** | 4.50 | 2.91 | 1.11 | 0.59 | 0.58 | 0.61 | 4.07 | 0.29 | 0.00 | 0.019 |
| **Mean** | 4.60 | 2.93 | 1.11 | 0.65 | 0.57 | 0.60 | 4.14 | 0.19 | -0.11 | |
| **Mean T** | 7.25 | 3.21 | 1.24 | 0.64 | 0.60 | 0.60 | 7.16 | 7.16 | | |

Genetic diversity indices for each locality are shown in Table 3. Inbreeding coefficients across all markers were not significantly different from zero in any locality. The average $N_e$ was 2.93 ranging from 2.79 in locality BA to 3.13 in locality JT. The lowest observed heterozygosity $H_o$ was found in SI (0.59) and the highest in BLN (0.72). All localities showed slight deviation from the mean $H_e$ (0.57), and varied between 0.55 (JS) and 0.59 (JT). These values are comparable with unbiased $H_e$ values indicating no effect due to the imbalance of the sampling size. Allelic richness ($A_r$) calculated with the rarefaction method, varied between 3.95 in BA and 4.31 in BLN, while private allele richness ($PA_r$) values ranged from 0.09 (JS) to 0.33 (JT). Generally, an increase in genetic parameters was observed when considering the metapopulation (Mean T, Table 3), where $N_a$ (7.25), $A_r$ (7.16) and $PA_r$ (7.16) showed higher values compared to mean diversity of the seven localities.

The results of the program BOTTLENECK mode-shift analysis showed that the allele frequencies of all localities followed a normal L-shaped distribution, which is typical of populations at mutation-drift equilibrium. Additionally, the heterozygosity excess tests under the Two-Phase Model (TPM, $p > 0.05$), and the Stepwise Mutation Model (SMM, $p > 0.05$) did not show a consistent or significant excess of heterozygosity across loci or localities (Table S3). These results indicate that there is no evidence of a recent bottleneck event in any of the studied localities.

## *Population structure*

The Principal Coordinates Analysis (PCoA) based on Nei's unbiased genetic distance matrix (Table 4, Fig 2) suggested the presence of a nonrandom association of the localities under study. The combined two axes explained 87.70 % of the variation. Additionally, the PCoA plot based on individuals (Fig S1) suggests that individuals from all localities do not show strong genetic structure, and are not clearly differentiated from each other.

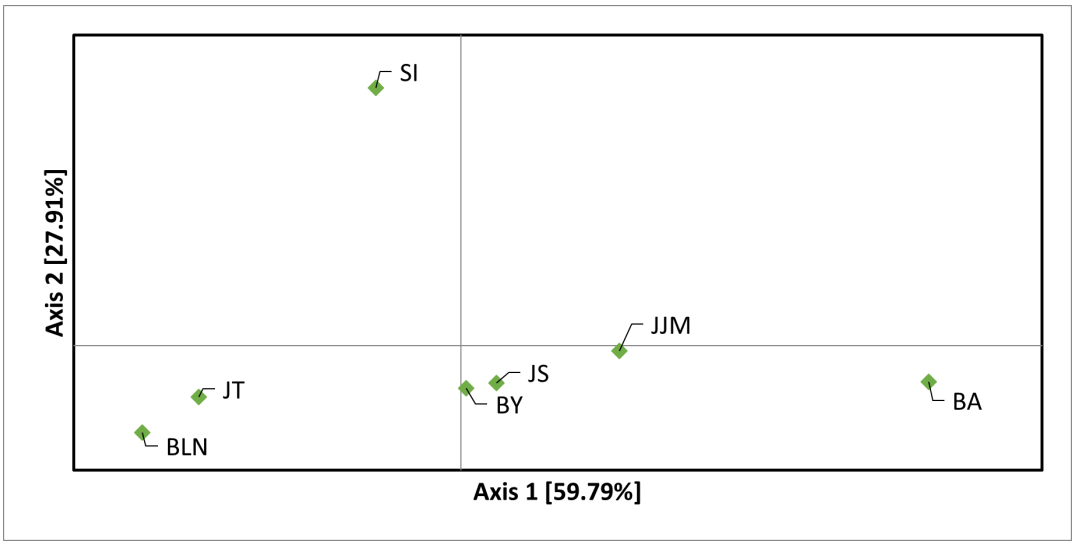


Fig 2.
Principal coordinates analysis (PCoA) of seven *Q. afares* localities based on a genetic distance matrix of 16 SSR loci. Note: The first and second axes extracted 59.79 % and 27.91 % of the total genetic variance, respectively.

