



## Weak genetic differentiation but strong climate-induced selective pressure toward the rear edge of mountain pine in north-eastern Spain



Belén Méndez-Cea<sup>a,\*</sup>, Isabel García-García<sup>a</sup>, Antonio Gazol<sup>b</sup>, J. Julio Camarero<sup>b</sup>, Ester González de Andrés<sup>b</sup>, Michele Colangelo<sup>b,c</sup>, Cristina Valeriano<sup>b</sup>, Francisco Javier Gallego<sup>a</sup>, Juan Carlos Linares<sup>d</sup>

<sup>a</sup> Dpto. Genética, Fisiología y Microbiología, Unidad de Genética, Facultad de CC Biológicas, Universidad Complutense de Madrid, 28040, Spain

<sup>b</sup> Pyrenean Institute of Ecology (IPE-CSIC), Zaragoza E-50059, Spain

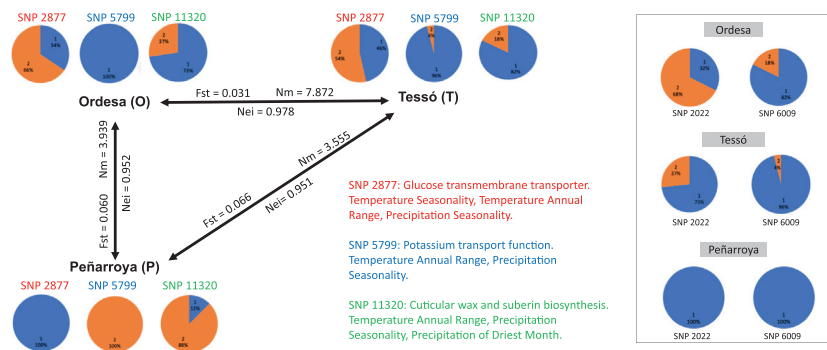
<sup>c</sup> School of Agricultural, Forest, Food and Environmental Sciences (SAFE), University of Basilicata, Potenza 85100, Italy

<sup>d</sup> Dpto. Sistemas Físicos, Químicos y Naturales, Universidad Pablo de Olavide, 41013 Sevilla, Spain

### HIGHLIGHTS

- Low heterozygosity and genetic differentiation were not observed in the relict population.
- Genetic data support gene flow among populations
- Genotype/environment analyses revealed effects of thermal seasonality and soil organic matter.
- Leaf mass area showed association with several single nucleotide polymorphisms.
- *P. uncinata* shows significant risk of non adaptedness based on expected changes in climate.

### GRAPHICAL ABSTRACT



### ARTICLE INFO

Editor: Manuel Esteban Lucas-Borja

Keywords:

Rear edge

Treeline

Selection signature

Genotype-environment associations

Genotype-phenotype associations

Risk of non-adaptedness

### ABSTRACT

Local differentiation at distribution limits may influence species' adaptive capacity to environmental changes. However, drivers, such as gene flow and local selection, are still poorly understood. We focus on the role played by range limits in mountain forests to test the hypothesis that relict tree populations are subjected to genetic differentiation and local adaptation. Two alpine treelines of mountain pine (*Pinus uncinata* Ram. ex DC) were investigated in the Spanish Pyrenees. Further, an isolated relict population forming the species' southernmost distribution limit in north-eastern Spain was also investigated. Using genotyping by sequencing, a genetic matrix conformed by single nucleotide polymorphisms (SNPs) was obtained. This matrix was used to perform genotype-environment and genotype-phenotype associations, as well as to model risk of non-adaptedness. Increasing climate seasonality appears as an essential element in the interpretation of SNPs subjected to selective pressures. Genetic differentiations were overall weak. The differences in leaf mass area and radial growth rate, as well as the identification of several SNPs subjected to selective pressures, exceeded neutral predictions of differentiation among populations. Despite genetic drift might prevail in the isolated population, the  $F_{st}$  values (0.060 and 0.066) showed a moderate genetic drift and  $Nm$  values (3.939 and 3.555) indicate the presence of gene flow between the relict population and both treelines. Nonetheless, the SNPs subjected to selection pressures provide evidences of possible selection in treeline ecotones. Persistence in range boundaries seems to involve several selective pressures in species' traits, which were significantly related to enhanced drought seasonality at the limit of *P. uncinata* distribution range. We conclude that gene flow is unlikely to constrain adaptation in the *P. uncinata* rear edge, although this species shows vulnerability to future climate change scenarios involving warmer and drier conditions.

\* Corresponding author.

E-mail addresses: [belenmen@ucm.es](mailto:belenmen@ucm.es) (B. Méndez-Cea), [isabel06@ucm.es](mailto:isabel06@ucm.es) (I. García-García), [agazol@ipe.csic.es](mailto:agazol@ipe.csic.es) (A. Gazol), [jjcamarero@ipe.csic.es](mailto:jjcamarero@ipe.csic.es) (J.J. Camarero), [ester.gonzalez@ipe.csic.es](mailto:ester.gonzalez@ipe.csic.es) (E.G. de Andrés), [cvaleriano@ipe.csic.es](mailto:cvaleriano@ipe.csic.es) (C. Valeriano), [fjgalleg@ucm.es](mailto:fjgalleg@ucm.es) (F.J. Gallego), [jlincal@upo.es](mailto:jlincal@upo.es) (J.C. Linares).

<http://dx.doi.org/10.1016/j.scitotenv.2022.159778>

Received 26 July 2022; Received in revised form 22 October 2022; Accepted 24 October 2022

Available online 26 October 2022

0048-9697/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

## 1. Introduction

Current climate change enhances the interest of the evolutionary significance of species' geographic range limits and peripherally isolated populations (Angert et al., 2020). Elucidating the processes determining the limits of a species range is particularly vital in predicting how species will respond to warming climate, as range-edge populations are assumed to be more sensitive to environmental changes (Hoffmann and Blows, 1994; Sexton et al., 2009; Holt, 2003; Angert et al., 2020). Hence, the fate of some species may depend upon the capacity of populations to cope with current climate trends, where the southernmost range margin often represents the point beyond drier climatic conditions exceed the tolerance limits of a given species (Bridle and Vines, 2007). Furthermore, range-edge populations are particularly important, in the face of adaptive management to rapid climate change, because dispersal, establishment, and extinction processes determine theoretically the pace of range shifts and whether a species' geographic range contracts or expands (Alexander et al., 2018).

Global warming provides a strong directional environmental change. This selective pressure may either enhance or reduce genetic differentiation and local adaptation at the limits of species' geographical ranges, depending on the magnitude and directionality of gene flow (Hoffmann and Sgrò, 2011; Shaw and Etterson, 2012; Alberto et al., 2013). Adaptive capacity to climate change in range-edge populations is influenced, among others, by the availability of genetic variation, the strength of natural selection and the extent of gene flow (Hoffmann and Sgrò, 2011).

In mountain forested landscapes, long-term isolation and genetic drift may limit or preclude to more extent the adaptation of relict sparser tree populations at the periphery, even if they locally experience intense directional selection (Hampe and Petit, 2005; Hampe and Jump, 2011; Kottler et al., 2021). By opposite, the altitudinal treelines provide a clinally shift, for instance in the temperature optimum for tree growth (Camarero and Gutiérrez, 2004). Differentiation might occur at these ecotonal tree populations (Bontrager et al., 2021). Meanwhile this divergence from the main forest population depends not only on gene flow and selection, but it might involve genetic drift or any other stochastic component, such founder event mechanisms (Angert et al., 2020).

Changes of treeline positions have occurred worldwide, matching with recent climate warming and leading to upward migrating in some treelines limited by low temperature (e.g., Harsch et al., 2009; Du et al., 2018; Sigdel et al., 2018). Besides, densification processes within the treeline ecotones have been also widely observed (Camarero et al., 2017; Feuillet et al., 2019; Davis et al., 2020). Hence, alpine treeline ecotones may originate a relatively fast component of genetic variation, ranging from dense forest to sparse individuals located at higher elevation, essentially similar to that ordinarily considered by genetic models of evolution at the edge of geographical ranges. Despite landscape features may influence patterns of gene flow and spatial genetic structuring, comparisons between the fine-scale genetic structures of contrasting range-limit populations are scarce. This knowledge is relevant, as fine-scale genetic structures may be reflecting the limitations for dispersal from the forests and lags in the establishment of recruits in the treeline ecotone (Alexander et al., 2018).

Hence, contrasting rear-edge populations may provide a demographical framework where the selective pressure of climate change might be detected (Bontrager et al., 2021). We focus on the role played by range limits in mountain forests comparing relict and treeline populations to investigate patterns of genetic differentiation and local adaptation using mountain pine (*Pinus uncinata* Ram. ex DC) as experimental system (Gazol et al., 2022). We compare a relict population forming the southernmost species' distribution limit in Europe and two treelines located in the Spanish Pyrenees.

