

Glacial refugia and mid-Holocene expansion delineate the current distribution of *Castanea sativa* in Europe

Roces-Díaz, Jose V.^{1,2,3*}, Jiménez-Alfaro, Borja^{4,5}, Chytrý, Milan⁶, Díaz-Varela, Emilio R⁷, Álvarez-Álvarez, Pedro³

1: Agroscope, Reckenholzstr. 191, CH-8046 Zurich (Switzerland)

2: CREAM, E-08193 Bellaterra (Cerdanyola del Vallès), Barcelona (Spain)

3: GIS-Forest Research Group, Department of Organisms and Systems Biology, University of Oviedo, E-33600 Mieres (Spain)

4: iDiv - German Centre for Integrative Biodiversity Research, Halle-Jena-Leipzig (Germany)

5: Geobotany and Botanical Garden, Institute of Biology, Martin Luther University, Am Kirchtor 1, D-06108 Halle (Saale) (Germany)

6: Department of Botany and Zoology, Masaryk University, Kotlářská 2, CZ-61137 Brno (Czech Republic)

7: ECOAGRASOC Research Group, Department of Plant Production and Engineering Projects, University of Santiago de Compostela. E-27002 Lugo (Spain)

*corresponding author José V. Rocés-Díaz (jvroces@gmail.com), GIS-Forest Research Group, Department of Organisms and Systems Biology, University of Oviedo. E-33600 Mieres (Spain)

Abstract

Areas of Quaternary refugia for tree species have been mainly delineated based on fossil records and phylogeography, but niche modelling can provide useful complementary information. Here we use niche modelling to test the main hypotheses about the past distribution of *Castanea sativa* for the Last Glacial Maximum (LGM) and the mid-Holocene in Europe. We computed distribution models for current climatic conditions using different methods, and projected them onto three climatic scenarios for the LGM and the mid-Holocene. The projections were validated with pollen and charcoal records. LGM refugia were suggested in the north of the Iberian, Italian and Balkan Peninsulas, and in northern Anatolia. The projections for the mid-Holocene indicated high climatic suitability and geographic expansion of the species range across southern Europe, including some areas where the species is nowadays considered as non-native. In general, our niche modelling results are consistent with the patterns proposed based on pollen and charcoal records, and partially also on phylogeographic information inferred from genetic data, suggesting that the most suitable areas for *C. sativa* were extended significantly during the mid-Holocene, but declined afterwards and lost connectivity. The projected patterns were compatible with existing palaeobotanical records of *C. sativa* and provide a spatially-explicit picture of its past distribution during last 22,000 years.

Keywords

Generalized Linear Models; Last Glacial Maximum; Mahalanobis distance; Maximum Entropy; mid-Holocene; Niche models; Species distribution modelling; Sweet chestnut

This is the author's version of a work that was accepted for publication in *Palaeogeography, palaeoclimatology, palaeoecology* (Ed. Elsevier). Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Rocés-Díaz, J.V. et al. "Glacial refugia and mid-Holocene expansion delineate the current distribution of *Castanea sativa* in Europe" in *Palaeogeography, palaeoclimatology, palaeoecology*, vol. 491 (Feb. 2018), p. 152-160. DOI 10.1016/j.palaeo.2017.12.004

1. Introduction

The geographical ranges of species are not static in time, showing expansions and contractions that depend on changes in climate variability and anthropic influence. Range contractions during the glaciations were characterized by species retreat to suitable areas commonly described as refugia (Stewart et al. 2010). Current species' distributions are therefore influenced by past climatic conditions (Svenning et al. 2015) and particularly by the drastic climatic events of the Quaternary (Bennett and Provan 2008). The idea of refugia is particularly relevant for understanding postglacial species dispersal from the areas where a given species had persisted (e.g. Nekola 1999; Médail and Diadema 2009; Keppel et al. 2012; Gavin et al. 2014). Knowledge of native distribution ranges and location of refugia of forest tree species is also important for the management, for example when defining conservation strategies or reforestation programmes (Krebs et al. 2004; Vesella et al. 2013; 2015). Although the classical concept of refugia generally refers to large areas with suitable macroclimatic conditions (macrorefugia), cryptic refugia (microrefugia) can occur in small areas with specific microclimate (Ashcroft 2010; Birks & Willis 2008; Rull 2009). Indeed, large regions traditionally considered as refugia are climatically complex and the past occurrences of relict species within them are difficult to assess (Gómez and Lunt 2007).

A remarkable example of climatic macro- and micro-refugia is the persistence of forest trees along the European Mediterranean region during the Last Glacial Maximum (LGM; approx. 18,000-22,000 yr BP, i.e. before present; Médail and Diadema 2009). Fossil pollen records have been the main source of information for the reconstruction of tree distributions after the LGM (Huntley and Birks 1983; Davis et al. 2003; 2013). In the last decades this information has been complemented with genetic and phylogeographical analyses (e.g. Rodríguez-Sánchez et al. 2010; Nieto-Feliner 2014) that developed hypotheses about possible connections of species lineages (e.g. Mattioni et al. 2013). More recently, species distribution models (SDMs) emerged as complementary tools for assessing the spatial extent of refugia (e.g. Svenning et al. 2011; Alba-Sánchez et al. 2010; Schorr et al. 2012; Vessella et al. 2015). SDMs were used to identify historically suitable areas for tree species in Central and Northern Europe and North Africa (Svenning et al. 2008; Alba-Sánchez et al. 2015), providing complementary evidence to fossil data (Tzedakis et al. 2013; Feurdean et al. 2014). Since all approaches based on pollen/fossil data, phylogeography or SDMs are subject to limitations, assessing areas of species refugia should integrate different sources of information (Gavin et al. 2014).

Castanea sativa Mill. (Fagaceae; sweet chestnut) is a tree species subject to controversial opinions about the history of its distribution in Europe. For a long time, it was supposed to have

been introduced in Western Europe by the Roman Empire from northern Turkey and the Caucasus (Zohary and Hopf 1988), where the species showed higher levels of genetic differentiation (Villani et al. 1994; 1999). However, genetic data from different European populations has suggested the existence of glacial refugia in Western Europe, assuming also a strong influence of humans on its historical distribution (Fineschi et al. 2000). Recent studies based on pollen records have reported the occurrence of *C. sativa* in Western Europe before the end of the Pleistocene (e.g. Muñoz-Sobrino et al. 2004; Kaltenrieder et al. 2009), also suggesting glacial refugia. Krebs et al. (2004) identified LGM refugia along the northern coast of Spain, in southern France and Italy, based on macrofossil and pollen data. Mattioni et al. (2008; 2013) found that the populations from the westernmost areas were genetically differentiated from those occurring in Greece and Turkey, while Fernández-Cruz and Fernández-López (2016) found a clear genetic differentiation among *C. sativa* populations from Western Europe. All these results reinforced the hypothesis of glacial refugia for *C. sativa* in Western Europe, in agreement with the general patterns of European biogeography that established the role of the three large Mediterranean peninsulas – Balkan, Italian and Iberian – as the key macrorefugia during the LGM (Médail and Diadema 2009).

Summarizing the pollen and genetic data, the most probable refugial areas for *C. sativa* during the LGM are the Transcaucasia, the Sea of Marmara area, the Italian Apennines and the North Atlantic and Cantabrian coasts of the Iberian Peninsula (Huntley and Birks 1983; Krebs et al. 2004; Mattioni et al. 2013). Moreover, it has been hypothesized that the shrink of the species' distribution during the LGM was followed by an expansion during the mid-Holocene (Davis et al. 2003). This period, also called the climatic optimum (Allen et al. 1996), was characterized by a warm and moist climate with the expansion of mixed deciduous forests in the Mediterranean basin and the European temperate zone (Huntley 1990). Mountain areas of Southern Europe above 1,000 m were probably colonized by *C. sativa* and other tree species with similar climatic requirements (Muñoz-Sobrino et al. 2005). In the Iberian Peninsula, Benito-Garzon (2006) suggested that climatically suitable areas for *C. sativa* during the mid-Holocene were more widespread than nowadays, implying a Late Holocene contraction of its distribution. In this area sweet chestnut showed an expansion derived from its cultivation since Roman ages, and especially during the Middle Age (Conedera et al. 2004). However, the spatial dimensions of past contractions and expansions of *C. sativa* are difficult to assess with pollen and genetic data at the continental scale. The missing spatially-explicit information about past distributions of tree species can be provided by species distribution models based on palaeoclimatic models (Svenning et al. 2008; 2011), which can also identify potential new refugia not detected by palaeobotanical records or genetic-derived information.

In this study, we use species distribution models to test major hypotheses about the past distribution patterns of *Castanea sativa*. We use the current distribution range of the species to compute SDMs based on the present-day climatic conditions and different modelling algorithms, projecting the models to different climatic scenarios for the LGM and the mid-Holocene, and comparing model projections with fossil pollen and charcoal records.

