Original Article

Word count: abstract 304; main body of the text, 7173

Combining molecular and fossil data to infer demographic history of *Quercus cerris*: insights on European eastern glacial refugia.

F. Bagnoli[1#], Y. Tsuda[1, 2#], S. Fineschi[3], P. Bruschi[4], D. Magri[5], P. Zhelev[6], L. Paule[7], M.C. Simeone[8], S.C. González-Martínez[9,10], G.G. Vendramin[1*]

[1] Institute of Biosciences and Bioresources, National Research Council, Via Madonna del Piano 10, 50019 Sesto Fiorentino (FI), Italy

[2] present address: Department of Biology, Graduate School of Science, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba-shi, Chiba, 263-8522, Japan

[3] Institute for Sustainable Plant Protection, National Research Council, Via Madonna del Piano 10, 50019 Sesto Fiorentino (FI), Italy

[4] Dept of Agriculture, Food and Environmental Science, University of Florence, Piazzale delle Cascine 28, 50144 Florence, Italy

[5] Department of Environmental Biology, Sapienza University of Rome, Piazzale Aldo Moro, 5 00185 Roma, Italy

[6] University of Forestry, 1797 Sofia, Bulgaria

[7] Technical University in Zvolen, Slovakia

[8] Agriculture, Forest, Nature and Energy Department (DAFNE), University of Tuscia, via S. Camillo de' Lellis, 01100 Viterbo, Italy

[9] INRA, UMR1202 BIOGECO, 33610 Cestas, France

[10] University of Bordeaux, UMR 1202 BIOGECO, 33400 Talence, France

* Corresponding author: Giovanni G. Vendramin, giovanni.vendramin@ibbr.cnr.it

# contributed equally

Running Head: Demographic history of *Quercus cerris*

ABSTRACT

Aim Phylogeographic studies of eastern Mediterranean species are rare. We aim to fill a gap in the current understanding of the role of Eastern Mediterranean refugia, and their connections with other refugia across Europe. To such aim, we studied the genetic diversity distribution and the genetic structure for the modern population of *Quercus cerris* in relation to its Quaternary demographic history and to more ancient events.

Location Mediterranean Basin; Italian, Balkan, Anatolian peninsulas.

Methods A total of 192 populations were genotyped with six polymorphic chloroplast microsatellites, and the genetic diversity and differentiation of the populations were evaluated. The geographical structure of genetic variation was analyzed with a Bayesian clustering method using BAPS 5.2. The demographic history of *Quercus cerris* was explored by an Approximate Bayesian Computation procedure using DIYABC v2.0. To reconstruct the past distribution of Turkey oak, we also considered the chronology and geographical distribution of fossil records.

Results Thirty-five haplotypes were found, three of which (together) were found in 71.82% of individuals. Bayesian analysis resulted in three genetically and geographically distinct clusters: a Western group, a Central group, and an Eastern group. The Approximate Bayesian Computation analysis, together with fossil data, showed a possible bottleneck leading to the divergence of the Eastern and Central populations in the Early Pleistocene (Gelasian). The split into two groups of populations in the Italian and Balkan Peninsulas, respectively, was probably caused by a marked population contraction during a glacial phase of the Middle Pleistocene.

Main conclusions This study provides information on the potential role of Eastern Europe and the

Near East as refugia and as a source for ancient westward range expansions in the Mediterranean region. Our study covers a remarkable gap in European oak phylogeography, showing a putative eastern origin of *Q. cerris* and the presence of large amounts of genetic diversity in this region.

Keywords Approximate Bayesian Computation; BAPS; cpSSR; Mediterranean basin; palaeobotany; Turkey oak.

INTRODUCTION

Understanding how past evolutionary processes shaped the genetic diversity distribution of plant species represents an important inquiry to enable predictions of forest ecosystem dynamics under today's impending climate change (Alberto *et al.*, 2013). This is particularly relevant for our ability to forecast the response of European forest ecosystems from temperate continental and Mediterranean regions in response to climate models that predict an overall increase in global mean precipitation, and yet considerable dryness in some regions, such as the Mediterranean Basin (IPCC, 2007). More specifically, in the Mediterranean Basin, biodiversity is influenced by complex patterns associated with the heterogeneous historical biogeography and landscape complexity (Blondel, 2010). Despite the numerous recent investigations into the richness of genetic diversity in Mediterranean plants over substantial parts of their ranges, the genetic structure of many species and the role of key palaeobiogeographic regions, particularly in the eastern part of the Basin, are still far from being understood (Médail &Diadema, 2009).

The current patterns of genetic diversity found in the European flora reflects a clear influence of Quaternary climate oscillations. Studies accumulated during the past two decades in Europe inferred these patterns assuming a contraction/expansion model (Provan & Bennet, 2008). It is worth noting that, in the Southern European regions, the disappearance of both genotypes and populations was minimized due to less severe climate conditions during the glacial periods (Nieto-Feliner, 2011). In the case of European temperate tree species, the accepted hypothesis assumes the existence of main refugial areas in the three southern peninsulas (Iberia, Italy and the Balkans), from which the recolonisation of the rest of Europe took place (Bennett *et al.,* 1991). This continues to be the prevailing hypothesis, despite an increasing volume of evidence suggesting that middle- to high-latitude refugia may have also contributed to post-glacial recolonisation (Gavin *et al.,* 2014).

A novel approach based on the analysis of the spatial distribution of analogous and non-analogous climate conditions between the Last Glacial Maximum (LGM) and the present age,

quantified and mapped potential source and sink areas for species recolonisation in Europe. This method identified South-Eastern Europe as the region bearing the highest post-glacial sink potential, in particular the Balkan Peninsula and Turkey, which likely acted as both source and sink areas for species range shifts since LGM (Ohlemüller *et al.*, 2012). Despite the importance of Anatolia and the Balkan Peninsula as refuge areas and recognized sources of genetic diversity for the European flora, these regions have not been thoroughly investigated yet. This gap of knowledge needs to be filled because in the Mediterranean area there is an east-west gradient in genetic diversity for many plant and animal species, in addition to a complex demographic history related to latitudinal displacements (Conord *et al.*, 2012). The east-west gradient in genetic diversity may be the result of processes acting at different time scales, including the legacy of a Tertiary palaeotropical flora that was extended along the Tethyan Ocean (Chen *et al.*, 2014), the contraction of formerly continuous ranges in relation to Quaternary oscillations, and the successive westward or eastward waves of colonization fostered by east-west climate gradients (Nieto-Feliner, 2014).

In addition to the generally accepted LGM southern refugia hypothesis, Magri *et al.* (2006), examining genetic and palaeocological data of European beech (*Fagus sylvatica*), proposed that older temporal scales and repeated ice ages in the Quaternary should also be considered when discussing factors affecting modern phylogeographical structure in tree species. For example, González-Martínez *et al.* (2010) found patterns of genetic diversity and genetic structure in *Taxus baccata* that reflect processes at different spatial and temporal scales, including ancient ones. In addition, genetic resequencing studies on European tree species detected severe bottlenecks predating the LGM in pine, spruce and aspen (Heuertz *et al.* 2006; Pyhäjärvi *et al.* 2007; Ingvarsson 2008). Finally, the southern European regions did not represent homogeneous and continuous refugial areas during the ice ages. On the contrary, those areas were characterized by a set of different ecological and climatic conditions that embraced many glacial refugia, resulting in the so-called ‘refugia-within-refugia’ scenario (Gómez & Lunt, 2007). Increasing evidence from multiple species within the southern European regions supports this hypothesis (e.g. López de Heredia *et al.*,

2007).

All these studies suggest that it is important to consider not only post-LGM episodes but also the demographic history of more ancient species in the Quaternary and/or the Tertiary, and ask for more detailed spatial analysis within known southern refugia, in order to explain the current genetic biodiversity of Europe's forest species.

