



OPEN

Late Quaternary range shifts of marcescent oaks unveil the dynamics of a major biogeographic transition in southern Europe

Carlos Vila-Viçosa^{1,2,3}✉, João Gonçalves¹, João Honrado^{1,3}, Ângela Lomba¹, Rubim S. Almeida^{1,2,3}, Francisco Maria Vázquez⁴ & Cristina Garcia^{1,5}

Marcescent forests are ecotones distributed across southern Europe that host increased levels of biodiversity but their persistence is threatened by global change. Here we study the range dynamics of these forests in the Iberian Peninsula (IP) during the Late Quaternary, a period of profound climate and anthropic changes. We modeled and compared the distribution of eight oak taxa for the present and two paleoclimatic environments, the Last Glacial Maximum (LGM, ~ 21 kya) and the Mid-Holocene (MH, ~ 6 kya). Presence records were combined with bioclimatic and topographic data in an ensemble modelling framework to obtain spatial projections for present and past conditions across taxa. Substantial distribution shifts were projected between the three studied periods, that were explained by precipitation, winter cold and terrain ruggedness. Results were congruent with paleoclimatic records of the IP and showed that range shifts of these contact zones concurred with range dynamics of both Submediterranean and Temperate oaks. Notably, the distribution ranges of hybrid oaks and marcescent forests matched throughout the late Quaternary. This study contributes to unveil the complex Late-Quaternary biogeography of the ecotone belt occupied by marcescent forests and, more broadly, of Mediterranean oaks. Improved knowledge of species' responses to climate dynamics will allow us to anticipate and manage future range shifts driven by climate change.

Ecotones are transitional areas between contrasting adjacent ecological systems integrating ecological features from neighbor regions¹. They host complex and highly diverse communities that combine endemic species and those from adjacent neighboring zones that typically use these areas as refugia when environmental changes challenge their persistence across their native ranges². This combination of species makes ecotones hotspots of biodiversity and, therefore, they are of utmost importance if we are to preserve biodiversity in a changing world³⁻⁵. Although the distribution and extent of ecotones have changed widely over millennia in response to climate change⁶, we still ignore whether and past climate changes have shaped the current distribution of species inhabiting ecotones.

Marcescent forests are ecotones located in the transition between temperate areas, with cold winters and mild rainy summers and Mediterranean ones, with dry and hot summers. They occur in the European Southern Peninsulas, through the Mediterranean Basin, including North Africa and Middle East regions^{7,8}. Marcescence, i.e., the absence of leaf fall abscission, is a life-history trait characteristic of tree species from ecotones between contrasted climatic zones, such as those across Temperate and Mediterranean Europe^{8,9}. Leaf retention was firstly

¹CIBIO (Research Center in Biodiversity and Genetic Resources) - InBIO (Research Network in Biodiversity and Evolutionary Biology), University of Porto, Campus Agrário de Vairão, Rua Padre Armando Quintas, 4485-661 Vairão, Portugal. ²MHNC-UP - Museum of Natural History and Science of the University of Porto - PO Herbarium, University of Porto, Praça Gomes Teixeira, 4099-002 Porto, Portugal. ³Biology Department, Faculty of Sciences, University of Porto, Rua Do Campo Alegre, s/n, 4169-007 Porto, Portugal. ⁴Agricultural Research Centre, Finca La Orden, Valdesequera, CICYTEX - Centro de Investigaciones Científicas Y Tecnológicas de Extremadura, Ctra. A-V, Km 372, 06187 Guadajira, Badajoz, Spain. ⁵Department of Evolution, Ecology and Behaviour, Institute of Integrative Biology (IIB), University of Liverpool, Bioscience Building, Liverpool L69 7ZB, UK. ✉email: cvv@cibio.up.pt

interpreted as an adaptation to cold climates that would allow a direct uptake of soil minerals back to the tree¹⁰. Further, marcescence proved to be adaptive in areas with summer drought and winter frosts by protecting leaf buds from desiccation and extreme cold¹¹. Despite the unclear adaptive value of marcescence, leaf retention is a trait highly responsive to environmental conditions that control distributional range fluctuations in response to climate changes¹².

Previous biogeographic studies in the Iberian Peninsula, a well-known hotspot of European and global biodiversity, confine the distribution of marcescent woodlands to the submediterranean bioclimatic belt, where these forests are fed by moderate summer precipitation, at least during one month of the dry season^{2,8,11}. Such mild summers allow marcescent forests to host warm-temperate and even sub-tropical species¹³, acting as refugia for paleoclimatic relicts, including rare taxa that are otherwise common in the Atlantic islands¹⁴. Model projections show that future climatic changes will likely hard-hit marcescent forests¹¹ and therefore, anticipating distributional range shifts of marcescent species in response to future climate changes is an urgent research goal. Tracking past range shifts can also improve our ability to forecast future changes. In this sense, the severe climatic conditions exhibited during cold periods in the late Pleistocene triggered massive migrations of the European flora towards the southern peninsulas, where they become confined. The Last Glacial Maximum (LGM, ca. 20 Ky) was one of the most critical periods of the Quaternary for forest persistence because of their low temperatures and water availability¹⁵. Then, during the Bolling-Allerod period, the ice sheets retreated and the onset of the temperate and moister climate and mild winters triggered the expansion of Temperate and Mediterranean forest at the start of the late glacial (13 to 10 Ky), interrupted during the Younger Dryas (12.9 to 11.7 Ky). These forests recovered and achieved their maximum extent at the Early to Middle Holocene (ca. 6 Ky) depending on the geographical location¹⁶.

The magnitude of the distribution range shifts that marcescent forests have undergone in response to past climate changes is still poorly understood. We currently ignore the biogeographic trajectories that co-occurring oak species inhabiting this ecotone experienced in response to climate changes during the Late Quaternary. This is partly because most studies aimed to delineate past distributional ranges are based on fossil pollen records that provide a broad depiction of plant communities but fail to distinguish between deciduous oak species^{17,18}. Recent approaches based on species distribution modeling (SDM) can provide new insights to advance our understanding of the biogeographic past trajectories^{19–21}. This is especially important to transiently co-occurring oak species at the boundary of two biogeographic regions in Southern Europe. This study aims to elucidate the late-Quaternary dynamics of the Temperate-Mediterranean transition in southern Europe by applying correlative SDMs to hindcast the past distribution of eighth oaks (seven marcescent and one deciduous) distributed across the Iberian Peninsula as model taxa. Specifically, we:

- (i) Applied correlative species distribution models to identify the main climatic factors determining the current distribution ranges for each studied species;
- (ii) Hindcasted projections of each species' past distribution and quantify distribution range shifts in response to climate changes;
- (iii) Tracked the distribution shifts of oaks distributed across the submediterranean belt in the Iberian Peninsula based on the SDMs, and compare these shifts to those observed for temperate and mediterranean oak species; and
- (iv) Compared current and past distributional ranges of parental species with the known occurrence of hybrids to infer ancient contact zones and the persistence or turn-over of hybrid-species.

We further discuss the implications of our results for preserving the biodiversity of marcescent forests that occupy the Temperate-Mediterranean transition in southern Europe in the face of projected climate change.

Results

Model performance and selection of environmental predictors. Selected models showed high performance scores for all species (Table 1), especially for *Quercus robur*, *Q. canariensis* and their hybrids. Model performance values (for details see Material and Methods section: *Modelling approach, model fitting and evaluation*) show that the lowest TSS values (True skill statistic bounded between [-1, 1], with values closer to one depicting better models) corresponded to *Q. faginea* (0.79) and *Q. broteroi* (0.77) up to *Q. robur* (0.98) which shows overall good to excellent model performance. The AUC (Area Under the Curve: [0, 1], values closer to one depict better models) were always above 0.96, also suggesting an excellent performance of the models. In addition, sensitivity and specificity scores (bounded between [0, 100]) were always above 85, confirming good and well-balanced performance of the models.

