Received: 24 August 2022 | Revised: 27 February 2023 | Accepted: 16 March 2023

DOI: 10.1111/mec.16930

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

# Demographic and spatially explicit landscape genomic analyses in a tropical oak reveal the impacts of late Quaternary climate change on Andean montane forests

Joaquín Ortego[1] | Josep Maria Espelta[2,3] | Dolors Armenteras[4] | María Claudia Díez[5] | Alberto Muñoz[6] | Raúl Bonal[7]

[1]Department of Ecology and Evolution, Estación Biológica de Doñana, EBD-CSIC, Seville, Spain

[2]CREAF, E08193, Bellaterra (Cerdanyola del Vallès), Catalonia, Spain

[3]Universitat Autònoma de Barcelona, E08193 Bellaterra (Cerdanyola del Vallès), Catalonia, Spain

[4]Laboratorio de Ecología del Paisaje y Modelación de Ecosistemas ECOLMOD, Departamento de Biología, Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, Colombia

[5]Grupo de Investigación en Ecología y Silvicultura de Especies Forestales Tropicales. Departamento de Ciencias Forestales, Facultad de Ciencias Agrarias, Universidad Nacional de Colombia sede Medellín, Medellín, Colombia

[6]Departamento de Didáctica de Ciencias Experimentales, Sociales y Matemáticas, Facultad de Educación, Universidad Complutense de Madrid, Madrid, Spain

[7]Department of Biodiversity, Ecology and Evolution, Complutense University of Madrid, Madrid, Spain

**Correspondence**
Joaquín Ortego, Estación Biológica de Doñana, EBD-CSIC, Avda. Américo Vespucio 26, Seville E-41092, Spain.
Email: joaquin.ortego@csic.es

**Funding information**
European Regional Development Fund, Grant/Award Number: RYC-2013-12501; Spanish Ministry of Economy and Competitiveness

**Handling Editor:** Victoria L. Sork

## Abstract

The tropical Andes are one of the most important biodiversity hotspots on Earth, yet our understanding of how their biotas have responded to Quaternary climatic oscillations is extraordinarily limited and the alternative models proposed to explain their demographic dynamics have been seldom formally evaluated. Here, we test the hypothesis that the interplay between the spatial configuration of geographical barriers to dispersal and elevational displacements driven by Quaternary cooling–warming cycles has shaped the demographic trajectories of montane oak forests (*Quercus humboldtii*) from the Colombian Andes. Specifically, we integrate genomic data and environmental niche modelling at fine temporal resolution to test competing spatially explicit demographic and coalescent models, including scenarios considering (i) isotropic gene flow through the landscape, (ii) the hypothetical impact of contemporary barriers to dispersal (i.e., inter-Andean valleys), and (iii) distributional shifts of montane oak forests from the Last Glacial Maximum to the present. Although our data revealed a marked genetic fragmentation of montane oak forests, statistical support for isolation-with-migration models indicates that geographically separated populations from the different Andean Cordilleras regularly exchange gene flow. Accordingly, spatiotemporally explicit demographic analyses supported a model of flickering connectivity, with scenarios considering isotropic gene flow or currently unsuitable habitats as persistent barriers to dispersal providing a comparatively worse fit to empirical genomic data. Overall, these results emphasize the role of landscape heterogeneity on shaping spatial patterns of genomic variation in montane oak forests, rejecting the hypothesis of genetic continuity and supporting a significant impact of Quaternary climatic oscillations on their demographic trajectories.

**KEYWORDS**
Andean oak, ddRADseq, demographic modelling, distributional shifts, isolation by elevation, landscape genetics, Pleistocene, Quaternary climatic oscillations, *Quercus*, tropical Andes

## 1 | INTRODUCTION

The interplay between Quaternary climate fluctuations (i.e., glacial–interglacial cycles) and the spatial configuration and attributes of geographical features (i.e., topography and barriers to dispersal) has impacted the demography and present-day distribution of genetic variation of most species, and contributed to the diversification of numerous organism groups over relatively short evolutionary timescales (Hewitt, 2004; Kadereit & Abbott, 2021; Stewart et al., 2010). Phylogeographical and palaeoecological evidence strongly supports that temperate species displaced toward lower latitudes during glacial periods, which led to vicariance processes in isolated refugia that served as source populations for northward range expansions during more favourable postglacial climatic conditions (Hewitt, 2004; Ortego & Knowles, 2022; Stewart et al., 2010). However, although the consequences of Quaternary climatic oscillations in organisms from temperate regions are relatively well understood, our knowledge about their impacts on tropical biotas is still very limited, and expectations in terms of population expansions/contractions and distributional shifts are less intuitive than at higher latitudes (Bush & de Oliveira, 2006; Carnaval et al., 2009; Prates et al., 2016; Ramírez-Barahona & Eguiarte, 2013).

The tropical Andes are a cradle of speciation (Hazzi et al., 2018; Hughes & Eastwood, 2006; Nevado et al., 2018; Pennington et al., 2010), harbour high levels of local endemism (Kattan et al., 2004; Kier et al., 2009) and have been identified as one of the most important biodiversity hotspots on Earth (Bibby et al., 1992; Myers et al., 2000). Beyond the role of Andean uplift and orogenic processes in creating novel highland environments and triggering in situ diversification in numerous groups of organisms (Kattan et al., 2004; e.g., Muñoz-Ortiz et al., 2015; Sanín et al., 2022), there is mounting evidence about the role of Pleistocene glacial cycles as an important engine of speciation in the region (Flantua et al., 2019; Pérez-Escobar et al., 2022; e.g., Nevado et al., 2018; Ribas et al., 2007). Palaeoecological evidence supports that Quaternary climate fluctuations have recurrently displaced populations along elevational gradients, although the specific outcomes of such phenomena are expected to vary among species and communities linked to the different vegetation belts. High-elevation Andean páramos above the tree line (>3000 m at present) probably underwent important elevational shifts (~1600 m) in response to changing climate conditions through the Quaternary, resulting in highest connectivity during extreme glacial periods and considerable fragmentation during warm interglacials (i.e., "flickering connectivity model"; sensu Flantua & Hooghiemstra, 2018; Flantua et al., 2019) that created cyclical opportunities for vicariant diversification and secondary contact (i.e., "species pump"; e.g., Nevado et al., 2018). However, alternative distributional-demographic models have been proposed to explain the dynamics of tropical montane biotas (in Colombian Andes, ~1200–3000 m at present) during the coldest stages of the Quaternary (Ramírez-Barahona & Eguiarte, 2013). These include downslope migration with demographic expansions and connectivity ("Moist forest model"; Ramírez-Barahona & Eguiarte, 2013), downslope migration and compression with population contractions, local extinctions and null or limited historical connectivity ("Dry refugia model"; Ramírez-Barahona & Eguiarte, 2013), and in situ glacial persistence ("Long-term *in situ* persistence model"; Ornelas et al., 2019). Fossil pollen records suggest that montane forests displaced toward lower elevations and probably compressed by ~800–1000 m during glacial times with respect to its current distribution (Hooghiemstra & Flantua, 2019). Although these elevational shifts during the coldest stages of the Quaternary were inferred to reduce available surface area by ~40% (Hooghiemstra & Flantua, 2019), they might have also facilitated dispersal across geographical barriers (e.g., valley bottoms) and increased genetic connectivity among populations from different cordilleras (e.g., Muñoz-Ortiz et al., 2015). The integration of genomic data and palaeodistribution modelling within spatiotemporally explicit process-based inference frameworks might help to distinguish among competing hypotheses about the demographic responses of Andean montane forests to Quaternary climatic oscillations (Hooghiemstra et al., 2022; Ramírez-Barahona & Eguiarte, 2013, 2014; e.g., He et al., 2013; Bemmels et al., 2016).

The Andean oak *Quercus humboldtii* Bonpl. (1805) (section *Lobatae*)—the only South American oak and the southernmost representative of the genus—is a characteristic species of the montane forest from the Colombian Cordilleras (Hooghiemstra et al., 2022; Torres et al., 2013). It occurs in the Lower and Upper Montane Forest belts (780–3600 m; Hooghiemstra et al., 2022; Zorrilla-Azcué et al., 2021), occupying a wide spectrum of climatological conditions and forming multiple phytosociological associations (Hooghiemstra et al., 2022; Hooghiemstra & Flantua, 2019; Pulido et al., 2006; Rangel & Avella, 2011). The broad ecological range of the species has been hypothesized to facilitate its southward expansion through the Panamanian Isthmus and the colonization of the three northern Andean Cordilleras (Hooghiemstra et al., 2022; Rangel & Avella, 2011). Palynological records date the arrival of the species to central Colombia in the mid-Pleistocene (~430 thousand years ago [ka]; Torres et al., 2013; see also Hooghiemstra, 2006) and support its progressive spread and continuous presence in the region since then, probably replacing other trees and often becoming a dominant species in the Andean montane forest (Hooghiemstra & Flantua, 2019). The fact that *Q. humboldtii* is a characteristic element of the Andean montane forest, together with its Pleistocene arrival in the region, makes this species an excellent model system to infer the consequences of Quaternary climatic oscillations on the dynamics of tropical montane forests and tease apart their impacts from the genetic imprints left by other geological events (e.g., Andean uplift; Boschman, 2021; Gregory-Wodzicki, 2000) that have been also documented to play a paramount role on the diversification of many neotropical organisms (Chaves et al., 2011; Muñoz-Ortiz et al., 2015; Nevado et al., 2018; Sanín et al., 2022).

Populations of *Q. humboldtii* probably migrated downslope during glacial times. Such elevational shifts might have promoted gene flow among populations from the different cordilleras through transient corridors across currently unsuitable habitats

characterizing the Cauca and Magdalena valley bottoms. However, it has been also hypothesized that the lower part of the elevational distribution of *Q. humboldtii* (~800 m) might have not changed through glacial–interglacial cycles (Hooghiemstra et al., 2022; Hooghiemstra & Flantua, 2019; Torres et al., 2013). Under this scenario, gene exchange across lowlands (<800 m) separating Western, Central and Eastern Cordilleras (i.e., lowest elevations of Cauca and Magdalena valleys; Figure 2) might have been limited during both glacial and interglacial periods in line with the "Dry refugia model" (Ramírez-Barahona & Eguiarte, 2013). So far, microsatellite-based genetic studies on *Q. humboldtii* at both local and phylogeographical scales suggest that the species is resilient to habitat fragmentation (Fernández & Sork, 2007), presents limited genetic structure (Zorrilla-Azcué et al., 2021) and experienced a bottleneck followed by a demographic expansion during the Last Glacial Maximum (LGM; Zorrilla-Azcué et al., 2021). Thus, in line with palaeoecological evidence pointing to the continuous presence of *Q. humboldtii* in the pollen record since its arrival 430 ka in Colombia (Torres et al., 2013), available genetic evidence would support a limited impact of unsuitable lowlands on gene flow and considerable connectivity among forest patches through the Quaternary (Hooghiemstra et al., 2022; Zorrilla-Azcué et al., 2021).