This research investigates the modern genetic structure of Turkey oak (*Quercus cerris*) while also investigating its Quaternary demographic history, considering east-to-west patterns of genetic diversity in the Mediterranean region and the possibility of more ancient events. Moreover, our case study aims to provide new and relevant insight into the role of Eastern Mediterranean refugia on the current populations, as well as on their connection to refugia across Europe. Turkey oak stands prominently among the broadleaf trees with Central-Eastern Mediterranean distribution. Native to Central Europe and Asia Minor, its distribution is centred on the Italian and Balkan peninsulas, from the western Alps to the Black Sea, Anatolia and Lebanon (Gellini & Grossoni, 1997), and overlaps with those of other oaks, some of which are suggested to hybridize with Turkey oak (e.g. *Q. suber* Camus, 1938-39). Our study combines palaeobotanical information with the current distribution of *Q. cerris* genetic variation, using large-scale chloroplast SSR analyses. Recent advances in population demographic analysis provide a flexible methodology to infer past demographic history (Bertorelle *et al*., 2010). In our study, palaeoecological data and demographic inference using Approximate Bayesian Computation (ABC) were combined to obtain insights into the origin, migration routes, and persistence of this important species in glacial refugia. These findings, increase the biogeographical knowledge of poorly studied regions in the Balkans and Anatolia. Moreover, we provide indications for core diversification areas, contact zones, and gene flow patterns between these regions and the rest of Europe.

MATERIALS AND METHODS

142 **Plant material, DNA extraction, and cpSSR genotyping**

143

144 Leaf samples of *Q. cerris* were collected from 192 populations (1,175 individuals) throughout the
145 whole natural species range (Table S1 in Appendix S3 in Supporting Information, Fig. 1). Fresh
146 leaves were stored at –80°C prior to DNA extraction. Total DNA was extracted using Invitek
147 Invisorb Spin Plant Mini Kit (following the manufacturers' instructions), from approximately 100
148 mg of fresh leaves, ground in a Retsch MM200 grinding mill. Six polymorphic chloroplast
149 microsatellite (cpSSR) markers were used: ccmp6 (Weising & Gardner, 1999), cmcs1, cmcs6,
150 cmcs7, cmcs8 and cmcs9 (Sebastiani *et al.*, 2004). Polymerase chain reactions (PCR) details and
151 sequencing are described in Appendix S1 in Supporting Information. To test whether differences in
152 SSR allele lengths were due to simple changes in the number of repeats or to mutations in the
153 flanking regions (i.e. size homoplasy) the amplified cpSSR fragments of at least two individuals for
154 each haplotype were directly sequenced from both ends using an automatic sequencer MegaBACE
155 (GE Healthcare), (methods are given in Appendix S1).

156 To rule out the possibility that oak species different from *Q. cerris* had been sampled in
157 Anatolia, at least two individuals for each haplotype found in this region were sequenced for the
158 *trnH-psbA* region. This region represents a rapidly evolving intergenic spacer currently considered
159 the most informative for fast species-level identification in oaks (Simeone *et al.*, 2013). Details are
160 provided in Appendix S1.

161

162 **Genetic diversity and population structure**

163 Gene diversity (*h*), number of haplotypes and number of population-specific (i.e. private or
164 endemic) haplotypes were determined for each population using GenAlEx 6.5 (Peakall & Smouse,
165 2006, 2012). An haplotype frequency map was constructed using MapInfo Pro Version 12.5
166 (MapInfo Corporation, New York, NY, USA). Population genetic differentiation was evaluated by
167 AMOVA (Excoffier *et al.*, 1992) using GenAlEx 6.5, G_{ST} for unordered (Pons & Petit, 1996) and

168 R_{ST} for ordered (Slatkin, 1995) alleles using SPAGeDi 1.4g (Hardy & Vekemans, 2002).
169 Permutation (1,000) of individuals among populations was used to test for deviation of G_{ST} or R_{ST}
170 from zero. The impact of mutations on among-population differentiation (test for $R_{ST} > R_{ST}$
171 [permuted]) was assessed using SPAGeDi by 1,000 permutations of pairwise haplotype distances
172 among pairs of haplotypes to evaluate whether distribution of haplotypes shows phylogeographic
173 structure, according to Pons & Petit (1996). Moreover, G'_{ST} (Hedrick, 2005), a standardized G_{ST}
174 estimate, was calculated using haplotype frequencies obtained by GenAlEx 6.5. The overall
175 geographical structure of genetic variation was analysed with a group clustering Bayesian method
176 as implemented in BAPS 5.2 (Corander & Marttinen, 2006), see details in Appendix S1.
177
178 **Demographic history inferred by Approximate Bayesian Computation (ABC)**
179
180 To trace the demographic history of *Q. cerris*, an Approximate Bayesian Computation (ABC)
181 procedure (Beaumont *et al.*, 2002) implemented in DIYABC v2.0 (Cornuet *et. al.*, 2014) was
182 performed, using the three clusters identified by BAPS (see Results, Fig. 2): Western Group (WG;
183 green dots in Fig. 2), Central Group (CG; blue dots) and Eastern Group (EG; red dots). DIYABC
184 implementation of ABC does not consider gene flow (but only admixture), however, our case study
185 showed high genetic differentiation and low haplotype sharing among the geographical groups used
186 for demographical inference (see Results) and, thus, no serious bias is expected. Two sets of
187 preliminary simulations were used to produce final competing scenarios (see Appendix S1), then
188 five relatively simple population demography scenarios (Fig. 3) were examined. Details of each
189 competing scenario, mutation rate, summary statistics, and model checking are given in Appendix
190 S1. In all scenarios, t# refers to timescale (scaled by generation time) and N# refers to effective
191 population size of the corresponding populations (i.e. WG, CG, EG, ancestral population "b", and
192 intermediate populations "a", "c" and "d") during each time period (i.e., $0 - t_1$ and $t_1 - t_2$) (Fig. 3).
193 Finally, because genetic diversity was the lowest in WG, all scenarios assumed that this group split

194 from CG or EG at t1 or t_2.

195

196 **Fossil data**

197

198 Fossil data were obtained through a careful bibliographic search. Different fossil types have been
199 considered to reconstruct the past distribution of *Q. cerris*. Based on pollen data, the presence of *Q.*
200 *cerris* may only be detected in areas where the distributions of *Q. cerris* and *Q. suber* do not
201 overlap, as the morphology and ornamentation of the pollen grains of the two species are
202 indistinguishable (van Benthem *et al.*, 1984). Bark has some distinctive anatomical characters (Şen
203 *et al.*, 2011), but it is seldomly found as a fossil. Wood of Quercus section *Cerris* is very similar to
204 that of the white oak group, although it can be distinguished by some features (Akkemik & Yaman,
205 2012). Several species of *Q.* Section *Cerris*, including *Q. cerris*, *Q. ithaburensis* subsp. *macrolepis*,
206 *Q. brantii*, *Q. libani*, and *Q. trojana*, currently grow in Asia Minor and cannot be distinguished by
207 fossil wood (Emery-Barbier & Thiébault, 2005). Fossil leaves and cupules are the best evidence for
208 the local presence of *Q. cerris* in an area, although identification of leaf specimens without acorns
209 may be controversial because of reported high variations in leaf morphology (Uslu & Bakiş, 2014).
210 Going back in time, the attribution of fossil leaves to *Q. cerris* may be problematic and so it is also
211 often limited to *Quercus* Section *Cerris*.

212

213 **RESULTS**

214

215 **Haplotype sequencing**

216 Allele sequencing for each cpSSR revealed no sequence differences in the microsatellite flanking
217 regions, except for cmcs9 (see details in Appendix S2 in Supporting Information).

218

219 **Genetic diversity and population structure**

The cpSSR loci ccmp6, cmcs1, cmcs6, cmcs7, cmcs8, and cmcs9 (after correction for variation in the flanking regions, see Appendix S2) were polymorphic in *Q. cerris*, and displayed 4, 2, 11, 5, 6, and 6 allele size variants, respectively. They combined into 35 haplotypes (Table S2 in Appendix S3 and Fig.1), resulting in a total haplotypic diversity of h=0.795 and a mean gene diversity over populations of h=0.109. Most populations (80.2%) were fixed for one haplotype, and only 38 populations out of 192 (19.8%) showed two or three haplotypes (35 and 3, respectively). A total of 16 private haplotypes were identified, 68.7% of which were detected in Turkey. Analysis of sequence variation at the *trnH-psbA* marker assigned all Turkish haplotypes to *Q. cerris*, although genetic distance with sister species *Q. trojana* and *Q. ithaburensis* was very low (data not shown). Genetic differentiation, as evaluated by G_{ST} and R_{ST}, was very high, 0.8807 and 0.8034, respectively, and significantly different from zero ($P < 0.05$). A phylogeographical structure was not observed (i.e., R_{ST} was not significantly larger than the permuted value).