Generally, the genetic distance between all localities was not high (Table 4). BY and JJM had no genetic distance, while BLN and BY, JJM and JS, BY and JT were closely related. However, BA was the most divergent from the two localities JT and BLN. SI showed moderate genetic distance from all other localities, ranging from 0.02 to 0.05. This suggests SI is distinct but not as isolated as BA.

Table 4
Unbiased Nei's genetic distance matrix for the seven *Q. afares* localities.

| | BA | BLN | BY | JJM | JS | JT | SI |
|---|---|---|---|---|---|---|---|
| **BA** | - | | | | | | |
| **BLN** | 0.065 | - | | | | | |
| **BY** | 0.016 | 0.005 | - | | | | |
| **JJM** | 0.009 | 0.027 | 0.000 | - | | | |
| **JS** | 0.027 | 0.015 | 0.006 | 0.003 | - | | |
| **JT** | 0.056 | 0.000 | 0.005 | 0.015 | 0.021 | - | |
| **SI** | 0.054 | 0.035 | 0.021 | 0.023 | 0.029 | 0.029 | - |

As indicated by $F_{st}$ values, very low but significant genetic differences were found between BA and all other localities except JJM, the same observation has been noted with SI and other localities except BY. We also noted a significant difference between BLN and JJM as well as between JS and JT (Table 5). However, estimates of the global $F_{st}$ values ($F_{st} = 0.016$) (Table 6) showed little genetic differentiation for all loci. Moreover, based on $F_{st}$ estimates strong genetic flow between the localities was deduced, highest between JS and JJM (8333.08) and lowest between BLN and BA (5.76) (Table 5).

Table 5:
Pairwise population $F_{st}$ values (Weir and Cockerham, 1984) (lower triangle) and gene flow estimates ($N_m$) (upper triangle) for comparisons of *Q. afares* localities among the seven sampling locations. (* $p < 0.05$; ** $p < 0.001$)

| | BA | BLN | BY | JJM | JS | JT | SS |
|---|---|---|---|---|---|---|---|
| **BA** | - | 5.76 | 20.58 | 41.42 | 13.64 | 7.10 | 7.33 |
| **BLN** | **0.042**** | - | 83.08 | 14.45 | 27.52 | 0.00 | 12.91 |
| **BY** | **0.012*** | 0.003 | - | 0.00 | 124.75 | 83.08 | 22.48 |
| **JJM** | 0.006 | **0.017*** | 0.000 | - | 8333.08 | 31.00 | 18.98 |
| **JS** | **0.018*** | 0.009 | 0.002 | 0.00003 | - | 20.58 | 12.91 |
| **JT** | **0.034**** | 0.000 | 0.003 | 0.008 | **0.012*** | - | 18.98 |
| **SI** | **0.033**** | **0.019*** | 0.011 | **0.013*** | **0.019*** | **0.013*** | - |

Table 6:
Hierarchical AMOVA (Excoffier et al. 2005) and F-statistics analysis.

| Source of variation | df | % of variation | F-statistics |
|---|---|---|---|
| Among localities | 6 | 1.63 | 0.016 |
| Among individuals within localities | 91 | 0.00 | -0.107 |
| Within individuals | 98 | 98.37 | -0.089 |

The overall AMOVA results showed that only 1.63 % of the variation was distributed among localities, while 98.37 % was allocated within individuals (Table 6). This was further supported by global pairwise $F_{st}$ comparisons of *Q. afares* localities extracted from AMOVA, indicating that individuals from each location are not genetically different from each other.

Pairwise Lin$F_{st}$ values were calculated between the seven localities and no correlation was observed between lin$F_{st}$ and geographic distance (Mantel test: RXY = -0.029, $p = 0.49$) (Fig 3). Hence, geographical distance does not significantly influence the genetic differentiation among localities.