2. Material and Methods

2.1. Species and climate data

The study focuses on the Mediterranean Basin and adjacent regions: the south of the European Atlantic Region, the Black Sea region and the Middle East (Fig. 1A). The hypothetical distribution range of *Castanea sativa* was obtained from the European Atlas of Tree Species (Conedera et al. 2016) which defines the native and introduced species distribution ranges based on various data sources, including pollen and charcoal records. This map represents the most updated distribution of *C. sativa* at continental level. To develop a database of fossil records (mainly pollen and charcoal) of *C. sativa* in Europe, we followed the recommendations by Moreno-Amat et al. (2017) to firstly compile information from general sources, European Pollen Database (Fyfe et al. 2009; Davids et al. 2013; EPD 2017) and Neotoma database (Goring et al. 2015). Then we added information for the studied periods (LGM and mid-Holocene) from a series of more specific sources, as for example Krebs et al. (2004) who provided a detailed revision of past chestnut records, and Arambarri et al. (2016) who compiled several existing data for the north of the Iberian Peninsula. These data were used to provide a statistical validation of the SDMs calculated for the LGM and mid-Holocene periods (see *Model calibration and validation* section for more details). For the records of the EPD, the minimum percentage of chestnut pollen used in our database was 0.1%. Main part of the records used for validation showed at least 0.5%, although some records represented percentages between 0.01-0.5%. In the compiled data, 79 records were from the period 5,000-7,500 yr BP (mid-Holocene; calibrated years) and 27 records from the period 18,000-22,000 yr BP (Last Glacial Maximum; calibrated years); these records were used in the current study (Fig. 1B).

We collected climatic data from WorldClim at 2.5 minute resolution (www.worldclim.org; Hijmans et al. 2005), consisting of spatial extrapolations from climatic stations for the period 1950-2000. We computed Pearson's correlations to test for co-linearity among the 19 bioclimatic variables of WorldClim, and selected six of them with correlations < 0.75 and with an expected impact on the distribution of temperate trees. Mean of the daily minimum temperatures of the coldest month (TMIN) identifies winter-cold areas (Sanchez de Dios et al. 2009; Franklin et al.

2013); mean temperature of the warmest quarter (TWARMEST) defines areas with the warmest summer (Schorr et al. 2012; Franklin et al. 2013); precipitation of the driest month (PDRIEST) defines the areas with a marked drought (Sanchez de Dios et al. 2009; Schorr et al. 2012); precipitation of the wettest quarter (PWETTEST) is a useful predictor for distribution of species that require a large amount of water in some part of the year (Franklin et al. 2013); and seasonality of precipitation (PSEASON) and temperature (TSEASON) provide information about within-year variations (Svenning et al. 2008). Maps of the selected climatic variables are shown in Supplementary information (Fig. S1).

The same bioclimatic variables were also obtained from www.worldclim.org for the past conditions of the mid-Holocene and the Last Glacial Maximum, according to the palaeoclimatic scenarios generated by three different sources: the Max-Planck-Institute Earth System Model (MPI), the Community Climate System Model (CCSM) and the Model for Interdisciplinary Research on Climate (MIROC).

2.2. Modelling framework

To account for potential differences in models obtained with different methods, we tested three algorithms based on distinct statistical principles, namely regression, machine learning and profile methods (Hijmans and Elith, 2016). We used generalized linear models (GLM) as the standard procedure for modelling species distributions with presence and absence data. GLM was computed to transform the mean of the responses to a linear function of the predictor variables using a binary logistic function, with an equal number of absences and presences obtained from the native distribution range of *C. sativa*, and a variable selection procedure using the Backward Wald criterion in R (R Core Team 2015). We also used Maxent version 3.3.3k (Dudik et al. 2011) as a machine learning method that is widely used for modelling species distributions with presence data and a background of the study area (Phillips et al. 2006). Finally, we applied the Mahalanobis distance (Mahalanobis 1936) to select the most similar areas as reflected by the statistical distance to the occurrence data in the environmental space, as implemented in a Geographic Information System (Jenness 2003). This method is useful to complement the other two since it does not assume species response functions, therefore providing an independent assessment of climatic similarity between two areas or periods.

The three methods were first applied to the current climatic conditions for obtaining habitat suitability maps, using a minimum threshold of suitability that includes 75% of the occurrence points. The models were then projected to the mid-Holocene and LGM palaeoclimatic conditions using the three scenarios (MPI, CCSM and MIROC). Having three different outputs for each

period allows to detect the potential effect of different reconstructions, which is expected to be our major source of uncertainty. For a better visualization, the models computed for each scenario were transformed into binary maps (0-1) using the same threshold used for the predictions with contemporary climate, and combined to represent the areas of habitat suitability with agreement of one, two or three methods. These binary outputs are provided in Supplementary information (Figures S2 and S3). Since part of the study area was covered by ice (and permafrost) during the LGM (Tzedakis et al. 2013; and Krebs et al. 2004 specifically for *C. sativa*), the map projections for this period exclude areas under continental ice-sheets and mountain glaciers (Ehlers and Gibbard 2004).

2.3. Model calibration and validation

A set of points randomly distributed across the whole study area was generated as the source of presence and absence data using ArcGIS 10.2 (ESRI, Redlands, CA) and the Geospatial Modelling Environment software (Beyer 2012). To reduce the negative effect of spatial autocorrelation in SDMs (Dormann et al. 2007; Franklin 2009) these points were generated with a minimum distance of 20 km between each other. This distance was selected after testing the spatial autocorrelation with the Moran's I statistic for each environmental variable in the study area. For the six variables, the value of Moran's I was lower than 0.5 (its maximum value for a sample with a very high spatial autocorrelation was 1.0) and four of them showed z-scores lower than 1.96, so their distribution could be produced by random processes ($P > 0.05$). Thus we assumed that the sample points were spatially independent. This sample of randomly distributed points was intersected with the current distribution of *C. sativa* to obtain a total of 250 presences (located on the area where the species is considered to be native; Fig. 1), from which 190 were randomly selected as input data for model calibration and the remaining 60 were kept for validation.

Model validation was computed with the Area Under the Receiver Operating Characteristic (ROC) Curve (AUC, Fielding and Bell, 1997) as a standard method to evaluate model fitting (Elith and Leathwick 2009). In addition, two measures of predictive accuracy, sensitivity and specificity, were calculated using an independent sample of 120 points (60 presences and 60 absences). Sensitivity is the proportion of the observed presences that are predicted as presences, thus measuring the omission error. Specificity is the proportion of the observed absences that are predicted as absences, thus measuring the commission error (Allouche et al. 2006). For the validation of the LGM and mid-Holocene models, we also used independent data representing localities with dated fossil pollen and charcoal records of *C. sativa* as provided by previously mentioned data sources (Fyfe et al. 2009; Davids et al. 2013; EPD 2015; Goring et al. 2015; Krebs

et al. 2004; Arrambarri et al. 2016). In this case sensitivity was calculated but not specificity, because absence records are not available for the fossil data.

3. Results

3.1. Modelled distribution under current climatic conditions

Statistical validation suggested good model performance in the three modelling methods without large differences in AUCs (Table 1; from 0.750 for Mahalanobis to 0.916 for Maxent) or sensitivity (from 62% for GLM to 78% for Mahalanobis). However, specificity values showed a wider range (from 52% for Mahalanobis to 95% for Maxent). The two variables with a higher contribution were those related with climate seasonality (PSEASON) and precipitation during the wettest quarter of year (PWETTEST). In the GLM, three variables showed positive relations (TWARMEST, PDRIEST AND PWESTTEST) while the other three showed negative values. The Wald criterion values for variable selection and their percentage contributions are presented in the Supplementary Information (Table S1).

The three modelling methods provided similar spatial patterns, showing the areas with higher habitat suitability under current climate (Fig. 2) in the Atlantic region of France and Spain and in the northern and central parts of the Mediterranean peninsulas. The models predicted suitable areas beyond the known distribution of *C. sativa*, e.g. in some areas of northern Africa or Central Europe. There were differences among the models at finer scales. High suitability was also predicted along the western coast of the Balkan Peninsula where there are only isolated populations of *C. sativa*. Maxent was the only model that showed high suitability on the islands of Corsica and Sardinia and in southern Italy, where *C. sativa* is a common and widespread species. GLM predicted low suitability in the Sea of Marmara which is within the current distribution range of the species, whereas Maxent and Mahalanobis predicted higher values there.

3.2. Habitat suitability during the Last Glacial Maximum and the mid-Holocene

Habitat suitability predicted for the LGM scenarios (Fig. 3A and Fig. S1 in Supplementary information) differed from the models for current climatic conditions. The zones with suitable conditions were mainly located in the Mediterranean Basin, the Atlantic region from the British Isles to Portugal, and the southern and eastern Black Sea region. These models showed a high variation in sensitivity computed with the pollen records for this period (Table 2), ranging from 11.1% (Mahalanobis) to 85.2% (for the GLM) in both cases for MIROC scenario. The results were best for the GLM models (74.1-85.2% sensitivity) and less accurate for Mahalanobis (11.1-33.3%). Figure 3A shows the spatial distribution of the refugia areas according to the different

SDMs used. The MIROC models were the most restrictive while the MPI models suggested refugia in areas that were identified by the other models (e.g. close to the Atlantic coast). The areas with higher spatial overlap for the three methods were located in the Italian Peninsula, southern France, the western Balkan Peninsula and the southern and eastern coast of the Black Sea. Other areas, such as north-western Iberian Peninsula or central France, were identified as suitable by two methods.

The models for the mid-Holocene conditions (Fig. 3B) had higher sensitivity (between 59.5% and 81.0%) than the models for the LGM conditions, and the results based on the three scenarios were relatively similar. In general they showed an expansion of the suitable areas for *C. sativa* from the glacial refugia mainly to the nearby areas. Although the total amount of suitable areas did not show high increases (or decreases) from the LGM to the mid-Holocene (Table 2), there was a higher level of spatial agreement between the three methods used for the mid-Holocene (Fig. 3). In all cases, the suitable areas were located mainly in the Atlantic region of Europe, the Iberian, Italian and Balkan peninsulas, northern Turkey, Transcaucasia, southern half of France and Central Europe.