Studies about how forest species respond to environmental variations are essential to understand how they cope with climate change. Several studies on the response of *P. uncinata* to climate are available (e.g., Camarero and Gutiérrez, 2007; Camarero et al., 2021; Sánchez-Salguero et al., 2017; Sanmiguel-Vallelado et al., 2019). However, they focus on morphological traits, with few exceptions which determine genetic variation between populations (Dzialuk et al., 2009) using chloroplast microsatellites (cpSSRs)

or describe evolutionary history (Heuertz et al., 2010; Zaborowska et al., 2019) using mitochondrial DNA markers. Others have used a small amount of neutral nuclear RAPD (Random Amplified Polymorphic DNA) markers (Monteleone et al., 2007), and have established links between genetic diversity and fitness (González-Díaz et al., 2020) using cpSSRs.

Huge genome sizes, absence of reference genomes and annotations, and massive amounts of data to analyze are the main reasons limiting genome-wide studies within conifer species (García-García et al., 2022) such as *P. uncinata*. Nonetheless, current genotyping by sequencing (GBS) methods enable the implementation of single nucleotide polymorphisms (SNPs) analysis, which provides a genome-wide coverage, even for non-model organisms, allowing to check for adaptive *loci*, as well as neutral ones (Unamba et al., 2015). Specifically, as tree density decreases upwards, from the forest limit to the treeline, climate, and micro-site conditions (mainly temperature, wind, light, and soil properties) shift gradually (Batllori et al., 2009; Gazol et al., 2022). According to that, we expect a clinal pattern of selection that favors different values of quantitative traits, derived by local selection pressure and gene flow. Even with random establishment of individuals, we hypothesized a net flux of migrants from the forest to the treeline, whose genetic differentiation reflects the balance between selection/local adaptation and gene flow. In contrast, isolated relict tree populations on the rear edge might reflect genetic divergence likewise as a result of genetic drift, to more extent than by responses to local selection pressures. Despite this challenging framework, next-generation sequencing (NGS) GBS techniques, which are based on using restriction enzymes to fragment the genomes and reduce their complexity, provide an innovative tool. This work proves the usefulness of double digest restriction-site associated DNA sequencing (ddRAD-seq) technique to obtain single nucleotide polymorphisms (SNP) markers holding relevant information about the genetics of this species.

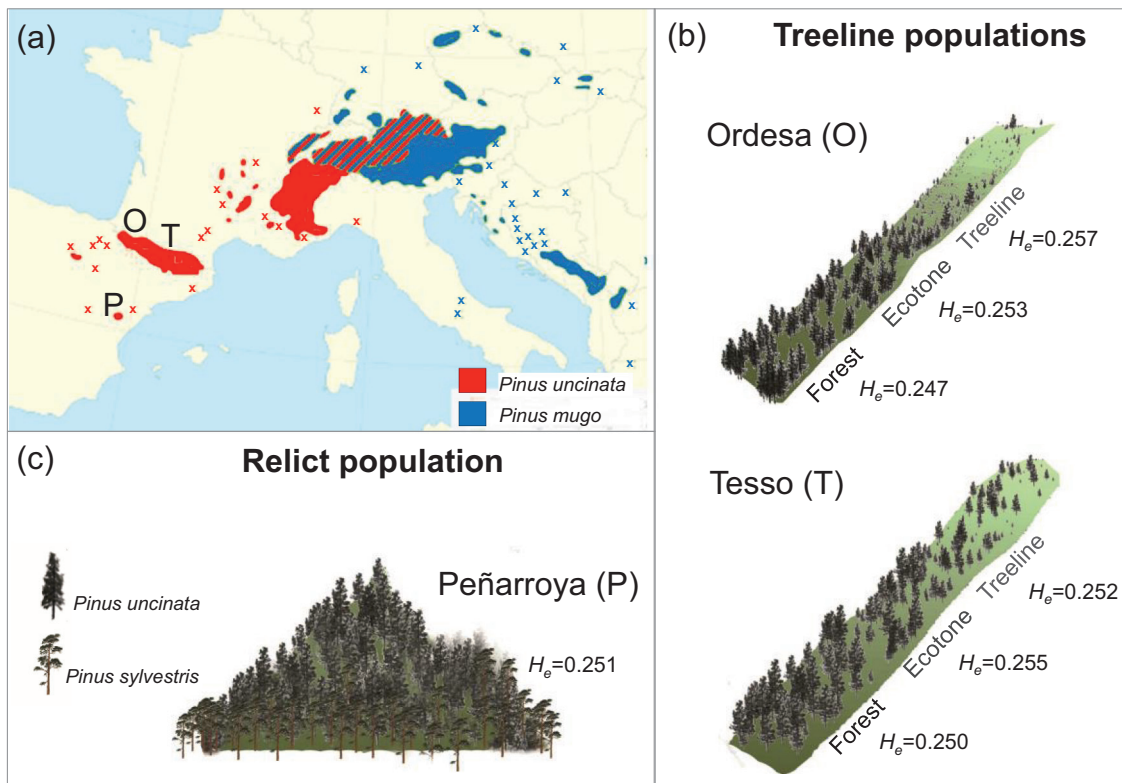
Hence, we obtained a genetic matrix based on SNP data, which allow us to determine diversity and differentiation among- and within-populations. Thereafter, SNPs subjected to selection were detected to determine the presence of associations between SNPs and bioclimatic variables, providing genotype-environment associations (GEA). On the other hand, the availability of functional traits data for the study individuals (Gazol et al., 2022), allows to investigate the genotype-phenotype associations. Finally, we modeled the risk of non-adaptedness (RONA) to estimate changes in allele frequency required within a given population to cope with the expected changes in climate under different emission scenarios (Pina-Martins et al., 2018).

We hypothesize lower genetic diversity and higher differentiation in the relict population, compared to the treeline ecotones, as a legacy of long term-isolation. We also hypothesize that neutral, non-adaptive, differentiation prevails in the relict population, according to genetic drift. By opposite, SNPs subjected to selection are hypothesized to be more frequent in the treeline ecotones, according to temperature rise. Finally, we hypothesize higher risk of non-adaptedness to the expected changes in climate under different emission scenarios in the relict population, according to drought increase.

## 2. Material and methods

### 2.1. Study sites description

We compared three *Pinus uncinata* Ram. ex DC sites corresponding to two structurally different treelines and a relict population (Fig. 1). The two treeline ecotones were located in the Central Spanish Pyrenees, in 'Ordesa y Monte Perdido' and 'Aigüestorters i Estany de Sant Maurici' National Parks, referred as Ordesa (O) and Tessó (T), thereafter. The O site (42.63° N, 0.08° W, 2100–2110 m a.s.l.) represents an abrupt treeline with strong differences in tree cover and height from the forest to the treeline. The aspect of this site is south, and the slope is 10–20°. Soils are rocky and mixed with calcareous and acid spots. The T site (42.58° N, 1.03° E, 2330–2360 m a.s.l.) represents a diffuse treeline with a gradual decrease in tree cover and height upwards (Camarero and Gutiérrez, 2004).



**Fig. 1.** (a) The map depicts the range of *Pinus uncinata* and the relative *Pinus mugo*, as well as the location of the three study sites in south-western Europe. Study sites: (b) sharp treeline (Ordesa, O), gradual treeline (Tesso, T); and (c) a relict population, restricted to the upper elevation of the mountain range (Peñarroya, P) where are populations of *Pinus sylvestris*; see more details about the study sites in Gazol et al. (2022). The values of expected heterozygosity ( $H_e$ ) are noted.

The treeline was defined as the highest elevation of 2-m tall trees where cover was below 20 %, whilst the forest limit corresponded to the highest elevation of continuous forest patches with tree cover >40 %. The aspect of the site is north-eastern, the slope is 25–30° and soils are calcareous and developed on shales. The rear-edge, relict population represents the southernmost distribution limit of *P. uncinata*, located in Peñarroya (site P, thereafter; 40.39° N, 0.57° W, 2010 m a.s.l.). This relict population grows in a flat terrain, where soils are developed on limestones (Gazol et al., 2022).

The climate is continental in the three sites. In the treelines O and T sites the mean annual temperatures have a range between 3° and 5 °C and total annual precipitation is ca. 1200–1660 mm (Camarero and Gutiérrez, 2002, 2004). Meanwhile the mean annual temperature in site P is 9 °C and its total annual precipitation is ca. 680 mm (Camarero and Gutiérrez, 2002).

## 2.2. Tree growth, functional traits and soil characteristics

Tree radial growth patterns were obtained by using dendrochronological methods, which allowed calculating the mean growth rate (either as mean ring width or as mean basal area increment) and the tree age at 1.3 m (see Gazol et al., 2022) top. Two cores were extracted at 1.3 m and perpendicular to the slope using 5-mm increment borers (Haglof, Sweden). Cores were air dried, glued onto supports, sanded, scanned at 2400 dpi resolution and visually cross-dated. Then, tree rings were measured with a 0.001 mm resolution using the CDendro software (Larsson and Larsson, 2018). In the treelines, we selected ten trees near the forest, ten in the ecotone and ten in the treeline. In the relict population, 15 trees were sampled and similarly measured. In total, 75 trees were sampled and processed (see Gazol et al., 2022 for details).