Bayesian analysis of population structure (BAPS) resulted in three genetically distinct clusters with a well-defined geographical pattern, and which, from here on, are referred to as Western group (WG, green dots in Fig. 2), Central group (CG, blue dots) and Eastern group (EG, red dots). Each cluster distinguished a well-defined part of the natural range of Turkey oak. In particular, EG was located only in Turkey and Lebanon, with the exception of one population in North Greece. WG extended from Italy to the northern countries of the Balkan Peninsula, and CG included all remaining Balkan populations and the westernmost population from Turkey. EG had the highest genetic diversity (h=0.922; SE 0.010) compared to CG (h=0.788; SE 0.012) and WG (h=0.459; SE 0.017). G_{ST} among groups was 0.4364 and R_{ST}=0.3881, both significantly different from zero ($P < 0.05$). The standardized G_{ST} (G'_{ST}) (Hedrick, 2005), which avoids biases due to high intra-population genetic variation, was 0.955 and pairwise-G'_{ST} ranged from 0.901 (WG vs CG) to 1.000 (WG vs EG), revealing extremely high population differentiation among the three BAPS-defined groups. Furthermore, AMOVA analysis showed that a high proportion of the total genetic

variation (33% $P < 0.001$) was explained by differences among the three groups.

Demographic history inferred by Approximate Bayesian Computation (ABC)

In ABC models, the highest posterior probability was found for scenario 1 (0.7316, 95%CIs: 0.7141 - 0.7491), which posterior probability was significantly higher than that for scenario 2 (0.0036), scenario 3 (0.0197), scenario 4 (0.0237), and scenario 5 (0.2213) (Table S3 in Appendix S3). This result suggested that the divergence of EG from a hypothetical ancestral population occurred before the divergence of WG and CG, as also indicated by genetic differentiation estimates (see above). The median values of effective population sizes for this scenario were 10,500, 180,000, 249,000, 39,600 for N_1 (WG), N_2 (CG), N_3 (EG), N_a (the branch conducting to WG between t_1 and t_2), and 44,700, for N_b (ancestral population) respectively (Table 1). The posterior parameters showed that the effective population size of the ancestral population was 5.5 times lower than that of EG, thus suggesting an expansion event at time t_2. At time t1, a bottleneck and an expansion event gave rise respectively to WG and CG. The median values of the divergence times, t_1 and t_2, were 8,700 (95%CIs: 1,360-25,400) and 55,800 (95%CIs: 17,000-94,100) generations ago, respectively (Table 1). Assuming 30 years as generation time, divergence times scaled to 261,000 years (95%CIs: 40,800-762,000) and 1,674,000 years (95%CIs: 510,000-2,823,000) for t_1 and t_2, respectively. Considering a longer generation time of 50 years, these values increased to 435,000 years (95%CIs: 68,000-1,270,000) and 2,790,000 years (95%CIs: 850,000-4,705,000). All observed summary statistics were not significantly different from the simulated values (Table S6 in Appendix S3), based on parameter values drawn from the posterior distributions for scenario 1, suggesting, together with the PCA comparison implemented in DIYABC (Fig. S1), a high goodness-of-fit of observed data to scenario 1. Type I error rate was 41.0%, and the average type II error rate was 10%.

Fossil data

Fossil leaves and fruits of *Q.* Sect. *Cerris* were found in several mid-Miocene to Pliocene deposits in Germany, Poland, Slovakia, Alsace, Georgia, Moldavia, Greece, and Abkashia (cf. Song *et al.*, 2000 for a review; Teodoridis *et al.*, 2009; Worobiec *et al.*, 2012; Velitzelos *et al.*, 2014). They cannot be attributed to *Q. cerris* yet, as they belong to various fossil taxa in the Section, including *Q. cerresaecarpa* Kol., *Q. sapperi* (Menzel) Mai ex Hummel, *Q. microcerrisaecarpa* Kol. (fossil fruits), *Q. gigas* Goepp. emend. Walther & Zastawniak, and *Q. czeczottiae* Hummel and *Q. pseudocastanea* Goepp. (fossil leaves).

No Pliocene or Pleistocene fossils are available from Turkey that can demonstrate the local presence of *Q. cerris*, but fossil leaves from an Early Pleistocene deposit in Armenia (Ollivier *et al.*, 2010) suggest that the distribution of the Anatolian populations of *Q. cerris* could have been more extensive in the past than at present. Moreover, a long pollen record from the Black Sea, documenting the history of vegetation in Anatolia from 6 Ma to Present (Site 380; Biltekin *et al.*, 2015), shows that from approx. 6 Ma to 3.37 Ma deciduous oaks were very abundant within a forest-dominated vegetation with meso-megathermic taxa (e.g., Taxodiaceae) (Fig. 4). Between 3.37 Ma and ca. 1.76 Ma there was a clear reduction of oaks (including *Q. cerris*), culminating with their near-disappearance around 2.62 Ma, in correspondence with the earliest glacial cooling in the Northern Hemisphere. After 1.76 Ma deciduous oaks are found again in significant amounts; their frequencies fluctuate following the climatic cycles, although since 1.22 Ma there was a supremacy of steppe vegetation over mesophilous forests during both glacial and interglacial periods (Biltekin *et al.*, 2015).

Starting from the Pliocene, in Greece and Southeast Albania extinct taxa with Quercus foliage co-occur with species markedly similar to modern western Eurasian species of Quercus section *Cerris* (*Q. cerris* and *Q. castaneifolia*) (Velitzelos *et al.*, 2014). In Bulgaria, macrofossils of *Q. cerris* have been reported from Pliocene deposits (Bozukov & Tsenov, 2012). In various sites of

298 the Balkan Peninsula, pollen records of Middle (Fletcher *et al.*, 2013) to Late Pleistocene and
299 Holocene age (Panagiotopoulos *et al.*, 2014; Tonkov *et al.*, 2014) support the local persistence of *Q.*
300 *cerris* through the Pleistocene glacial-interglacial cycles, as also documented for undifferentiated
301 deciduous oaks (Tzedakis, 2006).

302 In Italy, macrofossils of Miocene and Pliocene age (Kovar-Eder *et al.*, 2006) are barely
303 compatible with modern *Q. cerris* (E. Martinetto, personal communication), while leaves attributed
304 to *Q. cerris* are reported from Early and Middle Pleistocene deposits (Martinetto & Sami, 2001;
305 Martinetto *et al.*, 2014). In addition, pollen data from various sites of Early Pleistocene age support
306 the presence of *Q. cerris* since at least 2 Ma (Fusco, 2007; Sadori *et al.*, 2010; Magri *et al.*, 2010).

307 In summary, fossil and pollen record evidence suggest the long-term continued presence of
308 *Q. cerris* in the Anatolian, Balkan, and Italian peninsulas. Turkey oak populations in these regions
309 probably increased in both number and size during the Pleistocene interglacial phases and
310 diminished – without ever disappearing – during the glacial periods, similarly to other deciduous
311 oaks (Fig. 4) (Tzedakis, 2006; Magri & Palombo, 2013; Biltekin *et al.*, 2015).

312

313 **DISCUSSION**

314

315 **CpDNA phylogeography and introgression with *Q. suber***

316

317 Turkey oak showed a high degree of chloroplast variation, similar to that detected in other European
318 oak species (Petit *et al.*, 2002), although not equally distributed throughout its range, with much
319 higher levels of genetic diversity found in the easternmost populations. A recent meta-analysis on a
320 high number of plant and animal species demonstrated the overall increase of genetic diversity
321 eastward in the Mediterranean basin (Conord *et al.*, 2012). The role of longitude in shaping genetic
322 diversity is particularly interesting because in many species longitudinal variation affects transitions
323 between discrete sets of genotypes, while latitudinal patterns typically involve refugial diversity

(Rokas *et al.*, 2003). This is clearly evidenced in the distribution of *Q. cerris*' private haplotypes, which have been detected mainly (about 69%) in the Anatolian peninsula. Haplotype sharing, involving QC10 and QC27, between the Italian and the Balkan regions was also observed. In particular, the presence of haplotype QC27 in the Balkans suggests possible contacts between Balkan and Italian populations of *Q. cerris* through land bridges across the Adriatic Sea. Geological studies put in evidence that the Apulian Platform connected twice, during the Early Oligocene, and then again during the latest Miocene, to the Balkans mainland by a trans-Adriatic land-bridge, approximately where the Tremiti islands are today (Patacca *et al*., 2008). In white oaks (Section Quercus), the presence of the same haplotype in the Balkan and Italian Peninsulas supports such connection (Fineschi *et al.*, 2002). The clear genetic relationship between southern Italian and Balkan populations of *Fagus sylvatica* L. (Vettori *et al.*, 2004), was also attributed to such an amphi-Adriatic connection.