4. Discussion

4.1. Distribution of *Castanea sativa* in the LGM and the mid-Holocene

Our models show distinct patterns of potential distribution of *Castanea sativa* in periods of contrasting extreme climatic conditions. For the Last Glacial Maximum, different methods and climatic scenarios showed spatial agreement with pollen records (Fig. 3A), reinforcing the hypothesis of refugia for *C. sativa* in Western Europe and also delineating potential distribution of refugia at the continental scale. This spatial model is congruent with the expected patterns of Mediterranean glacial refugia (Médail and Diadema 2009) but also with other studies that have analyzed the past distribution of *C. sativa* in Europe (Krebs et al. 2004). The most relevant potential refugia identified in this study include (Fig. 3A; from west to east): i) the Cantabrian coast of Spain and nearby areas (in agreement with Mattioni et al. 2013; Fernández-Cruz and Fernández-López 2016; Benito-Garzón et al. 2006) and Western Iberia (middle zone of Tagus River valley, in agreement with Martin-Arroyo 1998); ii) southern France and the French Alps (Uzquiano 1992; Nakagawa 1998); iii) central and northern Italian Peninsula (Lowe et al. 1996; Paranelli and Miola 1991; Kaltenrieder et al. 2009; Mattioni et al. 2013); and iv) the Anatolian Peninsula, especially the Marmara Sea (Caner and Algan 2002) and the southern coast of the Black Sea (Euxinian biogeographic region). In contrast, some areas modelled as climatically suitable were not considered as glacial refugia in previous works, for example north-eastern Spain, an Atlantic corridor from north-western to south-western Spain, the northern half of France or southern Britain, and different islands in the Western Mediterranean Sea.

In the mid-Holocene, the European climatic conditions changed considerably as a result of a general increase in temperature until 7,000 years before present (Davis et al. 2003; Heiri et al. 2003). In Southern and Central Europe, this warming event was linked with the well-documented expansion of different species, especially deciduous trees (Huntley 1990; Brewer et al. 2002). Several studies using different approaches and focusing on different regions have suggested that the mid-Holocene was particularly favourable for the expansion of *C. sativa* and other species with similar ecological requirements (Ramil-Rego et al. 1998; Muñoz-Sobrino et al. 2005; Benito-Garzón et al. 2007; Kaltenrieder et al. 2010). However, the spatial dimension of this expansion at the continental scale has not been assessed until now. Besides supporting the idea of general expansion of *C. sativa* to the south in the mid-Holocene, our results suggest that this expansion may have occurred mainly in the southern half of Europe (southern Spain, Italy and Greece), This agrees with the fossil records dated for the mid-Holocene in areas where *C. sativa* is currently not present or its presence is scarce (Fig. 1). We highlight the occurrence of suitable areas for *C. sativa* in the Atlantic coast of France since the LGM and especially during the mid-Holocene (Fig. 3), which is in agreement with the high density of pollen records for this period, despite the species is not being considered as native in most of this region (Fig. 1). However, as *C. sativa* fruits were probably used in the diet of prehistoric humans, it cannot be excluded that humans, in addition to climate suitability and natural dispersal, may have already contributed (passively or actively) to its dispersal in the Early Holocene, as suggested for the dispersal of *Corylus avellana* by Mesolithic hunter-gatherers in Central Europe (Kuneš et al. 2008).

We note that the spatial patterns provided in this study are subjected to the limitations of palaeodistribution modelling in terms of uncertainties of predictors and projections to past conditions (Svenning et al. 2011; Janská et al. 2017). In addition, despite European distribution of *C. sativa* is partially driven by soil conditions (preferring well-drained and acidic to neutral soils; Conedera et al. 2016), this information was not included on our analysis. Although we have used an updated review of the native and non-native distribution of *C. sativa*, there is also uncertainty in the current native range of the species. For example, recent palaeobotanical records suggested a native distribution more extended for some specific areas (e.g., Central Iberian Peninsula in López-Sáez et al. 2017). In addition it should be highlighted that our models showed only those climatically suitable areas where the species could be present. We also note that other relevant drivers of dynamics of species ranges (such as species dispersal capacity, Svenning and Skov 2007; Lehsten et al. 2014; or species competence, Lurgi et al. 2015) were not included in our models since they are based exclusively on climate. Nevertheless, the combination of different methods and palaeoclimatic scenarios and the validation with independent fossil data allowed us to define the strengths and weaknesses of our modelling approach. In general, we found relatively

high spatial agreement among modelling methods and palaeoclimatic scenarios, with a weaker performance of Mahalanobis in the LGM, and warmer conditions predicted by the MPI scenario than by the MIROC scenario. Despite this, the spatial predictions supported the patterns suggested by previous studies based on fossil records (e.g. Krebs et al. 2004) and phylogeography (e.g. Mattioni et al. 2013). The spatial resolution (2.5 minutes, approximately 4.5 km at the equator) reflects an intermediate level between macro- and micro-refugia at the continental scale, at least under their general use for referring to large and small refugial areas, respectively (Ashcroft 2010; Gavin et al. 2014).

4.2. Biogeographic history and the present-day distribution of *Castanea sativa*

The cold and dry climate of the last glacial period and post-glacial climatic warming are major drivers of the distribution of *C. sativa* (Krebs et al., 2004; Mattioni et al., 2008; 2013) in combination with human impact during the last 2,000 years (Conedera et al. 2004; Bradshaw 2004). The modelled climatic suitability suggests that the expansion of this species in the first half of the Holocene was followed by a contraction of its distribution range during the last 5,000 years, resulting in the current fragmented range. However, humans also have participated in the expansion of *C. sativa* across European landscapes during the last two millennia (Conedera et al. 2004; Bradshaw 2004; Vescovi et al. 2010; Kaltenrieder et al. 2010; Morales-Molino et al. 2015; López-Sáez et al. 2017). The human influence could have resulted in the similarity of some genotypes from Italy and Spain (Mattioni et al. 2013). Nevertheless, evidence from pollen records, genetic data and SDMs suggest that human influence has modulated the regional distributions of *C. sativa* in Europe, rather than influencing its broad-scale distribution pattern.

Our results provide a strong support for the hypothesis of the glacial persistence of *C. sativa* in the northern and central Iberian Peninsula, where the species has been considered native only recently (López-Sáez et al. 2017). In this region the species shows a high genetic diversification in comparison with other European populations (Mattioni et al. 2008; Pereira-Lorenzo et al. 2010; Fernández-Cruz and Fernández-López 2016). There is a millennial tradition in the production of sweet chestnut (Conedera et al. 2004). The populations of *C. sativa* in this region also show good adaptation to limiting environmental conditions, e.g. drought, in comparison with the southern Iberian populations (Ciordia et al. 2012). This supports the idea that the species colonized central and southern Iberia in the warm and humid period of the mid-Holocene, but later on it was negatively affected by dry summers that characterize this region, particularly by the so-called 4200 cal BP arid event, as well as by the increase of anthropic activities since the Chalcolithic period (López-Sáez et al. 2014). The persistence of *C. sativa* in north-western Iberia may be explained by suitable climatic conditions that probably occurred there since the LGM, with

regional contractions and expansions from lowlands to highlands and vice versa. This idea is in line with the expected role of the climatically stable areas and the refugia-within-refugia scenario in Europe (Abellan and Svenning 2014).

Observing the current distribution of *C. sativa* and the palaeobotanical evidence of the two studied periods, the presence of isolated populations in continental areas with temperate (Romania, Bulgaria or Hungary) and Mediterranean (Iberian and Balkan peninsulas, Anatolia) climates should be highlighted. There is archaeological evidence of *C. sativa* in Germany during the Roman Period (Conedera et al. 2004) and the species is considered as a Roman-introduced archaeophyte in Slovakia (Medvecká et al. 2012). The forests of *C. sativa* in this area were planted in the High Middle Ages and the early Modern Period while the species is unknown in other areas such as the Czech Republic before the 16th century (Košnovská 2013). Similar cases could have occurred along the Mediterranean Basin, where ancient cultures probably spread the species in suitable areas. Nevertheless, evidence from the paleo-distribution modelling conducted here, together with fossil pollen and charcoal records, suggest that at least part of the scattered populations currently occurring in southern Central Europe and the Mediterranean coast are climatic relicts from the mid-Holocene, when *C. sativa* probably reached its maximum distribution range since the last full glacial period. This fact is also clear in some areas of Western Europe (such as the west and south of France) where the species found large suitable areas during both the LGM and the mid-Holocene.

5. Conclusions

Using a palaeodistribution modelling approach, this study provides a spatially-explicit picture of the two crucial periods of the biogeographical history of *Castanea sativa* in the Mediterranean Basin and the Atlantic region of Europe. For the LGM, we detected potential refugia on the three large Mediterranean peninsulas (Balkan, Italian and Iberian), along the Atlantic coast of France and Spain, and along the southern coast of the Black Sea. This distribution was validated by fossil records, and is also in agreement with genetic analyses of this species. Our results also suggest a postglacial expansion of *C. sativa* in the mid-Holocene, when climatic conditions were more favourable for it than current climatic conditions in Europe. We therefore hypothesize that some of the isolated populations in southern Central Europe and on the Mediterranean peninsulas could represent current refugia from the Holocene expansion. Our study provides a climatic framework for understanding potential refugia of *C. sativa* in Europe, despite later influence of humans also has to be considered. Moreover, our proposed model of expansion and contraction of the species range during the last 21,000 years provides a basis for

planning of appropriate management strategies for conservation of genetic resources and for reforestation projects.

Supplementary information

Additional Supporting Information may be found in the online version of this article: i) Climatic variables description and statistics; and ii) Individual distribution models for *C. sativa*.