Leaf area (LA), specific leaf area (SLA), leaf mass area (LMA), leaf dry matter content (LDMC), and wood specific gravity (WSG) were used as basic functional traits characterizing competition for light and structural costs (Niinemets, 2001; Wright et al., 2004; Nardini, 2022). These traits

were measured in the same tree. Dendro-phenotypic traits such as resistance and recovery to extreme events, using estimate of secondary growth (basal area increments) resistance and recovery to 2012 drought event were estimated. Relationships with climate sensitivity were also investigated using climate variables significantly related to secondary growth (tree-ring width indexes) by dendrochronological analyses (April and May mean temperatures; Gazol et al., 2022).

Finally, soil samples were collected around the trunk (<50 cm in distance) and below the canopy projection of each sampled tree, to obtain soil texture, soil pH, soil carbon and nitrogen concentrations, and organic matter (see Gazol et al., 2022 for details).

## 2.3. SNP genotyping

Fresh leaf samples (needles) were collected from 75 *P. uncinata* individuals. 100 mg of each sample were lyophilized, and total genomic DNA was extracted using DNeasy Plant Mini Kit (Qiagen®, Germany) following the manufacturer's instructions with some modifications. DNA concentration was measured on a NanoDrop™ spectrophotometer (Thermo Scientific) and an electrophoresis in agarose gel was carried out to determine DNA quality. 72 samples had the required quality to proceed with NGS. ddRAD-seq (Peterson et al., 2012) libraries were constructed using ApeKI and PstI restriction enzymes and sequenced by LGC Genomics (Germany). ddRAD-seq is a technique based on NGS technology which requires no prior genomic knowledge about study species. It is suitable for huge genome size species due to restriction enzymes cutting reduces the complexity of the genome. The ddRAD-seq allows us to describe a large amount of new molecular markers such as SNPs which could be used to carry out multiple types of studies as genotype-phenotype or genotype-environment associations studies (Peterson et al., 2012).

Paired-end reads were obtained with a read depth of 1 M. Quality of raw reads was checked using FastQC v0.11.9 (Andrews, 2010). Then, adapter sequences were trimmed, and low-quality reads were removed using fastp



v0.12.4 (Chen et al., 2018). A de novo assembly of the retained reads and a SNP calling were performed using ipyrad v.0.9.65 (Eaton and Overcast, 2020). The VCFtools v0.1.16 program package (Danecek et al., 2011) was used to filter for high-quality, informative SNPs for genetic structure, selection, and association analyses (biallelic, minimum allele frequency of 5 %, maximum missingness of 50 %, and 1 per locus to avoid linkage disequilibrium). Similarly, individuals with <50 % of the retained SNPs were removed. In addition, a more restrictive set of SNPs was created to perform the statistical genetic analysis, retaining SNPs that met the previous criteria but with a maximum missingness value of 25 %.

#### 2.4. Genetic structure of tree populations

Two different methods were applied to study the genetic structure of populations with the aim to obtain robust results: (1) principal component analysis (PCA) and (2) sparse non-negative matrix factorization analysis (sNMF). For the first approach, the plink2 2.00a2.3 software (Chang et al., 2015) --pca option was used. For the second, admixture coefficients were obtained using the snmf function of the LEA package v3.4.0 (Frichot and François, 2015) in R software (version 4.1.2 R Core Team, 2020). In total, 10 repetitions were run for different K values, number of ancestry populations, ranging from 1 to 6 and represented using the software package called pong (Behr et al., 2016). The cross-entropy criterion was used to determine the value of K that best explains the obtained structure.

The GenAlEx v6.5 software (Peakall and Smouse, 2006, 2012) was used to perform the statistical analyses of the genetic matrix. Population differentiation and gene flow were estimated by fixation indexes (Fst) and migration rate (Nm) statistics, respectively. The Nei's genetic distance among populations was also estimated. Heterozygosity, private alleles, and polymorphic loci were calculated. Shannon index was calculated to infer genetic diversity. Lastly, allele frequencies of those SNPs which were relevant for our study were calculated.

A molecular variance analysis (AMOVA) based on Fst coefficient was performed to determine the proportion of genetic variation attributable to differences among and within populations. AMOVA was estimated based on 9999 permutations.

#### 2.5. Detection of selection signatures

Detection of selection signatures was carried out with the BayeScan 2.1 software (Foll and Gaggiotti, 2008) using default parameters. Fixation indexes (Fst) with their respective q-values were obtained. This version of BayeScan directly calculates q-values using the false discovery rate (FDR) correction. Those SNPs with a q-value <0.05 were considered candidate genetic variants under selective pressure. To look for possible biological functions underlying their potential importance, the scaffolds with SNPs under selection (obtained from the ipyrad assembly) were queried against the BLAST (NCBI; Altschul et al., 1990) nucleotide database. When a match was obtained, the scaffolds were also queried against the non-redundant protein sequences database using BLASTx (NCBI; Gish and States, 1993). The E-value obtained in each BLAST describes the random noise of the background and give us information about the accuracy of the similarity of the protein functions found in BLAST.

#### 2.6. Associations studies

Two types of environmental variables were tested for association with genotype. For the first one the 19 bioclimatic variables from the WorldClim database (Fick and Hijmans, 2017), with a grid cell resolution of 30 s were used at population level. And the second one employed individual-level measurements of 8 soil variables (sand, lime and clay soil percentage, soil pH, soil organic carbon, organic matter, soil nitrogen, and carbon-to-nitrogen ratio) obtained as described in the soil data section.

After downloading the variables from WorldClim, the freeware QGIS 3.18 (Quantum Geographic Information System) (QGIS Development Team, 2022) was used to extract the bioclimatic variables for each population's

geographic position. Point sampling tool, included in QGIS, allowed us to take out information about those geographical points where are located the three study areas.

Genotype-environment associations (GEA) were assessed using the lfm (latent factor mixed models) function of the LEA R package v3.4.0 (Frichot and François, 2015) in R software (version 4.1.2 R Core Team, 2020). After an imputation step to fill in the missing data, each run was repeated a total of 20 times with 100,000 iterations and a burn-in of 50,000. Again, p-values were transformed into q-values with FDR correction and the significance threshold was 5 %. Then, the significant scaffolds were analyzed looking for homologies with the nucleotide and protein NCBI databases as it was described previously in the detection of selection signatures section.

A genome-wide association study (GWAS) was carried out to determine the associations between the imputed SNPs matrix and functional and dendro-phenotypic traits. The measurement of each trait is described in tree growth, functional traits and soil characteristics section.

GWAS was performed with the rrBLUP v4.6.1 R package (Endelman, 2011; Endelman and Jannink, 2012). A FDR correction was applied with the significance threshold set to 5 % and the SNPs which showed associations were analyzed looking for homologies with the nucleotide and protein NCBI databases as it was described previously.

#### 2.7. Risk of non-adaptedness

The value of RONA shows the theoretical percentage of change in allele frequency at loci associated with environmental variables required for a given population to survive changes in that variable (Pina-Martins et al., 2018). Consequently, the lower the RONA is, the more likely a population is to be able to adapt to the given new environmental conditions.

We used pyRona v0.3.6 (Pina-Martins et al., 2018) to calculate the RONA of each tree population to two different climate scenarios, predicted to happen by the end of this century (2081–2100): low emissions (RCP2.6), which limit the increase of global mean temperature to 2 °C, and high emissions (RCP8.5), whose limitation is 4.9 °C. The information was obtained using WorldClim database and QGIS (for more detail, see genotype-environment association section). We used the lfm results from the GEA analysis as input, with their q-values.

### 3. Results

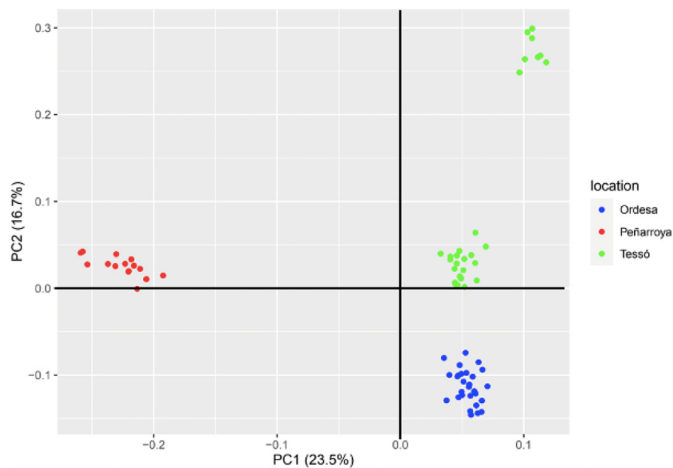
#### 3.1. Genetic structure of tree populations

A total of 11,904 SNPs and 71 individuals were retained after the filtering steps. These molecular markers were used to create the genetic matrix of *P. uncinata*. PCA showed, with a 23.5 % of variance attributed to the first principal component and 16.7 % to second principal component, 3 different groups: O and T (core treeline populations) and P (rear-edge population). One of the PCA axes separated O and T treelines from the P rear-edge site (Fig. 2). In addition, it is remarkable that 8 individuals of site T are separated of their own population: 5 of them belonging to forest, 2 to treeline and 1 to the treeline ecotone separating the forest from the treeline.