Bioclimatic predictors showed the highest importance scores when compared to topographic ones (Supplementary Fig. S1). Exceptions were found for *Q. faginea* and *Q. canariensis*, for which the average topographic ruggedness index (TRI_AVG) was the most important variable. The contribution of variables related to the precipitation regime contributed the most to explain the present distribution of the focal species (Supplementary Fig. S1), especially the precipitation of the warmest quarter (BIO_18), the annual precipitation (BIO_12), and precipitation seasonality (BIO_15). Winter cold temperatures also held substantial explanatory power, especially the Annual mean temperature (BIO_01), the Mean Temperature of Coldest Quarter (BIO_11), and Isothermality (BIO_03).

Changes in predicted distributions of marcescent oak species and their hybrids. When taking into account all taxa past-present dynamics (Fig. 3a), we observed an average loss of suitable areas from the LGM to the present of ca. 25%, offset by an overall gains of ca. 38% while about 37% of the range remained stable. From

Species	TSS	AUC	Sensitivity	Specificity
<i>Quercus robur</i>	0.98	1.00	98.88	99.09
<i>Quercus canariensis</i>	0.94	0.99	100.00	94.01
<i>Quercus × coutinhoi</i>	0.93	0.99	97.67	95.53
<i>Quercus × marianica</i>	0.91	0.99	97.67	93.34
<i>Quercus lusitanica</i>	0.90	0.99	95.65	94.72
<i>Quercus estremadurensis</i>	0.88	0.98	95.52	92.66
<i>Quercus faginea</i>	0.79	0.97	85.71	92.87
<i>Quercus broteroi</i>	0.77	0.96	89.20	87.50

Table 1. Evaluation scores per species for the ensemble of 21 top-models (TSS: True skill statistic; AUC: Area Under the Curve; Sensitivity also called true positive rate or recall; Specificity also defined as the true negative rate).

the MH to the present, all selected oak species maintained 60% of their area, with ca. 16% of gains and ca. 23% of losses in terms of suitable area (Figs. 1, 2 and 3).

Considering taxonomic groups, we observed that subsection *Galliferae* species expanded 43% of their overall area from LGM to MH, with losses of ca. 23% and 34% of their range remaining stable (Figs. 2 and 4). Since the MH to the present, the range of these species remained stable for the most part (58%) (Figs. 1, 2, 3 and 4). Regarding roburoid oaks, they showed limited latitudinal shifts, while the overall extension of their distribution area remained generally stable (Figs. 2 and 3).

Regarding individual species, *Quercus broteroi* had an expansion of 52% from LGM to the present, expanding from southern basins and mountains, towards the Central Iberian Peninsula and the Northern basins. This species showed the highest stability in its distribution range from MH to the present (76%) (Figs. 2 and 3). In contrast, the distribution ranges of *Q. canariensis* contracted from LGM to the present (51%) with a fast-pace contraction from MH to present. Yet, this species experienced a notorious expansion of 39% towards the north-eastern part of the Iberian Peninsula from LGM to MH (Figs. 1 and 3). *Q. faginea* had the highest range expansion from LGM to MH (70%), mainly from Southern and Eastern Iberian Peninsula to the Northern Mountains, Central System and both Douro and Tagus basins, while 50% of the distribution area remained generally stable from MH to the present. The distribution range of *Q. lusitanica* contracted 45% from LGM to MH followed by an expansion of 44% from MH to present (Figs. 1 and 3) when the species did not contract its distribution area. The distribution of *Q. estremadurensis* expanded (46%) from LGM to MH but ca. 40% of peripheral locations disappeared from MH to present. Finally, the distribution of *Q. robur* remained stable, particularly from the MH to the present (65%) (Figs. 2 and 3).

As for hybrids, changes in the distribution range of *Q. × marianica* were very similar to those observed for the parental *Q. broteroi* (Fig. 2). Changes in the distribution of *Q. × coutinhoi* resembled those observed for *Q. estremadurensis* expanding ca. 40% northwards and contracting ca. 30% in the southern areas (Figs. 2 and 3).

Complementarily, the hierarchical clustering analysis, reflects the similarity of species projections in the three studied periods (Supplementary Fig. S3).

Identification of contact zones. The contact zone between species of the two major groups (Roburoid oaks – Section *Quercus* vs. *Gall* oaks – Subsection *Galliferae*) was located across the Tagus basin in the LGM (with ca. 11% of overlap between distribution areas) and it expanded North in the MH, followed by a contraction of its area (ca.20%) towards the present (Figs. 3 and 5). The current distribution of *Q. × marianica* tracks the distribution of its parental species (*Q. canariensis* and *Q. broteroi*) that were predicted to co-exist during the LGM and MH periods mainly across southern and western areas (Figs. 2 and 5).

The distribution areas of *Q. broteroi* and *Q. faginea* showed a residual overlap (3%) in the LGM located in southern Spanish mountains. This contact zone expanded in the MH to reach 12% from the expansion of *Q. faginea* towards North, which originated a second contact zone in the Douro basin (Figs. 1 and 5). The distribution of the two roburoid oaks (*Q. robur* and *Q. estremadurensis*) also overlapped in the central and northern regions of Portugal. The area of this contact zone increased from 9 to 12% from LGM to the MH period and it remains generally stable since then (Figs. 2 and 5).

Discussion

Our results reveal substantial distribution range shifts across Iberian Peninsula during the latest 20 ky for all studied species, suggesting a rather dynamic transition zone between two major biogeographic regions in Europe (Eurosiberian and Mediterranean). Overall, changes in precipitation and temperature during the late Quaternary were the main drivers of inferred distributional ranges. This also implies that future projected changes in precipitation regime and water balance will likely have strong effects on species ranges, thus calling for improvements in regional and local scenarios based on global climate models.

We complemented previous projections on Iberian trees [22] and other studies that underline a broad association of marcescent forests with the submediterranean bioclimatic belt [2,8,11]. In fact, while the distribution range obtained for the deciduous *Quercus robur* closely tracked the Eurosiberian region, the distribution ranges obtained for marcescent species (*Q. broteroi*, *Q. canariensis*, *Q. faginea* and *Q. marianica*) broadly corresponded to this submediterranean transition zone (Fig. 4).