Acknowledgments

We thank JA López-Saéz for general comments which helped us improve the quality of the manuscript. We also thank two anonymous reviewers for their useful suggestions. JVRD was supported by “Severo Ochoa” PhD Grant (BP 12-093) and by funding through “Ayuda para Estancias Breves” (EB15-12) for a research stay at Masaryk University (Brno, Czech Republic) in 2015. Both grants were provided by the “Plan de Ciencia, Tecnología e Innovación” (PCTI) Government of Principado de Asturias. BJA was supported by the project Employment of Best Young Scientists for International Cooperation Empowerment (CZ.1.07/2.3.00/30.0037) co-financed from the European Social Fund and the state budget of the Czech Republic. MC was supported by the Czech Science Foundation (Centre of Excellence PLADIAS, 14-36079G).

References

- Abellan, P., Svenning, J.C., 2014. Refugia within refugia - patterns in endemism and genetic divergence are linked to Late Quaternary climate stability in the Iberian Peninsula. *Biol. J. Linn. Soc.* 113, 13–28. doi:10.1111/bij.12309
- Alba Sánchez, F., López Sáez, J.A., De Benito Pando, B., Linares, J.C., Nieto, D., López Merino, L. 2010. Past and present potential distribution of the Iberian *Abies* species: a phytogeographic approach using fossil pollen data and species distribution models. *Divers. Distributions* 16, 214–228. doi:10.1111/j.1472-4642.2010.00636.x
- Alba, F., López Sáez, J.A., Nieto, D. Svenning, J.C., 2015. Long-term climate forcings to assess vulnerability in North Africa dry argan woodlands. *Appl. Veg. Sci.* 18, 283–296. doi:10.1111/avsc.12133
- Allen, R.M., Huntley, B., Watts, W.A., 1996. The vegetation and climate of northwest Iberia over the last 14000 yr. *J. Quat. Sci.* 11, 125–147.
- Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* 43, 1223–1232. doi:10.1111/j.1365-2664.2006.01214.x
- Aranbarri, J., Bartolomé, M., Alcolea, M., Sancho, C., Celant, A., González-Sampériz, P. Rodríguez-Lázaro, J. 2016. Palaeobotanical insights from early-mid Holocene fluvial tufas in the Moncayo natural park (Iberian range, NE Spain): regional correlations and biogeographic implications. *Review of Palaeobotany and Palynology*, 234, 31–43.
- Ashcroft, M.B., 2010. Identifying refugia from climate change. *J. Biogeogr.* 37, 1407–1413. doi:10.1111/j.1365-2699.2010.02300.x
- Benito Garzón, M., Sánchez De Dios, R., Sáinz Ollero, H., 2007. Predictive modelling of tree species distributions on the Iberian Peninsula during the Last Glacial Maximum and Mid-Holocene. *Ecography*. 30, 120–134. doi:10.1111/j.2006.0906-7590.04813.x
- Benito-Garzón, M., 2006. El efecto del cambio climático sobre las distribuciones de los bosques ibéricos: pasado, presente y futuro. Memoria de Tesis Doctoral. Madrid. 194 pp. Unpublished.
- Bennett, K., Provan, J., 2008. What do we mean by “refugia”? *Quat. Sci. Rev.* 27, 2449–2455. doi:10.1016/j.quascirev.2008.08.019
- Beyer, H.L., 2012. Geospatial Modelling Environment (Version 0.7.2.1). (software). URL: <http://www.spatialecology.com/gme>

- Birks, H.J.B., Willis, K.J., 2008. Alpines, trees, and refugia in Europe. *Plant Ecology & Diversity*. 1, 147–160
- Bradshaw, R.H.W. 2004. Past anthropogenic influence on European forests and some possible genetic consequences. *Forest Ecology and Management* 197, 203–212
- Caner, H., Algan, O., 2002. Palynology of sapropelic layers from the Marmara Sea. *Marine Geology*. 3156, 1–12.
- Ciordia, M., Feito, I., Pereira-Lorenzo, S., Fernández, a., Majada, J., 2012. Adaptive diversity in *Castanea sativa* Mill. half-sib progenies in response to drought stress. *Environ. Exp. Bot.* 78, 56–63. doi:10.1016/j.envexpbot.2011.12.018
- Conedera, M., Krebs, P., Tinner, W., Pradella, M., Torriani, D., 2004. The cultivation of *Castanea sativa* (Mill.) in Europe, from its origin to its diffusion on a continental scale. *Veg. Hist. Archaeobot.* 13, 161–179. doi:10.1007/s00334-004-0038-7
- Conedera, M., Tinner, W., Krebs, P., de Rigo, D., Caudullo, G., 2016. *Castanea sativa* in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., et al (Eds.) *European Atlas of Forest Tree Species*. Publ. Off. EU, Luxembourg, pp. e0125e0+
- Davis, B.A.S., Brewer, S., Stevenson, A.C., Guiot, J., 2003. The temperature of Europe during the Holocene reconstructed from pollen data. *Quat. Sci. Rev.* 22, 1701–1716. doi:10.1016/S0277-3791(03)00173-2
- Davis, B.A.S., Zanon, M., Collins, P., Mauri, A., Bakker, J., Barboni, D., Barthelmes, A., Beaudouin, C., Bjune, A.E., Bozilova, E., Bradshaw, R.H.W., Brayshay, B.A., Brewer, S., Brugiapaglia, E., Bunting, J., Connor, S.E., de Beaulieu, J.-L., Edwards, K., Ejarque, A., Fall, P., Florenzano, A., Fyfe, R., Galop, D., Giardini, M., Giesecke, T., Grant, M.J., Guiot, J., Jahns, S., Jankovská, V., Juggins, S., Kahrmann, M., Karpińska-Kołaczek, M., Kołaczek, P., Köhl, N., Kuneš, P., Lapteva, E.G., Leroy, S.A.G., Leydet, M., Guiot, J., Jahns, S., Jankovská, V., Juggins, S., Kahrmann, M., Karpińska-Kołaczek, M., Kołaczek, P., Köhl, N., Kuneš, P., Lapteva, E.G., Leroy, S.A.G., Leydet, M., López Sáez, J.A., Masi, A., Matthias, I., Mazier, F., Meltsov, V., Mercuri, A.M., Miras, Y., Mitchell, F.J.G., Morris, J.L., Naughton, F., Nielsen, A.B., Novenko, E., Odgaard, B., Ortu, E., Overballe-Petersen, M.V., Pardoe, H.S., Peglar, S.M., Pidek, I.A., Sadori, L., Seppä, H., Severova, E., Shaw, H., Święta-Musznicka, J., Theuerkauf, M., Tonkov, S., Veski, S., van der Knaap, W.O., van Leeuwen, J.F.N., Woodbridge, J., Zimny, M., Kaplan, J.O., 2013. The European Modern Pollen Database (EMPD) project. *Veg. Hist. Archaeobot.* 22, 521–530. doi:10.1007/s00334-012-0388-5
- Dormann, C.F., McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., G. Davies, R., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., R. Peres-Neto, P., Reineking, B., Schröder, B., M. Schurr, F., Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*. 30, 609–628. doi:10.1111/j.2007.0906-7590.05171.x
- Dudik, M., Phillips, S., Schapire, R., 2011. Maximum Entropy Modeling of Species Geographic Distributions” version 3.3.3k. <https://www.cs.princeton.edu/~schapire/maxent/>
- Ehlers, J., Gibbard, L., 2004. *Quaternary Glaciations: Extent and Chronology*. Elsevier, Amsterdam.
- Elith, J., Leathwick, J.R., 2009. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annu. Rev. Ecol. Evol. Syst.* 40, 677–697. doi:10.1146/annurev.ecolsys.110308.120159
- EPD. European Pollen Database, 2017. Records for *Castanea sativa* Mill. On-line query (27/10/2017). URL: <http://www.europeanpollendatabase.net/index.php>
- Fernández-Cruz, J., Fernández-López, J., 2016. Genetic structure of wild sweet chestnut (*Castanea sativa* Mill.) populations in northwest of Spain and their differences with other European stands. *Conserv. Genet.* 17, 949–967. doi:10.1007/s10592-016-0835-4
- Feurdean, A., Perşoiu, A., Tanţău, I., Stevens, T., Magyari, E.K., Onac, B.P., Marković, S., Andrič, M., Connor, S., Fărcaş, S., Gałka, M., Gaudeny, T., Hoek, W., Kołaczek, P., Kuneš, P., Lamentowicz, M., Marinova, E., Míchezyńska, D.J., Perşoiu, I., Płóciennik, M., Słowiński, M., Stancikaite, M., Sumegi, P., Svensson, A., Tămaş, T., Timar, A., Tonkov, S., Toth, M., Veski, S., Willis, K.J., Zernitskaya, V., 2014. Climate variability and associated vegetation response throughout Central and Eastern Europe (CEE) between 60 and 8 ka. *Quat. Sci. Rev.* 106, 206–224. doi:http://dx.doi.org/10.1016/j.quascirev.2014.06.003
- Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* 24, 38–49. doi: 10.1017/S0376892997000088