In order to explore those isolated points of site T, a PCA with only this population was carried out. This analysis revealed a cluster consisting mainly of ecotone and treeline samples and, again, the group with 8 individuals that was seen in the previous PCA, consisting mainly of trees located in the forest limit.

The cross-entropy analysis determined that K = 2 is the value that best explained the obtained genetic structure (Fig. S1). In contrast to the PCA, this analysis showed a marked genetic differentiation between a group formed by O and T treelines, and the remaining P relict population. Hence, the following analyses were carried out taking into consideration this genetic structure.

On the other hand, the statistical analyses were carried out with the most restrictive set of SNPs (75 %) which is composed of 72 individuals and 5374 SNPs.



**Fig. 2.** First two principal components of PCA based on genetic data (SNP matrix) obtained from ddRAD-seq. It is shown three different groups one for each population of study. Colors are related with the location of each point: blue for Ordesa, red for Peñarroya and green for Tessó.

The AMOVA indicated large differences within populations with 95 % of the total (Table S4). This is related with the  $F_{st}$  value obtained (0.047) which showed, with highly significance ( $p < 0.0001$ ), that there are no significant differences ( $p\text{-value} > 0.05$ ) between populations. Pairwise population  $F_{st}$  values are between 0.031 and 0.066 (Fig. 4; Table S1). The highest values are between P and the other two populations. However, those values are near 0, pointing that all the populations are in complete panmixia.

The gene flow which is measure with migration rate,  $Nm$ , showed a range value from 3.555 to 7.872. The higher value is found between O and T which indicates a panmictic behavior. On the other hand, P shows the lower values which indicates the presence of gene flow with the other two populations (Fig. 4; Table S2).

The average percentage of polymorphic alleles is  $90.47 \pm 5.21$  % being lower in site P (80.05 %). The number of private alleles is very similar in the O and T treelines ( $0.012 \pm 0.001$  and  $0.011 \pm 0.001$ , respectively), but lower in site P ( $0.007 \pm 0.001$ ). In terms of Shannon index values, genetic diversity is higher in O and T than in P. Observed heterozygosity ( $H_o$ ) has a range from 0.240 to 0.245 and expected heterozygosity ( $H_e$ ) ranges from 0.251 to 0.267. The lowest values for both parameters derive from the P relict populations. The average values of  $H_o$  and  $H_e$  are  $0.243 \pm 0.001$  and  $0.261 \pm 0.001$ , respectively.

Several differences among populations are showed in allele frequencies (Fig. 4). It is remarkable that site P has the highest percentage of fixated alleles followed by site T.

Finally, Nei's genetic distance allows us to explain the differences which are found between our populations. All of them are close to 1 which indicates that our 3 populations are very similar among them (Fig. 4; Table S3).

### 3.2. Detection of selection signatures

The BayeScan analyses showed 39 *loci* under selection (Table S5). The results indicate that these *loci* showed significant differences ( $q\text{-value} 5$  %) in their  $F_{st}$  values. Homology with previously described sequences was found for 17 of these sequences, and only 10 of them could be defined at protein level. It is important to emphasize that these matches with protein sequences are a first approach to the probable function of our study sequences. Table S6 shows sequence type, the name and functions of the proteins, and the  $E\text{-value}$  obtained for each match.

The most interesting proteins obtained were NADP-specific glutamate dehydrogenase, which is involved in the ammonium and nitrogen assimilation, aquaglyceroporin, which mediates water, glycerol, and other molecules flow through membranes, GABA transporter whose functions are related to stress response, and a protein related to the response to light stimuli.

### 3.3. Associations studies

The LEA package found 129 associations with 5 of the 8 soil variables (Table S7). Sand soil percentage was associated with 17 *loci*, soil pH with 15 and lime soil percentage with 3. Soil nitrogen and carbon-to-nitrogen ratio showed the highest number of associations with 40 and 54 *loci*, respectively. Additionally, one of these *loci* which are associated with soil nitrogen, is under selection too.

It is remarkable the presence of several *loci* associated with more than one variable. Such is the case of the soil nitrogen and carbon-to-nitrogen ratio variables, which shared 5 *loci*. Lime, sand, and pH soil shared 3 *loci*. Sand and lime percentage shared 2 associations. Finally, lime percentage and soil pH shared 1 *locus*. The alignments against the nucleotide database gave 42 matches, and 20 of them also got hits for protein homology (Table S8). Some of the identified proteins were related to stress response, such as glutathione peroxidase,  $\beta\text{-1,3-glucanasa}$ , E3 ubiquitin-protein ligase listerin and GABA transporter. It is highlighted that the *locus*, which is similar to GABA transporter, is under selection.

On the other hand, 395 associations were observed with the 19 bioclimatic variables (Table S9). The largest number of associations were found with variables related with temperature seasonality BIO4 (34 *loci*), minimum temperature of coldest month BIO6 (37 *loci*), temperature annual range BIO7 (44 *loci*), and precipitation seasonality BIO15 (43 *loci*). The majority of *loci* were shared with two or more variables (for more details, see Table S9). Considering these results, the relation found between genetic data and temperature gives us evidence of the important role which is performed by temperature in the survival of this species.

Moreover, it is important to highlight that 25 of the *loci* associated with bioclimatic variables were under selection. In terms of the function of some of these *loci*, it should be emphasized those related with nitrogen and ammonium assimilation, potassium transport, and light response.

Once the alignments were performed, 44 hits with nucleotide database were obtained, and 23 showed protein homologies too (Table S10). It is remarkable that some of these proteins are related with: electrons exchange in mitochondrial, response to light stimulus, and ion transports.

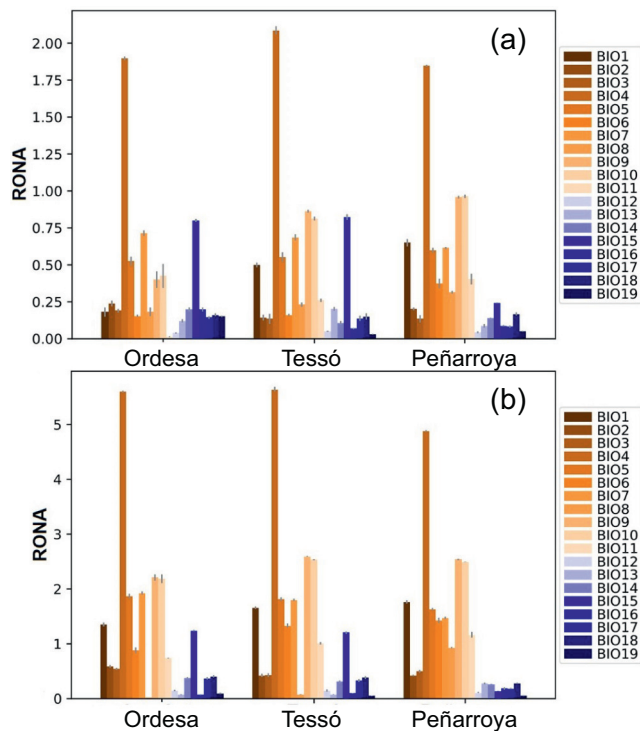
The results obtained with rrBLUP for GWAS study showed 22 associations with 7 of the 15 functional traits (Table S11) which were: mean tree-ring width (TRW), basal area increment (BAI), LDMC, LMA, SLA, WSG, and tree age. LDMC was associated with 8 *loci*, WSG with 5, BAI with 3, LMA with 3, TRW with 2, and tree age with 1 *locus*. As it is shown, LDMC had the higher number of associations. In this case only WGS shared 1 *locus* with LDMC and LMA shared other with LDMC. All of these associations had a  $q\text{-value}$  significance of 5 % (Table S11).

In total, 12 matches have been found against the nucleotide database and 7 of them obtained hits for protein homology (Table S12). The majority of the protein functions identified were related with metabolism or nucleic acid issues.

A total of 8 dendro-phenotypic variables were used to carry out another GWAS study with the same SNP matrix. In this case, the results showed no associations.

### 3.4. Risk of non-adaptedness

The low emission scenario (Fig. 3a) showed lower risk of non-adaptedness than the high emission one (Fig. 3b), as it could be expected, as the latter's variation in its BIOs values is higher. However, at least 4 environmental variables required changes in allele frequency higher than 50 % in both scenarios. Moreover, BIO4 (temperature seasonality) stands out, reaching values of around 200 % for the low emission scenario and higher than 500 % for the high emission one, becoming the most threatening environmental variable in both predictions. All populations showed similar RONA values, but the relict P populations was less capable to adapt to changes in variables related to temperature (BIO1-BIO11), while changes in precipitation-related variables (BIO12-BIO19) were not as challenging (Fig. 3).