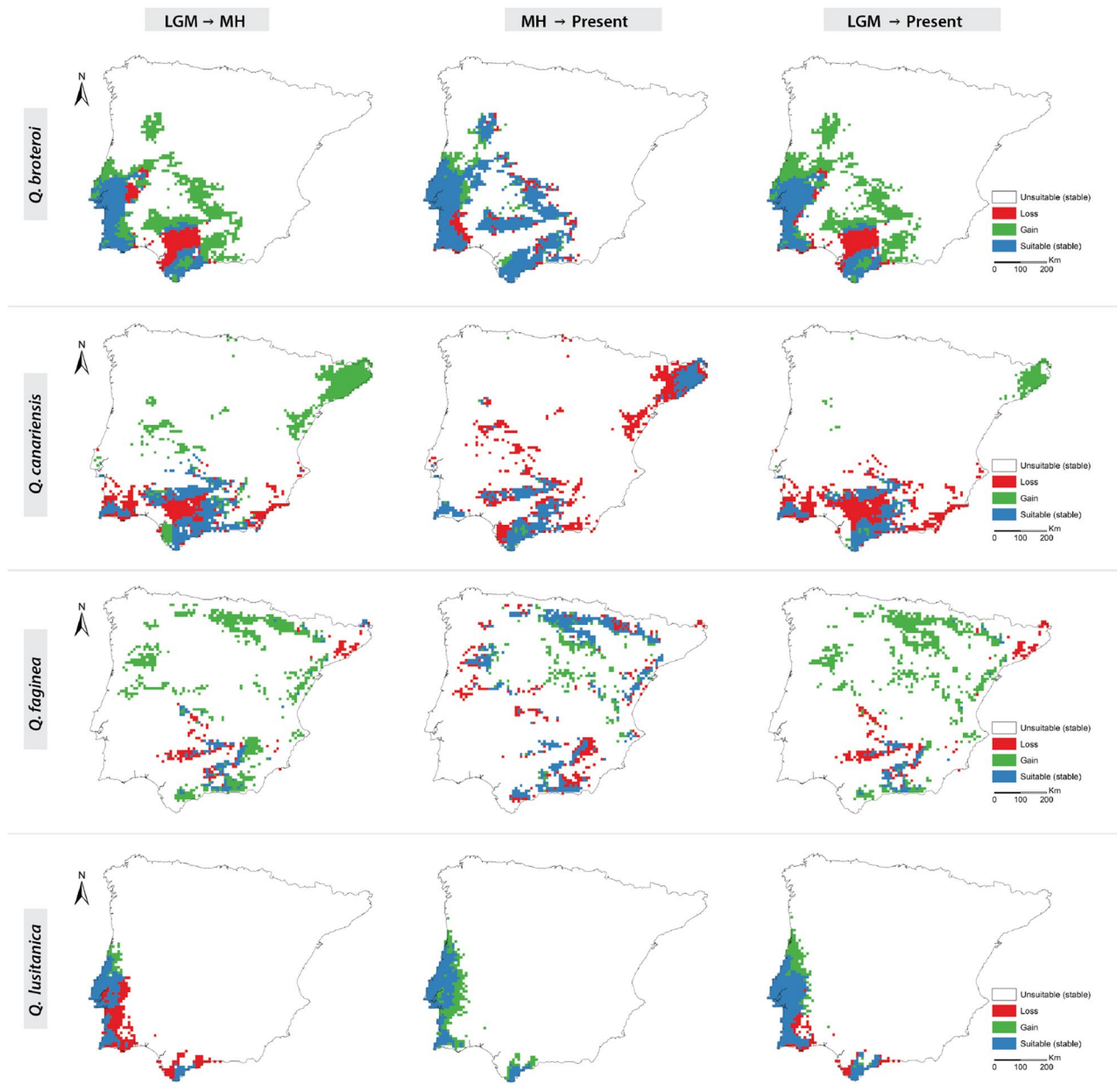


Figure 1. Changes in predicted distributions (gained, lost, or stable) for *Galliferae* oak species for the studied periods: LGM-MH (left); LGM-Present (center); MH-Present (right). Maps were generated by JG and CVV in R v.4.0.3 (<https://www.r-project.org>) and assembled in ArcMap 10.5 (<https://www.esri.com/en-us/arcgis>).

Our models also revealed that the distribution range of marcescent forests included the western coast of the Iberian Peninsula, an area with frequent advection fogs during summer that might compensate low annual precipitation values [11], thus favoring the expansion and establishment of marcescent forests. This overlooked source humidity has led previous biogeographic models to characterize *Q. broteroi* and *Q. canariensis* as thermophilic taxa inhabiting dry conditions [8]. On the contrary, they are maintained by frequent summer-fog regimes^{23,24} that increase annual and summer precipitation and confer a submediterranean character to these areas. Also, this climatic and topographic heterogeneities favoured by small mountain ranges play a crucial role as identified across major refugia areas throughout the Mediterranean Basin^{25,26}.

The geographic patterns obtained for the different study periods are consistent with the known biogeography and dynamics of Iberian forests since the last glacial maximum (LGM)^{16,27}. Overall, climatic hardness and increased annual precipitation in Southwest Iberian Peninsula during the LGM allowed the persistence of most marcescent oaks in the southern range limits of their present distribution. The distributional segregation between marcescent oaks and the deciduous *Q. robur* (Figs. 1, 2 and Supplementary Fig. S3) is evidenced by the range contraction of the latter between the LGM and MH, from the inland towards the Atlantic and Cantabrian coasts. This coastal expansion follows an increase in winter minimum temperatures, while the inland retreats are

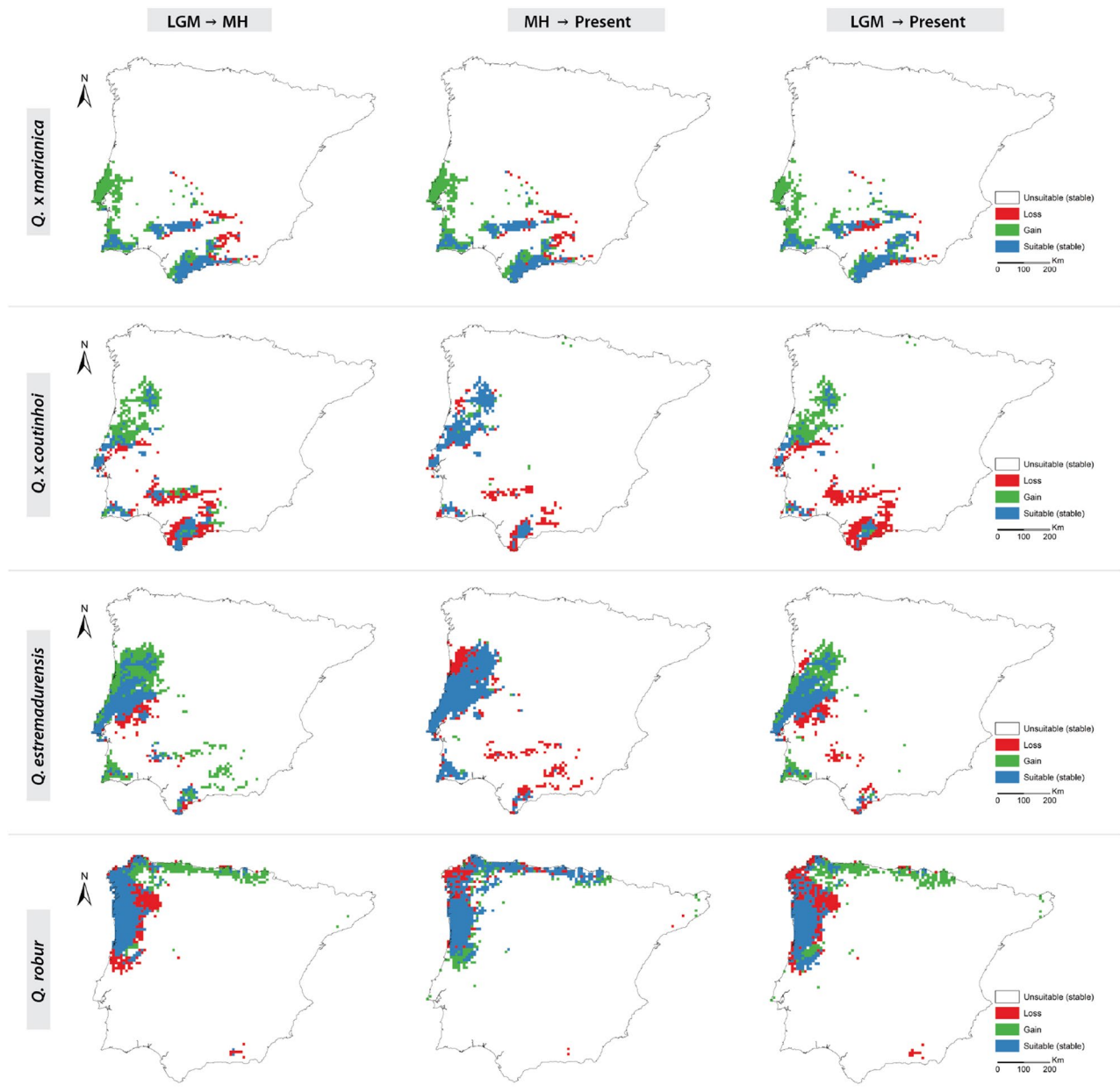


Figure 2. Changes in the predicted distributions (gained, lost, or stable) of hybrids (*Q. x marianica* and *Q. x coutinhoi*) and Roburoid oaks (*Q. estremadurensis* and *Q. robur*) for the studied periods: LGM-MH (left); MH-Present (center); LGM-Present (right). Maps were generated by JG and CVV in R v.4.0.3 (<https://www.r-project.org>) and assembled in ArcMap 10.5 (<https://www.esri.com/en-us/arcgis>).

a response to increasing summer drought (i.e. more Mediterranean climate). Other roburoid oaks (*Q. estremadurensis* and *Q. x coutinhoi*) and *Q. lusitanica* follow the same general pattern.