- Fineschi, S., Turchini, D., Villani, F., Vendramin, G.G., 2000. Chloroplast DNA polymorphism reveals little geographical structure in *Castanea sativa* Mill. (Fagaceae) throughout southern European countries. *Mol. Ecol.* 9, 1495–1503.
- Franklin, J., 2009. *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge University Press, Cambridge, UK. 320 pp.
- Franklin, J., Davis, F.W., Ikegami, M., Syphard, A.D., Flint, L.E., Flint, A.L., Hannah, L., 2013. Modeling plant species distributions under future climates: How fine scale do climate projections need to be? *Glob. Chang. Biol.* 19, 473–483. doi:10.1111/gcb.12051
- Fyfe, R.M., de Beaulieu, J.L., Binney, H., Bradshaw, R.H.W., Brewer, S., Le Flao, A., Finsinger, W., Gaillard, M.J., Giesecke, T., Gil-Romera, G., Grimm, E.C., Huntley, B., Kunes, P., Köhl, N., Leydet, M., Lotter, A.F., Tarasov, P.E., Tonkov, S., 2009. The European Pollen Database: Past efforts and current activities. *Veg. Hist. Archaeobot.* 18, 417–424. doi:10.1007/s00334-009-0215-9
- Gavin, D.G., Fitzpatrick, M.C., Guggen, P.F., Heath, K.D., Rodríguez-Sánchez, F., Dobrowski, S.Z., Hampe, A., Hu, F.S., Ashcroft, M.B., Bartlein, P.J., Blois, J.L., Carstens, B.C., Davis, E.B., de Lafontaine, G., Edwards, M.E., Fernandez, M., Henne, P.D., Herring, E.M., Holden, Z. a., Kong, W., Liu, J., Magri, D., Matzke, N.J., McGlone, M.S., Saltré, F., Stigall, A.L., Tsai, Y.E., Williams, J.W., 2014. Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytol.* 204, 37–54. doi:10.1111/nph.12929
- Gómez, A., Lunt, D.H., 2007. Refugia within refugia: Patterns of phylogeographic concordance in the Iberian Peninsula. Pp. 155–188 in: Weiss, S., Ferrand, N., (eds.) *Phylogeography of southern European refugia*. Dordrecht: Springer.
- Goring, S., Dawson, A., Simpson, G.L., Ram, K., Graham, R.W., Grimm, E.C., Williams, J.W., 2015. Neotoma: a programmatic interface to the Neotoma paleoecological database. *Open Quaternary*. 1, 1e17. <http://doi.org/10.5334/oq.ab>.
- Heiri, O., Lotter, A. F., Hausmann, S., Kienast, F. 2003. A chironomid-based Holocene summer air temperature reconstruction from the Swiss Alps. *The Holocene*, 13(4), 477-484. doi: 10.1191/0959683603hl640ft
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25, 1965–1978. doi:10.1002/joc.1276
- Hijmans, R.J., Elith, J., 2016. Species distribution modelling with R. Online [31/01/2016] url: <https://cran.r-project.org/web/packages/dismo/vignettes/sdm.pdf>
- Huntley, B., 1990. European post-glacial forests: compositional changes in response to climatic change. *J. Veg. Sci.* 1, 507–518. doi: 10.2307/3235785
- Huntley, B., Birks, H.J.B., 1983. *An atlas of past and present pollen maps for Europe: 0–13000 years ago*. Cambridge University Press, Cambridge, UK. 688 pp.
- Janská, V., Jiménez-Alfaro, B., Chytrý, M., Divišek, J., Anenkhonov, O., Korolyuk, A., Lashchinskyi, N., Culek, M., 2017. Palaeodistribution modelling of European vegetation types at the Last Glacial Maximum using modern analogues from Siberia: Prospects and limitations. *Quat. Sci. Rev.* 159, 103–115. doi:10.1016/j.quascirev.2017.01.011
- Jenness, J., 2003. Mahalanobis distances (Mahalanobis.avx) extension for ArcView 3.x, Jenness Enterprises. Available at: <http://www.jennessent.com/arcview/Mahalanobis.htm>.
- Kaltenrieder, P., Belis, C.A., Hofstetter, S., Ammann, B., Ravazzi, C., Tinner, W., 2009. Environmental and climatic conditions at a potential Glacial refugial site of tree species near the Southern Alpine glaciers. New insights from multiproxy sedimentary studies at Lago della Costa (Euganean Hills, Northeastern Italy). *Quat. Sci. Rev.* 28, 2647–2662. doi:10.1016/j.quascirev.2009.05.025
- Kaltenrieder, P., Procacci, G., Vannièrè, B., Tinner, W., 2010. Vegetation and fire history of the Euganean Hills (Colli Euganei) as recorded by Lateglacial and Holocene sedimentary series from Lago della Costa (northeastern Italy). *Holocene* 20, 679–695. doi:10.1177/0959683609358911
- Keppel, G., Van Niel, K.P., Wardell-Johnson, G.W., Yates, C.J., Byrne, M., Mucina, L., Schut, A.G.T., Hopper, S.D., Franklin, S.E., 2012. Refugia: identifying and understanding safe havens for biodiversity under climate change. *Glob. Ecol. Biogeogr.* 21, 393–404. doi:10.1111/j.1466-8238.2011.00686.x
- Košnovská, J., 2013. The Origin, Archaeobotany and Ethnobotany of Sweet Chestnut (*Castanea sativa* Miller) in the Czech Republic. *Interdiscip. Archaeol.* 4, 163–176.
- Krebs, P., Conedera, M., Pradella, M., Torriani, D., Felber, M., Tinner, W., 2004. Quaternary refugia of the sweet chestnut (*Castanea sativa* Mill.): an extended palynological approach. *Veg. Hist. Archaeobot.* 13, 145–160. doi:10.1007/s00334-004-0041-z

- Kuneš, P., Pokorný, P., Šída, P., 2008. Detection of the impact of early Holocene hunter-gatherers on vegetation in the Czech Republic, using multivariate analysis of pollen data. *Veg. Hist. Archaeobot* 17, 269–287
- Lehsten, D., Dullinger, S., Hülber, K., Schurgers, G., Cheddadi, R., Laborde, H., Lehsten, V., François, L., Dury, M., Sykes, M.T., 2014. Modelling the Holocene migrational dynamics of *Fagus sylvatica* L. and *Picea abies* (L.) H. Karst. *Global Ecology and Biogeography* 23,658-668. DOI: 10.1111/geb.12145
- López-Sáez, J.A., Abel Schaad, D., Pérez Díaz, S., Blanco González, A., Alba Sánchez, F., Dorado, M., Ruiz Zapata, B., Gil García, M.J., Gómez González, C., Franco Múgica, F., 2014. Vegetation history, climate and human impact in the Spanish Central System over the last 9,000 years. *Quat. Int.* 353, 98-122. doi:10.1016/j.quaint.2013.06.034
- López-Sáez, J.A., Glais, A., Robles-López, S., Alba-Sánchez, F., Pérez-Díaz, S., Abel-Schaad, D., Luelmo-Lautenschlaeger, R., 2017. Unraveling the naturalness of sweet chestnut forests (*Castanea sativa* Mill.) in central Spain. *Veg. Hist. Archaeobot.* 26, 167-182. doi:10.1007/s00334-016-0575-x
- Lowe, J.J., Accorsi, C.A., Bandini Mazzanti, M., Bishop, A., Van der Kaars, S., Forlani, L., Mercury, A.M., Rivalenti, C., Torri, P., Watson, C., 1996. Pollen stratigraphy of sediment sequences from lakes Albano and Nemi (near Rome) and from the central Adriatic, spanning the interval from oxygen isotope Stage 2 to the present day. *Mem. Ist. ital. Idrobiol.* 55, 71–98
- Lurgi, M., Brook, B.W., Salter, F., Fordham, D.A. 2015. Modelling range dynamics under global change: which framework and why?. *Method Ecol. Evolut.* 6, 247-256. DOI: 10.1111/2041-210X.12315
- Mahalanobis, P.C., 1936. On the generalised distance in statistics. *Proc Natl Inst Sci India* 2:49-55.
- Martín Arroyo T. 1998. Paleoclimatología y Paleoambiente durante el Pleistoceno Medio y Superior en el valle del Tajo. Tesis doctoral. Universidad de Alcalá de Henares, 173 pp. (In Spanish). Unpublished.
- Mattioni, C., Cherubini, M., Micheli, E., Villani, F., Bucci, G., 2008. Role of domestication in shaping *Castanea sativa* genetic variation in Europe. *Tree Genet. Genomes* 4, 563–574. doi:10.1007/s11295-008-0132-6
- Mattioni, C., Martin, M.A., Pollegioni, P., Cherubini, M., Villani, F., 2013. Microsatellite markers reveal a strong geographical structure in European populations of *Castanea sativa* (Fagaceae): evidence for multiple glacial refugia. *Am. J. Bot.* 100, 951–61. doi:10.3732/ajb.1200194
- Médail, F., Diadema, K., 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *J. Biogeogr.* 36, 1333–1345. doi:10.1111/j.1365-2699.2008.02051.x
- Medvecká, J., Kliment, J., Májková, J., Halada, L., Zaliberová, M., Gojdičová, E., Feráková, V., Jarolímek, I., 2012. Inventory of the alien flora of Slovakia. *Preslia* 84, 257–309.
- Morales-Molino, C., Vescovi, E., Krebs, P., Carlevaro, E., Kaltenrieder, P., Conedera, M., Tinner, W., Colombaroli, D., 2015. The role of human-induced fire and sweet chestnut (*Castanea sativa* Mill.) cultivation on the long-term landscape dynamics of the southern Swiss Alps. *The Holocene* 25, 482–494. doi:10.1177/0959683614561884
- Moreno-Amat, E., Rubiales, J.M., Morales-Molino, C., García-Amorena, I. et al. 2017. Incorporating plant fossil data into species distribution models is not straightforward: Pitfalls and possible solutions. *Quaternary Science Reviews* 170, 56-68. doi: org/10.1016/j.quascirev.2017.06.022
- Muñoz Sobrino, C., Ramil-Rego, P., Gómez-Orellana, L., 2004. Vegetation of the Lago de Sanabria area (NW Iberia) since the end of the Pleistocene: A palaeoecological reconstruction on the basis of two new pollen sequences. *Veg. Hist. Archaeobot.* 13, 1–22. doi:10.1007/s00334-003-0028-1
- Muñoz Sobrino, C., Ramil-Rego, P., Gómez-Orellana, L., Díaz Varela, R., 2005. Palynological data on major Holocene climatic events in NW Iberia. *Boreas* 34, 381–400. doi:10.1080/03009480510013006
- Nakagawa, T., 1998. Etudes palynologiques dans les Alpes françaises centrales et meridionales: histoire de la vegetation Tardiglaciaire et Holocene. PhD Thesis, Université de Marseille
- Nekola, J.C., 1999. Paleorefugia and neorefugia: the influence of colonization history on community pattern and process. *Ecology.* 80, 2459–2473
- NeotomaDB Neotoma Palaeoecology Database, 2015. Records for *Castanea* genre in Europe. On-line query (12/19/2015). URL: <http://www.neotomadb.org/about/category/about>
- Nieto Feliner, G., 2014. Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. *Perspect. Plant Ecol. Evol. Syst.* 16, 265–278. doi:10.1016/j.ppees.2014.07.002
- Paganelli, A., Miola, A., 1991. Chestnut (*Castanea sativa* Mill.) as an indigenous species in northern Italy. *Il Quat., It. J. Quat. Sci.* 4, 99-106.