**Fig. 3.** Risk of non-adaptedness (RONA) of each population to two different climatic scenarios. Environmental variables related to temperature are represented in orange colors (BIO1-BIO11), whilst environmental variables related to precipitation are represented in blue colors (BIO12-BIO19). (a) Low emission scenario RCP2.6. (b) High emission scenario RCP8.5.

#### 4. Discussion

Identifying constraints to adaptation at range edges remains a key challenge (Angert et al., 2020). In this work, we provide new insights regarding how evolutionary forces interact to shape adaptation at contrasting range margins of mountain pine. For this purpose, we focused on peripheral populations, which encompass the full altitudinal transition from closed forest to the treeline. This provides an analogy of species range expansion gradient while controlling for factors such as regional climate and colonization history. Furthermore, our sampling approach allows for fine-scale spatial analysis of genetic structuring at the rear distribution edge formed by the relict population.

Range limits are correlated with a number of abiotic and biotic factors, but underlying mechanisms are poorly understood (Sexton et al., 2009). It should be hypothesized a concomitant variation in the relative importance of genetic drift and gene flow due to limited size in the relict population (Peñarroya, P) and declining density upslope in the treelines (Ordesa, O and Tessó, T), respectively. Genetic drift leading to erosion of local genetic diversity and enhanced population differentiation should be expected to prevail at the relict P population and might occur near the upper margins of the treelines because of the smaller size and sparser distribution of tree populations, randomness associated with founding events, and limitations for seed and pollen flow (Hampe and Jump, 2011).

The results obtained in the cross-entropy study show two groups with a visible genetic differentiation: one consisting of O and T treeline sites, and another one consisting of the P relict, rear-edge populations on its own (e.g., Sjölund et al., 2019). On the contrary, PCA results suggested three separate populations, instead of two, indicating that some genetic differences do exist between the two treeline sites. This differentiation could be related with their constrain structures since O shows an abrupt treeline while T presents a gradual transition (Camarero and Gutiérrez, 2002, 2004), which implies a wider range of environmental conditions, so genetic variability could be expected along this gradient (Bontrager et al., 2021). It

is important to highlight that Peñarroya site represents the southernmost distribution limit of the species. Dzialuk et al. (2009) found that the populations from the southernmost range were the most genetically different by using cpSSRs, which is consistent with the results obtained in our study. Moreover, some evidence about the existence of genetic differences between O and P populations could be found in a previous study of González-Díaz et al. (2020) by using SSR as molecular markers.

In contrast to our preliminary hypothesis of genetic isolation, results support unambiguously a situation closer to a drift-gene flow equilibrium, where, despite the isolation and relative low populations sizes, gene exchange events via pollen or seeds seems not to have been infrequent, at least to the extent to determine significant genetic isolation. This interpretation is strongly supported by low values of genetic differentiation ( $F_{st}$ ) and high migration rates ( $N_m$ ), while within-population genetic diversity ( $H_e$ ) does not vary predictably along the forest-treeline gradient (O and T sites; Fig. 1), nor was significantly reduced in the relict population (P site; Fig. 1). A previous study showed low differentiations between populations of mountain pine (Dzialuk et al., 2009). However, the genetics studies found in the bibliography are performed using SSRs and we cannot make comparisons with the same kind of data obtained in the present study (SNPs). Predictions regarding trait divergence at relict populations were supported (see also Gazol et al., 2022). However, genetic variations across the leading edge defined by forest-to-treeline ecotones were not observed.

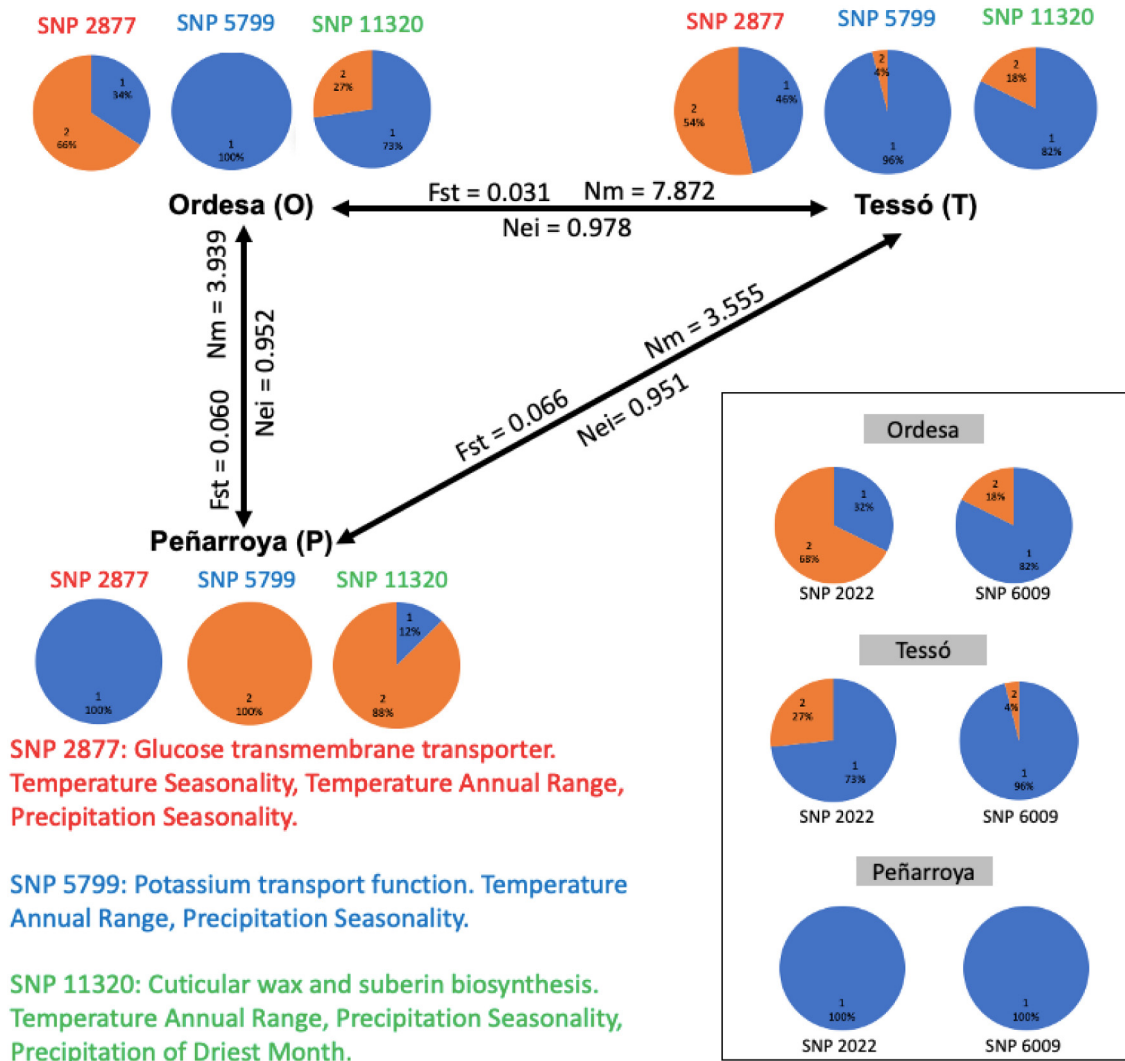
Nonetheless, we obtained significant evidence of possible selection in several SNPs. Some of these marks are in genes likely related with compounds involve in stress response such as transporters of ions, water, or GABA. When we obtained the alleles frequencies of the SNPs under selective pressures, some differences between populations are evidenced. For example, the SNP 2022, which is related with light gravitropic response (Fig. 4), showed the presence of both alleles in O and T sites, being more represented the reference one. It could be appreciated a decrease in the frequency of the alternative among these 2 populations, 18 % in O and 4 % in T. However, the P site presented only the reference allele (Fig. 4). This difference could be explained by the light variations which are found in P due to its drier Mediterranean climate with more radiation than northward treeline sites.

Other SNP under selective pressures is the SNP 6009 which is similar to an aquaglyceroporin (Table S3). Studies show modifications in the expression of this kind of proteins during drought stress periods (e.g., Mahdieh et al., 2008; Ding and Chaumont, 2020). However, the mechanisms which link drought stress response to aquaglyceroporin are poorly studied. In our study, the allele frequencies for this SNP showed differences among populations. Meanwhile O and T had a representation of both alleles, P showed only the reference allele (Fig. 4). These variations could indicate some advantages of having a higher proportion of the reference allele for the population with less amount of precipitation. Another example of these kind of SNPs is SNP 8051 which is a GABA transporter. Moreover, this SNP showed association with N variable in GEA study. Its allele frequencies show the same fixed allele in T and P sites, meanwhile the O site had more representation of the second allele. Due to GABA is involved in plant response to abiotic stresses, it increases the drought resistance in plants (Mekonnen et al., 2016), and this metabolite are related with C:N balance (e.g., Fait et al., 2011; Batushansky et al., 2014). This difference showed in allele frequencies may be caused by environmental variations, mainly in the soil composition, that are found in the distribution range of the T treeline and in the relict P population.