The climatic amelioration during the MH explains most of the expansion of marcescent oaks towards Center and North Iberian Peninsula, followed by a slight contraction as the climate became drier towards the present. The present distribution of *Quercus faginea* approaches the roburoid taxa, unlike the remaining *Galliferae* oaks (Fig. 4 and Supplementary Fig. S3). This is related with this species recent expansion through the Eurosiberian region, where it forms secondary forests in locally drier biotopes, in areas where the more favorable soils are occupied by deciduous primary forests of *Fagus sylvatica*, *Q. petraea* and *Q. pubescens*^{28,29}. Terrain ruggedness followed by summer precipitation plays a major role on *Q. faginea* and *Q. canariensis* distributions (Fig. 1 and Supplementary Fig. S1). Both are associated with increased summer precipitation, with the first expanding through interior mountainous areas, which is confirmed by the influence of precipitation and temperature seasonality, while *Q. canariensis* distribution is determined by a summer fog regime both in coastal and inland areas (Fig. 1 and Supplementary Fig. S1). The less fragmented distribution of *Q. canariensis* during the LGM, across southern Iberia, discloses an ancient contact between nowadays relict populations. This contact is supported by

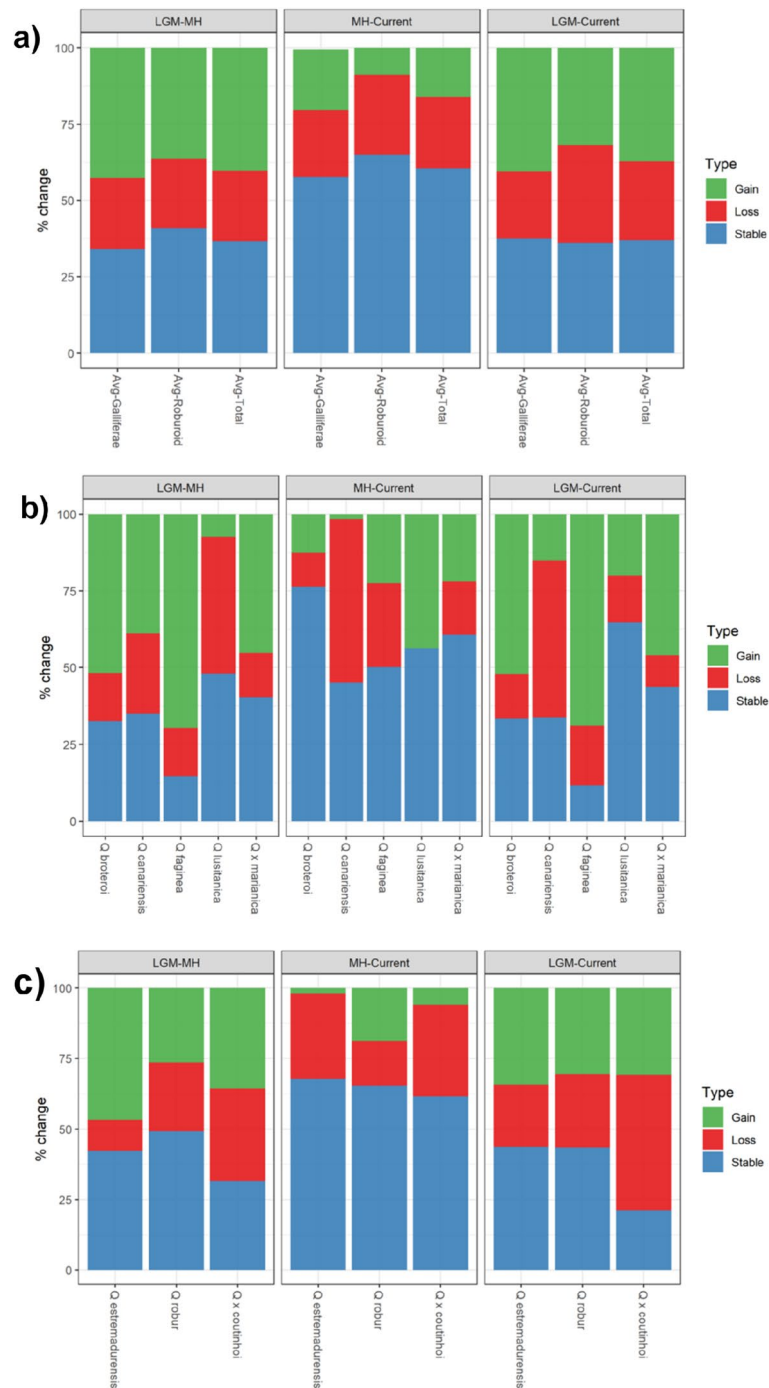


Figure 3. Percentage of change (% change) of the predicted distributions (gained, lost, or stable) in the focal time frames: LGM-MH (left), MH-Current (center) and LGM-Current (right) (a) Broad taxonomic groups (*Galliferae*, Roburoid, and Total); (b) Individual species within the *Galliferae* group; (c) Individual species within the Roburoid group and the hybrid *Q. x coutinhoi*. Graphics were generated by JG and CVV in R v.4.0.3 (<https://www.r-project.org>).

the current presence of vicariant assemblages of understory species, including relict taxa like *Frangula baetica*, *Cytisus baeticus*, *Davallia canariensis*, *Rhododendron ponticum* and *Myrica* spp.³⁰

Our results support ancient, but not current, sympatry between parental species during the LGM and MH at locations where their hybrids are currently present (Fig. 5). The distribution of *Q. x maritima* tracks the distribution of *Q. canariensis* in the LGM but it progressively tracks the distribution range of its other parental species (*Q. broteroi*) during the MH and towards the present (Fig. 5 and Supplementary Fig. S3). This shows both adaptation towards drier conditions in face of *Q. canariensis* and expansion towards equilibrium considering the

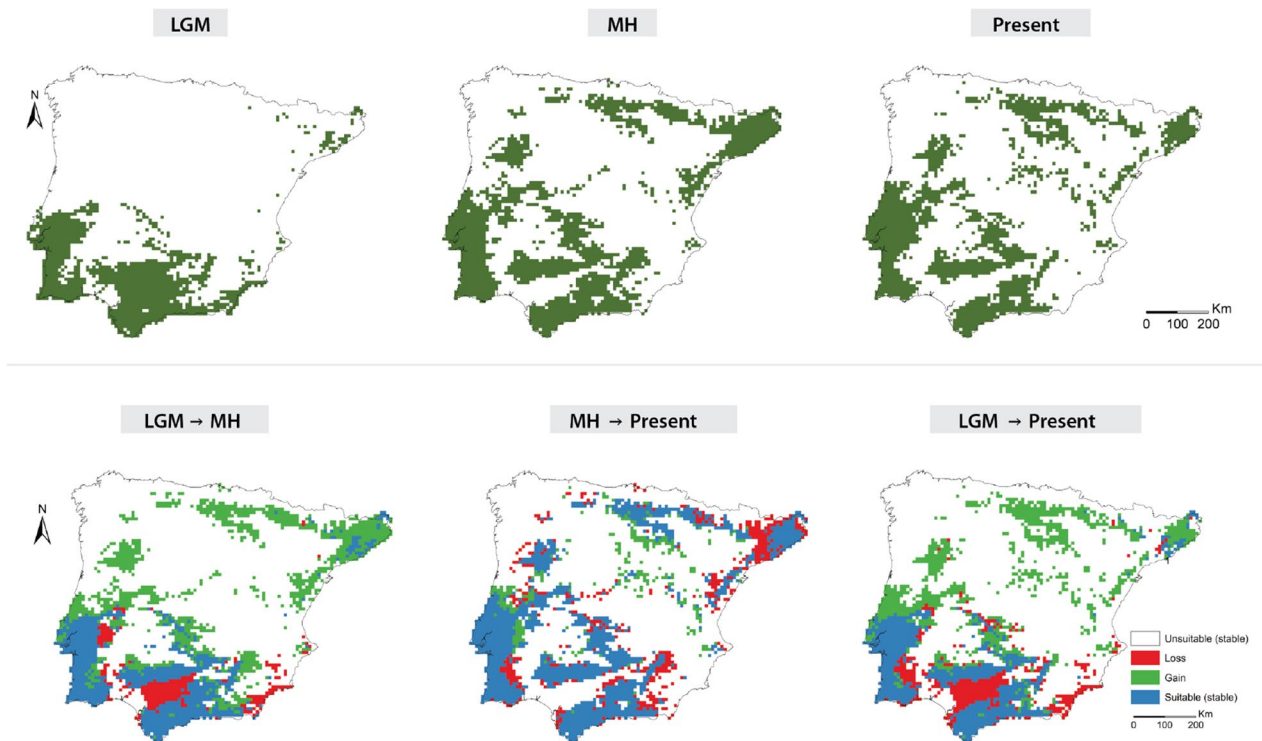


Figure 4. Predicted distribution of the submediterranean belt in Iberian Peninsula, based in the distribution of the strictly marcescent oaks (*Q. broteroi*, *Q. canariensis*, *Q. faginea* and *Q. marianica*) (top row) and their distributional range changes across the three studied periods (bottom row). Maps were generated by JG and CVV in R v.4.0.3 (<https://www.r-project.org>) and assembled in ArcMap 10.5 (<https://www.esri.com/en-us/arcgis>).