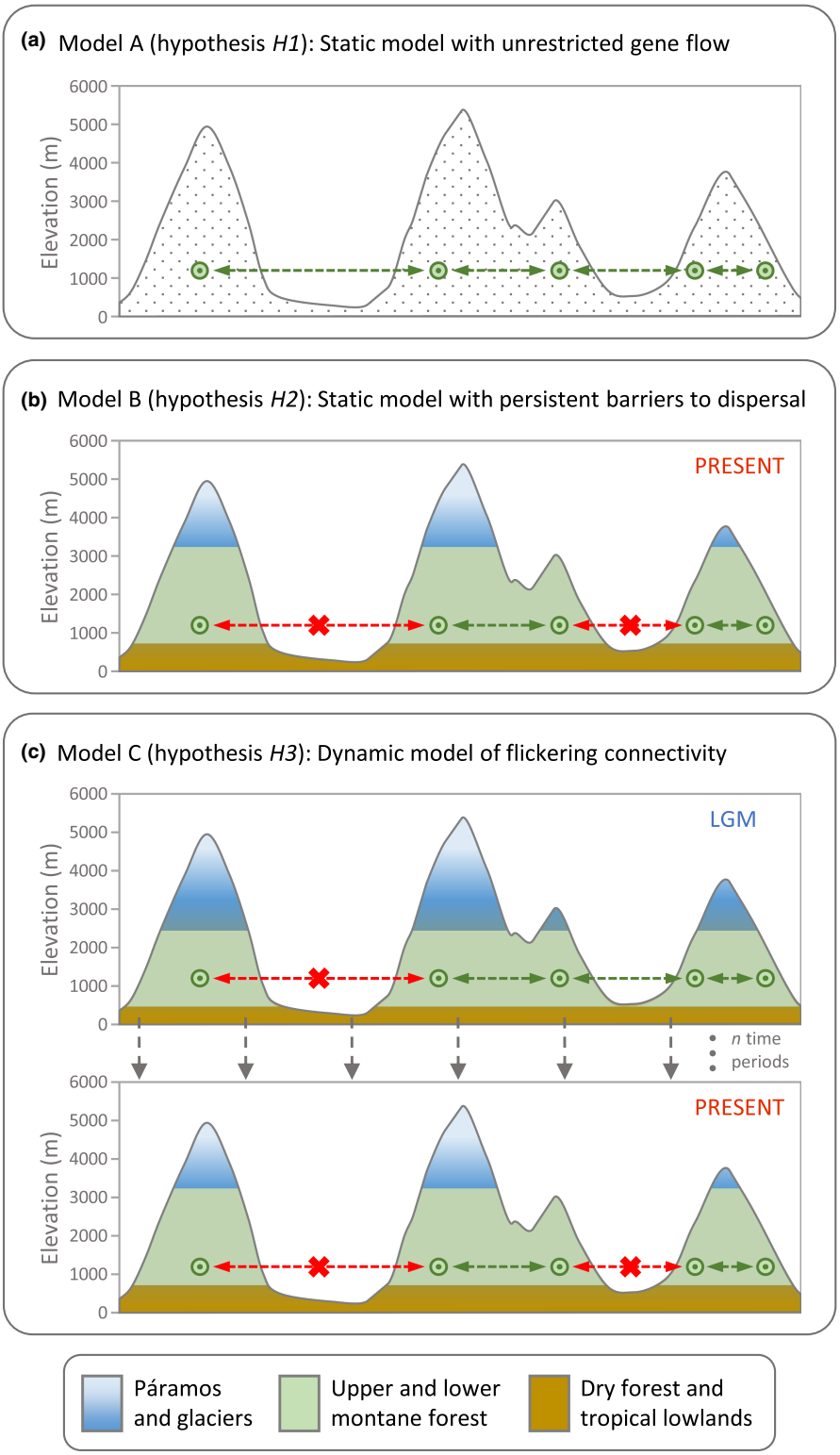
Here, we integrate genomic data (double digest restriction-site associated DNA sequencing [ddRADseq]) and environmental niche modelling (ENM) to test the hypothesis that the interplay between distributional shifts driven by Quaternary climatic oscillations and the spatial configuration of geographical barriers to dispersal (i.e., inter-Andean valleys separating montane forests) have shaped the demographic history and contributed to the genetic fragmentation of populations of *Q. humboldtii* from the different Andean Cordilleras (Zorrilla-Azcué et al., 2021). We first applied palaeodistribution modelling at fine temporal resolution (100-year time intervals) to reconstruct the range dynamics of *Q. humboldtii* since the LGM to the present, determine the spatial configuration of suitable habitats through time, and infer the hypothetical distributional shifts experienced by the species in response to cooling–warming cycles. Then, we genotyped populations across a representative landscape at the core of the species distribution in the Colombian Andes, quantified spatial patterns of genetic structure of populations separated by the Cauca and Magdalena valleys and encompassing the Western, Central and Eastern Cordilleras, and evaluated whether they have experienced parallel or idiosyncratic demographic trajectories compatible with local/regional distributional dynamics inferred from ENM. Finally, we applied a spatiotemporally demographic and coalescent modelling framework to test whether: (i) unrestricted gene flow across unsuitable habitats through the climatic fluctuations of the Quaternary has resulted in population connectivity, genetic homogenization and a lack of genetic structure in the species (i.e., a null model of isotropic gene flow; Hooghiemstra et al., 2022; Zorrilla-Azcué et al., 2021) (hypothesis *H1*; Figure 1a); (ii) currently unsuitable habitats act as persistent barriers to dispersal, disrupting gene flow and contributing to a marked genetic fragmentation among populations from the different Andean Cordilleras (hypothesis *H2*; Figure 1b); and (iii) distributional shifts toward lower elevations during the coldest stages of the Quaternary promoted secondary contact and genetic connectivity among geographically separated populations through intermittently permeable barriers to dispersal (i.e., ephemeral barriers sensu Araya-Donoso et al., 2022) (hypothesis *H3*; Figure 1c).

## 2 | MATERIALS AND METHODS

### 2.1 | Population sampling and genomic library preparation and processing

In 2015, we sampled seven populations of *Quercus humboldtii* Bonpl. (1805) ($n = 63$ individuals; Table 1) from the core of the species' distribution in the Colombian Andes (Table 1; Figure 2a). Our sampling included populations located at both the Lower and Upper Montane Forest belts (LMF and UMF, respectively) and from the Western, Central and Eastern Cordilleras and, thus, separated by the low valleys of the Cauca and Magdalena Rivers (i.e., hypothetical barriers to dispersal; Figure 2). All our sampling sites were located at elevations >1750 m (Table 1), as most populations of *Q. humboldtii* at the lower elevational range originally occupied by the species in the LMF have disappeared due to heavy deforestation and replacement by tropical crops (Hooghiemstra et al., 2022; see

**FIGURE 1** Schematic summary of the three spatially explicit demographic and coalescent models tested for *Quercus humboldtii* within the studied landscape in central Colombia. (a) Model A (hypothesis *H1*) is a static model in which carrying capacities of demes (*k*) are homogeneous across space and through time (i.e., a flat landscape, equivalent to an isolation-by-distance model). (b) Model B (hypothesis *H2*) is a static model analogous to the previous one but in which currently unsuitable habitats for the focal species were considered as persistent barriers to dispersal (i.e., $k = 0$) through time. Under this model, gene flow (green arrows) only takes place between populations (green bull eyes) connected by suitable habitat of montane oak forest (green surfaces) and it is disrupted (red arrows and crosses) among populations at both sides of the barrier (i.e., lowlands, brown surfaces). Cells with a probability of presence of the species below the maximum training sensitivity plus specificity (MTSS) logistic threshold for occurrence from the species-specific environmental niche model (ENM) built with MAXENT were considered as unsuitable habitat patches. (c) Model C (hypothesis *H3*) is a dynamic model incorporating distributional shifts resulting from the interaction between the species bioclimatic envelope and Quaternary climatic oscillations. Carrying capacities change through time according to habitat suitability maps obtained from projections of the ENM to bioclimatic conditions during the last 22,000 years (i.e., from the LGM to present, at 100-year time intervals). Under this model, certain populations recurrently merged or fragmented depending on mountain profile and the elevational displacements experienced by the species at each time period (i.e., flickering connectivity sensu Flantua et al., 2019). Note that (b) and (c, bottom) are identical because they represent current conditions, but the two models reflect different processes.




also table S1 in Zorrilla-Azcué et al., 2021). We used a mixer mill to grind ~50 mg of leaf tissue in tubes with a tungsten bead and performed DNA extraction and purification with NucleoSpin Plant II kits (Macherey-Nagel). We processed genomic DNA into one genomic library using the double-digestion restriction-fragment-based procedure (ddRADseq) described in Peterson et al. (2012) (Methods S1) and used the different programs distributed as part of the STACKS version 1.35 pipeline (Catchen et al., 2013) to filter and assemble our sequences into de novo RAD loci and call genotypes (Methods S2).

## 2.2 | Genetic structure

We analysed the genetic structure of the studied populations using the Bayesian Markov chain Monte Carlo (MCMC) clustering method

implemented in the program STRUCTURE version 2.3.3 (Pritchard et al., 2000). We ran STRUCTURE assuming correlated allele frequencies and admixture without using prior population information. We conducted 15 independent runs with 200,000 MCMC cycles, following a burn-in step of 100,000 iterations, for each of the different possible *K* genetic clusters (from $K=1$ to $K=8$). We retained the 10 runs having the highest likelihood for each value of *K*. As recommended by Gilbert et al. (2012) and Janes et al. (2017), we used two statistics to interpret the number of genetic clusters (*K*) that best describes our data: log probabilities of Pr(X|*K*) (Pritchard et al., 2000) and Δ*K* (Evanno et al., 2005). These statistics were calculated as implemented in STRUCTURE HARVESTER (Earl & vonHoldt, 2012). We used CLUMPP version 1.1.2 and the Greedy algorithm to align multiple runs of STRUCTURE for the same *K*-value (Jakobsson & Rosenberg, 2007) and DISTRUCT version 1.1 (Rosenberg, 2004) to visualize individual's probabilities of genetic cluster membership as bar plots. Complementary to STRUCTURE analyses, we performed a principal component analysis (PCA) as implemented in the R version 4.0.3 (R Core Team, 2022) package "adegenet" (Jombart, 2008). We also estimated genetic differentiation between populations calculating pairwise $F_{ST}$ values and testing their significance with Fisher's exact tests after 10,000 permutations, as implemented in ARLEQUIN version 3.5 (Excoffier & Lischer, 2010). Finally, we used Mantel and partial Mantel tests to analyse whether genetic differentiation ($F_{ST}$) between populations is explained by geographical distance (isolation-by-distance, IBD) and/or elevation dissimilarity (isolation-by-environment, IBE), which can provide key information about the specific processes shaping the spatial distribution of genetic variation in the species (Sexton et al., 2014; Wang & Bradburd, 2014).

## 2.3 | Environmental niche modelling

We built an environmental niche model (ENM) to reconstruct the geographical distribution of climatically suitable habitats for *Q. humboldtii* from the LGM (~22 ka) to the present. We used this information to test alternative demographic models considering distributional shifts experienced by the species in response to Quaternary climatic oscillations and/or population fragmentation resulted from the spatial configuration of hypothetical barriers to dispersal (see Section 2.6; Figure 1). To build the ENM, we used the maximum entropy algorithm implemented in MAXENT version 3.4.1 (Phillips et al., 2006; Phillips & Dudik, 2008), 687 occurrence records for the species (for details, see Methods S3) and the 19 bioclimatic layers (30-arcsec resolution; for variable description, see Table S1) from the CHELSA database (http://chelsa-climate.org/bioclim/; Karger et al., 2017). To estimate environmental suitability from the LGM to the present, we projected the ENM to bioclimatic conditions during the last 22,000 years at 100-year time intervals (i.e., from 1990 CE to the LGM, for a total of 220 snapshots) using bioclimatic layers available at a high resolution (30-arcsec) from the CHELSA-TraCE21k version 1.0 database (https://chelsa-climate.org/; Karger et al., 2023). Further details on ENM are presented in Methods S3.

## 2.4 | Testing alternative models of population divergence

We used the simulation-based approach implemented in FASTSIMCOAL2 (Excoffier et al., 2013) to test four hypothetical models of divergence among populations of *Q. humboldtii* from different cordilleras: strict isolation (SI), isolation-with-migration (IM), ancestral migration (AM) and secondary contact (SC) (Figure 3; e.g., Sanín et al., 2022). Statistical support for the SI model would indicate that persistent barriers to dispersal have disrupted post-divergence gene flow among populations from the different cordilleras, which aligns with expectations of spatially explicit hypothesis *H2* (Figure 1b). The remaining models (IM, AM and SC) would lend support to different scenarios of post-divergence gene flow connecting populations from the different cordilleras, in line with expectations from spatially explicit hypotheses *H1* and *H3* (Figure 1a,c; see also Section 2.6). For FASTSIMCOAL2 analyses, we considered one representative population from each Andean cordillera (BELM, PUEB and TIPA-ARCA-FUQU; Figure 2) and tested the alternative models for all possible pairwise comparisons. As populations TIPA, ARCA and FUQU from the Eastern Cordillera clustered together according to STRUCTURE and PCA analyses (see Results), they were pooled in the same population for FASTSIMCOAL2 analyses. We estimated the composite likelihood of the observed data given a specified model using the folded joint site frequency spectrum (SFS), which was calculated only considering one single nulcleotide polymorphism (SNP) per RAD locus to avoid the effects of linkage disequilibrium. Because we did not include invariable sites in the SFS, we fixed the effective population size for one of the populations (BELM or PUEB) to enable the estimation of other parameters in FASTSIMCOAL2 (Excoffier et al., 2013). The effective population size fixed in the model was calculated from the level of nucleotide diversity ($\pi$) and a mutation rate ($\mu$) of $5.92\times10^{-8}$ per site per generation (see Ortego et al., 2018), which was inferred from *Populus* (Tuskan et al., 2006) following Gossmann et al. (2012) and assuming the 50-year generation time considered in previous studies on long-lived oak trees (see Bemmels et al., 2016; Ortego et al., 2018; Ortego & Knowles, 2020; Sork et al., 2022). Note that this mutation rate is similar to that obtained for *Quercus robur* ($4.2–5.2\times10^{-8}$; Schmid-Siegert et al., 2017) and of the same order of magnitude as estimates calculated for *Quercus lobata* ($1.01\times10^{-8}$; Sork et al., 2022) based on genome sequencing data. To remove all missing data for calculation of the joint SFS, minimize errors in allele frequency estimates and maximize the number of variable SNPs retained, each population was projected down to eight individuals using a custom Python script written by Qixin He and available on Dryad (https://doi.org/10.5061/dryad.23hs1). Each model was run 100 replicated times considering 100,000–250,000 simulations for the calculation of the composite likelihood, 10–40 expectation-conditional maximization (ECM) cycles, and a stopping criterion of 0.001 (Excoffier et al., 2013). We used an information-theoretic model selection approach based on Akaike's information criterion