Therefore, the presence of possible selection pressure marks in this kind of genes, evidences the effect of climate variations on this species' genome and points out the relevance of keeping an eye on its genetics to be able to anticipate their response to climate change in the coming years.

The genetic differences observed among tree populations (Fig. 4) suggest that random genetic drift was not the prevailing force shaping the genetic structure of *P. uncinata* at its rear edge. It should be expected that random genetic drift increases the among-population component of diversity. This expectation is weakly supported by the genetic differentiation showed in PCA of the relict population, while this hypothesis appears not





**Fig. 4.** Examples of SNPs showing associations in GEA studies with our three *P. uncinata* populations. It is shown  $F_{st}$ ,  $N_m$  and Nei's genetic distance (Nei) values among populations. The allele percentages of these SNPs, whose functions are inferred for homology, are represented with circles: color blue for the reference allele and orange for the alternative allele. A brief description of the protein function which present homology for each SNP and of the bioclimatic variables which showed associations with these SNPs are indicated. The inset describes the allele percentages of two SNPs under selection.

reinforced at the reduced spatial scale of the forest-ecotone-treeline transitions. Furthermore, despite random genetic drift is also expected to reduce gene diversity within small populations (Hampe and Jump, 2011), our results support the maintenance of relatively high genetic diversity, which was in agreement with the conclusions reached by González-Díaz et al. (2020).

At a spatial scale shorter than those of the among-populations differences, as the population range spreads from the forest to the treeline, allele frequencies in each treeline transition might drift independently without any relation to the fine-scale spatial distances. Thus, random sampling of gametes may create a wide degree of scatter between points (Fig. S2). It must be noted, however, some relationship between genetic and treeline position was also observed in T. The presence of 8 individuals from T separated from their main nucleus in the PCA is remarkable. Since most part of forest samples can be found in this isolated group, a loss of genetic traits which were shared in the past with the ecotone and treeline areas could have happened.

Relict populations are assumed to have limited performance due to genetic drift and inbreeding depression (Hampe and Jump, 2011; Bontrager et al., 2021). Nonetheless, evidence for increased vulnerability to climate change in the relict population is lacking, based on risk of non-adaptedness

estimations (Fig. 3). Furthermore, the comparison of this relict population with the O and T treelines did not support the hypothesis that relict populations tend to have increased genetic isolation and loss of heterozygosity (Fig. 1). By opposite, adequacy between phenotypes and local environment (that is, local adaptation), might be hypothesized for the relict P population according to obtained SNPs frequencies and its relationships with climate seasonality and dryness (Fig. 4). However, the small population size may be limited to cope with changing environmental conditions if they are subjected to strong demographic stochasticity and competitive exclusion by neighboring species such *P. sylvestris* (Fig. 1). Gazol et al. (2022) found that Peñarroya did not show definite adaptation to cope with drier conditions in leaf traits or climate-growth sensitivity, despite the general trend of rear-edge populations being more vulnerable to drought, compared to populations located in the core distribution range (Camarero et al., 2021; Gazol et al., 2022).

Contrasting to the relict populations, the treelines show a stepping-stone pattern of density decrease from the forest to the upper ecotone (Camarero et al., 1998; Camarero and Gutiérrez, 2002, 2004), depicting a small-scale species range limit. Hence, an increased frequency of locally adapted genes might be expected, as climate warming acts as a selective pressure over tree recruitment and growth near or beyond

the tree line (Bontrager et al., 2021). Specifically, it should be expected that some genotypes boost population fitness at the treeline (Hargreaves and Eckert, 2019). However, we found no relationship between genetics and treeline position, suggesting that rapid climate change was not limiting gene flow from the forest to the expanding treeline. Notwithstanding, the role of epigenetics changes and phenotypic plasticity should not be discarded (Alberto et al., 2013; García-García et al., 2022).

Lack of evidence regarding climate-driven selection may be partially reflecting that land-use changes have played a more important role than climate in driving forest dynamics at a landscape scale over the last half century (e.g., Améztegui et al., 2010). The expansion of *P. uncinata* has been observed mainly by increasing canopy cover of pre-existing forests, either through enhanced growth of pre-existing individuals or the recruitment of new ones (Améztegui et al., 2010).

The associations found with GEA and GWAS studies allow us to identify the relationship between genetics and environmental and phenotypic variables, respectively. The presence of a high number of associations, in GEA studies, with nitrogen and C:N ratio reveals the importance of soil nutrients for tree growth. These associations could be related with less fertile soils in the treeline as compared with the forest, as was previously described by Gazol et al. (2022). Regarding the comparisons between sites, these authors reported that C:N ratio has the maximum in the O forest trees and the minimum in the P site.

Differences in allele frequencies allow us to have an idea about the relevance of the alleles and the homology protein (Fig. 4). For example, a potassium transporter (SNP 5799) showed huge differences in allele frequencies among populations. Mainly one of the alleles is fixed in site O and the other one in site P. Moreover, it is highlighted that this SNP is under selection too. It is relevant due to studies describing the important role of potassium in plant stress response to water shortage (Wang et al., 2013).

The ability to alter traits in response to a changing environment is particularly important for trees (Des Marais et al., 2013). Traits related to the carbon and water balance may define tree performance in a given environment, and therefore plasticity in these traits should provide adaptive capacity when conditions change (Wright et al., 2004). Leaf traits depict variations in resource investment widely associated with different evolutionary strategies across plant species (Reich et al., 2003; Wright et al., 2004). Low values of leaf mass area (LMA) reveal low investment in tissue density and nutrients, which has been related to high rates of photosynthesis and resource acquisition, but at the cost of longevity. On the other hand, high values of LMA reveal high investment in long-lived leaf material (Reich, 2014). The associations found with LMA and LDMC agree with the among-populations differences found in other leaf traits and soil properties (Gazol et al., 2022). In addition, in some cases, a relationship between genotype and LMA values was found (Fig. S3).

Several anatomical traits leading to high LMA are mechanistically correlated to physiological traits conferring tolerance to dehydration (Nardini, 2022). Given that the relict P population is subjected to drier conditions, a selective pressure for needles stiffness should be expected, as needles with high mass per unit area occur more frequently in water-limited habitats (Niinemets, 2001). However, the results regarding LMA were contrary to our expectations. In addition, Gazol et al. (2022) performed dendroecological analyses which revealed a stronger dependency of growth on water availability in this relict, rear-edge population than in the two alpine treeline sites. By opposite, as some of the anatomical modifications associated with high LMA and leaf mechanical stiffness provide leaf mechanical resistance, in addition to drought tolerance, the higher values of LMA obtained in the treelines (Fig. S2) might confer adaptive advantages to trees growing in this harsh environment, which is also cold- and nutrient-limited (Nardini, 2022).

Lastly, RONA values allowed us to predict the adaptation capacity of our populations to different climatic scenarios predicted for the end of the 21st

century. The results show that they are at risk of non-adaptedness to both the low and high emission scenarios. As pines generation times are long, taking about 20–25 years to reach sexual maturity (Camarero et al., 2017), the magnitude of change, which was obtained in this study, is extremely difficult to achieve in 60–80 years. Environmental variables related to temperature seem to pose a higher threat, which is consistent with the results obtained from GEA and selection analyses. The SNPs subjected to selection processes and associated with environmental variables were related to 7 out of 11 temperature BIOs and only 1 out of 8 precipitation BIOs. RONA studies have been performed with other species such as *Eucalyptus microcarpa* (Jordan et al., 2017), or *Quercus suber* (Pina-Martins et al., 2018).

Temperature seasonality (BIO4) is the most hazardous environmental variable for every population in both predictions. An increase of extreme meteorological events, including high and low temperature episodes, is expected in the future (Hoegh-Guldberg et al., 2018), leading to considerable changes in this BIO, such that any of our three populations' allele frequencies are likely to be able to match.

It is of particular concern that, while only 4–5 environmental variables seem to be threatening at the low emission scenario, this number rises to 8–10 when looking at the high emission expectation.

## 5. Conclusions

Forests dynamics in range boundaries seem to involve several selective pressures in species' traits, which were significantly related to the characteristics of harsh and marginal climates. Contrasting patterns of local adaptation among relict populations and treeline ecotones may be influenced by microsite conditions, while the dispersal and upward establishment in the treeline need to overcome several limitations to effectively result in further range shifts to higher elevation.

Increasing dryness and climate seasonality toward the species border appears as an essential element in the interpretation of regional genetic structuring of *P. uncinata* at its rear edge, and a factor which may influence the evolutionary potential of peripheral populations by local adaptation. Nevertheless, genetic differentiations were overall weak. Our results showed reduced among-populations genetic differences, while heterozygosity values were similar, contrasting to expectations of increasing differentiation and limited genetic diversity in relict populations. Further, among-population gen flow seems to prevail over any process of genetic drift or inbreeding.

By opposite, the differences in several phenotypic traits, such leaf mass area and growth rates, as well as the identification of several SNPs subjected to selection, exceeded neutral predictions of differentiation among populations. Despite limited differences among populations, the SNPs subjected to selection provide evidences of strong selection in marginal climates.