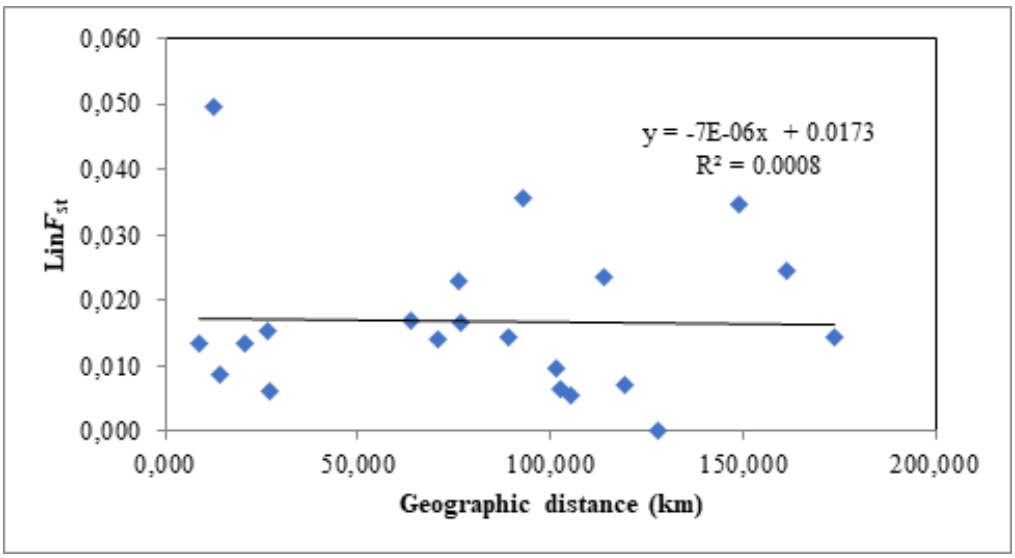


Fig 3.
Spatial genetic structure. No correlation was observed between the geographic distance and linearized $F_{st}$ (Mantel test: $R_{xy} = -0.29$, $p = 0.5$).

The application of the Evanno method to the Bayesian clustering analysis of the entire data set performed in STRUCTURE identified a peak at $K = 2$ in the plot of estimated $\Delta K$ versus $K$ (Fig S2, supplementary data). However, the $\Delta K$ value was very low suggesting minimal magnitude of the underlying genetic structure signal (Evanno et al. 2005). This weak genetic differentiation signal indicates a lack of distinct population clusters among the *Q. afares* samples. Furthermore, the clustering pattern with the use of LOCPRIOR was consistent with the analysis without LOCPRIOR (Fig S3, supplementary data), and the assignment tests in STRUCTURE demonstrated that the sampled locations likely represent one main population ($K = 1$) across all the samples.

## Discussion

### *Q. afares genetic diversity*

Species with small population size or limited geographic range are vulnerable due to reduced genetic diversity and demographic stochastic events. Therefore, genetic diversity surveys are essential for evaluating the evolutionary and adaptive potential of endangered species. Such studies are vital for developing effective conservation strategies to protect threatened species (Theissinger et al. 2023).

To the best of our knowledge, this is the first report on genetic diversity and structure of the vulnerable *Quercus afares* employing nuclear microsatellites. Besides located on different linkage groups, the markers selected for our study are dispersed rather regularly across the genome (Durand et al. 2010). In addition, the absence of null alleles and linkage disequilibrium between marker pairs implies the use of a representative marker set for analysis.

The absence of significant deviation from HWE implies that the observed genotype frequencies are not statistically distinguishable from those expected under HWE. A positive and significant $F_{is}$ value for a single marker (FIR043) might result from limited statistical power due to sampling effects (e.g., low sample sizes per population), insufficient allelic diversity, or stochastic processes such as genetic drift, mutation and gene flow (Greenbaum et al. 2014). Our results showed an absence of population structure and high level of gene flow inferred indirectly through model-based estimates of genetic differentiation (discussed below). Therefore, the significant $F_{is}$ could result from small sample size and biological factors, such as directional selection at closely linked loci (Kim and Maruki 2011), that the statistical test for HWE fails to detect, because of conservative thresholds or random genotyping errors (Sen and Burmeister 2008). It is also possible that the locus is influenced by micro-environmental selection pressures, causing localized homozygosity without affecting the overall equilibrium at the population level (Epperson 1992). On the other hand, PIE239 showed significant deviation from HWE with negative $F_{is}$ and low $F_{st}$ value. Besides sampling size effects (Lachance 2009), forces like balancing selection or gene flow can increase heterozygosity producing negative $F_{is}$ and lower $F_{st}$ values while disturbing overall genotype frequencies, causing HWE deviations (Meirmans and Hedrick 2011).