parental niches. Similarly, the spatial distribution of roburoid (*Q. estremadurensis* and *Q. robur*) and *Galliferae* parental species (*Q. broteroi* and *Q. faginea*) closely resembles the current distribution of the hybrid *Q. × coutinhoi* towards northern areas (Figs. 2 and 5). Here, *Q. robur* may have shifted between submediterranean and temperate areas, benefitting from less severe summers and winters during the Holocene³¹, which may explain the detected hybrid swarms during fieldwork in the Douro Basin and Central Portugal (Fig. 6), in populations where roburoid oaks are absent.

The species-level segregation of *Q. broteroi* and *Q. faginea* is congruent with their exiguous overlap during the LGM, followed by the posterior expansion (MH) of *Q. faginea* to NW Iberia (Figs. 1 and 5). The distribution of *Quercus lusitanica* is generally closer to roburoid oaks, especially in the LGM, but approaching other *Galliferae* species towards the present (Fig. 4 and Supplementary Fig. S3).

The projected distribution of *Q. estremadurensis* (Fig. 1) is consistent with its thermophilic preferences and resembles the one obtained for *Q. canariensis* (Fig. 1). This strengthens the hypothesis of *Q. estremadurensis* being a Tertiary relict oak that evolved under a subtropical paleo-environment^{32–34}. Herbaria review performed by the authors confirms the presence of this roburoid oak in northern Africa where it co-exists with other relicts such as *Q. canariensis* and *Prunus lusitanica*. In contrast, the projected distribution of *Q. estremadurensis* is opposed to the well-established Eurosiberian distribution of *Q. robur*³⁵. This projection dismisses any southern Iberian references of *Q. robur* enhancing the taxonomic dichotomy between these two roburoid oaks. Furthermore, the projected distribution of *Q. estremadurensis* in southwest Spain, emphasizes the role of Guadalquivir basin as a “dry” barrier for its expansion across southern Iberia (Figs. 2, 5 and 6).

Conclusions

We provided insights into the late-Quaternary distribution range shifts of a set of oak species in response to past climate changes. By combining hindcast and current projections obtained from species distribution models, we provide novel insights to disentangle taxonomic uncertainties. The performance of the models and the major importance of climatic variables shows that these oaks are rather suitable model species to undergo paleo-environmental and biogeographic studies along transitional bioclimatic and biogeographic areas in southern Europe. Further, our methodology proved valid to infer the distributional range shifts of hybrid taxa that showed to be congruent with the biogeographic trajectories of their parental species. This is quite important in a group with low interspecific barriers to gene flow and, therefore, characterized by complex evolutionary and speciation patterns^{36,37}. Our work addresses a major gap in paleo-vegetation studies that typically rely on palinological data, which tend to fail in delimitating the distribution of tree species accurately, particularly in the case of deciduous oaks (Sect. *Quercus*)^{17,18}. Our study suggests that conservation efforts should target *Q. canariensis* and *Q.*

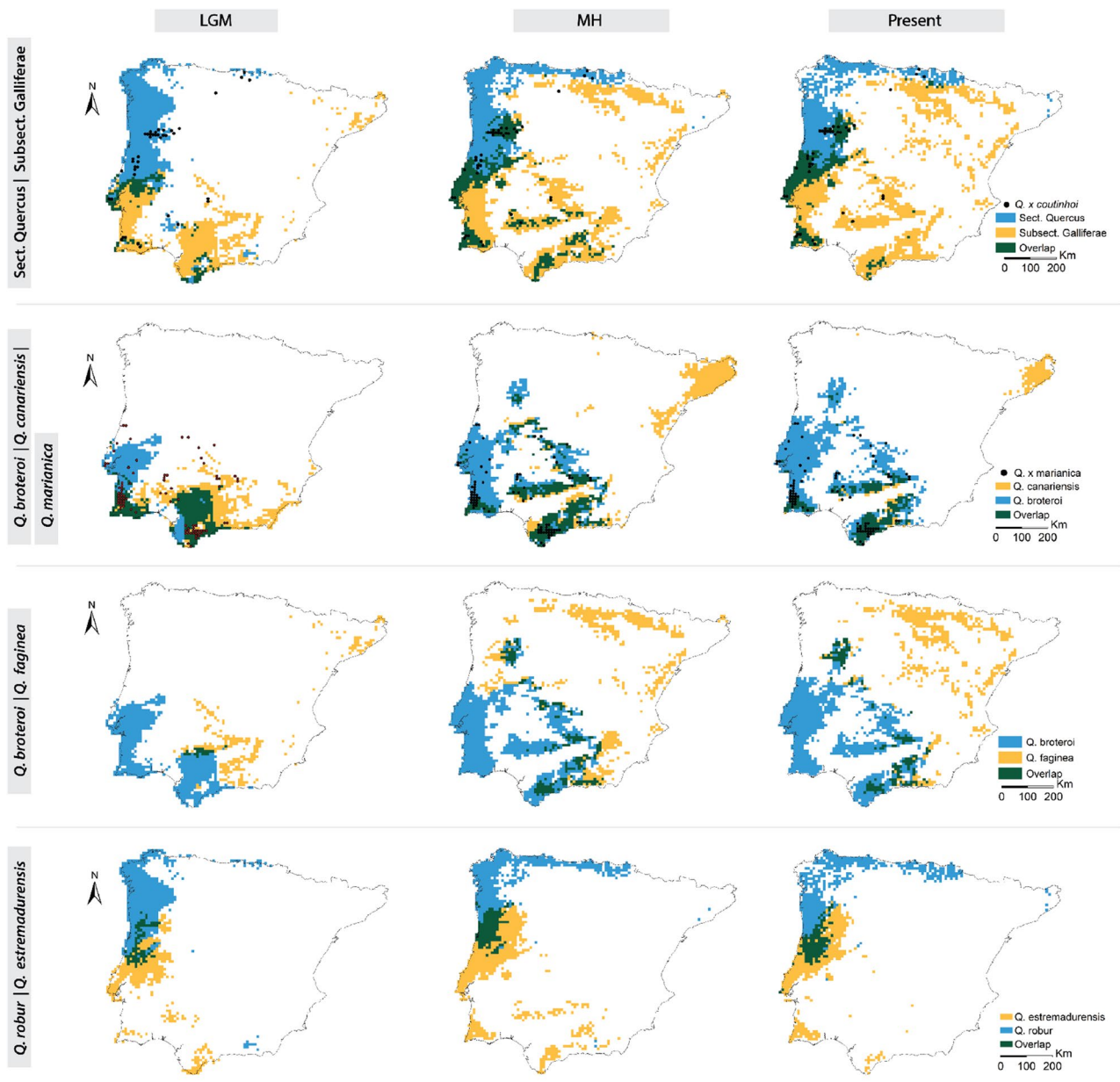


Figure 5. Distribution of the contact zones (green) of main targeted taxa across time periods. LGM (left), MH (center), and the present (right), and current distribution of hybrids ($Q. \times coutinhoi$ and $Q. \times marianica$). Sect. *Quercus* refers to the roburoid oaks ($Q. robur$ and $Q. estremadurensis$) and Subsection *Galliferae* refers to the Gall oaks ($Q. broteroi$ and $Q. faginea$) (Colours were chosen in way to enhance visual discrimination of different groups and their overlap). Maps were generated by JG and CVV in R v.4.0.3 (<https://www.r-project.org>) and assembled in ArcMap 10.5 (<https://www.esri.com/en-us/arcgis>).

estremadurensis because both taxa presently show a scattered (known and projected) distribution. In this regard, species distribution models prove useful to forecast the responses of these species to future climatic scenarios will contribute to prioritize conservation efforts.