Turkey oak distribution overlaps that of *Q. suber* in the Italian Peninsula and, interestingly, we found that they share some haplotypes in this region (haplotypes QC10 and QC27 in *Q. cerris* correspond to haplotypes QS2 and QS1 in *Q. suber*, described in Magri *et al.*, 2007), as shown in the haplotype network in Fig. S2. Sharing of haplotypes may result from species introgression (Petit, 2004), as reported for other oak species (Dodd & Afzal-Rafii, 2004; Eaton *et al*., 2015). Repeated asymmetrical reproductive events between species with weak reproductive barriers can lead to morphologically (and genetically) unambiguous individuals of a species with plastid signatures of another (Petit & Excoffier, 2009). This process has been suggested to be triggered (or favored) by environmental changes, disturbance of local communities and historical range fluctuations (Excoffier *et al*., 2009; Lagache *et al*., 2013). Since the pollen of *Q. cerris and Q. suber* is morphologically identical (van Benthem *et al.*, 1984), it is not possible to define the age of their coexistence from pollen records. However, the ABC analysis suggested that WG (including the Italian Peninsula), where the shared haplotypes QC10 and QC27 are found, split from CG (i.e. most of the Balkan Peninsula) at relatively recent times, even when considering generation time

uncertainty and the wide 95% CIs obtained for time divergence estimates (see below). Thus, haplotype sharing between *Q. suber* and *Q. cerris* in Pleistocene glacial refugia, as suggested by other oak species (see below), seems a reasonable hypothesis.

At the intra-generic level, a wide share of identical haplotypes has been repeatedly documented across sympatric oak taxa (Okaura *et al.*, 2007; Simeone *et al.*, 2009; 2013). With few exceptions, a stronger correlation between plastid differentiation and geographic distribution rather than taxonomic identity was always detected, which is indeed compatible with the well-known, extensive hybridization ability of oaks and its persistence in shared glacial refugia (Petit *et al.*, 2002). However, the possibility of shared ancestral polymorphism across related oak species cannot be fully ruled out (Denk & Grimm, 2010). Interestingly, interspecific genetic exchange similar to that suggested for *Q. cerris* and *Q. suber* in our study has been found between *Q. suber* and *Q. ilex* sympatric populations in the Western Mediterranean (Belahbib *et al.*, 2001; López de Heredia *et al.*, 2007). Additional studies in *Q. cerris*, *Q. suber* and *Q. ilex* using nuclear markers would be necessary to better evaluate the impact of introgression and hybridization on demographic history of Mediterranean oak species and to obtain deeper understanding on oak speciation in the Mediterranean region.

In agreement with the cpDNA haplotype distribution discussed above, BAPS analysis and genetic differentiation indexes reflected strong spatial genetic structure in *Q. cerris*, involving three distinct genetic clusters along a westward longitudinal gradient. The highest genetic differentiation was observed in Turkey, where 68.7% of private haplotypes were identified. This region has a complex geological history that influenced the biogeography of many plant and animal taxa. The Anatolian Peninsula contains many barriers to gene flow, such as mountain chains and the Central Anatolian Lake System, whose formation can be traced back to the Tertiary (Bilgin, 2011, and references therein). The presence of these geo-morphological barriers may explain the very high level of genetic differentiation and haplotype endemism in this region. Furthermore, these eastern populations may not only represent a current hotspot of genetic diversity, but also the origin of

populations now occurring in Europe (see also the discussion of ABC results below). A similar westward pattern was observed in the gallwasp *Andricus quercustozae* (Rokas *et al.*, 2003), an obligate parasite whose sexual generation develops in galls induced on oaks of the section Cerris, and whose distribution is strictly dependent on the hosts. Together with other gallwasps, *A. quercustozae* diverged in Asia Minor and expanded westwards prior to the Pleistocene glaciations.

Demographic history

A hierarchical population split model (scenario 1) was found to be the most likely explanation for the observed data using ABC, suggesting successive historical divergences of *Q. cerris* populations. However, this result has to be interpreted with caution given different sources of uncertainty in ABC inference, such as generation time, assumed models and wide confidential intervals (Tsuda *et al.*, 2015). Indeed, estimation of generation time is still controversial in forest tree species, because of their longevity, overlapping generations, variable age of first flowering, and forest replacement dynamics (Petit & Hampe, 2006). On the basis of field observations (P. Grossoni, personal communication), we considered a conservative generation time of 30-50 years for *Q. cerris*. According with these generation times, scenario 1 suggests that WG (Italy and northern Balkan Peninsula) diverged at t_1 from CG (mainly the Balkan Peninsula) 261,000 years (generation time of 30 years) to 435,000 years ago (generation time of 50 years). Although t_1 was inferred with wide 95% CIs, lower limits are still 40,800 years ago (generation time of 30 years) or 68,000 years ago (generation time of 50 years), and the divergence time clearly predates the LGM regardless of the generation time assumed. Therefore, the impact of the LGM on the demographical history of *Q. cerris* would have been limited even when the species European distribution was reduced to glacial refugia in the Italian and Balkan Peninsulas, as previously reported in other European oaks (Petit *et al.*, 2002).

The Italian and Balkan populations (WG and CG) diverged from the Eastern populations

(EG, mainly consisting of Turkish and Lebanese populations) at t_2, which was estimated 1,674,000 years to 2,790,000 years ago assuming a generation time of 30 or 50 years, respectively. The t_2 had also wide 95% CIs ranging from 510,000 years ago at lower limit (generation time of 30 years) to 4,705,000 years ago at upper limit (generation time of 50 years). Thus, despite divergence time uncertainty, these results still suggest that the ancient divergence of WG and CG from EG probably predated the Quaternary. These findings together with the higher richness of private haplotypes and higher haplotypic diversity suggest a possible origin of *Q. cerris* in the Anatolian Peninsula and/or close to the Middle East region.

The fossil data, although limited due to species identification issues, confirm the demographic history of *Q. cerris* suggested by the genetic data. In particular, the possible bottleneck leading to the divergence of the Eastern and Central populations (see Results) is also reflected in the pollen record of Site 380 in the Black Sea region, which suggests that around 2,620,000 years ago a considerable reduction in deciduous oak populations (including *Q. cerris*) occurred in Anatolia and the eastern Balkan Peninsula, followed by a very slow recovery (Fig. 4). Vigorous oak populations are found only after 2 million years ago. This historical bottleneck, which corresponds to the most important oak distribution reduction in the region during the last six million years, is in excellent accordance with the hierarchical population split model that suggested divergence times between 2,790,000 and 1,674,000 years ago (t_2). After that time, corresponding to the Gelasian, *Q. cerris* is well documented by fossils in both the Italian and Balkan Peninsulas, where it persisted up to Present. Since approx. 900,000 years, increased amplitude and duration of the glacial-interglacial cycles provoked range reductions of different magnitude, depending on the severity of climate conditions (Tzedakis, 2006; Magri & Palombo, 2013). The long pollen record from Tenaghi Philippon in Greece (Tzedakis, 2006) indicates that the most extreme tree population contractions in the region, in decreasing order, occurred during the glacial maxima of Marine Isotope Stages (MIS) 12, 16, 6, 2, 10, and 8 (Fig. 4). These stages were also the time periods when bottlenecks and genetic divergence were most likely to occur. Three of these glacial maxima,

including MIS 12 (ca. 430 kyr), MIS 10 (ca. 340 kyr), and MIS 8 (ca. 260 kyr), chronologically correspond to the divergence time t_1 suggested by scenario 1, within the range 435,000 to 261,000 years ago (Fig. 4), and support a Middle Pleistocene age for the split of a former unique population of *Q. cerris* into two populations now living in the Italian and Balkan Peninsulas, respectively.

After the genetic divergence of the Western and Central populations, new forest expansions may have led to the diffusion of haplotype QC10 into Slovenia and neighbouring countries. Extensive land exposure in the northern Adriatic, caused by low sea level during the glacial periods, may have acted as a landbridge for early-interglacial population migrations. The amphi-Adriatic distribution of haplotype QC27, supported by similar evidence in other taxa (see above), highlights the possibility that landbridges were formed also at much lower latitudes, in the southern Adriatic Sea. If so, the demographic history of *Q. cerris* populations would suggest a southern Adriatic landbridge with a more recent age than the divergence of the Western and Central populations.