Our data set exhibited high levels of genetic diversity at the SSR loci examined, as well as of $A_r$ and the Shannon index. These results are in agreement with those observed in other threatened *Quercus* species in Mexico (Pingarroni et al. 2020), the United States (Spence et al. 2021) and China (An et al. 2017). Compared to *Q. suber, Q. afares* exhibited similar (Assemar et al. 2024) to higher genetic diversity (Sousa et al. 2022; Laakili et al. 2023). Despite *Q. afares* IUCN status (Jerome et al. 2020), our results highlight the overall good conditions of the studied localities of *Q. afares* in terms of gene diversity. $F_{is}$ values were negative for all localities under study except for SI, indicating that the *Q. afares* localities are largely outbreeding, in contrast with the high inbreeding coefficients typically expected in plants with small population sizes and shrinking habitat (Spigler et al. 2017). In addition, heterozygosity excess in some species may result from population regeneration, a small number of populations with few individuals, or high allelic richness in the parent plants (Cao et al. 2022).

Absence of genetic structure and high levels of gene flow

Genetic population structure arises from the combined influence of mutation, migration, selection, and genetic drift, allowing important information about species ecology and evolutionary history, including dispersal ability, adaptation, and speciation. Genetic population structure can also help to determine conservation management strategies in rare and endemic species as well as geographically restricted species (Backs and Ashle 2021). Here, no distinct phylogeographic structure was detected across the seven *Q. afares* localities in either the pairwise genetic tests or the cluster analysis. We found a very low genetic differentiation ($G''_{st} = 0.044$; $F_{st} = 0.016$), additionally, negative values were recorded between BLN and JT and between BY and JJM localities. Negative Fst values can occur when the within-population genetic variation is high compared to between-population variation (Table 5). In this situation, $F_{st}$ values are not representative of population structure, but rather reflect the statistic limitations under particular conditions, such as small population size or high genetic diversity within populations (Table 2, 3) (Karlin 2023), similarly, very low $F_{st}$ values demonstrate minimal genetic differentiation and suggest substantial gene flow between populations (Meirmans and Hedrick 2011). The lack of clear population structure and evidence of outbreeding align with the assumptions of Mir et al. (2006) who suggested that *Q. afares* exhibits gene combinations that occur at high frequencies, essentially due to assortative mating among individuals with similar genomic compositions rather than inbreeding. However, gene migration could still influence *Q. afares* genomic structure (Degen et al. 2021).

Gene flow is an important mechanism that can counteract the effects of genetic drift and natural selection, helping maintain genetic diversity within populations. High gene flow reduces genetic differentiation, while low gene flow allows differentiation to occur due to drift or selection (Spieth 1974). In *Q. afares*, estimates of gene flow among the localities were very high, consistent with AMOVA results, which identified more

than 98 % of genetic variation within localities. This pattern aligns with the general gene flow behavior in oak populations, where high gene flow and low inbreeding levels are common due to wind pollination, outcrossing, and long-distance pollen dispersal (Buschbom et al. 2011). However, we did not directly estimate gene flow (e.g., through paternity or parentage analyses), but rather inferred it indirectly through genetic differentiation measures, which are model-based and depend on assumptions that may not always hold in natural populations (Whitlock and Mccauley 1999). Specifically, the use of $N_m$ as a proxy for gene flow is known to have important limitations especially when population structure may deviate from model assumptions such non-equilibrium conditions, unequal population sizes, or spatial structure (Whitlock and Mccauley 1999), which has not been observed in our results.

The Mantel test is often used to determine the relationship between genetic and geographic distances between populations, providing insights into the influence of geographic factors on genetic distance (Legendre et al. 2015). Our results showed an absence of correlation between the two matrices supporting our findings of population structure absence and strong gene flow across the studied localities. Moreover, there was no evidence of recent bottleneck events, further supporting the stability of these localities over time.

### *Sample Size Consideration*

While some of localities in this study were represented by sample sizes close to the recommended threshold (Miyamoto et al. 2008; Hale et al. 2012), a few localities had lower numbers due to limitations in available material. It is acknowledged that small sample sizes can reduce the power to detect rare alleles and subtle population structure, and may introduce bias in allele frequency estimates (Puechmaille, 2016). However, the consistency of results across loci, the use of a sufficient number of polymorphic markers (Puechmaille, 2016), and the overall patterns observed —such as high gene flow and absence of significant structure—support the reliability of the findings. Our results are interpreted cautiously, and the potential limitations associated with smaller samples are taken into account in the conclusions.

### *Implications for conservation*

Establishment of in situ dynamic conservation units is the most common strategy in conservation of genetic diversity of forest trees. Dynamic genetic conservation aims to capture and maintain genetic diversity while ensuring that populations continue to evolve and adapt to changing environmental conditions over time. This is crucial, especially in the context of climate change and anthropogenic pressures that may affect the resilience and survival of species. Additionally, high genetic variation enhances species adaptation to selective pressures. Thus, investigating structure and genetic variation among vulnerable taxa is essential for developing effective management practices (Kahilainen et al. 2014).