- Pereira-Lorenzo, S., Costa, R.M.L., Ramos-Cabrer, A.M., Ribeiro, C.A.M., Silva, M.F.S., Manzano, G., Barreneche, T., 2010. Variation in grafted European chestnut and hybrids by microsatellites reveals two main origins in the Iberian Peninsula. *Tree Genet. Genomes* 6, 701–715. doi:10.1007/s11295-010-0285-y
- Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. *Ecol. Modell.* 190, 231–259. doi:10.1016/j.ecolmodel.2005.03.026
- R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Ramil-Rego, P., Rodríguez-Gutián, M.A., Muñoz-Sobrino, C. 1998. Sclerophyllous vegetation dynamics in the north of the Iberian peninsula during the last 16.000 years. *Glob. Ecol. Biogeo. Lett.* 7, 335-351.
- Rodríguez-Sánchez, F., Hampe, A., Jordano, P., Arroyo, J., 2010. Past tree range dynamics in the Iberian Peninsula inferred through phylogeography and palaeodistribution modelling: A review. *Rev. Palaeobot. Palynol.* 162, 507–521. doi:10.1016/j.revpalbo.2010.03.008
- Rull, V., 2009. Microrefugia. *J. Biogeogr.* 361 381-384. doi:10.1111/j.1365-2699.2008.02023.x
- Sánchez de Dios, R., Benito-Garzón, M., Sainz-Ollero, H., 2009. Present and future extension of the Iberian submediterranean territories as determined from the distribution of marcescent oaks. *Plant Ecol.* 204, 189–205. doi:10.1007/s11258-009-9584-5
- Schorr, G., Holstein, N., Pearnan, P.B., Guisan, a., Kadereit, J.W., 2012. Integrating species distribution models (SDMs) and phylogeography for two species of alpine primula. *Ecol. Evol.* 2, 1260–1277. doi:10.1002/ece3.100
- Stewart, J.R., Lister, A.M., Barnes, I., Dalén, L., 2010. Refugia revisited: individualistic responses of species in space and time. *Proc. Biol. Sci.* 277, 661–71. doi:10.1098/rspb.2009.1272
- Svenning, J.-C., Eiserhardt, W.L., Normand, S., Ordonez, A., Sandel, B., 2015. The Influence of Paleoclimate on Present-Day Patterns in Biodiversity and Ecosystems. *Annu. Rev. Ecol. Evol. Syst.* 46, annurev-ecolsys-112414-054314. doi:10.1146/annurev-ecolsys-112414-054314
- Svenning, J.C., Fløjgaard, C., Marske, K. a., Nógues-Bravo, D., Normand, S., 2011. Applications of species distribution modeling to paleobiology. *Quat. Sci. Rev.* 30, 2930–2947. doi:10.1016/j.quascirev.2011.06.012
- Svenning, J.C., Normand, S., Kageyama, M., 2008. Glacial refugia of temperate trees in Europe: Insights from species distribution modelling. *J. Ecol.* 96, 1117–1127. doi:10.1111/j.1365-2745.2008.01422.x
- Svenning, J.C., Skov, F., 2007. Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? *Ecol. Lett.* 10, 453–460. doi:10.1111/j.1461-0248.2007.01038.x
- Tzedakis, P.C., Emerson, B.C., Hewitt, G.M., 2013. Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends Ecol. Evol.* 28, 696–704. doi:10.1016/j.tree.2013.09.001
- Uzquiano, P. 1992. Recherches anthracologiques dans le secteur Pyreneo-Cantabrique (Pays Basque, Cantabria et Asturias): environnements et relations homme-milieu au Pleistocene superieur et debut de l'Holocene. PhD Thesis, Universite Montpellier IIVescovi.
- Vessella, F., Schirone, B., 2013. Predicting potential distribution of *Quercus suber* in Italy based on ecological niche models: Conservation insights and reforestation involvements. *For. Ecol. Manage.* 304, 150–161. doi:10.1016/j.foreco.2013.05.006
- Vessella, F., Simeone, M.C., Schirone, B., 2015. *Quercus suber* range dynamics by ecological niche modelling: from the Last Interglacial to present time. *Quat. Sci. Rev.* 119, 85–93. doi:10.1016/j.quascirev.2015.04.018
- Villani, F., Pigliucci, M., Cherubini, M. 1994. Evolution of *Castanea sativa* Mill. in Turkey and Europe. *Gen. Res.* 63, 109-116.
- Villani, F., Sansotta, A., Cherubini, M., Cesaroni, D., Sbordoni, V. 1999 Genetic structure of natural populations of *Castanea sativa* in Turkey: Evidence of a hybrid zone. *Int. J. Evol. Biol.* 12, 233-244.
- Zohary, D., Hopf, M., 1988. Domestication of plants in the world. Clarendon Press, Oxford, 249 pp.

Tables

Table 1.

Validation of the models developed.

Model type	Validation (independent data subset with points n = 120)			Suitable area predicted	
	Area Under Curve	Sensitivity (%)	Specificity (%)	km ²	% of the total study area
GLM	0.811	62	70	1,773,946	34.5
Maxent	0.916	67	95	784,647	15.3
Mahalanobis	0.750	78	52	2,386,892	46.4

Table 2.

Suitable area, habitat suitability and sensitivity for the three modelling methods (GLM, Maxent and Mahalanobis) calculated and three scenarios (MPI, CCSM and MIROC) used.

MODEL	Last Glacial Maximum					Mid-Holocene					
	Suitable area		Habitat suitability		Sensitivity* (%)	Suitable area		Habitat suitability		Sensitivity* (%)	
	10 ³ km ²	%	Mean	SD		10 ³ km ²	%	Mean	SD		
MPI	GLM	3,232	52.3	0.271	0.263	74.1	1,374	26.7	0.140	0.177	70.9
	Maxent	1,399	22.6	0.316	0.224	37.0	856	16.6	0.239	0.219	74.7
	Mahalanobis	806	13.0	0.083	0.162	33.3	1,304	25.4	0.158	0.221	69.6
CCSM	GLM	3,491	56.5	0.303	0.271	81.5	1,317	25.6	0.142	0.186	75.9
	Maxent	1,421	22.9	0.295	0.249	51.9	900	17.6	0.236	0.227	68.4
	Mahalanobis	292	4.7	0.037	0.104	29.6	1,154	22.4	0.149	0.216	65.8
MIROC	GLM	3,385	54.7	0.281	0.259	85.2	1,631	31.7	0.170	0.208	81.0
	Maxent	998	16.1	0.251	0.258	44.4	1,196	23.3	0.259	0.238	75.9
	Mahalanobis	112	1.8	0.013	0.060	11.1	1,059	20.6	0.142	0.201	59.5

*: Calculated from fossil pollen and charcoal records for the Last Glacial Maximum (27 points from 18,000 to 22,000 cal. yr BP) and the Mid Holocene (79 points from 5,000 to 7,500 cal. yr BP).

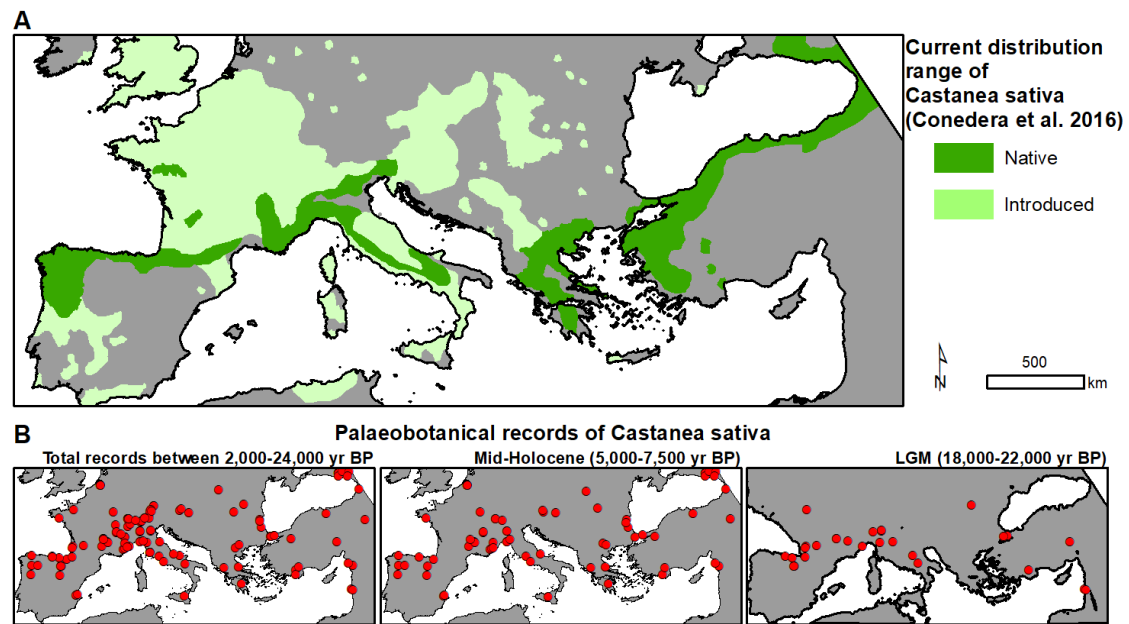


Figure 1

Figure 1. Assumed native and introduced distribution range of *Castanea sativa* (from Conedera et al. 2016) (A); and fossil, pollen and charcoal records for the pre-Roman period, mid-Holocene and LGM (B).