The current lack of knowledge on genetic diversity and population structure of European species with an eastern distribution have resulted in the underestimation of the role of eastern refugia for European plant biodiversity. Our extensive research of the modern genetic structure of Turkey oak, *Q. cerris*, covered a remarkable gap in European oak phylogeography, showing a putative eastern origin of the species and the presence of large amounts of genetic diversity in this region.

Acknowledgments

We warmly thank several co-workers and friends who made possible to realize a so comprehensive collection of Turkey oak material throughout its distribution range. Without their help, we would not have managed to complete this research: Dalibor Ballian (Bosnia and Herzegovina), Robert Brus (Slovenia), Vasilije Isajev and Vladislava Galović (Serbia), Davorin Kajba and Želimir Borzan (Croatia), Juraj Piecka, Dušan Gömöry, Erika Gömöryová, Karol Ujházy, Margaréta Tullová (Slovakia), Jaromír Koblížek, Václav Buriánek (Czech Republic), Csaba Mátyás (Hungary),

Raphael Klumpp (Austria), Hajri Haska (Albania), Tolga Ok, Sezgin Ayan, Hakan Şevik and Emrah Cicek (Turkey), Flaviu Popescu and Dragoş Postolache (Romania), Athanassios Rubos (Greece), Kiril Sotirovski and Jane Acevski (Macedonia).

REFERENCES

Akkemik, Ü. & Yaman, B. (2012) Wood anatomy of eastern Mediterranean species. Kessel Publishing House, Germany.

Alberto, F.J., Aitken, S.N., Alía, R., González-Martínez, S.C., Hänninen, H., Kremer, A., Lefèvre, F., Lenormand, T., Yeaman, S., Whetten, R. & Savolainen, O. (2013) Potential for evolutionary responses to climate change – evidence from tree populations. *Global Change Biology*, **19**, 1645-1661.

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

Belahbib, N., Pemonge, M.H., Ouassou, A., Sbay, H., Kremer, A. & Petit, R.J. (2001) Frequent cytoplasmic exchanges between oak species that are not closely related: *Quercus suber* and *Q. ilex* in Morocco. *Molecular Ecology*, **10**, 2003-2012.

Bennett, K.D., Tzedakis, P.C. & Willis, K.J. (1991) Quaternary refugia of north European tree species. *Journal of Biogeography*, **18**, 103-115.

Bertorelle, G., Benazzo, A. & Mona, S. (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*, **19**, 2609-2625.

Bilgin, R. (2011) Back to the suture: the distribution of intraspecific genetic diversity in and around Anatolia. *International Journal of Molecular Sciences*, **12**, 4080-4103.

Biltekin, D., Popescu, S. M., Suc, J. P., Quézel, P., Jiménez-Moreno, G., Yavuz, N. & Çağatay, M. N. (2015) Anatolia: A long-time plant refuge area documented by pollen records over the last 23million years. *Review of Palaeobotany and Palynology*, **215**, 1-22.

480 Blondel, J., Aronson, J., Bodiou, J.Y. & Boeuf, G. (2010) *The Mediterranean region: biological*
481 *diversity in space and time*, 2nd edn. Oxford University Press, New York.
482 Bozukov, V.S. & Tsenov, B.V. (2012) Catalogue of the Cenozoic plants of Bulgaria (Eocene to
483 Pliocene) Addendum and Corrigendum. *Phytologia Balcanica*, **18**, 237-261.
484 Camus, A. (1938–39) Les chênes: Monographie du genre Quercus, tome 2, Genre Quercus, sous-
485 genre Euquercus (sections Lepidobalanus et Macrobalanus) Texte. Paris: Lechevalier.
486 Chen, C., Qi, Z. C., Xu, X. H., Comes, H. P., Koch, M. A., Jin, X. J., Fu C. X. & Qiu, Y. X. (2014)
487 Understanding the formation of Mediterranean–African–Asian disjunctions: evidence for
488 Miocene climate‐driven vicariance and recent long‐distance dispersal in the Tertiary
489 relict *Smilax aspera* (Smilacaceae). *New Phytologist*, **204**, 243-255.
490 Conord, C., Gurevitch, J. & Fady, B. (2012) Large-scale longitudinal gradients of genetic diversity:
491 a meta-analysis across six phyla in the Mediterranean basin. *Ecology and Evolution*, **2**,
492 2595-2609.
493 Corander, J. & Marttinen, P. (2006) Bayesian identification of admixture events using multilocus
494 molecular markers. *Molecular Ecology*, **15**, 2833-2843.
495 Cornuet, J.M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.M. &
496 Estoup, A. (2014) DIYABC v2.0: a software to make approximate Bayesian computation
497 inferences about population history using single nucleotide polymorphism, DNA sequence
498 and microsatellite data. *Bioinformatics*, **30**, 1187-1189.
499 Denk, T. & Grimm, G.W. (2010) The oaks of western Eurasia: traditional classifications and
500 evidence from two nuclear markers. *Taxon*, **59**, 351-366.
501 Dodd, R.S. & Afzal-Rafii, Z. (2004) Selection and dispersal in a multispecies oak hybrid zone.
502 *Evolution*, **58**, 261-269.
503 Eaton, D.A.R., Gonzalez-Rodriguez, A., Hipp, A.L. & Cavender-Bares, J. (2015) Introgression
504 obscures and reveals historical relationships among the American live oaks. BioRxiv
505 DOI:10.1101/016238.

506 Emery-Barbier, A. & Thiébault, S. (2005) Preliminary conclusions on the Late Glacial vegetation in
507 south-west Anatolia (Turkey): the complementary nature of palynological and
508 anthracological approaches. *Journal of Archaeological Science*, **32**, 1232-1251.
509 Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992) Analysis of molecular variance inferred from
510 metric distances among DNA haplotypes: application to human mitochondrial DNA
511 restriction data. *Genetics*, **131**, 479–491.
512 Excoffier, L., Foll, M. & Petit, R.J. (2009) Genetic Consequences of Range Expansions. *Annual*
513 *Review of Ecology Evolution and Systematics*, **4,** 481-501.Fineschi, S., Taurchini, D.,
514 Grossoni, P., Petit, R.J. & Vendramin, G.G. (2002) Chloroplast DNA variation of white
515 oaks in Italy. *Forest Ecology and Management,* **156**, 103-114.
516 Fletcher, W.J., Mueller, U.C., Koutsodendris, A., Christanis, K. & Pross, J. (2013) A centennial-
517 scale record of vegetation and climate variability from 312 to 240 ka (Marine Isotope Stages
518 9c-a, 8 and 7e) from Tenaghi Philippon, NE Greece. *Quaternary Science Reviews*, **78**, 108-
519 125.
520 Fusco, F. (2007) Vegetation response to early Pleistocene climatic cycles in the Lamone valley
521 (Northern Apennines, Italy). *Review of Palaeobotany and Palynology*, **145**, 1-23.
522 Gavin, D.G., Blois, J., Dobrowski, S., *et al.* (2014) Climate refugia: joint inference from fossil
523 records, species distribution models and phylogeography. *New Phytologist*, **204**, 37-54.
524 Gellini, R. & Grossoni, P. (1997) *Quercus cerris.* In: Botanica Forestale- II. Angiosperme. Edited
525 by CEDAM-Padova pp 124-129.
526 Gómez, A. & Lunt, D.H. (2007) Refugia within refugia: Patterns of phylogeographic concordance
527 in the Iberian Peninsula. *Phylogeography of Southern European Refugia.* (ed. by S. Weiss
528 and N. Ferrand), pp. 155-188. Springer Verlag, Dordrecht.
529 González-Martínez, S.C., Dubreuil, M., Riba, M., Vendramin, G.G., Sebastiani, F. & Mayol, M.
530 (2010) Spatial genetic structure of *Taxus baccata* L. in the western Mediterranean Basin:
531 past and present limits to gene movement over a broad geographic scale. *Molecular*
532 *Phylogenetics and Evolution*, **55**, 805-815.