**TABLE 1** Geographical location, Andean Cordillera, vegetation belt (LMF, Lower Montane Forest; UMF, Upper Montane Forest; following Hooghiemstra & Flantua, 2019), number of genotyped individuals ($n$), and genetic diversity statistics ($H_O$, $H_E$ and $\pi$) for the studied populations of *Quercus humboldtii* in central Colombia.

| | | | | | | | All positions | | | Variant positions | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Code | Locality | Latitude | Longitude | Elevation (m) | Cordillera | Belt | $n$ | $H_O$ | $H_E$ | $\pi$ | $H_O$ | $H_E$ | $\pi$ |
| JARD | Jardín | 5.578180 | −75.788130 | 2200 | Western | LMF | 10 | 0.0007 | 0.0006 | 0.0007 | 0.0941 | 0.0836 | 0.0881 |
| PUEB | Pueblorrico | 5.791780 | −75.819480 | 2240 | Western | LMF | 10 | 0.0007 | 0.0007 | 0.0008 | 0.0944 | 0.0919 | 0.0968 |
| BELM | Belmira | 6.606360 | −75.663250 | 2580 | Central | UMF | 11 | 0.0007 | 0.0007 | 0.0007 | 0.0917 | 0.0904 | 0.0948 |
| QUIN | Quinini | 4.320260 | −74.487530 | 1750 | Eastern | LMF | 10 | 0.0007 | 0.0006 | 0.0007 | 0.0884 | 0.0809 | 0.0854 |
| FUQU | Fúquene | 5.453510 | −73.659660 | 2840 | Eastern | UMF | 7 | 0.0006 | 0.0006 | 0.0006 | 0.0833 | 0.0764 | 0.0825 |
| ARCA | Arcabuco | 5.726050 | −73.416040 | 2940 | Eastern | UMF | 8 | 0.0007 | 0.0006 | 0.0007 | 0.0878 | 0.0815 | 0.0870 |
| TIPA | Tipacoque | 6.397900 | −72.723230 | 2850 | Eastern | UMF | 7 | 0.0006 | 0.0006 | 0.0006 | 0.0812 | 0.0749 | 0.0812 |

*Note*: Average values across loci are presented for observed ($H_O$) and expected ($H_E$) heterozygosity, and nucleotide diversity ($\pi$). Genetic statistics were calculated in STACKS for all positions (polymorphic and nonpolymorphic) and only variant (polymorphic) positions (see Methods S2).

(AIC) to compare the set of candidate models and identify the one that is best supported by the data (Burnham & Anderson, 2002). After the maximum likelihood was estimated for each model in every replicate, we calculated the AIC scores as detailed in Thome and Carstens (2016). AIC values for each model were rescaled (ΔAIC) calculating the difference between the AIC value of each model and the minimum AIC obtained among all competing models (i.e., the best model has ΔAIC = 0). Point estimates of the different demographic parameters for the best supported model were selected from the run with the highest maximum composite likelihood. Finally, we calculated confidence intervals (based on the percentile method; e.g., de Manuel et al., 2016) of parameter estimates from 100 parametric bootstrap replicates by simulating the SFS from the maximum composite likelihood estimates and re-estimating parameters each time (Excoffier et al., 2013).

## 2.5 | Historical changes in effective population size

We inferred changes in effective population sizes ($N_e$) through time for each genetic population (as identified by STRUCTURE and PCA analyses; see Section 2.2) using STAIRWAY PLOT version 2.1, a flexible multi-epoch model approach based on the SFS that does not require a predefined model for estimating past demographic histories (Liu & Fu, 2020). To compute the folded SFS, each population was projected down to eight individuals as detailed in Section 2.4. We ran STAIRWAY PLOT considering a 50-year generation time and assuming a mutation rate per site per generation of $5.92 \times 10^{-8}$ (see Section 2.4).

## 2.6 | Testing alternative spatially explicit demographic models

We used the integrative distributional, demographic and coalescent (iDDC) approach (He et al., 2013) and an Approximate Bayesian Computation (ABC) framework (Beaumont et al., 2002) to test three spatially explicit models (Models A, B and C) corresponding to the three hypotheses presented in Section 1 (*H1*, *H2* and *H3*, respectively) and illustrated in Figure 1. These three models differ in the hypothetical demographic processes that have shaped present-day patterns of genetic variation in the studied populations of *Q. humboldtii*. (i) *Model A* (hypothesis *H1*) is a static model (sensu He et al., 2013) in which carrying capacities of demes ($k$) are homogeneous across space and through time (Figure 1a). This model is equivalent to a flat landscape or an IBD model. (ii) *Model B* (hypothesis *H2*) is a static model analogous to the previous one but in which unsuitable habitats for the focal species at the present time were considered as absolute and persistent barriers to dispersal (i.e., $k = 0$) (Figure 1b). Under this model, gene flow only takes place between populations connected by suitable habitat of montane oak forest and it is disrupted among isolated demes at both sides of the barriers (i.e., lowlands). Unsuitable

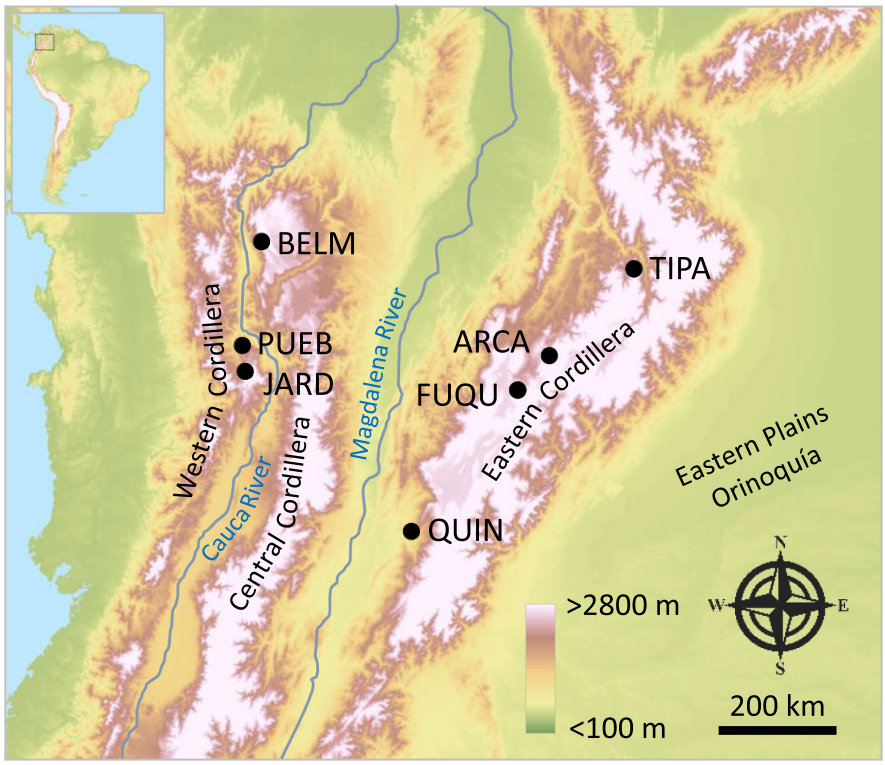



FIGURE 2 Map showing sampling localities (black dots) of *Quercus humboldtii* in central Colombia (see map inset for focal area) and main geographical features (mountain ranges and rivers). Population codes are described in Table 1.

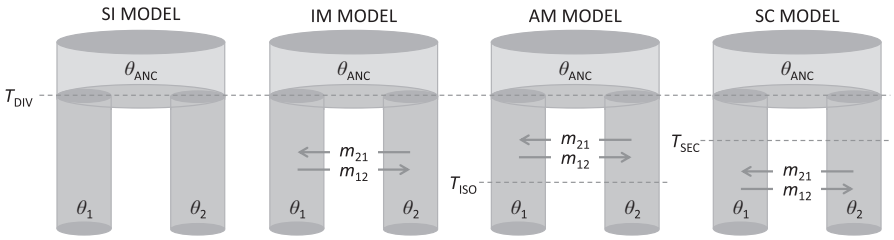



FIGURE 3 Population divergence models tested using FASTSIMCOAL2, including strict isolation (SI), isolation-with-migration (IM), ancestral migration (AM) and secondary contact (SC). Parameters include mutation-scaled ancestral ($\theta_{ANC}$) and contemporary effective population sizes ($\theta_1$ and $\theta_2$), effective migration rates per generation ($m_{12}$ and $m_{21}$), and timing of population divergence ($T_{DIV}$), interruption of gene flow ($T_{ISO}$) and secondary contact ($T_{SEC}$).

habitat patches (i.e., barriers) were identified as those cells with a probability of presence of the species below the maximum training sensitivity plus specificity (MTSS) logistic threshold for occurrence from MAXENT (Liu et al., 2005). (iii) *Model C* (hypothesis *H3*) is a dynamic model incorporating distributional shifts resulting from the interaction between the species bioclimatic envelope and Quaternary climatic oscillations (e.g., He et al., 2013; Massatti & Knowles, 2016) (Figure 1c). Carrying capacities (*k*) are scaled linearly to the logistic output from MAXENT (e.g., González-Serna et al., 2019; He et al., 2013; Massatti & Knowles, 2016) and change through time according to habitat suitability maps obtained from projections of the ENM to bioclimatic conditions during the last 22,000 years (as detailed in Section 2.3). Under this model, certain populations recurrently merged or fragmented (i.e., flickering connectivity sensu Flantua et al., 2019) depending on mountain profile and the elevational displacements experienced by the species at each time period. A detailed description of iDDC and ABC analyses, including demographic and genetic simulations under each scenario, model selection and estimation of demographic parameters, is presented in Methods S4.

## 3 | RESULTS

### 3.1 | Genomic data

After quality filtering, we retained a total of 100,303,535 reads (mean ± *SD* = 1,592,119 ± 328,146 reads per individual; Figure S1). On average, this represented 86% (range = 59%–92%) of the total number of reads recovered for each individual (Figure S1). After filtering loci (detailed in Methods S2), the final data set including all genotyped populations contained 3095 SNPs with an average proportion of missing data of 2% (range = 0%–43%) and a mean coverage depth of 36× (range = 9–56×).

### 3.2 | Genetic structure

STRUCTURE analyses identified $K=3$ as the most likely number of genetic clusters according to the $\Delta K$ criterion, but LnPr(X|$K$) steadily increased up to $K=5$ (Figure S2). Analyses for $K=2$–5 showed that the different populations structured hierarchically and presented different levels of genetic admixture in areas of geographical contact but also involving populations from different cordilleras (e.g., BELM-PUEB and ARCA-BELM; Figure 4a). For $K=3$, STRUCTURE analyses separated populations JARD (Western Cordillera), QUIN (southern Eastern Cordillera) and FUQU-ARCA-TIPA (northern Eastern Cordillera) (Figure 4a). The populations PUEB (Western Cordillera) and BELM (Central Cordillera) showed a considerable degree of admixed ancestry between the genetic clusters mostly represented in JARD and FUQU-ARCA-TIPA (Figure 4a). For $K=4$–5, the populations PUEB and BELM separated in different genetic clusters, but with considerable admixture between them and, to a lesser extent, with genetic clusters mostly represented in other populations (JARD and FUQU-ARCA-TIPA) (Figure 4a). All pairwise estimates of genetic differentiation ($F_{ST}$) were significantly different from zero ($p<.001$) and ranged between 0.042 (ARCA-TIPA) and 0.193 (TIPA-QUIN) (Table S2). In line with STRUCTURE analyses, PCA clustered together the populations from the northern part of the Eastern Cordillera (FUQU, ARCA and TIPA) and separated well the rest of the populations along the two main axes (PC1 and PC2; Figure 4b). Mantel tests showed that genetic differentiation between populations ($F_{ST}$) was significantly correlated with both geographical distance ($r=.527$, $p=.007$; Figure 5a) and elevation dissimilarity ($r=.840$, $p<.001$; Figure 5b). However, only elevation dissimilarity ($r=.788$, $p<.001$), but not geographical distance ($r=.265$, $p=.127$), remained significantly associated with genetic differentiation after controlling for the effects of each other in partial Mantel tests.