The seven relict localities of *Q. afares* studied were characterized by high levels of genetic diversity (e.g., expected heterozygosity), an absence of genetic structure and high gene flow estimates among them, and despite the fact that most fragmented populations have a restricted geographic range, Afares oak must be considered as a metapopulation and should be defined as one protected unit. *Q. afares* is still at high risk from anthropogenic pressures, habitat loss, degradation, fragmentation, water pollution, and global warming. In the last decade, wildfires have become more frequent and severe in oak forests, increasing the vulnerability of *Q. afares* (Curt et al. 2020). Furthermore, a weak to complete absence of regeneration has been noted in the seven sampling areas, mainly due to uncontrolled livestock and weather fluctuation (data not shown). Thus, habitat conservation and population management are urgently needed to preserve the species in its natural environment. The high genetic variability of Afares oak indicates the species' ability to recover if provided with effective protection. In fact, *Q. afares* was extensively planted in an arboretum near Constantine (northeastern Algeria) between 1954 and 1967. The region is characterized by a semiarid to subhumid bioclimate, which differs from the species' usual humid to sub-humid natural bioclimate. Despite this, the Afares oak forest remains healthy (personal observations).

In the light of our results, we suggest the following conservation and management strategies for this species: 1) habitat protection and restoration, including the preservation of existing habitats, the restoration of degraded areas, the improvement of habitat quality and enhancing connectivity between fragmented localities. Notably, from the seven localities studied only JS is within Taza National Park (Sissaoui et al. 2017). 2) ex situ conservation strategy. Currently, conservation through living collections in botanical gardens and arboreta is the principal conservation option for threatened oak species (Kramer and Pence 2012). Therefore, we strongly encourage the creation of new arboreta along a bioclimatic range to ensure living collections and genetic diversity preservation. 3) Local resident engagement and awareness, involving local communities in recognizing the ecological importance of preserving the species and promoting sustainable land-use practices and livestock management. Protection efforts can also be enhanced by setting up conservation warning signs and emphasizing the value of oak forest ecosystems to local residents.

## Conclusion

In this study, we estimated the genetic diversity and population structure of seven relict localities of the endemic and vulnerable Afares oak in northeastern Algeria. *Q. afares* localities showed high genetic diversity, which was unexpected given the conservation status of the species. An absence of population structure and low genetic differentiation were observed. Although gene flow was not directly measured, the genetic patterns suggest a high potential for connectivity among localities. Consequently, we suggest that *Q. afares* in Algeria should be considered as a single protected unit for management purposes. Conservation efforts should be based on both in situ

and ex situ conservation units, as well as the involvement of the nearby residents in recognizing the ecological importance of the species.

## Acknowledgements

We are grateful and thankful to Dr. Hicham Berrabah for his assistance in sampling across Afares forests, and Alexandra Dolynska for her valuable technical assistance.

### References

Aldrich PR, Cavender-Bares J (2011) Quercus. In: Kole C (eds) Wild Crop Relatives: Genomic and Breeding Resources: Forest Trees. Springer Berlin Heidelberg, pp 89–129, ISBN 978-3-642-21249-9

An M, Deng M, Zheng SS, Jiang XL, Song YG (2017) Introgression threatens the genetic diversity of *quercus austrocochinchinensis* (Fagaceae), an endangered oak: A case inferred by molecular markers. Frontiers in Plant Sciences 8:1–15. https://doi:10.3389/fpls.2017.00229

Assemar FE, Alami M, Rabeh K, Medraoui L, El Antri S, Filali-Maltouf A, Belkadi B (2024) Genetic diversity and population structure in *Quercus suber* L. revealed by nuclear microsatellite markers and generation of a core collection. Tree Genetics & Genomes 20(1). https://doi.org/10.1007/s11295-024-01638-w.