533 Hardy, O.J. & Vekemans, X. (2002) SPAGeDi: a versatile computer program to analyse spatial
534 genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618-
535 620.
536 Hedrick, P.W. (2005) A standardized genetic differentiation measure. *Evolution*, **59,** 1633-1638.
537 Heuertz, M., De Paoli, E., Källman, T., Larsson, H., Jurman, I., Morgante, M., Lascoux, M. &
538 Gyllenstrand, N. (2006) Multilocus patterns of nucleotide diversity, linkage disequilibrium
539 and demographic history of Norway Spruce (*Picea abies* (L.) Karst.). *Genetics*, **174**, 2095-
540 2105.
541 Ingvarsson, P.K. (2008) Multilocus patterns of nucleotide polymorphism and the demographic
542 history of *Populus tremula*. *Genetics*, **180,** 329-340.
543 IPCC (2007) *Climate change 2007: The Physical Science Basis*. Cambridge University Press,
544 Cambridge.
545 Kovar-Eder, J., Kvaček, Z., Martinetto, E. & Roiron, P. (2006) Late Miocene to Early Pliocene
546 vegetation of southern Europe (7-4 Ma) as reflected in the megafossil plant record.
547 *Palaeogeography, Palaeoclimatology, Palaeoecology*, **238**, 321-339.
548 Lagache L., Klein E.K., Guichoux E. & Petit R.J. (2013) Fine-scale environmental control of
549 hybridization in oaks. *Molecular Ecology*, **22**, 423-436.
550 Lisiecki, L.E. & Raymo, M.E. (2005) A Pliocene-Pleistocene stack of 57 globally distributed
551 benthic δ18O records. *Paleoceanography*, 20 (PA1003).
552 López de Heredia, U.L., Carrion, J.S., Jimenez, P., Collada, C. & Gil, L. (2007) Molecular and
553 palaeoecological evidence for multiple glacial refugia for evergreen oaks on the Iberian
554 Peninsula. *Journal of Biogeography*, **34**, 1505-1517.
555 Magri, D. & Palombo, M.R. (2013) Early to Middle Pleistocene dynamics of plant and mammal
556 communities in South West Europe. *Quaternary International*, **288**, 63-72.
557 Magri, D., Vendramin, G.G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D., Latałowa, M.,
558 Litt, T., Paule, L., Roure, J.M., Tantau, I., Van Der Knaap, W.O., Petit, R.J. & De Beaulieu,

559 J.-L. (2006) A new scenario for the Quaternary history of European beech populations:
560 palaeobotanical evidence and genetic consequences. *New Phytologist*, **171**, 199-221.
561 Magri, D., Fineschi, S., Bellarosa, R., Buonamici, A., Sebastiani, F., Schirone, B., Simeone, M.C. &
562 Vendramin, G.G. (2007) The distribution of *Quercus suber* chloroplast haplotypes matches
563 the palaogeographical history of the western Mediterranean. *Molecular Ecology*, **16**, 5259-
564 5266.
565 Magri, D., Di Rita, F. & Palombo, M.(2010) An Early Pleistocene interglacial record from an
566 intermontane basin of Central Italy (Scoppito, L'Aquila). *Quaternary International*, **225**,
567 106-113.
568 Martinetto, E. & Sami, M. (2001) Paleoflora delle "sabbie gialle" mediopleistoceniche di Oriolo,
569 presso Faenza (Ravenna). *Quaderno di Studi e notizie di storia Naturale della Romagna*, **14**,
570 1-28.
571 Martinetto, E., Bertini, A., Basilici, G., Baldanza, A., Bizzarri, R., Cherin, M., Gentili S. & Pontini,
572 M. R. (2014) The plant record of the Dunarobba and Pietrafitta sites in the Plio-Pleistocene
573 palaeoenvironmental context of Central Italy. *Alpine and Mediterranean Quaternary*, **27**,
574 29-72
575 Médail, F. & Diadema, K. (2009). Glacial refugia influence plant diversity patterns in the
576 Mediterranean Basin . *Journal of Biogeography* **36**, 1333-1345.
577 Nieto Feliner, G. (2011) Southern European glacial refugia: A tale of tales. *Taxon*, **60**, 365-372.
578 Nieto Feliner, G. (2014) Patterns and processes in plant phylogeography in the Mediterranean
579 Basin. A review. *Perspectives in Plant Ecology, Evolution and Systematics*, **16**, 265-278.
580 Ohlemüller, R., Huntley, B., Normand, S. & Svenning, J.C. (2012) Potential source and sink
581 locations for climate-driven species range shifts in Europe since the Last Glacial Maximum.
582 *Global Ecology and Biogeography*, **21,** 152-163.
583 Okaura, T., Duc Quang, N., Ubukata, M. & Harada, K. (2007) Phylogenetic structure and late
584 Quaternary population history of the Japanese oak *Q. mongolica* var. *crispula*, and related

585 species revealed by chloroplast DNA variation. *Genes and Genetic Systems*, **82**, 465-477.
586 Ollivier, V., Nahapetyan, S., Roiron, P., Gabrielyan, I., Gasparyan, B., Chataigner, C., Joannin, S.,
587 Cornée, J.J., Guillou, H., Scaillet, S., Munch, P. & Krijgsman, W. (2010) Quaternary
588 volcano-lacustrine patterns and paleobotanical data in South Armenia. *Quaternary*
589 *International*, **223**, 312-326.
590 Panagiotopoulos, K., Böhm, A., Leng, M. J., Wagner, B. & Schäbitz, F. (2014) Climate variability
591 over the last 92 ka in SW Balkans from analysis of sediments from Lake Prespa. *Climate of*
592 *the Past*, **10**, 643-660.
593 Patacca, E., Scandone, P. & Mazza, P.(2008) Oligocene migration path for Apulia macromammals,
594 the Central-Adriatic bridge. *Bollettino della Società Geolologica Italiana*, **127**, 337-355.
595 Peakall, R. & Smouse, P.E (2006) GenAlEx 6: genetic analysis in Excel. Population genetic
596 software for teaching and research. *Molecular Ecology Notes*, **6**, 288-295.
597 Peakall, R. & Smouse, P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic
598 software for teaching and research-an update. *Bioinformatics*, **28**, 2537-2539.
599 Petit, R.J. (2004) Biological invasions at the gene level. *Diversity and Distributions*, **10**, 159-165.
600 Petit, R.J. & Excoffier, L. (2009) Gene flow and species delimitation. *Trends in Ecology &*
601 *Evolution*, **24**, 386-393.
602 Petit, R.J. & Hampe, A. (2006) Some evolutionary consequences of being a tree. *Annual Review of*
603 *Ecology, Evolution and Systematics*, **37**, 187-214.
604 Petit, R.J., Csaikl, U.M., Bordacs, S. *et al.* (2002) Chloroplast DNA variation in European white
605 oaks: phylogeography and patterns of diversity based on data from over 2600 populations.
606 *Forest Ecology and Management*, **156**, 5-26.
607 Pons, O. & Petit, R. (1996) Measuring and Testing Genetic Differentiation With Ordered Versus
608 Unordered Alleles. *Genetics*, **144**, 1237-1245.
609 Provan, J. & Bennett, K.D. (2008) Phylogeographic insights into cryptic glacial refugia.*Trends in*
610 *Ecology & Evolution*, **23**, 564-571.

611 Pyhäjärvi, T., García-Gil, M.R., Knürr, T., Mikkonen, M., Wachowiak, W. & Savolainen, O. (2007)
612 Demographic history has influenced nucleotide diversity in European *Pinus sylvestris*
613 populations. *Genetics*, **177**, 1713-1724.
614 Rokas, A., Atkinson, R.J., Webster, L.M.I., Csóka, G. & Stone, G.N. (2003) Out of Anatolia:
615 longitudinal gradients in genetic diversity support an eastern origin for a circum-
616 Mediterranean oak gallwasp *Andricus quercustozae*. *Molecular Ecology*, **12**, 2153-2174.
617 Sadori, L., Giardini, M., Chiarini, E., Mattei, M., Papasodaro, F. & Porreca, M. (2010) Pollen and
618 macrofossil analyses of Pliocene lacustrine sediments (Salto river valley, Central Italy).
619 *Quaternary International*, **225**, 44-57.
620 Sebastiani, F., Carnevale, S. & Vendramin, G.G. (2004) A new set of mono- and dinucleotide
621 chloroplast microsatellites in Fagaceae. *Molecular Ecology Notes*, **4**, 259-261.
622 Şen, A., Quilhó, T. & Pereira, H. (2011) Bark anatomy of *Quercus cerris* L. var. *cerris* from
623 Turkey. *Turkish Journal of Botany*, **35,** 45-55.
624 Simeone, M.C., Papini, A., Vessella, F., Bellarosa, R., Spada, F. & Schirone, B. (2009) Multiple
625 genome relationships and a complex evolutionary history in the eastern range of *Quercus*
626 *suber* L. (Fagaceae), implied by nuclear and chloroplast DNA variation. *Caryologia*, **62**,
627 236-252.
628 Simeone, M.C., Piredda, R., Papini, A., Vessella, F. & Schirone, B. (2013) Application of plastid
629 and nuclear markers to DNA barcoding of Euro-Mediterranean oaks (Quercus, Fagaceae):
630 problems, prospects and phylogenetic implications. *Botanical Journal of the Linnean*
631 *Society*, **172**, 478-499.
632 Slatkin, M. (1995) A measure of population subdivision based on microsatellite allele frequencies.
633 *Genetics*, **139**, 457-462.
634 Song, Y., Krajewska, K. & Wang, Y.F. (2000) The first occurrence of the *Quercus* section *Cerris*
635 Spach fruits in the Miocene of China. *Acta Palaeobotanica*, **40**, 153-163.
636 Teodoridis, V., Kvaček, Z. & Uhl, D. (2009) Pliocene palaeoenvironment and correlation of the