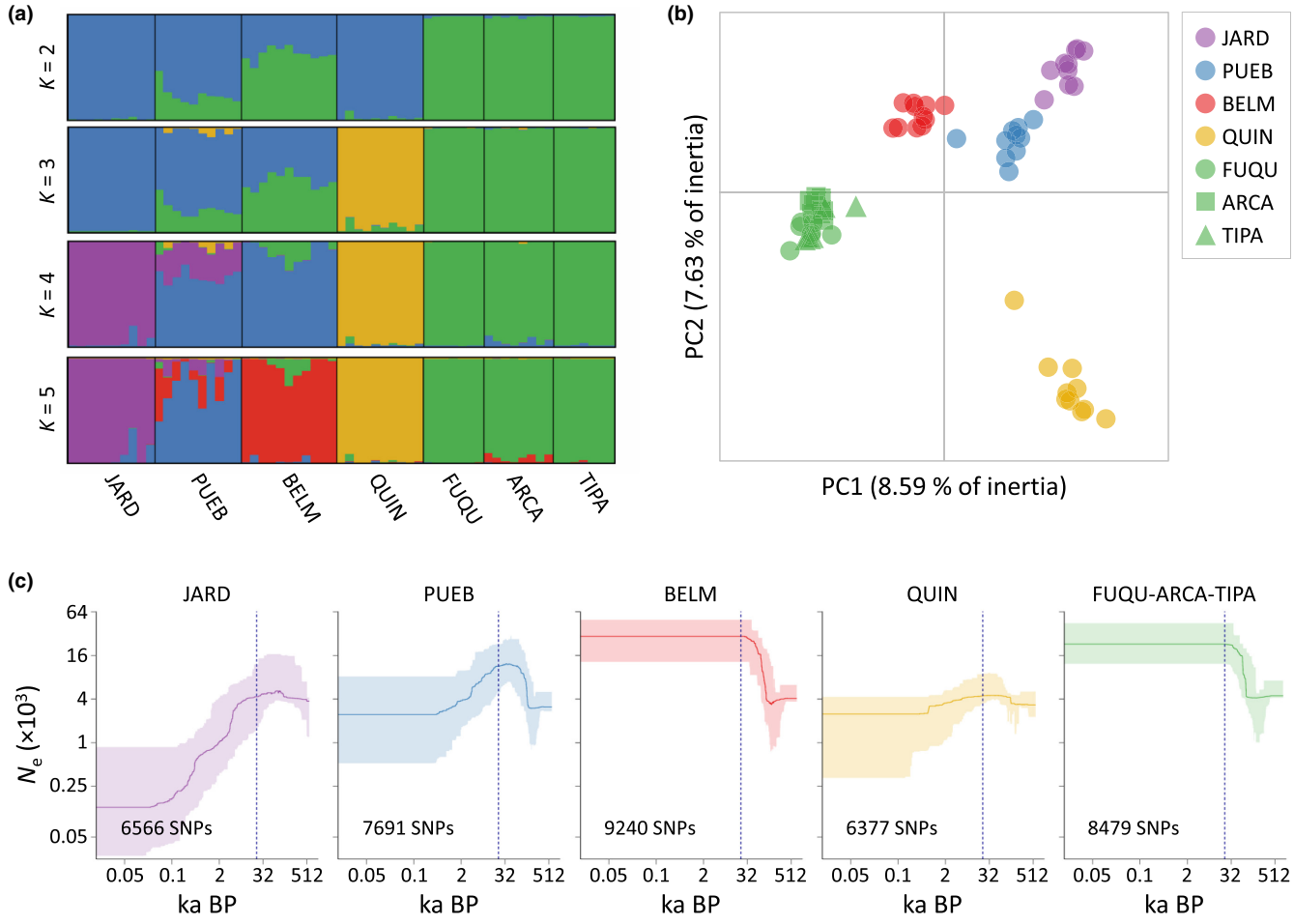



FIGURE 4 (a and b) Genetic structure and (c) demographic reconstructions for the studied populations of *Quercus humboldtii* in central Colombia. Genetic structure was inferred using (a) the Bayesian method implemented in STRUCTURE and (b) a principal component analysis (PCA) of genetic variation. Both analyses are based on a data set of 3095 SNPs. In STRUCTURE barplots each individual is represented by a vertical bar, which is partitioned into $K$ coloured segments showing the individual's probability of membership to each genetic cluster. Thin vertical black lines separate individuals from different populations. (c) Demographic reconstructions are based on the multi-epoch model implemented in STAIRWAY PLOT. Panels show the median (solid lines) and 2.5 and 97.5 percentiles (shaded areas) of effective population size ($N_e$) through time (both axes on a logarithmic scale). Vertical dashed lines indicate the LGM (~22 ka BP). The number of polymorphic SNPs used to calculate the site frequency spectrum (SFS) in each analysis is indicated. Populations FUQU, ARCA and TIPA, assigned to the same genetic cluster (see panels a and b), were grouped for STAIRWAY PLOT analyses. Population codes are described in Table 1.

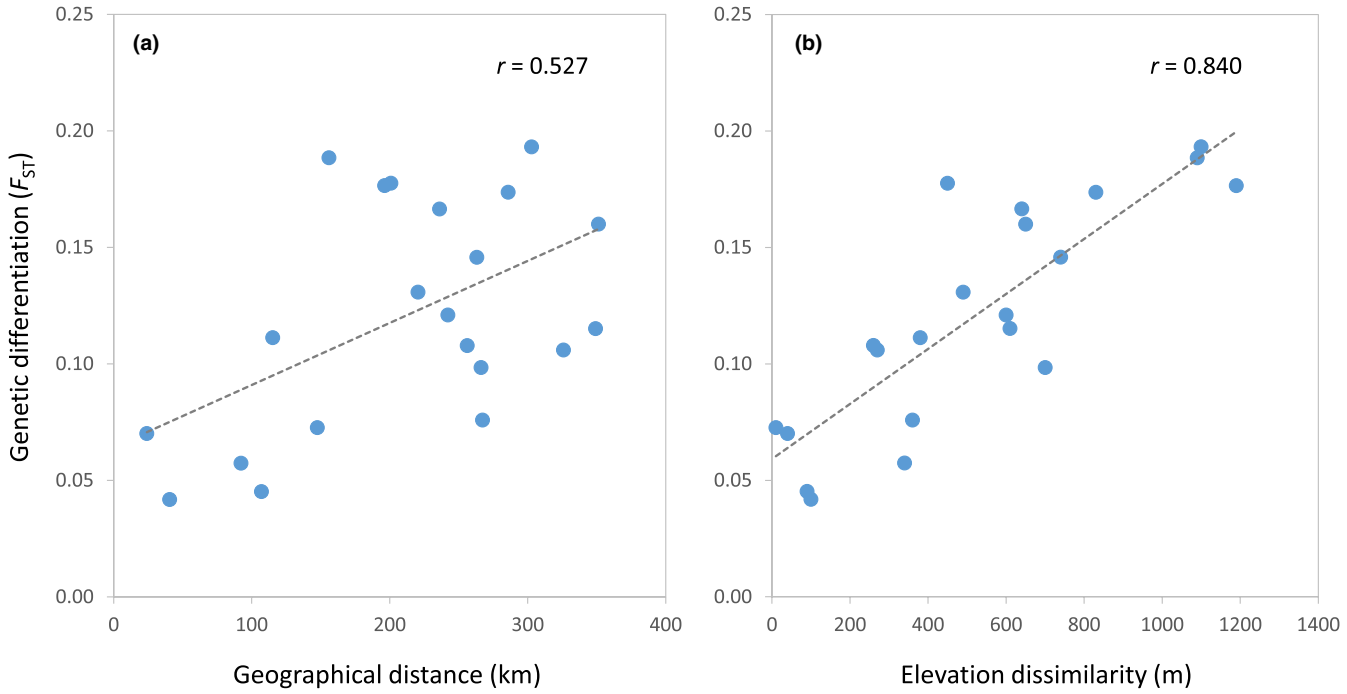



FIGURE 5 Relationship between genetic differentiation ($F_{ST}$) and (a) geographical distance and (b) elevation dissimilarity among the studied populations of *Quercus humboldtii* in central Colombia.

### 3.3 | Environmental niche modelling

After parameter tuning, the best-supported ENM (i.e., lowest AICc) was the one built considering a regularization multiplier of 2 and an LQHPT feature class. After removing highly correlated variables ($r \geq .9$), nine bioclimatic layers were retained to build the final ENM (sorted by percentage contribution, BIO1: 74.4%; BIO15: 7.6%; BIO12: 6.5%; BIO2: 4.0%; BIO4: 2.8%; BIO19: 2.1%; BIO14: 1.6%; BIO3: 0.8%; BIO18: 0.1%; for variable description, see Table S1). The high area under the receiver-operating characteristic plot on the testing data ($AUC_{TEST}$=0.881) and the low minimum training presence omission rate ($OR_{MTP}$=0.003) estimates for the best-supported model, evaluated using the "block" method for data partitioning, indicate that it has high discriminatory power and a low degree of overfitting, respectively (Peterson et al., 2011; Radosavljevic & Anderson, 2014). Climatically suitable areas predicted by the ENM are highly congruent with the present-day observed distribution of *Q. humboldtii* in the Colombian Andes (Western, Central and Eastern Cordilleras) and distributional gaps in lowlands (<700 m; e.g., Cauca and Magdalena valleys) (Figure 6d). As found in previous studies, some high-elevation areas (e.g., Cordillera de Mérida and Ecuadorian Andes) outside the current distribution range of *Q. humboldtii* were predicted as climatically suitable for the species (Figure 6d; e.g., Zorrilla-Azcué et al., 2021). Our reconstructions since the LGM showed that distributional shifts were mostly limited to elevational displacements, alternating periods of population fragmentation and connectivity at local/regional scales, but with the presence of at least some refugial populations persisting in the three cordilleras or adjacent areas (Figure 6). The total surface area of climatically suitable habitats for *Q. humboldtii* was on average smaller than today, peaking across multiple time periods since the LGM (e.g., ~18, 14 and 6–0 ka) and reaching a minimum around the onset of the Holocene (~12 ka) (Figure 6c). The mean elevation of environmentally suitable habitats displaced downslope ~500 m during the LGM with respect to its present distribution (i.e., from ~2400 to ~1900 m, respectively; Figure 6b). The minimum elevational range during the LGM was displaced to 600 m, with sporadic pulses of downward migration to even 250 m, implying a downslope shift of the lower limit of 200–500 m with respect to the minimum elevational range occupied by the species today (~800 m) (Figure 6a). Population-specific analyses within an area of 1 km radius centred on each sampling locality revealed a similar general profile for estimates of mean elevation (Figure S3a–c). However, estimates of total surface area through time were more population-specific, although generally similar among sites within the same cordillera (Figure S3d–f).

### 3.4 | Testing alternative models of population divergence

FASTSIMCOAL2 analyses showed that an IM scenario was the best-supported model of divergence among populations of *Q. humboldtii* from the three Andean Cordilleras (Table 2; Figure 3), rejecting the hypothesis that currently unsuitable habitats act as persistent barriers to dispersal (i.e., hypothesis *H2*; Figure 1b). Assuming the 50-year generation time considered in previous studies on long-lived oak trees (e.g., Bemmels et al., 2016; Ortego et al., 2018; Ortego & Knowles, 2020; Sork et al., 2022), the split between populations ($T_{DIV}$) from the three cordilleras was estimated to have taken place during the last glacial period, ~12–71 ka before the present (BP) (Table 3). Estimates of effective migration rates per generation ($m$) were low ($<\times 10^{-3}$) and asymmetric