Materials and methods

Study area. The Iberian Peninsula, located in southwestern Europe and with a total area of 583,832 km² (Fig. 6), holds a highly diverse geological history³⁸ and a well-documented late Holocene climatic influence³⁹. The present climate ranges from dry Mediterranean in the Southeast to wet Temperate Atlantic in the North⁷. The former is considered to have a Submediterranean variant that represents 65.64% of its total area in the peninsula, where at least in one summer month, the average of rainfall in millimeters is 2.8 times lower than the average temperature. Under these conditions marcescent forests constitute the natural potential vegetation^{7,8,11}. Due to its high edaphoclimatic heterogeneity, the Iberian Peninsula is a major hotspot of European plant diver-

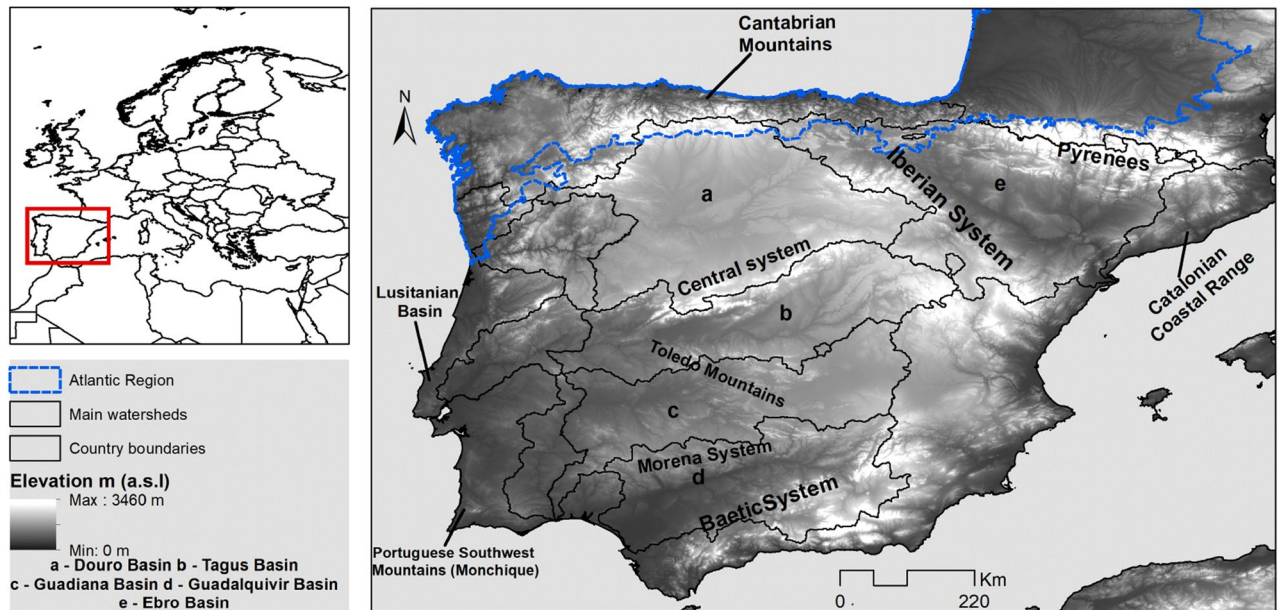


Figure 6. Study area (Iberian Peninsula) in the European context (left) and with the boundary of the Atlantic region highlighted in blue (right) and major river basins and mountain systems. Source: EEA. (European Environment Agency, Copenhagen, Denmark) 2012.

sity, hosting 54% of all European species, with 22.7% percentage of endemism, and a remarkable number of 21 endemic genera⁴⁰.

Focal taxa and occurrence data. The European white oaks (Section *Quercus*)⁴¹ include a group traditionally segregated as a subsection (*Galliferae* (Spach) Guerke)^{9,42} that contains a group of oak species, considered strictly marcescent, with a broad (though disjoint) distribution across the Mediterranean Basin, from the Middle East to the Western Mediterranean region⁹. Five taxa belonging to this group (*Quercus broteroi*, *Q. canariensis*, *Q. faginea*, *Q. lusitanica*, including one hybrid *Q. × marianica*) plus two roburoid oaks (*Q. estremadurensis* and *Q. robur*) and the hybrid between them and the previous group (*Q. × coutinhoi*) were selected as focal taxa totaling eight taxa (Supplementary Note S1, Fig. S2).

Occurrence data was gathered through extensive fieldwork in Western Iberian Peninsula (since 2005), plus literature and herbaria review (Supplementary Table S1). Specimens were assigned to each taxa based on the authors' experience and supported in relevant literature^{32,34,43–45}. To tackle taxonomic uncertainty, a thorough examination of more than 5000 vouchers of oak specimens was performed in 16 reference herbaria. This info was assessed and harmonized with other herbaria online collections and databases (Supplementary Information S1). Records for the target species (presence-only data) were aggregated to a 10 × 10 km grid (Supplementary Table S2).

Environmental variables (predictors). Model development was based on regional/coarse-scale factors potentially influencing the target species distribution, related to bioclimatic and topographic descriptors of the niche environmental space.

For bioclimatic data we used the WorldClim v.1.4 dataset (URL: <http://worldclim.org/version1>), containing 19 bioclimatic indices, available between for the period 1960–1990⁴⁶. For obtaining model-based hindcasts, we used WorldClim's downscaled paleoclimate reconstructions for the Last Glacial Maximum (LGM; ca. 22 Ky) and the Mid-Holocene (MH; ca. 6 Ky). WorldClim infers the mean values of bioclimatic variables of the LGM based on three Global Circulation Models (GCMs) whereas it uses nine GCMs to estimate those values for current distributions. Therefore we opted for calculating the average of all available models as to reduce uncertainties linked to each specific GCM⁴⁷. All bioclimatic data were later re-projected and re-sampled (average) to a common reference grid at 10 km spatial resolution (Datum WGS1984/UTM30N).

For portraying the 'static' influence of terrain morphology, wetness and its spatial heterogeneity on species distributions, we used the Topographic Ruggedness Index (TRI; as a proxy of slope and terrain complexity) and the Topographic Wetness Index (TWI; as a proxy of soil moisture and flow accumulation) both calculated from the SRTM (v.4) elevation data at 100 m of spatial resolution. For quantifying both variables, we aggregated/upscaled TRI and TWI values from their original resolution using the average (TRI_AVG) and the standard-deviation (TWI_STD).

We performed the pre-selection of bioclimatic variables based on preliminary models, using an exhaustive approach (similar to the one suggested by Cobos⁴⁸ which assessed all combinations of two variables for temperature and two variables for precipitation. This selection aimed to maximize per species model performance and parsimony (i.e. combinations with greater predictive performance where screened), balance variables linked

to both temperature and precipitation, assess and prevent prediction issues related to non-analogous climatic conditions, and take advantage of previous knowledge regarding each species ecological requirements. Finally, six environmental variables were used for model fitting per species (Supplementary Table S3).

Modelling approach, model fitting and evaluation. Species Distribution Models (SDMs) were developed using the *biomod2* R package, which applies a multi-technique ensemble forecasting approach to analyze species-environment relations and to obtain spatiotemporal predictions⁴⁹. Models were fitted using ten modelling techniques: GLM (Generalized Linear Models); GBM (Generalized Boosted Models); GAM (Generalized Additive Models); CTA (Classification Tree Analysis); ANN (Artificial Neural Networks); FDA (Flexible Discriminant Analysis); MARS (Multivariate Adaptive Regression Splines); RF (Random Forests); MAXENT.Phillips and MAXENT.Tsuruoka (Maximum Entropy Models) currently available in *biomod2*. Default parameters were used (with the exception of the smoothing degree term in GAM for *Q. canariensis*, which was set to $k = 2$ to avoid over-fitting issues⁵⁰). We only counted for presence data for focal species and therefore we obtained ten sets of randomly generated pseudo-absences (PA), each one with ten times the number of presences to increment the representativeness of the environmental space of the study area⁵¹. Since no previous information was assumed about species prevalence (p), model weights were adjusted to $p = 0.5$ (*biomod2* default) thus giving similar weight to presences and pseudo-absences. We also verified that all species (either narrow or widespread ones) had sufficient presence records to train models following guidelines in previous research^{52,53}. Holdout cross-validation was employed to evaluate the models, with 80% of the input records used for model fitting and 20% for model evaluation at each round. Twenty rounds were performed for model evaluation (plus one additional ‘full calibration’ round which uses all input records). For assessing model performance, the Area Under the Receiver-Operating Curve (AUC), the True-skill Statistic (TSS), and the Sensitivity and the Specificity values were calculated in *biomod2*⁴⁹.