Relict populations are assumed to have limited performance due to genetic drift and inbreeding depression. Nonetheless, evidence for increased vulnerability to climate change in the relict population was not higher than in the treelines. Finally, it should be noted that, despite we obtained several evidences of selective pressures, likely driven by climate, land-use changes have also played a determinant role driving recent *P. uncinata* forest dynamics at a landscape scale.

## Funding

Belén Méndez-Cea is recipient of a UCM Santander predoctoral fellowship (CT42/18-CT43/18), Isabel García-García is recipient of a predoctoral fellowship (FPU18/01153). This research was funded by projects LMP242\_18 (Gobierno de Aragón, co-financed by FEDER European Funds), FORMAL (ref. RTI2018-096884-B-C31; Spanish Ministry of Science, Innovation and Universities) and the project "CLU-2019-05 – IRNASA/CSIC Unit of Excellence", funded by the Junta de Castilla y León and co-financed by the European Union (ERDF "Europe drives our growth").



## CRedit authorship contribution statement

**Belén Méndez-Cea:** Investigation, Formal analysis, Writing – original draft, Writing – review & editing. **Isabel García-García:** Investigation, Writing – review & editing. **Antonio Gazol:** Resources. **J. Julio Camarero:** Conceptualization, Funding acquisition, Resources, Writing – review & editing. **Ester González de Andrés:** Resources, Writing – review & editing. **Michele Colangelo:** Resources. **Cristina Valeriano:** Resources. **Francisco Javier Gallego:** Conceptualization, Funding acquisition, Writing – review & editing. **Juan Carlos Linares:** Conceptualization, Funding acquisition, Writing – review & editing.

## Data availability

Data will be made available on request.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2022.159778>.

## References

- Alberto, F.J., Aitken, S.N., Alia, R., Gonzalez-Martinez, S.C., Hanninen, H., Kremer, A., et al., 2013. Potential for evolutionary responses to climate change evidence from tree populations. *Glob. Change Biol.* 19, 1645–1661. <https://doi.org/10.1111/gcb.12181>.
- Alexander, J.M., Chalmardrier, L., Lenoir, J., et al., 2018. Lags in the response of mountain plant communities to climate change. *Glob. Change Biol.* 24, 563–579. <https://doi.org/10.1111/gcb.13976>.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Améztegui, A., Brotons, L., Coll, L., 2010. Land-use changes as major drivers of mountain pine (*Pinus uncinata* ram.) expansion in the Pyrenees. *Glob. Ecol. Biogeogr.* 19, 632–641. <https://doi.org/10.1111/j.1466-8238.2010.00550.x>.
- Andrews, S., 2010. FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Angert, A.L., Bontrager, M.G., Ågren, J., 2020. What do we really know about adaptation at range edges? *Ann. Rev. Ecol. Evol. Syst.* 51, 341–361. <https://doi.org/10.1146/annurev-ecolsys-012120-091002>.
- Batllo, E., Camarero, J.J., Ninot, J.M., Gutiérrez, E., 2009. Seedling recruitment, survival and facilitation in alpine *Pinus uncinata* treeline ecotones. Implications and potential responses to climate warming. *Glob. Ecol. Biogeogr.* 18, 460–472.
- Batushansky, A., Kirma, M., Grillich, N., et al., 2014. Combined transcriptomics and metabolomics of *Arabidopsis thaliana* seedlings exposed to exogenous GABA suggest its role in plants is predominantly metabolic. *Mol. Plant* 7, 1065–1068. <https://doi.org/10.1093/mp/ssu017>.
- Behr, A.A., Liu, K.Z., Liu-Fang, G., Nakka, P., Ramachandran, S., 2016. Pong: fast analysis and visualization of latent clusters in population genetic data. *Bioinformatics* 32, 2817–2823. <https://doi.org/10.1093/bioinformatics/btw327>.
- Bontrager, M., Usui, T., Lee-Yaw, J.A., Anstett, D.N., Branch, H.A., Hargreaves, A.L., et al., 2021. Adaptation across geographic ranges is consistent with strong selection in marginal climates and legacies of range expansion. *Evolution* 75, 1316–1333. <https://doi.org/10.1111/evo.14231>.
- Bridle, J.R., Vines, T.H., 2007. Limits to evolution at range margins: when and why does adaptation fail? *trends ecol. Evol.* 22, 140–147.
- Camarero, J.J., Gutiérrez, E., 2002. Plant species distribution across two contrasting treeline ecotones in the Spanish Pyrenees. *Plant Ecol.* 162, 247–257. <https://doi.org/10.1023/A:1020367918521>.
- Camarero, J.J., Gutiérrez, E., 2004. Pace and pattern of recent treeline dynamics: response of ecotones to climatic variability in the Spanish Pyrenees. *Clim. Chang.* 63, 181–200. <https://doi.org/10.1023/B:CLIM.0000018507.71343.46>.
- Camarero, J.J., Gutiérrez, E., 2007. Response of *Pinus uncinata* recruitment to climate warming and changes in grazing pressure in an isolated population of the Iberian system (NE Spain). *Arct. Ant. Alp. Res.* 39, 210–217.
- Camarero, J.J., Guerrero-Campo, J., Gutiérrez, E., 1998. Tree-ring growth and structure of *Pinus uncinata* and *Pinus sylvestris* in the central Spanish Pyrenees. *Arct. Antarct. Alp. Res.* 30, 1–10. <https://doi.org/10.2307/1551739>.
- Camarero, J.J., Linares, J.C., García-Cervigón, A.I., Batllori, E., Martínez, I., Gutiérrez, E., 2017. Back to the future: the responses of alpine treelines to climate warming are constrained by the current ecotone structure. *Ecosystems* 20, 683–700. <https://doi.org/10.1007/s10021-016-0046-3>.
- Camarero, J.J., Gazol, A., Sangüesa-Barreda, G., et al., 2021. Tree growth is more limited by drought in rear-edge forests most of the times. *For. Ecosyst.* 8, 25. <https://doi.org/10.1186/s40663-021-00303-1>.
- Chang, C.C., Chow, C.C., Tellier, L.C.A.M., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4.
- Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., 1000 Genomes Project Analysis Group, 2011. The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>.
- Davis, E.L., Brown, R., Daniels, L., Kavanagh, T., Gedalof, Z.E., 2020. Regional variability in the response of alpine treelines to climate change. *Clim. Chang.* 162, 1365–1384. <https://doi.org/10.1007/s10584-020-02743-0>.
- Des Marais, D.L., Hernandez, K.M., Juenger, T.E., 2013. Genotype-by-environment interaction and plasticity: exploring genomic responses of plants to the abiotic environment. *Ann. Rev. Ecol. Evol. Syst.* 44, 5–29.
- Ding, L., Chaumont, F., 2020. Are aquaporins expressed in stomatal complexes promising targets to enhance stomatal dynamics? *Front. Plant Sci.* 11, 458.
- Du, H., Liu, J., Li, M.H., Buntgen, U., Yang, Y., Wang, L., et al., 2018. Warming-induced upward migration of the alpine treeline in the Changbai Mountains, Northeast China. *Glob. Change Biol.* 24, 1256–1266. <https://doi.org/10.1111/gcb.13963>.
- Działuk, A., Muchewicz, E., Boratyński, A., Montserrat, J.M., Boratynska, K., Burczyk, J., 2009. Genetic variation of *Pinus uncinata* (Pinaceae) in the Pyrenees determined with cpSSR markers. *Plant Syst. Evol.* 277, 197–205. <https://doi.org/10.1007/s00606-008-0123-y>.
- Eaton, D.A.R., Overcast, I., 2020. Ipyrad: interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36, 2592–2594. <https://doi.org/10.1093/bioinformatics/btz966>.
- Endelman, J.B., 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome* 4, 250–255. <https://doi.org/10.3835/plantgenome2011.08.0024>.
- Endelman, J.B., Jannink, J.L., 2012. Shrinkage estimation of the realized relationship matrix. *G3: genes, genom. Genet.* 2, 1405–1413. <https://doi.org/10.1534/g3.112.004259>.
- Fait, A., Nesi, A.N., Angelovici, R., Lehmann, M., Pham, P.A., Song, L., et al., 2011. Targeted enhancement of glutamate-to- $\gamma$ -aminobutyrate conversion in *Arabidopsis* seeds affects carbon-nitrogen balance and storage reserves in a development-dependent manner. *Plant Physiol.* 157, 1026–1042. <https://doi.org/10.1104/pp.111.179986>.
- Feuillet, T., Birre, D., Milian, J., Godard, V., Clauzel, C., Serrano-Notivol, R., 2019. Spatial dynamics of alpine tree lines under global warming: what explains the mismatch between tree densification and elevational upward shifts at the tree line ecotone? *J. Biogeogr.* 47, 1056–1068. <https://doi.org/10.1111/jbi.13779>.
- Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315. <https://doi.org/10.1002/joc.5086>.
- Foll, M., Gaggiotti, O.E., 2008. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180, 977–993. <https://doi.org/10.1534/genetics.108.092221>.
- Frichot, E., François, O., 2015. LEA: an R package for landscape and ecological association studies. *Methods Ecol. Evol.* 6, 925–929. <https://doi.org/10.1111/2041-210X.12382>.
- García-García, I., Méndez-Cea, B., Martín-Gálvez, D., Seco, J.I., Gallego, F.J., Linares, J.C., 2022. Challenges and perspectives in the epigenetics of climate change-induced forests decline. *Front. Plant Sci.* 12, 797958. <https://doi.org/10.3389/fpls.2021.797958>.
- Gazol, A., Camarero, J.J., Igual, J.M., González de Andrés, E., Colangelo, M., Valeriano, C., 2022. Intraspecific trait variation, growth, and altered soil conditions at tree species distribution limits: from the alpine treeline to the rear edge. *Agric. For. Meteorol.* 315, 18811. <https://doi.org/10.1016/j.agrformet.2022.108811>.
- Gish, W., States, D.J., 1993. Identification of protein coding regions by database similarity search. *Nature Genet.* 3, 266–272. <https://doi.org/10.1089/cmb.1994.1.39>.
- González-Díaz, P., Gazol, A., Valbuena-Carabaña, M., Sangüesa-Barreda, G., Moreno-Urbano, A., Zavala, M.A., Camarero, J.J., 2020. Remaking a stand: links between genetic diversity and tree growth in expanding mountain pine populations. *For. Ecol. Manag.* 472, 118244. <https://doi.org/10.1016/j.foreco.2020.118244>.
- Hampe, A., Jump, A.S., 2011. Climate relicts: past, present, future. *Ann. Rev. Ecol. Evol. Syst.* 42, 313–333.
- Hampe, A., Petit, R.J., 2005. Conserving biodiversity under climate change: the rear edge matters. *Ecol. Lett.* 8, 461–467.
- Hargreaves, A.L., Eckert, C.G., 2019. Local adaptation primes cold-edge populations for range expansion but not warming-induced range shifts. *Ecol. Lett.* 22, 78–88. <https://doi.org/10.1146/annurev-ecolsys-012120-091002>.
- Harsch, M.A., Hulme, P.E., McGlone, M.S., Duncan, R.P., 2009. Are treelines advancing? a global meta-analysis of treeline response to climate warming. *Ecol. Lett.* 12, 1040–1049. <https://doi.org/10.1111/j.1461-0248.2009.01355.x>.
- Heuert, M., Teufel, J., González-Martínez, S.C., Soto, A., Bruno, F., Alía, R., Vendramin, G.G., 2010. Geography determines genetic relationships between species of mountain pine (*Pinus mugo* complex) in western Europe. *J. Biogeogr.* 37, 541–556.
- Hoegh-Guldberg, O., Jacob, D., Taylor, M., et al., 2018. Impacts of 1.5°C Global Warming on Natural and Human Systems. Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.)]. In Press.
- Hoffmann, A.A., Blows, M.W., 1994. Species borders: ecological and evolutionary perspectives. *Trends Ecol. Evol.* 9, 223–227.