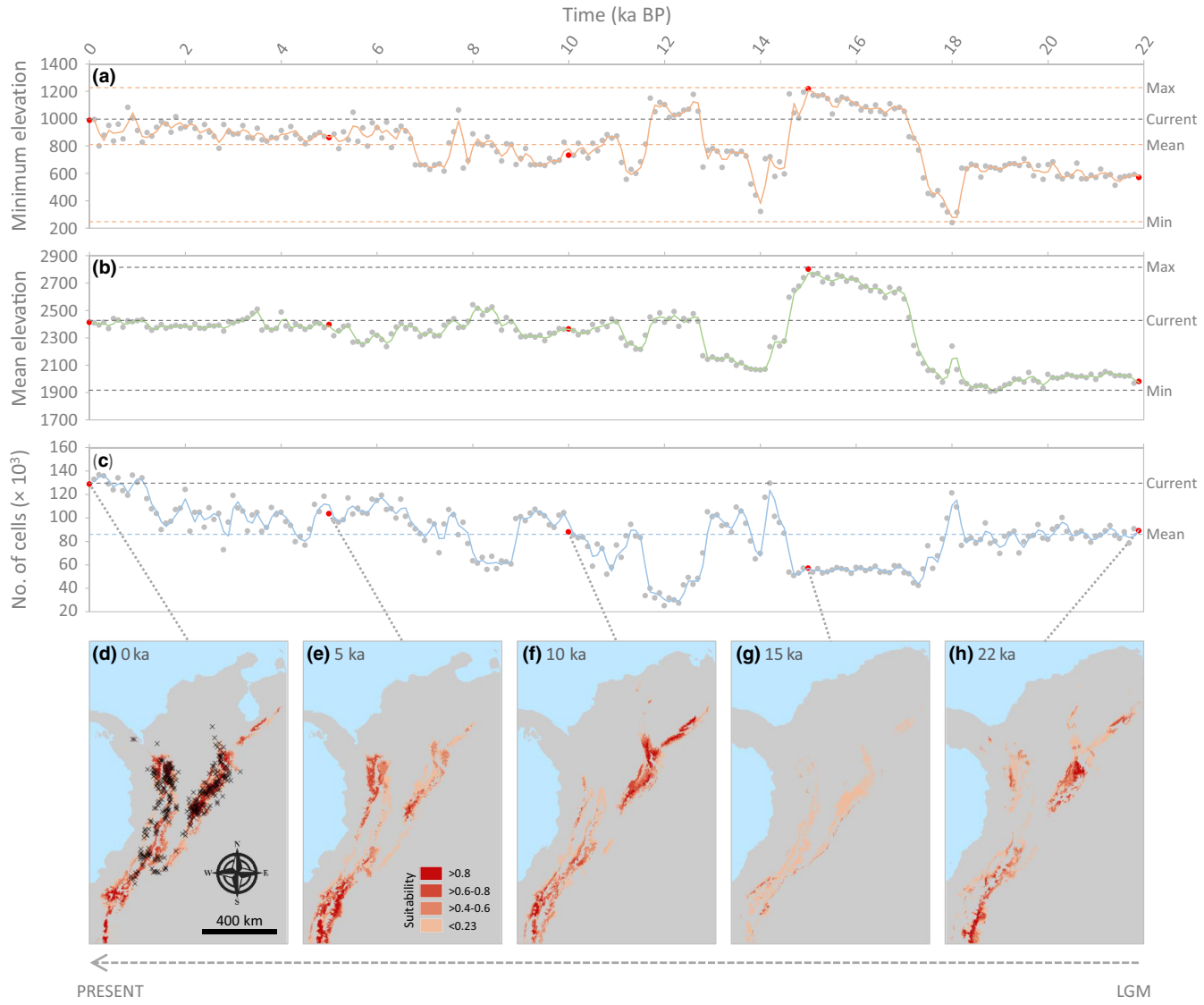



**FIGURE 6** (a) Minimum and (b) mean elevation and (c) total surface area of climatically suitable habitats for *Quercus humboldtii* as inferred from projections of the species-specific environmental niche model (ENM) to bioclimatic conditions during the last 22,000 years (i.e., from 1990 CE to the LGM) at 100-year time intervals. Cells with a probability of presence of the species above the maximum training sensitivity plus specificity (MTSS) logistic threshold for occurrence from MAXENT were considered suitable habitats and used for estimating surface area (number of suitable cells) and minimum and mean elevation occupied by *Q. humboldtii* at each time period. (d–h) Maps showing the distribution of climatically suitable habitats for the species at five temporal snapshots (red dots in a–c), including (d) the present (0 ka; crosses show occurrence points used for ENM), (f) Holocene Climate Optimum (~10 ka) and (h) LGM (~22 ka). Grey areas were considered as impassable barriers to dispersal in spatially explicit demographic and coalescent analyses (Model B; see Section 2.6).

(i.e., 95% confidence intervals did not overlap). The effective population size of ancestral populations ($\theta_{ANC}$) was lower than estimates for contemporary populations ($\theta_1$ and $\theta_2$), suggesting post-divergence demographic expansions (Table 3).

## 3.5 | Historical changes in effective population size

STAIRWAY PLOT analyses revealed that most populations have experienced idiosyncratic demographic dynamics, which include (i) sudden expansions during the last glacial period followed by either demographic stability (BELM and FUQU-ARCA-TIPA) or marked declines of $N_e$ after the LGM (PUEB); (ii) a relatively constant $N_e$ through time (QUIN); and (iii) reductions of $N_e$ from the LGM to present (JARD) (Figure 4c). Glacial expansions increased $N_e$ by >75% (range = 79%–88%), whereas postglacial declines reduced $N_e$ at slower rates but to a similar magnitude (range 80%–97%) (Figure 4c).

## 3.6 | Testing alternative demographic models

Based on marginal densities calculated from the 1500 simulations retained for each model, the best-fitting scenario was the dynamic model (Model C, corresponding to hypothesis *H3*; Figure 1c; Table 4). The null model considering a flat landscape (Model A, corresponding to hypothesis *H1*; Figure 1a) was highly unlikely (Bayes factor [BF] > 1000;

**TABLE 2** Comparison of alternative models of divergence between Colombian populations of *Quercus humboldtii* from the three Andean Cordilleras tested using FASTSIMCOAL2.

| Model | lnL | *k* | AIC | ΔAIC | $\omega_i$ |
|---|---|---|---|---|---|
| (A) Central–Western Cordilleras (*n* = 6781 SNPs) | | | | | |
| SI | −9399.04 | 3 | 18,804.08 | 2.45 | 0.23 |
| **IM** | **−9395.81** | **5** | **18,801.63** | **0.00** | **0.77** |
| AM | −9490.30 | 6 | 18,992.61 | 190.98 | 0.00 |
| SC | −9504.43 | 6 | 19,020.87 | 219.24 | 0.00 |
| (B) Central–Eastern Cordilleras (*n* = 7244 SNPs) | | | | | |
| SI | −10,188.73 | 3 | 20,383.47 | 5.71 | 0.05 |
| **IM** | **−10,183.88** | **5** | **20,377.76** | **0.00** | **0.95** |
| AM | −10,221.23 | 6 | 20,454.47 | 76.71 | 0.00 |
| SC | −10,242.26 | 6 | 20,496.52 | 118.77 | 0.00 |
| (C) Western–Eastern Cordilleras (*n* = 6280 SNPs) | | | | | |
| SI | −9177.45 | 3 | 18,360.89 | 12.08 | 0.00 |
| **IM** | **−9169.40** | **5** | **18,348.81** | **0.00** | **1.00** |
| AM | −9187.37 | 6 | 18,386.74 | 37.93 | 0.00 |
| SC | −9202.85 | 6 | 18,417.69 | 68.89 | 0.00 |

*Note*: These models (illustrated in Figure 3) include strict isolation (SI), isolation-with-migration (IM), ancestral migration (AM) and secondary contact (SC). The best-supported scenario for each pairwise comparison is highlighted in bold. The number of variable SNPs retained for the calculation of the SFS in each comparison is indicated in parentheses.

Abbreviations: ΔAIC, difference in AIC value from that of the strongest model; *k*, number of parameters in the model; lnL, maximum likelihood value of the model; $\omega_i$, AIC weight.

Table 4), which is in line with the relatively weak pattern of IBD observed among the studied populations (see Section 3.2). The static model incorporating persistent barriers to dispersal through time (Model B, corresponding to hypothesis *H2*; Figure 1b) also fitted the data and was statistically indistinguishable from the best-supported model (BF < 20; Table 4; see Methods S4). However, only Model C was capable of generating simulated data comparable with the empirical data (Wegmann's $p = .527$; Table 4). The two other models (Models A and B) presented a substantial difference between the likelihoods of the simulated data compared with the empirical data (Wegmann's $p < .05$; Table 4). Posterior distributions of parameters under the most probable model were considerably distinct from the prior, indicating that the simulated data contained information relevant to estimating the parameters (Figure S4). Comparison of the posterior distributions before and after the ABC-general linear model (ABC-GLM; see Methods S4) also showed the improvement that this procedure had on parameter estimates (Figure S4). The coefficients of determination ($R^2$) between each demographic parameter and the four extracted partial least squares (PLS) (Figure S5; see Methods S4) indicated that the summary statistics we used had a high potential to correctly estimate all the parameters (Table 4). However, the histograms of the posterior quantiles of *m* and $N_{ANC}$ significantly deviated from a uniform distribution, suggesting a potential bias in the estimation of these parameters (Figure S6).

## 4 | DISCUSSION

Palaeodistribution modelling showed that *Quercus humboldtii* had a quite stable range through the late Quaternary, with distributional shifts limited to elevational displacements and local- to regional-scale fragmentation and expansion dynamics in response to cooling–warming cycles. Genomic data revealed a marked genetic structure in present-day populations of the species. However, signatures of admixture among most genetic clusters and statistical support for IM models indicate that geographically separated populations from the three Andean Cordilleras regularly exchange gene flow, albeit at low rates in absolute terms. Along the same lines, we found support for a dynamic model of flickering population connectivity since the LGM, with scenarios considering isotropic gene flow or currently unsuitable habitats as persistent barriers to dispersal (i.e., Cauca and Magdalena valleys) providing a comparatively worse fit to empirical genomic data. Overall, these results emphasize that spatial patterns of genomic variation in *Q. humboldtii* have been shaped by the interplay of landscape heterogeneity and past climate changes, supporting a significant impact of Quaternary climatic oscillations on the demographic dynamics of Andean montane forests (Ramírez-Barahona & Eguiarte, 2013) and rejecting the hypothesis of uninterrupted genetic connectivity of populations throughout the Pleistocene proposed in previous studies (Zorrilla-Azcué et al., 2021; see also Hooghiemstra et al., 2022).

### 4.1 | Distributional and demographic trajectories during the Quaternary

Palaeodistribution modelling indicates that *Q. humboldtii* experienced marked elevational displacements (~800–1000 m) in response to glacial–interglacial cycles, with a probable continuous persistence

**TABLE 3** Parameters inferred from coalescent simulations with FASTSIMCOAL2 under the best-supported model of divergence (i.e., isolation-with-migration model in all cases, illustrated in Figure 3) between Colombian populations of *Quercus humboldtii* from the three Andean Cordilleras.

| Parameter | Point estimate | Lower bound | Upper bound |
|---|---|---|---|
| (A) Central–Western Cordilleras | | | |
| $\theta_{ANC}$ | 2096 | 1542 | 2623 |
| $\theta_{1\text{–BELM}}$ | 6761 | — | — |
| $\theta_{2\text{–PUEB}}$ | 6987 | 4719 | 9731 |
| $T_{DIV}$ | 25,500 | 20,574 | 43,023 |
| $m_{12}$ | $1.65\times10^{-3}$ | $1.24\times10^{-3}$ | $2.01\times10^{-3}$ |
| $m_{21}$ | $1.94\times10^{-10}$ | $4.38\times10^{-11}$ | $6.61\times10^{-4}$ |
| (B) Central–Eastern Cordilleras | | | |
| $\theta_{ANC}$ | 1126 | 1044 | 1643 |
| $\theta_{1\text{–BELM}}$ | 6761 | — | — |
| $\theta_{2\text{–TIPA-ARCA-FUQU}}$ | 1866 | 1773 | 3283 |
| $T_{DIV}$ | 12,050 | 12,243 | 24,024 |
| $m_{12}$ | $1.81\times10^{-3}$ | $1.38\times10^{-3}$ | $2.36\times10^{-3}$ |
| $m_{21}$ | $5.63\times10^{-4}$ | $4.59\times10^{-11}$ | $1.28\times10^{-3}$ |
| (C) Western–Eastern Cordilleras | | | |
| $\theta_{ANC}$ | 3254 | 2690 | 3836 |
| $\theta_{1\text{–PUEB}}$ | 7606 | — | — |
| $\theta_{2\text{–TIPA-ARCA-FUQU}}$ | 4955 | 3935 | 6724 |
| $T_{DIV}$ | 43,200 | 39,151 | 70,824 |
| $m_{12}$ | $7.96\times10^{-4}$ | $5.96\times10^{-4}$ | $9.92\times10^{-4}$ |
| $m_{21}$ | $1.44\times10^{-4}$ | $3.32\times10^{-10}$ | $4.78\times10^{-4}$ |

*Note*: The table shows point estimates and lower and upper 95% confidence intervals for each parameter, which include mutation-scaled ancestral ($\theta_{ANC}$) and contemporary effective population sizes ($\theta_1$ and $\theta_2$), effective migration rates per generation ($m_{12}$ and $m_{21}$), and timing of population divergence ($T_{DIV}$). Estimates of time are given in years, considering a generation time of 50 years for long-lived oak trees. Effective population size of one population ($\theta_1$) was calculated from the level of nucleotide diversity (π; Table 1) and fixed in FASTSIMCOAL2 analyses to enable the estimation of other parameters (see Section 2.4 for further details). Population codes are described in Table 1.

of populations in the three Andean Cordilleras since, at least, the last glacial period (Figure 6; see also Hooghiemstra et al., 2022; Zorrilla-Azcué et al., 2021). Accordingly, sediment cores from Funza and Lake Fúquene in the Eastern Cordillera indicate an uninterrupted presence of *Q. humboldtii* in the pollen record since it colonized the region ~430 ka (Bogotá-A et al., 2011; Hooghiemstra, 2006; Torres et al., 2013; Van der Hammen, 1974). Both Funza (2550 m) and Lake Fúquene (2540 m) are located at elevations where *Q. humboldtii* was indeed uninterruptedly present since the LGM, which is an agreement with the predictions of our ENM. In line with previous studies, our reconstructions indicate that the total surface of oak forests in the Colombian Andes probably peaked during short warm interstadials and, thus, it was smaller than today through most of the Quaternary (Figure 6c; Hooghiemstra et al., 2022; Zorrilla-Azcué et al., 2021). Although *Q. humboldtii* maintained a relatively stable range-wide distribution (i.e., limited latitudinal/longitudinal shifts), the species experienced intermittent population connectivity and fragmentation at local to regional scales depending on the interplay between varying climatic conditions and mountain range profile (Figure 6; e.g., Flantua et al., 2019; Zorrilla-Azcué et al., 2021). Downslope shifts during the coldest stages of the Quaternary probably brought into closer geographical contact populations that today remain separated by an extensive matrix of unsuitable habitat for Andean montane species (e.g., dry forest in inter-Andean valleys; Hooghiemstra & Flantua, 2019). This might have promoted gene flow via pollen dispersal and contributed to demographic expansions (i.e., genetic admixture) in certain populations despite the lower total surface occupied by the species during the LGM and cool interstadials in the last 22,000 years. In this regard, our ENM suggests that populations of *Q. humboldtii* might have reached a minimum elevational limit at ~600 m during the LGM, with sporadic pulses of downward migration to elevations as low as 250 m, implying a displacement of the lower limit of the oak forest of at least 200 m with respect to the minimum elevational range of 800 m in present-day populations (Figure 6a). However, it is noteworthy that all reconstructions indicate the long-term persistence of extensive unsuitable habitats that have probably contributed to limit gene flow between populations of *Q. humboldtii* on both sides of the wide Magdalena Valley

**TABLE 4** Statistics from the ABC procedure used to evaluate the relative support of each spatially explicit demographic and coalescent model tested for *Quercus humboldtii* in Colombia.

| Model | Dynamic | Barrier | Marginal density | Wegmann's *p*-value | Bayes factor | $R^2$ | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | $K_{MAX}$ | $m$ | $N_{ANC}$ |
| A | No | No | $7.28\times10^{-11}$ | <.001 | $2.07\times10^{7}$ | .752 | .949 | .931 |
| B | No | Yes | $1.82\times10^{-4}$ | .032 | 8.27 | .922 | .969 | .948 |
| **C** | **Yes** | **Yes** | $\mathbf{1.51\times10^{-3}}$ | **.527** | **1** | **.881** | **.969** | **.939** |

*Note*: A higher marginal density corresponds to a higher model support, and a high (i.e., nonsignificant) Wegmann's *p*-value (>.05) indicates that the model is able to generate data in agreement with empirical data. Bayes factors represent the degree of relative support for the model with the highest marginal density (in bold) over the other models. Bayes factors >20 indicate strong support, while those >150 indicate very strong support (Kass & Raftery, 1995). $R^2$ is the coefficient of determination from a regression between each demographic parameter ($K_{MAX}$, $m$, $N_{ANC}$) and the four partial least squares (PLS) extracted from all summary statistics.