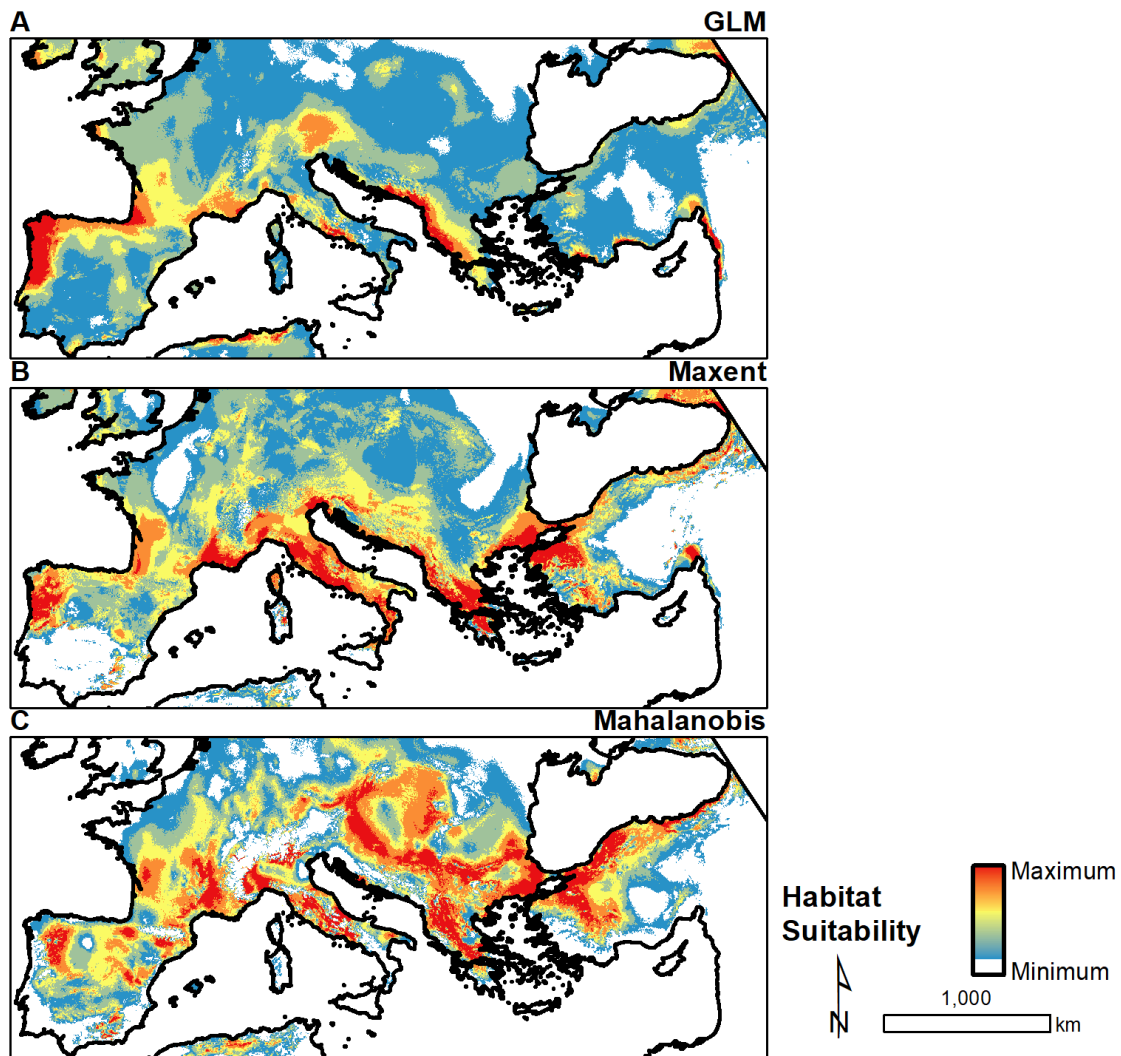


Figure 2

Figure 2. Distribution models for *Castanea sativa* based on three modelling methods (GLM, Maxent and Mahalanobis) for the recent climatic conditions (1950-2000).

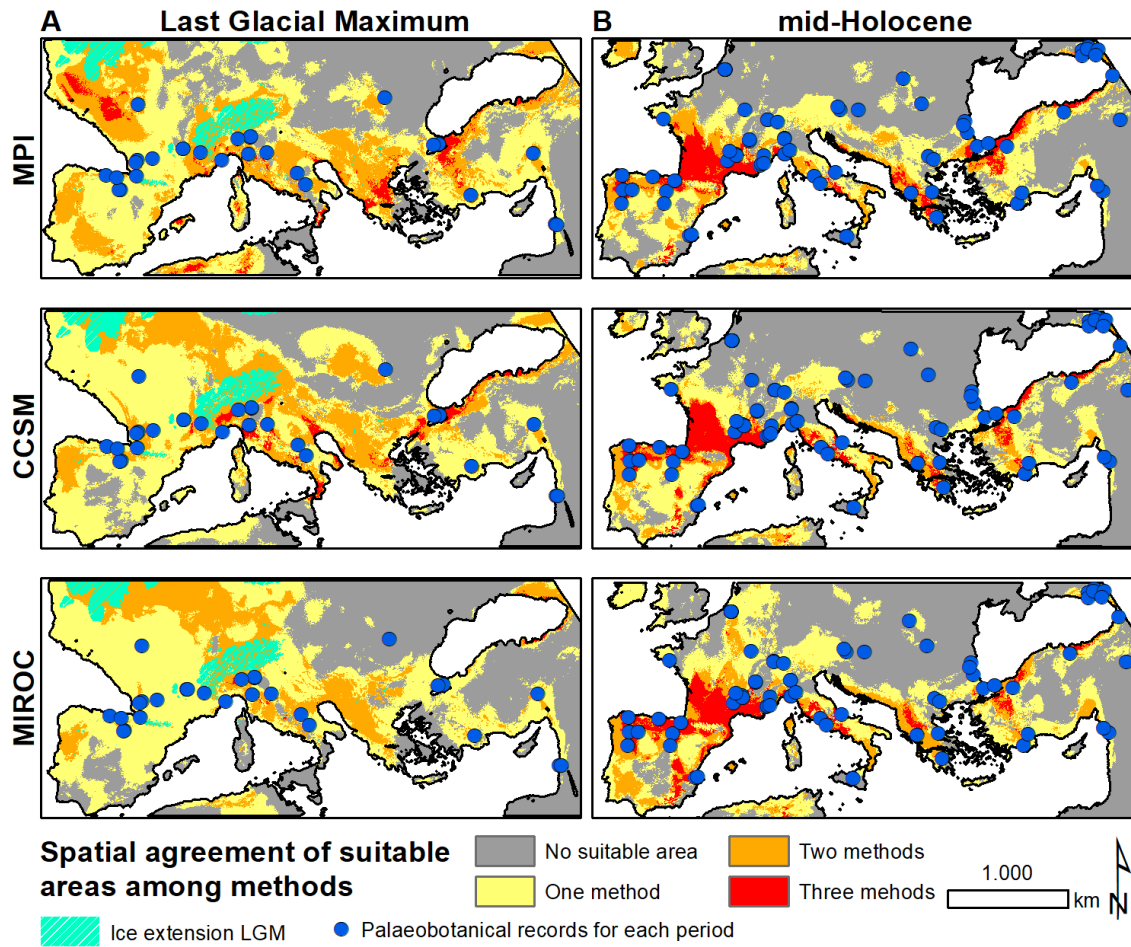


Figure 3

Figure 3. Suitable areas for *Castanea sativa* in the Last Glacial Maximum and mid-Holocene for three different climatic scenarios (MPI; CCSM and MIROC) based on the combined results of three modelling methods (GLM, Maxent and Mahalanobis). Different colours show the degree of agreement among the three methods. Black dots reflect fossil pollen and charcoal records.

SUPPLEMENTARY INFORMATION**Glacial refugia and mid-Holocene expansion delineate the current distribution of *Castanea sativa* in Europe**

Roces-Díaz, Jose V., Jiménez-Alfaro, Borja, Chytrý, Milan, Díaz-Varela, Emilio R., Álvarez-Álvarez, Pedro

1. Geographic patterns of the climatic variables used for species distribution modelling and their statistic relevance on the models.