Backs JR, Ashle M (2021) *Quercus* conservation genetics and genomics: Past, present, and future. Forests 12(7):882. https://doi:10.3390/f12070882

Balloux F, Brunner H, Lugon-Moulin N, Hausser J, Goudet J (2000) Microsatellites can be misleading: An empirical and simulation study. Evolution 54(4):1414–1422. https://doi:10.1111/j.0014-3820.2000.tb00573.x

Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF (2018) Present and future köppen-geiger climate classification maps at 1-km resolution. Scientific Data 5:1–12. https://doi:10.1038/sdata.2018.214

Burger K, Gailing O. 2022. Genetic variability of indigenous (Quercus robur L.) and late flushing oak (*Quercus robur* L. subsp. *slavonica* (Gáyer) Mátyás) in adult stands compared with their natural regeneration. European Journal of Forest Research 141(6):1073–1088. https://doi:10.1007/s10342-022-01491-3

Buschbom J, Yanbaev Y, Degen B (2011) Efficient long-distance gene flow into an isolated relict oak stand. Journal of Heredity 102(4):464–472. https:// https://doi:10.1093/jhered/esr023

Cao Y, Ma Y, Li Z, Liu X, Liu D, Qu S, Ma H (2022) Genetic diversity and population structure of *Rhododendron longipedicellatum*, an endangered species. Tropical Conservation Science 15: 1–10. https://doi:10.1177/19400829221078112

Curt T, Aini A, Dupire S (2020) Fire activity in Mediterranean forests (The Algerian case). Fire 3(4):1–17. https://doi:10.3390/fire3040058

Degen B, Yanbaev Y, Mader M, Ianbaev R, Bakhtina S, Schroeder H, Blanc-Jolivet C (2021) Impact of gene flow and introgression on the range wide genetic structure of *Quercus robur* (L.) in Europe. Forests 12(10). https://doi:10.3390/f12101425

Denk T, Grimm GW, Hipp AL, Bouchal JM, Schulze ED, Simeone MC (2023) Niche evolution in a northern temperate tree lineage: biogeographical legacies in cork oaks (*Quercus* section *Cerris*). Annals of Botany 131(5):769–787. https:// https://doi:10.1093/aob/mcad032

Dobeš C, Konrad H, Geburek T (2017) Potential population genetic consequences of habitat fragmentation in central European forest trees and associated understorey species-An introductory survey. Diversity 9(1). https://doi:10.3390/d9010009

Durand J, Bodénès C, Chancerel E, Frigerio JM, Vendramin G, Sebastiani F, Buonamici A, Gailing O, Koelewijn HP, Villani F, et al. (2010) A fast and cost-effective approach to develop and map EST-SSR markers: Oak as a case study. BMC Genomics 11(1):570. https://doi:10.1186/1471-2164-11-570 http://www.biomedcentral.com/1471-2164/11/570.

Epperson BK (1992) Spatial structure of genetic variation within populations of forest trees. New Forests 6:257–278. https://doi:10.1007/BF00120648

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular Ecology 14(8):2611–2620. https://doi:10.1111/j.1365-294X.2005.02553.x

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. Genetics 131:479–491.

Feng J, Dan X, Cui Y, Gong Y, Peng M, Sang Y, Ingvarsson PK, Wang J (2024) Integrating evolutionary genomics of forest trees to inform future tree breeding amid rapid climate change. Plant Communications 5(10):101044. https://doi:10.1016/j.xplc.2024.101044

Gailing O, Curtu AL (2014) Interspecific gene flow and maintenance of species integrity in oaks. Annals of Forest Research 57(1):5–18. https://doi:10.15287/afr.2014.171

Greenbaum G, Templeton AR, Zarmi Y, Bar-David S (2014) Allelic richness following population founding events - A stochastic modeling framework incorporating gene flow and genetic drift. PLoS One 9(12):1–23. https://doi:10.1371/journal.pone.0115203

Hale ML, Burg TM, Steeves TE (2012) Sampling for microsatellite-based population genetic studies: 25 to 30 Individuals per Population Is Enough to Accurately Estimate Allele Frequencies. PLoS One 7(9). https://doi:10.1371/journal.pone.0045170

Hipp AL, Manos PS, Hahn M, Avishai M, Bodénès C, Cavender-Bares J, Crowl AA, Deng M, Denk T, Fitz-Gibbon S, et al. (2020) Genomic landscape of the global oak phylogeny. New Phytolist 226(4):1198–1212. https://doi:10.1111/nph.16162

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9(5):1322–1332. https://doi:10.1111/j.1755-0998.2009.02591.x

Jerome D, Carrero C, Gorener V (2020) *Quercus afares*. IUCN Red List Threat Species 2020 eT78802767A85409923. doi:https://dx.doi.org/10.2305/IUCN.UK.2020- 2.RLTS.T78802767A85409923.en Copyright:

Journal officiel de la republique algerienne (2012) Décret exécutif n° 12-03 du 10 Safar 1433 correspondant au 4 janvier 2012 fixant la liste des espèces végétales non cultivées protégées. N° 03:1–11