Abbreviations: $K_{MAX}$, carrying capacity of the deme with highest suitability; $m$, migration rate per deme per generation; $N_{ANC}$, ancestral population size.

(see also Hooghiemstra & van der Hammen, 1993, 2004; Van der Hammen, 1974).

Genomic-based reconstructions of changes in effective population size ($N_e$) through time revealed idiosyncratic responses of populations to Quaternary climatic oscillations (Figure 4c). These inferences point to population-specific demographic trajectories, which probably reflect complex processes linked to distributional shifts of montane forests in the different cordilleras (Figure S3) and a role of limited gene flow among populations in the maintenance of the genetic imprints left by their idiosyncratic demographic dynamics (Figure 4a). Note that there is always considerable uncertainty around mutation rates and generation times of long-lived trees and woody plants in general (see Sork et al., 2022), which can impact the scale of both time and $N_e$ in demographic reconstructions (i.e., the scale of *x*- and *y*-axes in STAIRWAY PLOT or similar methods), albeit in a predictable manner (Nadachowska-Brzyska et al., 2015). For instance, assuming a halved generation time (i.e., 25 years) or a halved mutation rate per year would double the estimates of $N_e$, but will not change the shape of $N_e$ trajectories (Nadachowska-Brzyska et al., 2015; Sork et al., 2022). Thus, independently of the mutation rates and generation times considered, the contrasting demographic dynamics inferred for the different studied populations of *Q. humboldtii* are expected to be genuine (Figure 4c). Using microsatellite markers and pooling data for populations sampled across the entire distribution of the species, Zorrilla-Azcué et al. (2021) inferred that *Q. humboldtii* experienced a bottleneck during the last glacial period followed by a demographic expansion after the LGM. In line with downslope shifts and probable population merging predicted by ENM and palaeoecological data during the coldest stages of the Quaternary (Hooghiemstra & Flantua, 2019; Hooghiemstra & Van der Hammen, 2004), our genomic-based reconstructions supported sudden demographic expansions during the last glacial period for certain populations but did not provide any evidence that these were preceded by demographic bottlenecks (Figure 4c). Palaeodistribution modelling suggests that the elevational range and surface area occupied by the species today was probably achieved ~8 ka BP (Figure 6a–c), which is consistent with the demographic stability of most populations since the onset of the Holocene (Figure 4c). Notably, the demographic stability of the QUIN population pre-dates the LGM (Figure 4c). The lower elevation of this population might have resulted in it having been little impacted by the compression of the upper part of the montane forest during the coldest stages of the Quaternary, which could ultimately explain the high stability of local effective population sizes through time in QUIN (Hooghiemstra & Van der Hammen, 2004).

## 4.2 | Genetic fragmentation with limited gene flow

Bayesian clustering analyses supported a marked genetic fragmentation of present-day populations of *Q. humboldtii*, which structured into five well-defined genetic clusters with different degrees of admixture among them (Figure 4a). These results are in line with a recent microsatellite-based study on *Q. humboldtii* identifying eight genetic groups across the entire distribution range of the species (Zorrilla-Azcué et al., 2021) and congruent with those reported for other congeneric tropical taxa (e.g., Deacon & Cavender-Bares, 2015; Ortego et al., 2015) and peripheral oak populations (e.g., Götz et al., 2022; Moracho et al., 2016). In agreement with palynological evidence supporting the Pleistocene arrival of *Q. humboldtii* in Colombia (Torres et al., 2013), our coalescent-based analyses estimate that the divergence between populations from the three cordilleras took place during the last glacial period (~12–71 ka BP; Table 3). This suggests that observed patterns of genetic structure pre-date the present-day distribution of dispersal corridors for the species (Figure 6d). Different ecological and abiotic factors can explain observed patterns of genetic structure in *Q. humboldtii*. On the one hand, the limited horizontal range shifts experienced by the species during the climatic oscillations of the Quaternary could have contributed not only to the development of genetic structure in refugial populations during periods of isolation and demographic declines, but also to the persistence of such distinctive genetic pools through extended periods of time. Thus, limited or absent latitudinal shifts together with elevation displacements tracking optimal climate during the Quaternary could explain both marked genetic structure within cordilleras and observed signatures of genetic admixture (i.e., pulses of gene flow) among populations from different cordilleras located at the same latitudinal range. On the other hand, the humid environment and high rainfall characterizing the tropical Andes is expected to negatively impact wind-pollination by reducing the availability and efficient transportation of airborne pollen (Regal, 1982; Turner, 2001; e.g., Deacon & Cavender-Bares, 2015; Moracho et al., 2016). Long-term isolation of populations (i.e., in situ persistence) and inefficient wind-pollination might have constrained the rates of gene flow across the landscape in *Q. humboldtii* and other tropical oaks, leading to the development of a marked genetic structure. This situation contrasts with the weak or absent genetic structure charactering congeneric taxa from Mediterranean or temperate regions (e.g., Craft & Ashley, 2007; Gugger et al., 2013, 2021; Ortego et al., 2017), which show high rates of pollen-mediated dispersal (e.g., Buschbom et al., 2011; Ortego et al., 2014) and have probably experienced considerable gene flow and admixture during recent post-glacial range expansions at both regional and continental-wide scales (e.g., Merceron et al., 2017; Ramírez-Valiente et al., 2023). In the same line, strong genetic fragmentation and limited gene flow among populations of *Q. humboldtii* could explain why local effective population sizes in this species are much smaller than those estimated for temperate oaks (Sork et al., 2022).

Despite a marked genetic structure, Bayesian clustering analyses also revealed certain degree of genetic admixture among populations assigned to the different genetic clusters and provided empirical evidence for gene flow across Cauca and Magdalena valleys that separate oak populations in the Eastern, Central and Western Cordilleras (e.g., BELM-PUE and ARCA-BELM; Figure 4a; see also Zorrilla-Azcué et al., 2021). According to STRUCTURE analyses, testing of alternative models of population divergence supported IM scenarios over those

considering strict isolation, ancestral migration or secondary contact (Table 2). Thus, populations from the different cordilleras have exchanged gene flow since the onset of their divergence, although this should not be strictly interpreted as evidence for uninterrupted gene flow through time. Clear signatures of admixed ancestry but limited evidence for first-generation hybrids (i.e., STRUCTURE *q*-values ~0.5) between geographically distant populations from different cordilleras (e.g., ARCA-BELM, separated by >250 km; Figure 4a) suggests that cross-valley gene flow was probably circumscribed to intermittent pulses of population connectivity intermixed with periods of isolation (i.e., ephemeral barriers; Araya-Donoso et al., 2022), in line with inferences from ENM (Figure 6) and available palaeoecological data (Hooghiemstra & Van der Hammen, 2004; Van der Hammen, 1974). High genetic connectivity among certain populations is also evidenced by the assignment to the same genetic cluster of the three northernmost populations from the Eastern Cordillera (FUQU, ARCA and TIPA) (Figure 4a,b). These three populations are located at high elevations (2840–2940 m; Table 1) and have been uninterruptedly connected by continuous suitable habitats for the species from the LGM to present (Figure 6), which is expected to have contributed to their genetic homogenization despite being separated by up to 140 km (Figure 2). Remarkably, our analyses also revealed that genetic differentiation among populations is best explained by elevational dissimilarity (i.e., IBE; Figure 5b) than geographical distances (i.e., IBD; Figure 5a). This suggests that differences in phenology (e.g., asynchronous flowering times) and/or local adaptations (e.g., selection against immigrant genotypes) in populations experiencing contrasting environmental conditions have also contributed to observed patterns of genetic structure (Hendry & Day, 2005; Sexton et al., 2014; Wang & Bradburd, 2014).

### 4.3 | Insights from spatially explicit model testing

Spatially explicit demographic and coalescent analyses rejected the hypothesis of isotropic gene flow, strongly supporting the role of landscape heterogeneity on structuring patterns of genomic variation in *Q. humboldtii*. The low statistical support of the flat landscape scenario (i.e., null-model; Model A, hypothesis *H1*) relative to models incorporating either the spatial distribution of contemporary barriers to dispersal (Model B, hypothesis *H2*) or distributional shifts of the species from the LGM to the present (Model C, hypothesis *H3*) indicates the role of landscape heterogeneity on shaping patterns of genetic structure in the species. The two most likely models incorporate the role of unsuitable habitats as hypothetical barriers to dispersal, which have dominated inter-Andean lowlands since the LGM (i.e., Magdalena Valley and, to a lesser extent, Cauca Valley; Figure 2) and were fitted either as persistent barriers (Model B) or as habitat patches sustaining minimal carrying capacities through time (Model C). However, although Model B was statistically well supported (BF < 10), only Model C was able to generate data comparable with observed data and, thus, fully compatible with the demographic dynamics experienced by the species (Table 4). Indeed, genetic admixture revealed by Bayesian clustering analyses (Figure 4a) and statistical support for IM scenarios of divergence (Table 2) indicate that populations from the different Andean Cordilleras have regularly exchanged gene flow, suggesting that hypothesis *H2* (i.e., the disruption of gene flow by the presence of persistent barriers to dispersal) is highly unlikely (Figure 1). Thus, in line with the "moist forest" (Ramírez-Barahona & Eguiarte, 2013) and "flickering connectivity" (Flantua et al., 2019) models, our results lend empirical support to the scenario hypothesizing downslope migrations during the coldest stages of the Quaternary and punctuated connectivity between geographically separated populations depending on climatic conditions and mountain profile (Figure 1c; see also Hooghiemstra & Van der Hammen, 2004; Hooghiemstra et al., 2022). Future analyses including a denser population sampling—our sampling was admittedly low in the Central Cordillera (Figure 2)—and covering the entire distribution of the species could provide a higher statistical power to distinguish among competing scenarios (e.g., hypotheses *H2* and *H3*) and would allow testing more refined hypotheses considering different spatial replicates (e.g., analyses focused on different cordilleras or latitudinal sections along the species range).

## 5 | CONCLUSIONS

Our study revealed an important role of landscape heterogeneity and restricted dispersal through unsuitable habitats in determining the marked genetic fragmentation of montane oak forests in the Andean Cordilleras, rejecting the hypothesis of persistent connectivity proposed in previous studies (Zorrilla-Azcué et al., 2021). Our genomic data support that elevational displacements and corresponding pulses of population merging and fragmentation at local to regional scales were not sufficient to erode the genetic structure of present-day oak populations and, on the contrary, probably contributed to fuel their geographical diversification (e.g., Gutiérrez-Rodríguez et al., 2022; Nevado et al., 2018). Future genomic-based demographic reconstructions for a representative number of taxa of the different vegetation belts in a comparative framework might help to gain further insights into the dynamics of Andean biotas in response to Quaternary climatic oscillations (e.g., see Helmstetter et al., 2020 for a forest community in tropical Africa). Collectively, our study exemplifies how the interdisciplinary integration of genomic data, spatial modelling and palaeoecological information can contribute to advance our understanding of the responses of tropical systems to Quaternary climatic fluctuations (see also Flantua et al., 2019; Ornelas et al., 2019; Ramírez-Barahona & Eguiarte, 2013), which is still very limited compared to the well-documented glacial–interglacial dynamics experienced by temperate biotas (Hewitt, 2004; Stewart et al., 2010). Ongoing biogeographical research in the tropical Andes guarantees future opportunities to further study the consequences of Quaternary climate change in this biodiversity hotspot (e.g., Flantua et al., 2019; Hazzi et al., 2018; Muñoz-Ortiz et al., 2015; Nevado et al., 2018; Sanín et al., 2022) and to gain further insights into the proximate ecological and

evolutionary processes that have contributed to its extraordinary levels of endemism (Kier et al., 2009; Myers et al., 2000).