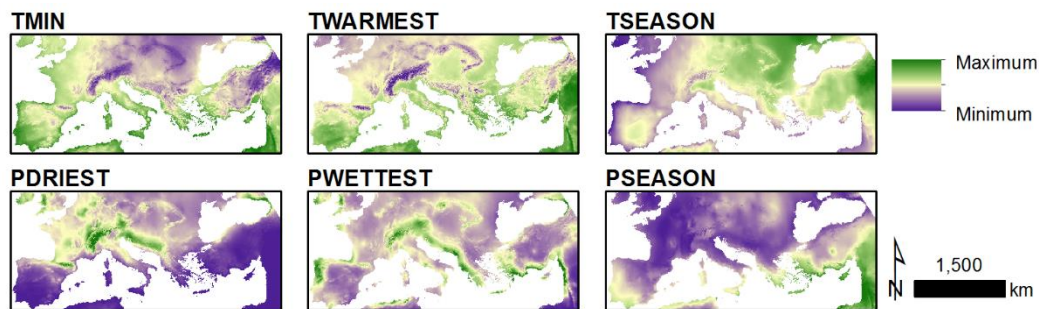


Figure S1. Representation of current climatic variables used in this work (source: www.worldclim.org).

Table S1.

GLM (coefficients and Wald criterion) and Maxent (% contribution and permutation importance of variables).

MODEL	GLM		Maxent	
	Variables	Coefficient	Wald criterion	Percent contribution
TMIN	-	-	20.9	26.4
TWARMEST	0.015	8.627	17.9	29.7
TSEASON	0.000	11.103	33.1	5.8
PDRIEST	-0.050	6.891	17.2	22.2
PWETTEST	0.015	17.927	0.9	11.7
PSEASON	-0.111	22.034	10	4.2
Constant	1.056	0.574	-	-

2. Individual models conducted for *Castanea sativa* in the Last Glacial Maximum and Mid-Holocene.

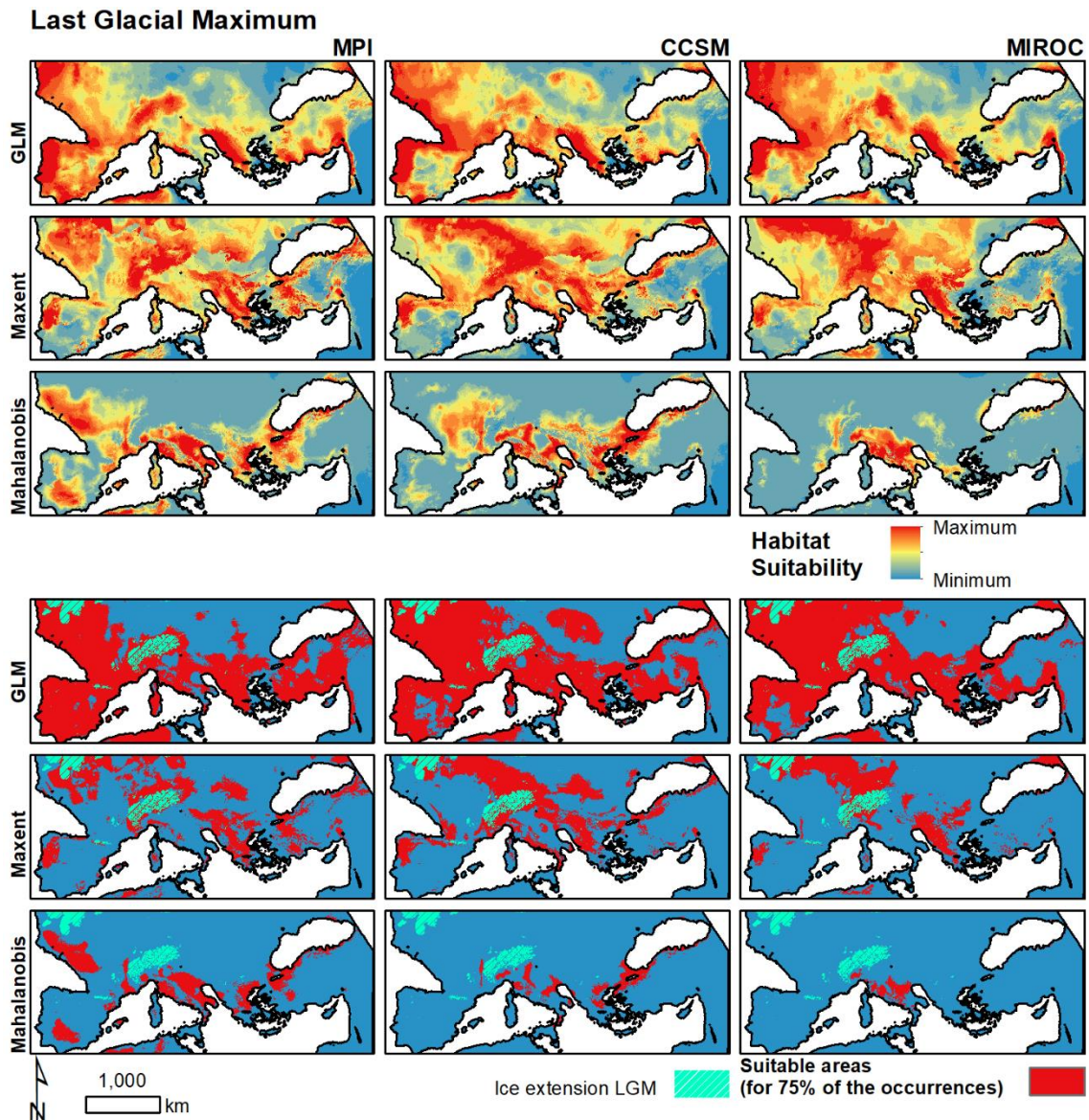


Figure S2. Representation of the habitat suitability models and of the suitable areas using the threshold of the 75% of current probability of occurrence of the three types of models (GLM, Maxent and Mahalanobis) for the three climatic scenarios (MPI, CCSCM and MIROC) for the Last Glacial Maximum.

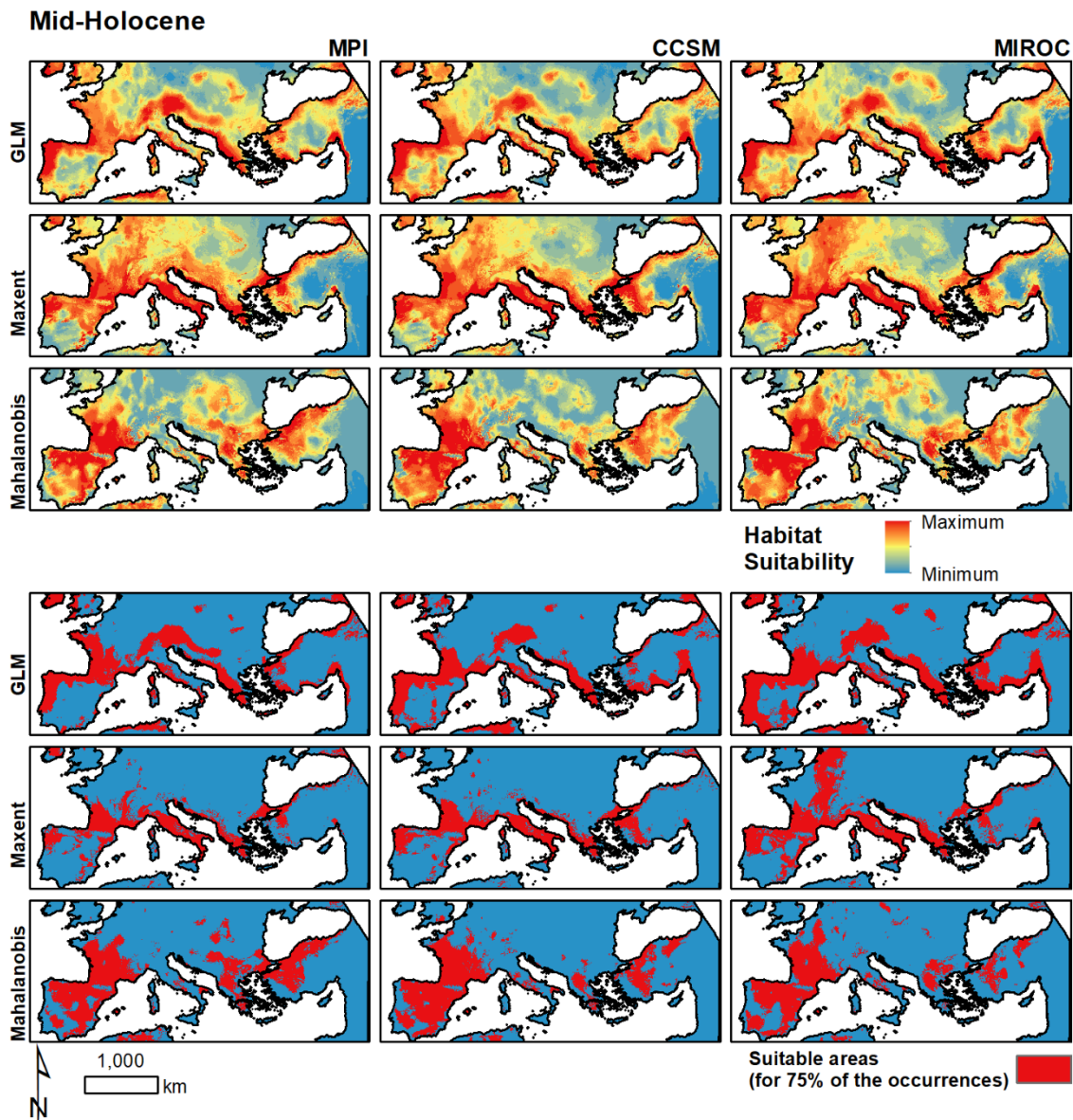


Figure S3. Representation of the habitat suitability models and of the suitable areas using the threshold of the 75% of current probability of occurrence of the three types of models (GLM, Maxent and Mahalanobis) for the three climatic scenarios (MPI, CCSCM and MIROC) for the Mid-Holocene.