- Hoffmann, A., Sgrò, C., 2011. Climate change and evolutionary adaptation. *Nature* 470, 479–485. <https://doi.org/10.1038/nature09670>.
- Holt, R.D., 2003. On the evolutionary ecology of species' ranges. *Evol. Ecol. Res.* 5, 159–178.
- Jordan, R., Hoffmann, A.A., Dillon, S.K., Prober, S.M., 2017. Evidence of genomic adaptation to climate in eucalyptus microcarpa: implications for adaptive potential to projected climate change. *Mol. Ecol.* 26 (21), 6002–6020. <https://doi.org/10.1111/mec.14341>.
- Kottler, E.J., Dickman, E.E., Sexton, J.P., Emery, N.C., Franks, S.J., 2021. Draining the swamping hypothesis: little evidence that gene flow reduces fitness at range edges. *Trends Ecol. Evol.* 36, 533–544. <https://doi.org/10.1016/j.tree.2021.02.004>.
- Larsson, L.A., Larsson, P.O., 2018. CDendro and Coorecorder (v. 9.3.1). [software]. Cybis Elektronik and Data AB, Saltsjobaden, Sweden Available from.
- Mahdieh, M., Mostajeran, A., Horie, T., Katsuhara, M., 2008. Drought stress alters water relations and expression of PIP-type aquaporin genes in *Nicotiana tabacum* plants. *Plant & Cell Physiology* 49, 801–813.
- Mekonnen, D.W., Flügel, U.I., Ludewig, F., 2016. Gamma-aminobutyric acid depletion affects stomata closure and drought tolerance of *Arabidopsis thaliana*. *Plant Sci.* 245, 25–34. <https://doi.org/10.1016/j.plantsci.2016.01.005>.
- Monteleone, I., Ferrazzini, D., Belletti, P., 2007. Effectiveness of neutral RAPD markers to detect genetic divergence between the subspecies *uncinata* and *mugo* of *Pinus mugo turra*. *Silva Fenn.* 40, 391–406. <https://doi.org/10.14214/sf.476>.
- Nardini, A., 2022. Hard and tough: the coordination between leaf mechanical resistance and drought tolerance. *Flora* 288, 152023. <https://doi.org/10.1016/j.flora.2022.152023>.
- Niinimets, Ü., 2001. Global-scale climatic controls of leaf dry mass per area, density, and thickness in trees and shrubs. *Ecology* 82, 453–469.
- Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* 6, 288–295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>.
- Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28, 2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>.
- Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: an inexpensive method for De novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* 7, e37135. <https://doi.org/10.1371/journal.pone.0037135>.
- Pina-Martins, F., Baptista, J., Pappas Jr., G., Paulo, O.S., 2018. New insights into adaptation and population structure of cork oak using genotyping by sequencing. *Glob. Change Biol.* 25, 337–350.
- QGIS Development Team, 2022. QGIS geographic information system. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>.
- Reich, P.B., 2014. The world-wide 'fast-slow' plant economics spectrum: a traits manifesto. *J. Ecol.* 102, 275–301.
- Reich, P.B., Buschena, C., Tjoelker, M.G., Wrage, K., Knops, J., Tilman, D., Machado, J.L., 2003. Variation in growth rate and ecophysiology among 34 grassland and savanna species under contrasting N supply: a test of functional group differences. *New Phytol.* 157, 617–631. <https://doi.org/10.1046/j.1469-8137.2003.00703.x>.
- Sánchez-Salguero, R., Camarero, J.J., Gutiérrez, E., González Rouco, F., Gazol, A., Sangüesa-Barreda, G., et al., 2017. Assessing forest vulnerability to climate warming using a process-based model of tree growth: bad prospects for rear-edges. *Glob. Change Biol.* 23, 2705–2719. <https://doi.org/10.1111/gcb.13541>.
- Sanmiguel-Valladolid, A., Camarero, J.J., Gazol, A., Moran-Tejeda, E., Sangüesa-Barreda, G., Alonso-González, E., et al., 2019. Detecting snow-related signals in radial growth of *Pinus uncinata* mountain forests. *Dendrochronologia* 57, 125622.
- Sexton, J.P., McIntyre, P.J., Angert, A.L., Rice, K.J., 2009. Evolution and ecology of species range limits. *Ann. Rev. Ecol. Syst.* 40, 415–436. <https://doi.org/10.1146/annurev.ecolsys.110308.120317>.
- Shaw, R.G., Etterson, J.R., 2012. Rapid climate change and the rate of adaptation: insight from experimental quantitative genetics. *New Phytol.* 195, 752–792. <https://doi.org/10.1111/j.1469-8137.2012.04230.x>.
- Sigdel, S.R., Wang, Y., Camarero, J.J., Zhu, H., Liang, E., Penuelas, J., 2018. Moisture-mediated responsiveness of treeline shifts to global warming in the Himalayas. *Glob. Change Biol.* 24, 5549–5559. <https://doi.org/10.1111/gcb.14428>.
- Sjölund, M.J., González-Díaz, P., Moreno-Villena, J.J., Jump, A.S., 2019. Gene flow at the leading range edge: the long-term consequences of isolation in european beech (*Fagus sylvatica* L. Kuhn). *J. Biogeogr.* 46, 2787–2799.
- Team, R.Core, R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Unamba, C.I.N., Nag, A., Sharma, R.K., 2015. Next generation sequencing technologies: the doorway to the unexplored genomics of non-model plants. *Front. Plant Sci.* 6, 1074. <https://doi.org/10.3389/fpls.2015.01074>.
- Wang, M., Zheng, Q., Shen, Q., Guo, S., 2013. The critical role of potassium in plant stress response. *Int. J. Mol. Sci.* 14, 7370–7390. <https://doi.org/10.3390/ijms14047370>.
- Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., et al., 2004. The worldwide leaf economics spectrum. *Nature* 428, 821–827.
- Zaborowska, J., Labiszak, B., Wachowiak, W., 2019. Population history of european mountain pines *Pinus mugo* and *Pinus uncinata* revealed by mitochondrial DNA markers. *J. Syst. Evol.* 58, 474–486. <https://doi.org/10.1111/jse.12520>.