## AUTHOR CONTRIBUTIONS

Joaquín Ortego and Raúl Bonal conceived the study. Josep Maria Espelta, Dolors Armenteras, María Claudia Díez, Alberto Muñoz and Raúl Bonal collected the samples. Joaquín Ortego generated the genomic data and designed and performed the analyses. Joaquín Ortego wrote the manuscript, with inputs from all the authors.

## ACKNOWLEDGEMENTS

We thank to Amparo Hidalgo-Galiana for genomic library preparation and Sergio Pereira (The Centre for Applied Genomics) for Illumina sequencing. We also thank Suzette Flantua, María José Sanín, Victoria L. Sork and an anonymous reviewer for their constructive and valuable comments on an earlier version of the manuscript. Logistical support was provided by Laboratorio de Ecología Molecular (LEM-EBD) from Estación Biológica de Doñana (CSIC). We also thank Centro de Supercomputación de Galicia (CESGA) and Doñana's Singular Scientific-Technical Infrastructure (ICTS-RBD) for access to computer resources and Universidad Nacional de Colombia for logistical support and providing sampling permits (Permiso marco de colecta de especímenes, Resolución 0255 Marzo 14, 2014 de la Autoridad Nacional de Licencias Ambientales). This study was funded by the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund (ERDF) (RYC-2013-12501).

## CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

## DATA AVAILABILITY STATEMENT

Raw Illumina reads have been deposited at the NCBI Sequence Read Archive (SRA) under BioProject PRJNA926014. Input files for all analyses (STRUCTURE, PCA, environmental niche modelling, FASTSIMCOAL2, STAIRWAY PLOT and SPLATCHE2) are available for download from Figshare (https://doi.org/10.6084/m9.figshare.21936888).

## ORCID

*Joaquín Ortego* https://orcid.org/0000-0003-2709-429X
*Josep Maria Espelta* https://orcid.org/0000-0002-0242-4988
*Dolors Armenteras* https://orcid.org/0000-0003-0922-7298
*María Claudia Díez* https://orcid.org/0000-0003-0066-9866
*Alberto Muñoz* https://orcid.org/0000-0002-1152-7444
*Raúl Bonal* https://orcid.org/0000-0002-6084-1771

## REFERENCES

Araya-Donoso, R., Baty, S. M., Alonso-Alonso, P., Sanín, M. J., Wilder, B. T., Munguia-Vega, A., & Dolby, G. A. (2022). Implications of barrier ephemerality in geogenomic research. *Journal of Biogeography*, *49*(11), 2050–2063. https://doi.org/10.1111/jbi.14487

Beaumont, M. A., Zhang, W. Y., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, *162*(4), 2025–2035.

Bemmels, J. B., Title, P. O., Ortego, J., & Knowles, L. L. (2016). Tests of species-specific models reveal the importance of drought in post-glacial range shifts of a Mediterranean-climate tree: Insights from integrative distributional, demographic and coalescent modelling and ABC model selection. *Molecular Ecology*, *25*(19), 4889–4906. https://doi.org/10.1111/mec.13804

Bibby, C. J., Collar, N. J., Crosby, M. J., Heath, M. F., Imboden, C., Johnson, T. H., Long, A. J., Stattersfield, A. J., & Thirgood, S. J. (1992). *Putting biodiversity on the map: Priority areas for global conservation*. International Council for Bird Preservation.

Bogotá-A, R. G., Groot, M. H. M., Hooghiemstra, H., Lourens, L. J., Van der Linden, M., & Berrio, J. C. (2011). Rapid climate change from north Andean Lake Fuquene pollen records driven by obliquity: Implications for a basin-wide biostratigraphic zonation for the last 284 ka. *Quaternary Science Reviews*, *30*(23–24), 3321–3337. https://doi.org/10.1016/j.quascirev.2011.08.003

Boschman, L. M. (2021). Andean mountain building since the late cretaceous: A paleoelevation reconstruction. *Earth-Science Reviews*, *220*, 103640. https://doi.org/10.1016/j.earscirev.2021.103640

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. Springer.

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://doi.org/10.1093/jhered/esr023

Bush, M. B., & de Oliveira, P. E. (2006). The rise and fall of the refugial hypothesis of Amazonian speciation: A paleoecological perspective. *Biota Neotropica*, *v6*(n1), bn00106012006.

Carnaval, A. C., Hickerson, M. J., Haddad, C. F. B., Rodrigues, M. T., & Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, *323*(5915), 785–789. https://doi.org/10.1126/science.1166955

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). STACKS: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Chaves, J. A., Weir, J. T., & Smith, T. B. (2011). Diversification in *Adelomyia* hummingbirds follows Andean uplift. *Molecular Ecology*, *20*(21), 4564–4576. https://doi.org/10.1111/j.1365-294X.2011.05304.x

Craft, K. J., & Ashley, M. V. (2007). Landscape genetic structure of bur oak (*Quercus macrocarpa*) savannas in Illinois. *Forest Ecology and Management*, *239*(1–3), 13–20.

de Manuel, M., Kuhlwilm, M., Frandsen, P., Sousa, V. C., Desai, T., Prado-Martinez, J., Hernandez-Rodriguez, J., Dupanloup, I., Lao, O., Hallast, P., Schmidt, J. M., Heredia-Genestar, J. M., Benazzo, A., Barbujani, G., Peter, B. M., Kuderna, L. F., Casals, F., Angedakin, S., Arandjelovic, M., … Marques-Bonet, T. (2016). Chimpanzee genomic diversity reveals ancient admixture with bonobos. *Science*, *354*(6311), 477–481. https://doi.org/10.1126/science.aag2602

Deacon, N. J., & Cavender-Bares, J. (2015). Limited pollen dispersal contributes to population genetic structure but not local adaptation in *Quercus oleoides* forests of Costa Rica. *PLoS One*, *10*(9), e0138783. https://doi.org/10.1371/journal.pone.0138783

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing structure output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

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.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genetics*, *9*(10), e1003905. https://doi.org/10.1371/journal.pgen.1003905

Excoffier, L., & Lischer, H. E. L. (2010). ARLEQUIN suite ver 3.5: A new series of programs to perform population genetics analyses under Linux

and windows. *Molecular Ecology Resources*, *10*(3), 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Fernández, J. F., & Sork, V. L. (2007). Genetic variation in fragmented forest stands of the Andean oak *Quercus humboldtii* Bonpl. (Fagaceae). *Biotropica*, *39*(1), 72–78.

Flantua, S. G. A., & Hooghiemstra, H. (2018). Historical connectivity and mountain biodiversity. In C. Hoorn, A. Perrigo, & A. Antonelli (Eds.), *Mountains, climate and biodiversity* (1st ed., pp. 171–185). Wiley-Blackwell.

Flantua, S. G. A., O'Dea, A., Onstein, R. E., Giraldo, C., & Hooghiemstra, H. (2019). The flickering connectivity system of the north Andean paramos. *Journal of Biogeography*, *46*(8), 1808–1825. https://doi.org/10.1111/jbi.13607

Gilbert, K. J., Andrew, R. L., Bock, D. G., Franklin, M. T., Kane, N. C., Moore, J. S., Moyers, B. T., Renaut, S., Rennison, D. J., Veen, T., & Vines, T. H. (2012). Recommendations for utilizing and reporting population genetic analyses: The reproducibility of genetic clustering using the program STRUCTURE. *Molecular Ecology*, *21*(20), 4925–4930. https://doi.org/10.1111/j.1365-294X.2012.05754.x

González-Serna, M. J., Cordero, P. J., & Ortego, J. (2019). Spatiotemporally explicit demographic modelling supports a joint effect of historical barriers to dispersal and contemporary landscape composition on structuring genomic variation in a red-listed grasshopper. *Molecular Ecology*, *28*(9), 2155–2172. https://doi.org/10.1111/mec.15086

Gossmann, T. I., Keightley, P. D., & Eyre-Walker, A. (2012). The effect of variation in the effective population size on the rate of adaptive molecular evolution in eukaryotes. *Genome Biology and Evolution*, *4*(5), 658–667. https://doi.org/10.1093/gbe/evs027

Götz, J., Rajora, O. P., & Gailing, O. (2022). Genetic structure of natural northern range-margin mainland, peninsular, and Island populations of northern red oak (*Quercus rubra* L.). *Frontiers in Ecology and Evolution*, *10*, 907414. https://doi.org/10.3389/fevo.2022.907414

Gregory-Wodzicki, K. M. (2000). Uplift history of the central and northern Andes: A review. *Geological Society of America Bulletin*, *112*(7), 1091–1105. https://doi.org/10.1130/0016-7606(2000)112<1091:uhotca>2.0.co;2

Gugger, P. F., Fitz-Gibbon, S. T., Albarrán-Lara, A., Wright, J. W., & Sork, V. L. (2021). Landscape genomics of *Quercus lobata* reveals genes involved in local climate adaptation at multiple spatial scales. *Molecular Ecology*, *30*(2), 406–423. https://doi.org/10.1111/mec.15731

Gugger, P. F., Ikegami, M., & Sork, V. L. (2013). Influence of late Quaternary climate change on present patterns of genetic variation in valley oak, *Quercus lobata* Nee. *Molecular Ecology*, *22*(13), 3598–3612. https://doi.org/10.1111/mec.12317

Gutiérrez-Rodríguez, J., Nieto-Montes de Oca, A., Ortego, J., & Zaldívar-Riverón, A. (2022). Phylogenomics of arboreal alligator lizards shed light on the geographical diversification of cloud forest-adapted biotas. *Journal of Biogeography*, *49*(10), 1862–1876. https://doi.org/10.1111/jbi.14461

Hazzi, N. A., Moreno, J. S., Ortiz-Movliav, C., & Palacio, R. D. (2018). Biogeographic regions and events of isolation and diversification of the endemic biota of the tropical Andes. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(31), 7985–7990. https://doi.org/10.1073/pnas.1803908115

He, Q. X., Edwards, D. L., & Knowles, L. L. (2013). Integrative testing of how environments from the past to the present shape genetic structure across landscapes. *Evolution*, *67*(12), 3386–3402. https://doi.org/10.1111/evo.12159

Helmstetter, A. J., Bethune, K., Kamdem, N. G., Sonke, B., & Couvreur, T. L. P. (2020). Individualistic evolutionary responses of central African rain forest plants to Pleistocene climatic fluctuations. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(51), 32509–32518. https://doi.org/10.1073/pnas.2001018117

Hendry, A. P., & Day, T. (2005). Population structure attributable to reproductive time: Isolation by time and adaptation by time. *Molecular Ecology*, *14*(4), 901–916. https://doi.org/10.1111/j.1365-294X.2005.02480.x

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, *359*(1442), 183–195.

Hooghiemstra, H. (2006). Immigration of oak into northern South America: A paleo-ecological document. In M. Kappelle (Ed.), *Ecology and conservation of neotropical montane oak forests* (Vol. *185*, pp. 17–28). Springer.

Hooghiemstra, H., Cleef, A. M., & Flantua, S. G. A. (2022). A paleoecological context to assess the development of oak forest in Colombia: A comment on Zorrilla-Azcué, S., González-Rodríguez, A., Oyama, K., González, MA, & Rodríguez-Correa, H., The DNA history of a lonely oak: *Quercus humboldtii* phylogeography in the Colombian Andes. *Ecology and Evolution* 2021, doi: 10.100-2/ece3.7529. *Ecology and Evolution*, *12*(3), e8702. https://doi.org/10.1002/ece3.8702

Hooghiemstra, H., & Flantua, S. G. A. (2019). Colombia in Quaternary: An overview of environmental and climatic change (Colombia durante el Cuaternario; una síntesis sobre cambio climático y ambiental). In J. Gómez-Tapias (Ed.), *The geology of Colombia book* (Vol. *4 Quaternary, Chapter 2*, pp. 33–84). Servicio Geológico Colombiano. https://doi.org/10.2345/tgocb.35.4.2

Hooghiemstra, H., & Van der Hammen, T. (1993). Late Quaternary vegetation history and paleoecology of Laguna Pedro Palo (subandean forest belt, Eastern Cordillera, Colombia). *Review of Palaeobotany and Palynology*, *77*(3–4), 235–262. https://doi.org/10.1016/0034-6667(93)90006-g

Hooghiemstra, H., & Van der Hammen, T. (2004). Quaternary ice-age dynamics in the Colombian Andes: Developing an understanding of our legacy. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, *359*(1442), 173–180. https://doi.org/10.1098/rstb.2003.1420

Hughes, C., & Eastwood, R. (2006). Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(27), 10334–10339. https://doi.org/10.1073/pnas.0601928103

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., & Andrew, R. L. (2017). The *K*=2 conundrum. *Molecular Ecology*, *26*(14), 3594–3602. https://doi.org/10.1111/mec.14187

Jombart, T. (2008). 'Adegenet': A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Kadereit, J. W., & Abbott, R. J. (2021). Plant speciation in the Quaternary. *Plant Ecology & Diversity*, *14*(3–4), 105–142. https://doi.org/10.1080/17550874.2021.2012849

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, *4*, 170122. https://doi.org/10.1038/sdata.2017.122

Karger, D. N., Nobis, M. P., Normand, S., Graham, C. H., & Zimmermann, N. E. (2023). CHELSA-TraCE21k – High-resolution (1 km) downscaled transient temperature and precipitation data since the last glacial maximum. *Climate of the Past*, *19*(2), 439–456. https://doi.org/10.5194/cp-19-439-2023

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kattan, G. H., Franco, P., Rojas, V., & Morales, G. (2004). Biological diversification in a complex region: A spatial analysis of faunistic diversity and biogeography of the Andes of Colombia.