Kadi-Bennane S, Takilt D, Bedad K, Harchaoui-Bournine C, Mezaour N, Adjoud-Sadadou D (2020) Preliminary study of the mycorrhizal community according to the age of the Afares oak (*Quercus afares*) of the Akfadou forest, Algeria. Environmental and Experimental Biology 18(4):237–247. https:// https://doi:10.22364/eeb.18.24

Kahilainen A, Puurtinen M, Kotiaho JS (2014) Conservation implications of species-genetic diversity correlations. Global Ecology and Conservation 2:315–323. https://doi:10.1016/j.gecco.2014.10.013

Karlin EF (2023) A Comparison of Entropic Diversity and Variance in the Study of Population Structure. Entropy 25(3). https://doi:10.3390/e25030492

Kim Y, Maruki T (2011) Hitchhiking effect of a beneficial mutation spreading in a subdivided population. Genetics 189(1):213–226. https://doi:10.1534/genetics.111.130203

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources 15(5):1179–1191. https://doi:10.1111/1755-0998.12387

Kramer A, Pence V (2012) The challenges of ex situ conservation for threatened oaks. International Oak Journal 23(23):91–107.

Kremer A, Ronce O, Robledo-Arnuncio JJ, Guillaume F, Bohrer G, Nathan R, Bridle JR, Gomulkiewicz R, Klein EK, Ritland K, et al. (2012) Long-distance gene flow and adaptation of forest trees to rapid climate change. Ecology Letters 15(4):378–392. https://doi:10.1111/j.1461-0248.2012.01746.x

Laakili A, Medraoui L, Alami M, Pakhrou O, Antry S El, Makhloufi M, Lamarti A, Filali-maltouf A, Belkadi B (2023) Genetic Resources and Crop Evolution Genotypic and Genetic diversity of Maâmora's Cork oak forest using SSR and ISSR markers for conservation. Genetic Resources and Crop Evolution 71:2399–2414. https://doi:10.1007/s10722-023-01774-7

Lachance J (2009) Detecting selection-induced departures from Hardy-Weinberg proportions. Genetics Selection Evolution 41(1):1–6. https://doi:10.1186/1297-9686-41-15

Legendre P, Fortin MJ, Borcard D (2015) Should the Mantel test be used in spatial analysis? Methods in Ecology and Evolution 6(11):1239–1247. https://doi:10.1111/2041-210X.12425

Luikart G, Cornuet JM (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology 12(1):228–237. https://doi:10.1046/j.1523-1739.1998.96388.x

Meirmans PG, Hedrick PW (2011) Assessing population structure: FST and related measures. Molecular Ecology Resources 11(1):5–18. https://doi:10.1111/j.1755-0998.2010.02927.x

Mhamdi S, Brendel O, Montpied P, Ghouil-Amimi H, Hasnaoui I, Dreyer E (2013) Leaf morphology displays no detectable spatial organisation in the relict *Quercus afares Pomel* compared to the co-occurring parental species *Q. canariensis* Willd. and *Q. suber* L. Annals of Forest Science 70(7):675–684. https://doi:10.1007/s13595-013-0312-x

Mir C, Toumi L, Jarne P, Sarda V, Di Giusto F, Lumaret R (2006) Endemic North African *Quercus afares* Pomel originates from hybridisation between two genetically very distant oak species (*Q. suber* L. and *Q. canariensis* Willd.): Evidence from nuclea and cytoplasmic markers. Heredity 96(2):175–184. https://doi:10.1038/sj.hdy.6800782

Miyamoto N, Fernández-Manjarrés JF, Morand-Prieur ME, Bertolino P, Frascaria-Lacoste N (2008) What sampling is needed for reliable estimations of genetic diversity in *Fraxinus excelsior* L. (*Oleaceae*)?. Annals of Forest Science 65(4):403–403. https://doi:10.1051/forest:2008014

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89(3):583–590. https://doi:10.1093/genetics/89.3.583

Neophytou C, Heer K, Milesi P, Peter M, Pyhäjärvi T, Westergren M, Rellstab C, Gugerli F (2022) Genomics and adaptation in forest ecosystems. Tree Genetics and Genomes 18(2):1–7. https://doi:10.1007/s11295-022-01542-1

Ortego J, Bonal R, Muñoz A (2010) Genetic consequences of habitat fragmentation in long-lived tree species: The case of the mediterranean holm oak (Quercus ilex L.). Journal of Heredity 101(6):717–726. doi:10.1093/jhered/esq081

Peakall R, Smouse PE (2012) GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28(19):2537–2539. https://doi:10.1093/bioinformatics/bts460

Pingarroni A, Molina-Garay C, Rosas-Osorio C, Alfonso-Corrado C, Clark-Tapia R, Monsalvo-Reyes A, Campos JE (2020) Abundance and genetic diversity of Quercus mulleri, microendemic, thretened species from Oaxaca. Madera y Bosques 26(1):1–12. doi:10.21829/myb.2020.2611782.