637 Sessenheim-Auenheim floristic complex (Alsace, France). *Palaeodiversity*, **2**, 1-17.
638 Tonkov, S., Lazarova, M., Bozilova, E., Ivanov, D. & Snowball, I. (2014) A 30,000-year pollen
639 record from Mire Kupena, Western Rhodopes Mountains (south Bulgaria). *Review of*
640 *Palaeobotany and Palynology*, **209**, 41-51.
641 Tsuda, Y., Nakao, K., Ide, Y. & Tsumura, Y. (2015) The population demography of *Betula*
642 *maximowicziana*, a cool-temperate tree species in Japan, in relation to the last glacial period:
643 its admixture-like genetic structure is the result of simple population splitting not admixing.
644 *Molecular Ecology*, **24,** 1403-1418.
645 Tzedakis, P. C., Hooghiemstra, H. & Pälike, H. (2006) The last 1.35 million years at Tenaghi
646 Philippon: revised chronostratigraphy and long-term vegetation trends. *Quaternary Science*
647 *Reviews*, **25**, 3416-3430.
648 Uslu, E. & Bakiş, Y. (2014) Morphometric analyses of the leaf variation within *Quercus* L. Sect.
649 *Cerris* Loudon in Turkey. *Dendrobiology*, **71**, 109-117.
650 van Benthem, F., Clarke, G.C.S. & Punt, W. (1984) Fagaceae. *The Northwest European Pollen*
651 *Flora, IV* (ed. By Punt, W. and Clarke, G.C.S.), pp. 87–110. Elsevier, Amsterdam,
652 Velitzelos, D., Bouchal, J. M. & Denk, T. (2014) Review of the Cenozoic floras and vegetation of
653 Greece. *Review of Palaeobotany and Palynology*, **204**, 56-117.
654 Vettori, C., Vendramin, G.G., Anzidei, M., Pastorelli, R., Paffetti, D. & Giannini, R. (2004)
655 Geographic distribution of chloroplast variation in Italian populations of beech (*Fagus*
656 *sylvatica* L.). *Theoretical & Applied Genetics*, **109**, 1-9.
657 Weising, K. & Gardner, R.C. (1999) A set of conserved PCR primers for the analysis of simple
658 sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms.
659 *Genome*, **42**, 9-19.
660 Worobiec, G., Worobiec, E. & Szynkiewicz, A. (2012) Plant assemblage from the Upper Miocene
661 deposits of the Bełchatów Lignite Mine (Central Poland). *Acta Palaeobotanica*, **52**, 369-
662 413.

663

664 **Supporting Information**

665 Additional Supporting Information may be found in the online version of this article:

666

667 **Appendix S1** Materials and Methods details

668 **Appendix S2** Results of haplotype sequencing

669 **Appendix S3** Supplementary tables and figures

670

671 **BIOSKETCHES**

672 The authors belong to a research consortium dealing with population genetics of forest tree species,
673 with particular emphasis to the study of genetic structure at different temporal and spatial scales.
674 They have been involved in projects studying the range-wide phylogeography of several forest tree
675 species in Europe and the Mediterranean Basin.

676

677 Author contributions: F.B., S.F. and G.G.V. designed the study; F.B., S.F., P.B., P.Z., L.P., M.C.S.
678 and G.G.V. organized the sampling; F.B., S.F. and M.C.S. performed the molecular laboratory
679 analysis; F.B., Y.T. and G.G.V. performed the statistical analysis, with advices by S.C.G.M.; D.M.
680 compiled the palaeobotanical section; F.B. led the writing; Y.T., S.C.G.M., D.M. and G.G.V.
681 contributed to the writing; all authors revised and approved the manuscript.

682

683 Editor: Stephen Cavers

Table

Table 1. Parameter estimates for best demographic scenario based on ABC.
N1= WG effective population size; N2=CG effective population size; N3=EG effective population size; Na=WG intermediate effective population size; Nb= effective population size of the ancestral population; times are considered from present (0) backwards in time, t_1=divergence of CG from WG; t_2=divergence of WG from EG; µmic=mean mutation rate of microsatellites; pmic=mean parameter of geometric distribution (GSM, Generalized Stepwise Mutation Model); snimic=individual locus SNI (Single Nucleotide Insertion/deletion) rate.

parameter	**mean**	**median**	**mode**	**quantiles**			
				2.5%	**5%**	**95%**	**97.5%**
N1	1.22E+04	1.05E+04	9.10E+03	2.55E+03	3.36E+03	2.71E+04	3.17E+04
N2	1.79E+05	1.80E+05	1.77E+05	7.53E+04	8.89E+04	2.65E+05	2.75E+05
N3	2.39E+05	2.49E+05	3.00E+05	1.28E+05	1.49E+05	2.96E+05	2.98E+05
t_1	1.05E+04	8.70E+03	3.44E+03	1.36E+03	1.86E+03	2.54E+04	2.76E+04
Na	4.40E+04	3.96E+04	2.27E+04	7.82E+03	1.03E+04	9.09E+04	9.51E+04
t_2	5.67E+04	5.58E+04	5.54E+04	1.70E+04	2.09E+04	9.41E+04	9.70E+04
Nb	4.64E+04	4.47E+04	2.86E+04	2.34E+03	4.43E+03	9.39E+04	9.67E+04
µmic	2.25E-05	1.78E-05	1.11E-05	5.43E-06	6.80E-06	5.70E-05	7.26E-05
pmic	2.10E-01	2.13E-01	3.00E-01	1.10E-01	1.17E-01	2.93E-01	2.99E-01
snimic	1.16E-06	2.59E-07	1.00E-08	1.14E-08	1.36E-08	5.77E-06	7.33E-06

Figure legends

Figure 1. Geographic distribution of *Q. cerris* cpDNA haplotypes. Haplotypes are defined by numbers (on the right). In green *Quercus cerris* distribution range.

Figure 2. Results of the spatial Bayesian cluster analysis, BAPS, showing genetically homogeneous population groups in *Q. cerris*. Western Group (WG) = green; Central Group (CG) = blue; Eastern Group (EG) = red. In green *Quercus cerris* distribution range.

Figure 3. Demographic models compared in ABC. WG = Western group; CG = Central group; EG = Eastern group; t# = divergence time scaled by generation time; N# = effective population size; ra = admixture rate; db = duration of bottleneck.

Figure 4. The pollen records of deciduous oaks from Tenaghi Philippon (Tzedakis *et al.*, 2006) and site 380 (Biltekin *et al.*, 2015) are correlated with the oxygen isotope curve by Lisiecki & Raymo (2005). t_1 and t_2 are the divergence times suggested by DIYABC best demographic model (scenario 1).

Figure S1. Principal Component Analysis of summary statistics based on 10,000 simulated datasets in DIYABC (scenario1).