*Journal of Biogeography*, *31*(11), 1829–1839. https://doi.org/10.1111/j.1365-2699.2004.01109.x

Kier, G., Kreft, H., Lee, T. M., Jetz, W., Ibisch, P. L., Nowicki, C., Mutke, J., & Barthlott, W. (2009). A global assessment of endemism and species richness across Island and mainland regions. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(23), 9322–9327. https://doi.org/10.1073/pnas.0810306106

Liu, C. R., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, *28*(3), 385–393. https://doi.org/10.1111/j.0906-7590.2005.03957.x

Liu, X. M., & Fu, Y. X. (2020). STAIRWAY PLOT 2: Demographic history inference with folded SNP frequency spectra. *Genome Biology*, *21*(1), 280. https://doi.org/10.1186/s13059-020-02196-9

Massatti, R., & Knowles, L. L. (2016). Contrasting support for alternative models of genomic variation based on microhabitat preference: Species-specific effects of climate change in alpine sedges. *Molecular Ecology*, *25*(16), 3974–3986. https://doi.org/10.1111/mec.13735

Merceron, N. R., Leroy, T., Chancerel, E., Romero-Severson, J., Borkowski, D. S., Ducousso, A., Monty, A., Porté, A. J., & Kremer, A. (2017). Back to America: Tracking the origin of European introduced populations of *Quercus rubra* L. *Genome*, *60*(9), 778–790. https://doi.org/10.1139/gen-2016-0187

Moracho, E., Moreno, G., Jordano, P., & Hampe, A. (2016). Unusually limited pollen dispersal and connectivity of pedunculate oak (*Quercus robur*) refugial populations at the species' southern range margin. *Molecular Ecology*, *25*(14), 3319–3331. https://doi.org/10.1111/mec.13692

Muñoz-Ortiz, A., Velásquez-Alvárez, A. A., Guarnizo, C. E., & Crawford, A. J. (2015). Of peaks and valleys: Testing the roles of orogeny and habitat heterogeneity in driving allopatry in mid-elevation frogs (Aromobatidae: Rheobates) of the northern Andes. *Journal of Biogeography*, *42*(1), 193–205. https://doi.org/10.1111/jbi.12409

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853–858.

Nadachowska-Brzyska, K., Li, C., Smeds, L., Zhang, G. J., & Ellegren, H. (2015). Temporal dynamics of avian populations during Pleistocene revealed by whole-genome sequences. *Current Biology*, *25*(10), 1375–1380. https://doi.org/10.1016/j.cub.2015.03.047

Nevado, B., Contreras-Ortiz, N., Hughes, C., & Filatov, D. A. (2018). Pleistocene glacial cycles drive isolation, gene flow and speciation in the high-elevation Andes. *New Phytologist*, *219*(2), 779–793. https://doi.org/10.1111/nph.15243

Ornelas, J. F., Ortiz-Rodriguez, A. E., Ruiz-Sanchez, E., Sosa, V., & Perez-Farrera, M. A. (2019). Ups and downs: Genetic differentiation among populations of the *Podocarpus* (Podocarpaceae) species in Mesoamerica. *Molecular Phylogenetics and Evolution*, *138*, 17–30. https://doi.org/10.1016/j.ympev.2019.05.025

Ortego, J., Bonal, R., Muñoz, A., & Aparicio, J. M. (2014). Extensive pollen immigration and no evidence of disrupted mating patterns or reproduction in a highly fragmented holm oak stand. *Journal of Plant Ecology*, *7*(4), 384–395. https://doi.org/10.1093/jpe/rtt049

Ortego, J., Bonal, R., Muñoz, A., & Espelta, J. M. (2015). Living on the edge: The role of geography and environment in structuring genetic variation in the southernmost populations of a tropical oak. *Plant Biology*, *17*(3), 676–683. https://doi.org/10.1111/plb.12272

Ortego, J., Gugger, P. F., & Sork, V. L. (2017). Impacts of human-induced environmental disturbances on hybridization between two ecologically differentiated Californian oak species. *New Phytologist*, *213*(2), 930–943. https://doi.org/10.1111/nph.14182

Ortego, J., Gugger, P. F., & Sork, V. L. (2018). Genomic data reveal cryptic lineage diversification and introgression in Californian golden cup oaks (section *Protobalanus*). *New Phytologist*, *218*(2), 804–818. https://doi.org/10.1111/nph.14951

Ortego, J., & Knowles, L. L. (2020). Incorporating interspecific interactions into phylogeographic models: A case study with Californian oaks. *Molecular Ecology*, *29*(23), 4510–4524. https://doi.org/10.1111/mec.15548

Ortego, J., & Knowles, L. L. (2022). Geographical isolation versus dispersal: Relictual alpine grasshoppers support a model of interglacial diversification with limited hybridization. *Molecular Ecology*, *31*(1), 296–312. https://doi.org/10.1111/mec.16225

Pennington, R. T., Lavin, M., Sarkinen, T., Lewis, G. P., Klitgaard, B. B., & Hughes, C. E. (2010). Contrasting plant diversification histories within the Andean biodiversity hotspot. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(31), 13783–13787. https://doi.org/10.1073/pnas.1001317107

Pérez-Escobar, O. A., Zizka, A., Bermúdez, M. A., Meseguer, A. S., Condamine, F. L., Hoorn, C., Hooghiemstra, H., Pu, Y., Bogarín, D., Boschman, L. M., Pennington, R. T., Antonelli, A., & Chomicki, G. (2022). The Andes through time: Evolution and distribution of Andean floras. *Trends in Plant Science*, *27*(4), 364–378. https://doi.org/10.1016/j.tplants.2021.09.010

Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M., & Araújo, M. B. (2011). *Ecological niches and geographic distributions*. Princeton University Press.

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*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3–4), 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips, S. J., & Dudik, M. (2008). Modeling of species distributions with MAXENT: New extensions and a comprehensive evaluation. *Ecography*, *31*(2), 161–175. https://doi.org/10.1111/j.0906-7590.2008.5203.x

Prates, I., Xue, A. T., Brown, J. L., Alvarado-Serrano, D. F., Rodrigues, M. T., Hickerson, M. J., & Carnaval, A. C. (2016). Inferring responses to climate dynamics from historical demography in neotropical forest lizards. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(29), 7978–7985. https://doi.org/10.1073/pnas.1601063113

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Pulido, M. T., Cavelier, J., & Cortés, S. P. (2006). Structure and composition of Colombian montane oak forests. In M. Kappelle (Ed.), *Ecology and conservation of neotropical montane oak forests* (Vol. *185*, pp. 17–28). Springer.

R Core Team. (2022). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Radosavljevic, A., & Anderson, R. P. (2014). Making better MAXENT models of species distributions: Complexity, overfitting and evaluation. *Journal of Biogeography*, *41*(4), 629–643. https://doi.org/10.1111/jbi.12227

Ramírez-Barahona, S., & Eguiarte, L. E. (2013). The role of glacial cycles in promoting genetic diversity in the neotropics: The case of cloud forests during the last glacial maximum. *Ecology and Evolution*, *3*(3), 725–738. https://doi.org/10.1002/ece3.483

Ramírez-Barahona, S., & Eguiarte, L. E. (2014). Changes in the distribution of cloud forests during the last glacial predict the patterns of genetic diversity and demographic history of the tree fern *Alsophila firma* (Cyatheaceae). *Journal of Biogeography*, *41*(12), 2396–2407. https://doi.org/10.1111/jbi.12396

Ramírez-Valiente, J. A., Solé-Medina, A., Robledo-Arnuncio, J. J., & Ortego, J. (2023). Genomic data and common garden experiments reveal climate-driven selection on ecophysiological traits in two

Mediterranean oaks. *Molecular Ecology*, *32*(5), 983–999. https://doi.org/10.1111/mec.16816

Rangel, J. O. C., & Avella, A. (2011). Oak forests of *Quercus humboldtii* in the Caribbean region and distribution patterns related with environmental factors in Colombia. *Plant Biosystems*, *145*, 186–198. https://doi.org/10.1080/11263504.2011.602727

Regal, P. J. (1982). Pollination by wind and animals: Ecology of geographic patterns. *Annual Review of Ecology and Systematics*, *13*, 497–524. https://doi.org/10.1146/annurev.es.13.110182.002433

Ribas, C. C., Moyle, R. G., Miyaki, C. Y., & Cracraft, J. (2007). The assembly of montane biotas: Linking Andean tectonics and climatic oscillations to independent regimes of diversification in *Pionus* parrots. *Proceedings of the Royal Society B-Biological Sciences*, *274*(1624), 2399–2408. https://doi.org/10.1098/rspb.2007.0613

Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, *4*(1), 137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Sanín, M. J., Mejia-Franco, F. G., Paris, M., Valencia-Montoya, W. A., Salamin, N., Kessler, M., Olivares, I., Jaramillo, J. S., & Cardona, A. (2022). Geogenomics of montane palms points to Miocene-Pliocene Andean segmentation related to strike-slip tectonics. *Journal of Biogeography*, *49*(9), 1711–1725. https://doi.org/10.1111/jbi.14327

Schmid-Siegert, E., Sarkar, N., Iseli, C., Calderon, S., Gouhier-Darimont, C., Chrast, J., Cattaneo, P., Schütz, F., Farinelli, L., Pagni, M., Schneider, M., Voumard, J., Jaboyedoff, M., Fankhauser, C., Hardtke, C. S., Keller, L., Pannell, J. R., Reymond, A., Robinson-Rechavi, M., … Reymond, P. (2017). Low number of fixed somatic mutations in a long-lived oak tree. *Nature Plants*, *3*(12), 926–929. https://doi.org/10.1038/s41477-017-0066-9

Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2014). Genetic isolation by environment or distance: Which patterns of gene flow is most common? *Evolution*, *68*(1), 1–15. https://doi.org/10.1111/evo.12258

Sork, V. L., Cokus, S. J., Fitz-Gibbon, S. T., Zimin, A. V., Puiu, D., Garcia, J. A., Gugger, P. F., Henriquez, C. L., Zhen, Y., Lohmueller, K. E., Pellegrini, M., & Salzberg, S. L. (2022). High-quality genome and methylomes illustrate features underlying evolutionary success of oaks. *Nature Communications*, *13*(1), 2047. https://doi.org/10.1038/s41467-022-29584-y

Stewart, J. R., Lister, A. M., Barnes, I., & Dalen, L. (2010). Refugia revisited: Individualistic responses of species in space and time. *Proceedings of the Royal Society B-Biological Sciences*, *277*(1682), 661–671. https://doi.org/10.1098/rspb.2009.1272

Thome, M. T. C., & Carstens, B. C. (2016). Phylogeographic model selection leads to insight into the evolutionary history of four-eyed frogs. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(29), 8010–8017. https://doi.org/10.1073/pnas.1601064113

Torres, V., Hooghiemstra, H., Lourens, L., & Tzedakis, P. C. (2013). Astronomical tuning of long pollen records reveals the dynamic history of montane biomes and lake levels in the tropical high Andes during the quaternary. *Quaternary Science Reviews*, *63*, 59–72. https://doi.org/10.1016/j.quascirev.2012.11.004

Turner, I. M. (2001). *The ecology of trees in the tropical rain forest*. Cambridge University Press.

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R. R., Bhalerao, R. P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., … Rokhsar, D. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, *313*(5793), 1596–1604. https://doi.org/10.1126/science.1128691

Van der Hammen, T. (1974). The Pleistocene changes of vegetation and climate in tropical South America. *Journal of Biogeography*, *1*(1), 3–26.

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, *23*(23), 5649–5662. https://doi.org/10.1111/mec.12938

Zorrilla-Azcué, S., González-Rodríguez, A., Oyama, K., González, M. A., & Rodríguez-Correa, H. (2021). The DNA history of a lonely oak: *Quercus humboldtii* phylogeography in the Colombian Andes. *Ecology and Evolution*, *11*(11), 6814–6828. https://doi.org/10.1002/ece3.7529

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Ortego, J., Espelta, J. M., Armenteras, D., Díez, M. C., Muñoz, A., & Bonal, R. (2023). Demographic and spatially explicit landscape genomic analyses in a tropical oak reveal the impacts of late Quaternary climate change on Andean montane forests. *Molecular Ecology*, *32*, 3182–3199. https://doi.org/10.1111/mec.16930