Given that $N_T = 2100$ models were fitted per species (with $N_T = N_{techniques} \times N_{PA_sets} \times N_{evaluation_rounds}$), the less performant models were filtered out before the final ensemble forecasting. Hence, we selected the top 21 models (the top 1% percentile) considering the AUC rank. Based on these top-performing models, an ensemble using the average was implemented, thus reducing inter-model uncertainty. To transform the predicted probabilities into binary predictions of suitable/unsuitable areas, we used a threshold minimizing the straight-line distance between the receiver operating curve and the upper-left corner of the unit square⁵⁴. To assess the importance of each variable in model fitting, we used *biomod2*'s internal method that calculates $1 - \text{Pearson's correlation}$ between reference predictions and predictions for a ‘randomized’ version of each variable. The highest the score the greater is the influence of a variable in model predictions. A value of zero assumes no influence of a given variable. Variable importance scores were averaged across the top-performing models.

Evaluating species range shift and overlap across time. After fitting the models and obtaining ensemble predictions for present conditions, we used the LGM or MH paleoclimate reconstructions of WorldClim (see above) to obtain model hindcasts for these periods. Niche conservatism (i.e., “(...) retention of niche-related ecological traits over time.”⁵⁵) was assumed which holds that models are reasonably transferable and thus allowing to obtain spatiotemporal projections for past reference frames (see e.g.^{55–57}). Then, we performed a spatial overlap between present and past projections to evaluate shifts in species distributions. This allowed us to identify stable areas (that remained suitable or unsuitable between time periods), gains (areas that gained environmental suitability between periods) and losses (areas that lost environmental suitability between periods). Afterwards, we spatially stacked all species distribution maps and obtained a binary matrix by extracting and transposing map values, later used to compute the Sorensen distance (Sd). Based on the Sd distance matrix, we then performed the hierarchical cluster analysis using the complete linkage method for each time step⁵⁸. This allowed to identify differences and similarities of distributions between species through time.

Data availability

The authors agree on sharing this study data and its deposit in public repositories, upon reasonable request.

Received: 23 February 2020; Accepted: 26 November 2020

Published online: 09 December 2020

References

- Hufkens, K., Scheunders, P. & Ceulemans, R. Ecotones in vegetation ecology: Methodologies and definitions revisited. *Ecol. Res.* **24**, 977–986 (2009).
- De Dios, R. S., Benito-Garzón, M. & Sainz-Ollero, H. Hybrid zones between two European oaks: A plant community approach. *Plant Ecol.* **187**, 109–125 (2006).
- Anderson, E. & Stebbins, G. L. Jr. Hybridization as an evolutionary stimulus. *Evolution* **8**, 378–388 (1954).
- Arnold, M. L. Natural hybridization as an evolutionary process. *Annu. Rev. Ecol. Syst.* **23**, 237–261 (1992).
- Risser, P. G. Ecotones at local to regional scales from around the world. *Ecol. Appl.* **3**, 367–368 (1993).
- Weaver, J. E. & Albertson, F. W. *Grasslands of the Great Plains* (Johnsen Pub Co, Lincoln, 1956).
- Rivas-Martínez, S., Penas, A., del Río, S., González, T. & Rivas-Sáenz, S. Bioclimatology of the Iberian Peninsula and the Balearic Islands. In *The Vegetation of the Iberian Peninsula* (ed. Loidi, J.) 29–80 (Springer, Berlin, 2017).
- García-Mijangos, I., Campos, J. A., Biurrun, I., Herrera, M. & Loidi, J. Marcescent forests of the Iberian Peninsula floristic and climatic characterization. In *Warm-Temperate Deciduous Forests around the Northern Hemisphere* (eds Box, E. O. & Fujiwara, K.) 119–138 (Springer, Berlin, 2015).
- Tschan, G. F. & Denk, T. Trichome types, foliar indumentum and epicuticular wax in the Mediterranean gall oaks, *Quercus* subsection *Galliferae* (Fagaceae): Implications for taxonomy, ecology and evolution. *Bot. J. Linn. Soc.* **169**, 611–644 (2012).
- Otto, C. & Nilsson, L. M. Why do beech and oak trees retain leaves until spring?. *Oikos* **37**, 387–390 (1981).