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity 90(4):502–503. https://doi:10.1093/jhered/90.4.502

Puechmaille SJ (2016) The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Molecular Ecologt Resources 16(3):608-27. https://doi: 10.1111/1755-0998.12512

Rice W (1989) Analyzing tables of statistical tests. Evolution 43(1):223–225.

Sakka H, Baraket G, Abdessemad A, Tounsi K, Ksontini M, Salhi-Hannachi A (2015) Molecular phylogeny and genetic diversity of Tunisian Quercus species using chloroplast DNA CAPS markers. Biochemical Systematics and Ecology 60:258–265. https://doi:10.1016/j.bse.2014.09.025

Schneider S, Kueffer JM, Roessli D, Excoffier L. Arlequin ver 1.1: a software package for population genetics. Genetics and Biometry Laboratory, University of Geneva, Switzerland. 2000.

Sen S, Burmeister M (2008) Hardy-Weinberg analysis of a large set of published association studies reveals genotyping error and a deficit of heterozygotes across multiple loci. Human genomics 3(1):36–52. https://doi:10.1186/1479-7364-3-1-36

Simeone MC, Cardoni S, Piredda R, Imperatori F, Avishai M, Grimm GW, Denk T (2018) Comparative systematics and phylogeography of Quercus Section Cerris in western Eurasia: Inferences from plastid and nuclear DNA variation. PeerJ 2018(10) 6:e5793. https://doi:10.7717/peerj.5793

Sissaoui A, Baba R, Chabi N (2017) Protected areas for conservation of biodiversity and sustainable development in Algeria : the case of Taza National Park – Jijel. Journal of New Sciences C(32):2840–2852.

Slatkin M (1995) A measure of population subdivision based on micro-satellite allele frequencies. Genetics 139(1):457–462. https://doi.org/10.1093/genetics/139.1.457

Smouse PE, Banks SC, Peakall R (2017) Converting quadratic entropy to diversity: Both animals and alleles are diverse, but some are more diverse than others. PLoS One 12(10):1–19. https://doi:10.1371/journal.pone.0185499

Sousa F, Costa J, Ribeiro C, Varandas M, Pina-Martins F, Simões F, Matos J, Glushkova M, Miguel C, Veloso MM, et al. (2022) Population structure in Quercus suber L. revealed by nuclear microsatellite markers. PeerJ 10:1–22. https://doi:10.7717/peerj.13565

Spence ES, Fant JB, Gailing O, Griffith MP, Havens K, Hipp AL, Kadav P, Kramer A, Thompson P, Toppila R, et al. (2021) Comparing genetic diversity in three threatened oaks. Forests 12(5):1–18. https://doi:10.3390/f12050561

Spieth PT. 1974. Gene flow and genetic differentiation. Genetics 78(3):961–965. https://doi:10.1093/genetics/78.3.961

Spigler RB, Theodorou K, Chang SM (2017) Inbreeding depression and drift load in small populations at demographic disequilibrium. Evolution 71(1):81–94. https://doi:10.1111/evo.13103

Theissinger K, Fernandes C, Formenti G, Bista I, Berg PR, Bleidorn C, Bombarely A, Crottini A, Gallo GR, Godoy JA, et al. (2023) How genomics can help biodiversity conservation. Trends in Genetics 39(7):545–559. https://doi:10.1016/j.tig.2023.01.005

de Vries SMG, Alan M, Bozzano M, Burianek V, Collin E, Cottrell J, Ivankovic M, Kelleher CT, Koskela J, Rotach P, et al. (2015) Pan-European strategy for genetic conservation of forest trees and establishment of a core network of dynamic conservation units. European Forest Genetic Resources Programme (EUFORGEN); Bioversity International: Rome, Italy, 2015. https://doi.org/10.1021/jo00237a048

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. Evolution 38(6):1358–1370.

Whitlock MC, Mccauley DE (1999) Indirect measures of gene flow and migration: FST ≠ 1/(4Nm + 1). Heredity 82(2):117–125. https://doi:10.1038/sj.hdy.6884960