11. De Dios, R. S., Benito-Garzón, M. & Sainz-Ollero, H. Present and future extension of the Iberian submediterranean territories as determined from the distribution of marcescent oaks. *Plant Ecol.* **204**, 189–205 (2009).
12. Greenwood, D. R. Leaf form and the reconstruction of past climates. *New Phytol.* **166**, 355–357 (2005).
13. Svenning, J. C. Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. *Ecol. Lett.* **6**, 646–653 (2003).
14. Rodríguez-Sánchez, F. & Arroyo, J. In *Climate change, ecology and systematics* Vol. 78 (eds T. R. Hodgkinson, M. B. Jones, S. Waldren, & J. A. Parnell) Ch. 13, 280–303 (Cambridge University Press, 2011).
15. Lowe, J. J. & Walker, M. J. *Reconstructing Quaternary Environments* (Longman, London, 1984).
16. Carrión, J. S. *et al.* Expected trends and surprises in the Lateglacial and Holocene vegetation history of the Iberian Peninsula and Balearic Islands. *Rev. Palaeobot. Palynol.* **162**, 458–475 (2010).
17. Muthreich, F., Zimmermann, B., Birks, H. J. B., Vila-Viçosa, C. M. & Seddon, A. W. R. Chemical variations in *Quercus* pollen as a tool for taxonomic identification: Implications for long-term ecological and biogeographical research. *J. Biogeogr.* <https://doi.org/10.1111/jbi.13817> (2020).
18. Tzedakis, P. Vegetation change through glacial—Interglacial cycles: A long pollen sequence perspective. *Phil. Trans. R. Soc. Lond. B* **345**, 403–432 (1994).
19. Svenning, J.-C., Flojgaard, C., Marske, K. A., Nógues-Bravo, D. & Normand, S. Applications of species distribution modeling to paleobiology. *Quatern. Sci. Rev.* **30**, 2930–2947 (2011).
20. Maguire, K. C., Nieto-Lugilde, D., Fitzpatrick, M. C., Williams, J. W. & Blois, J. L. Modeling species and community responses to past, present, and future episodes of climatic and ecological change. *Ann. Rev. Ecol. Evol. System.* **46**, 343–368 (2015).
21. Forester, B. R., DeChaine, E. G. & Bunn, A. G. Integrating ensemble species distribution modelling and statistical phylogeography to inform projections of climate change impacts on species distributions. *Divers. Distrib.* **19**, 1480–1495 (2013).
22. Benito-Garzón, M., Sánchez de Dios, R. & Sainz Ollero, H. Predictive modelling of tree species distributions on the Iberian Peninsula during the Last Glacial Maximum and Mid-Holocene. *Ecography* **30**, 120–134 (2007).
23. Ward, A. Fog at North Front, Gibraltar. *Meteorol. Mag* **81**, 272–279 (1952).
24. Wheeler, D. Factors governing sunshine in south-west Iberia: A review of Western Europe's sunniest region. *Weather* **56**, 189–197 (2001).
25. Médail, F. & Diadema, K. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *J. Biogeogr.* **36**, 1333–1345 (2009).
26. Walas, Ł. *et al.* Past, present, and future geographic range of an oro-Mediterranean Tertiary relict: The juniperus drupacea case study. *Reg. Environ. Change* **19**, 1507–1520 (2019).
27. Blanco-Castro, E. *et al.* *Los bosques ibéricos. Una interpretación geobotánica*, 597 (Editorial Planeta, Barcelona, Spain, 2005).
28. Díaz-González, T. E. & Penas, Á. The High Mountain Area of Northwestern Spain: The Cantabrian Range, the Galician-Leonese Mountains and the Bierzo Trench. In *The Vegetation of the Iberian Peninsula* (ed. Loidi, J.) 251–321 (Springer, Berlin, 2017).
29. Amigo, J., Rodríguez-Gutián, M. A., Honrado, J. J. P. & Alves, P. The Lowlands and Midlands of Northwestern Atlantic Iberia. In *The Vegetation of the Iberian Peninsula* (ed. Loidi, J.) 191–250 (Springer, Berlin, 2017).
30. Vila-Viçosa, C. *et al.* Syntaxonomic update on the relict groves of Mirbeck's oak (*Quercus canariensis* Willd. and *Q. marianica* C. Vicioso) in southern Iberia. *Plant Biosyst.* **149**, 512–526. <https://doi.org/10.1080/11263504.2015.1040484> (2015).
31. Mauri, A., Davis, B., Collins, P. & Kaplan, J. The climate of Europe during the Holocene: a gridded pollen-based reconstruction and its multi-proxy evaluation. *Quatern. Sci. Rev.* **112**, 109–127 (2015).
32. Vila-Viçosa, C., Vázquez, F. M., Meireles, C. & Pinto-Gomes, C. Taxonomic peculiarities of marcescent oaks (*Quercus*, Fagaceae) in southern Portugal. *Lazaroa* **35**, 139–153. https://doi.org/10.5209/rev_LAZA.2014.v35.42555 (2014).
33. Vicioso, C. *Revisión del Género Quercus en España* (Tipografía Artística, Caligrafía, 1950).
34. Vázquez, F. M. *et al.* Anotaciones a la nomenclatura del género *Quercus* L., (FAGACEAE) en la Península Ibérica y NW de África. *Folia Botánica Extremadurensis* **12**, 5–79 (2018).
35. Ülker, E. D., Tavşanoğlu, Ç. & Perktas, U. Ecological niche modelling of pedunculate oak (*Quercus robur*) supports the 'expansion–contraction' model of Pleistocene biogeography. *Biol. J. Lin. Soc.* **123**, 338–347 (2017).
36. Hipp, A. L. Should hybridization make us skeptical of the oak phylogeny. *Int. Oak J.* **26**, 9–18 (2015).
37. McVay, J. D., Hipp, A. L. & Manos, P. S. A genetic legacy of introgression confounds phylogeny and biogeography in oaks. *Proc. R. Soc. B* **284**, 20170300 (2017).
38. Loidi, J. Introduction to the Iberian Peninsula, General Features: Geography, Geology, Name, Brief History, Land Use and Conservation. In *The Vegetation of the Iberian Peninsula* (ed. Loidi, J.) 3–27 (Springer, Berlin, 2017).
39. Rodríguez-Sánchez, F., Hampe, A., Jordano, P. & Arroyo, J. Past tree range dynamics in the Iberian Peninsula inferred through phylogeography and palaeodistribution modelling: a review. *Rev. Palaeobot. Palynol.* **162**, 507–521 (2010).
40. Aedo, C., Buiria, A., Medina, L. & Fernández-Albert, M. The Iberian vascular Flora: Richness, endemism and distribution patterns. In *The Vegetation of the Iberian Peninsula* (ed. Loidi, J.) 101–130 (Springer, Berlin, 2017).
41. Denk, T., Grimm, G. W., Manos, P. S., Deng, M. & Hipp, A. L. An updated infrageneric classification of the oaks: Review of previous taxonomic schemes and synthesis of evolutionary patterns. *bioRxiv* <https://doi.org/10.1101/168146> (2017).
42. Gürke, M. In *Plantae Europaeae [Enumeratio systematica et synonymica plantarum phanerogamicarum in Europa sponte crescentium vel mere inquilinarum]* Vol. 2 (ed K. Richter) Ch. XXIX, 54–72 (Engelmann, 1897).
43. Rivas-Martínez, S. & Saénz, C. Enumeración de los *Quercus* de la Península Ibérica. *Rivasgodaya* **6**, 101–110 (1991).
44. Vasconcellos, J. C. & Franco, J. A. Os *Carvalhos* de Portugal. *Anais do Instituto Superior de Agronomia* **21**, 1–135 (1954).
45. Schwarz, O. In *Flora Europaea* Vol. 1 (ed Tutin TG *et al.*) Ch. 36, 72–76 (Cambridge University Press, 1993).
46. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978. <https://doi.org/10.1002/joc.1276> (2005).
47. Goberville, E., Beaugrand, G., Hautekèete, N. C., Piquot, Y. & Luczak, C. Uncertainties in the projection of species distributions related to general circulation models. *Ecol. Evol.* **5**, 1100–1116 (2015).
48. Cobos, M. E., Peterson, A. T., Osorio-Olvera, L. & Jiménez-García, D. An exhaustive analysis of heuristic methods for variable selection in ecological niche modeling and species distribution modeling. *Ecol. Inform.* **53**, 100983 (2019).
49. Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M. B. BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography* **32**, 369–373 (2009).
50. Guisan, A., Edwards, T. C. Jr. & Hastie, T. Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecol. Model.* **157**, 89–100 (2002).
51. Barbet-Massin, M., Jiguet, F., Albert, C. H. & Thuiller, W. Selecting pseudo-absences for species distribution models: How, where and how many?. *Methods Ecol. Evol.* **3**, 327–338 (2012).
52. Elith, J., Graham, C., Anderson, R., Dudík, M. & Ferrier, S. Novel methods improve prediction of species' distribution models. *Ecography* **32**, 66–77 (2006).
53. van Proosdij, A. S., Sosef, M. S., Wieringa, J. J. & Raes, N. Minimum required number of specimen records to develop accurate species distribution models. *Ecography* **39**, 542–552 (2016).
54. Freeman, E. A. & Moisen, G. G. A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecol. Model.* **217**, 48–58. <https://doi.org/10.1016/j.ecolmodel.2008.05.015> (2008).
55. Wiens, J. J. *et al.* Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol. Lett.* **13**, 1310–1324 (2010).

56. Peterson, A. T. Ecological niche conservatism: A time-structured review of evidence. *J. Biogeogr.* **38**, 817–827 (2011).
 57. Wiens, J. J. & Graham, C. H. Niche conservatism: integrating evolution, ecology, and conservation biology. *Annu. Rev. Ecol. Evol. Syst.* **36**, 519–539 (2005).
 58. Boccad, D., Gillet, F. & Legendre, P. *Numerical Ecology with R* (Springer, Berlin, 2011).

Acknowledgements

We thanks to Sandra Gomes for the insights on paleoclimatic and paleopalynological information, that validated this work inferences on the paleovegetation of Iberian Peninsula.

Author contributions

C.V.V. and J.G. conceived the ideas; C.V.V. and F.M.V. contributed with fieldwork, herbaria review, taxonomic data filtering and curation; J.G. and C.V.V. developed the methodological and analytical framework; J.G., C.V.V., J.H., C.G. and A.L. interpreted the results; C.V.V. led the writing of the manuscript with assistance of J.G., J.H., C.G., A.L. and R.S.A.. All authors contributed and commented on the manuscript drafts.

Funding

CVV received support from the Portuguese Ministry of Education and Science and the European Social Fund, through the Portuguese Foundation of Science and Technology (FCT) (contract PD/BD/52607/2014, within the Doctoral Program in Biodiversity, Genetics & Evolution (BIODIV) from the Faculty of Sciences of Porto University. JG was supported by FCT Scientific Employment Stimulus 2018—CEECIND/02331/2017. AL was supported by national funds through FCT—Fundação para a Ciência e a Tecnologia, I.P., in the context of the Transitory Norm—DL57/2016/CP1440/CT0001. CG was funded by the Investigator FCT Programme (IF/01375/2012) and PLANTSHIFTS (PTDC/BIA-BIC/5233/2014—POCI-01-0145-FEDER-016817).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-78576-9>.

Correspondence and requests for materials should be addressed to C.V.-V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020