Figures

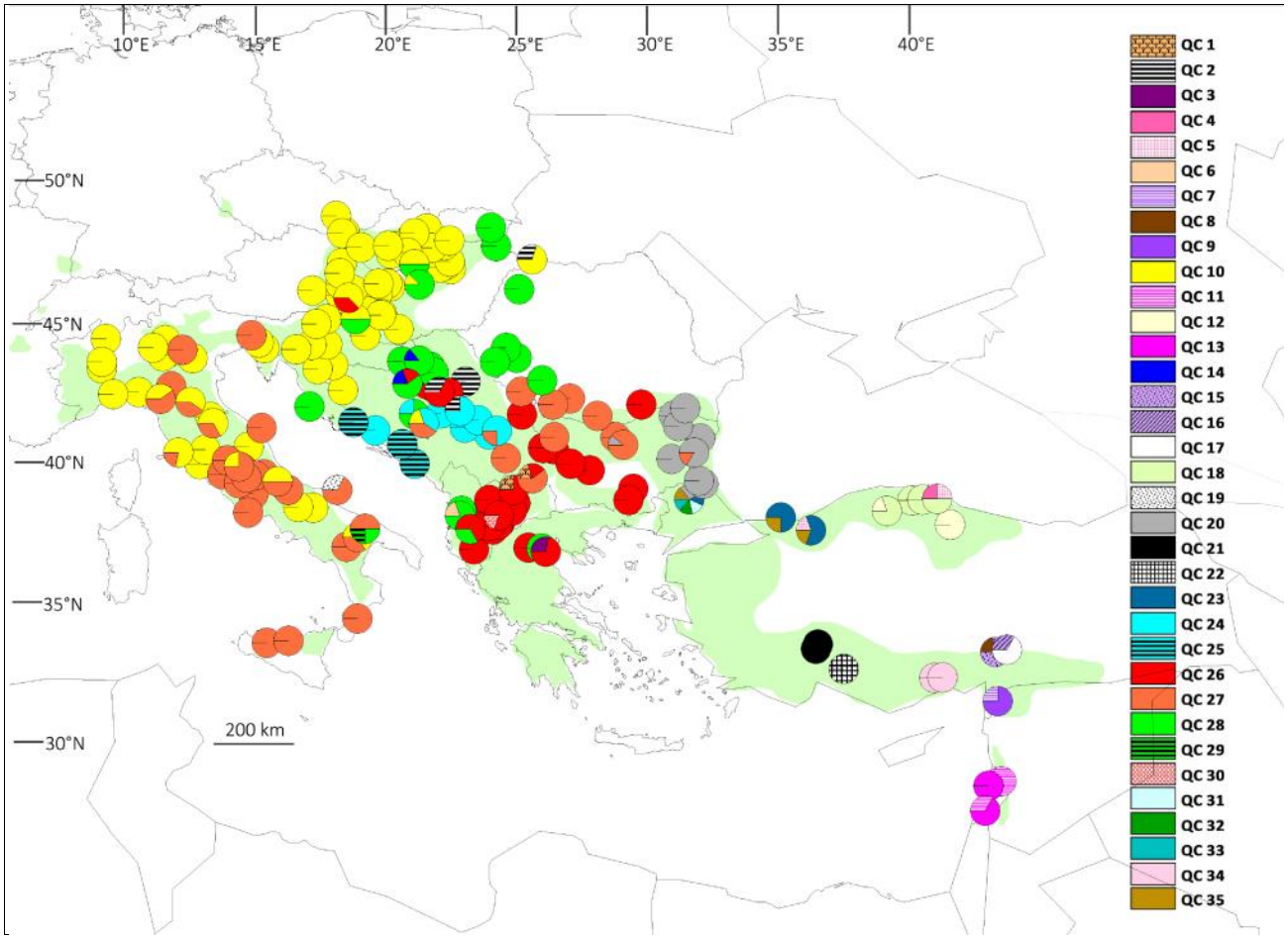
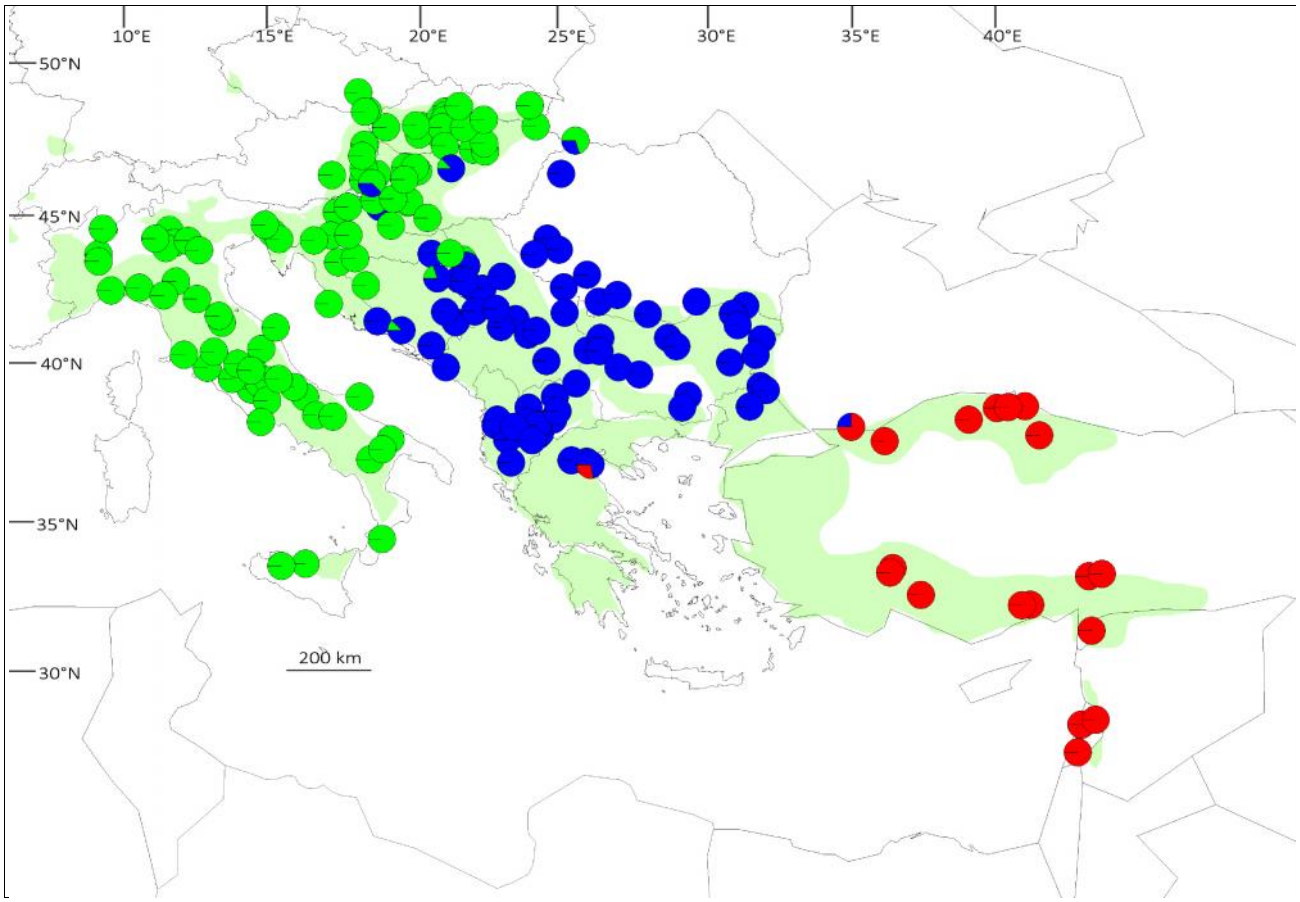


Figure 1

Figure 2



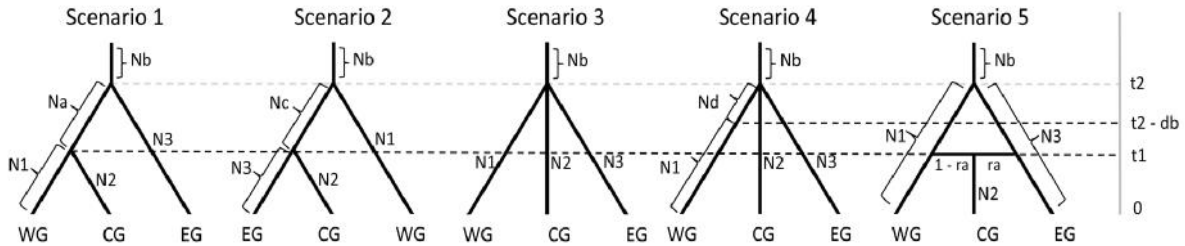
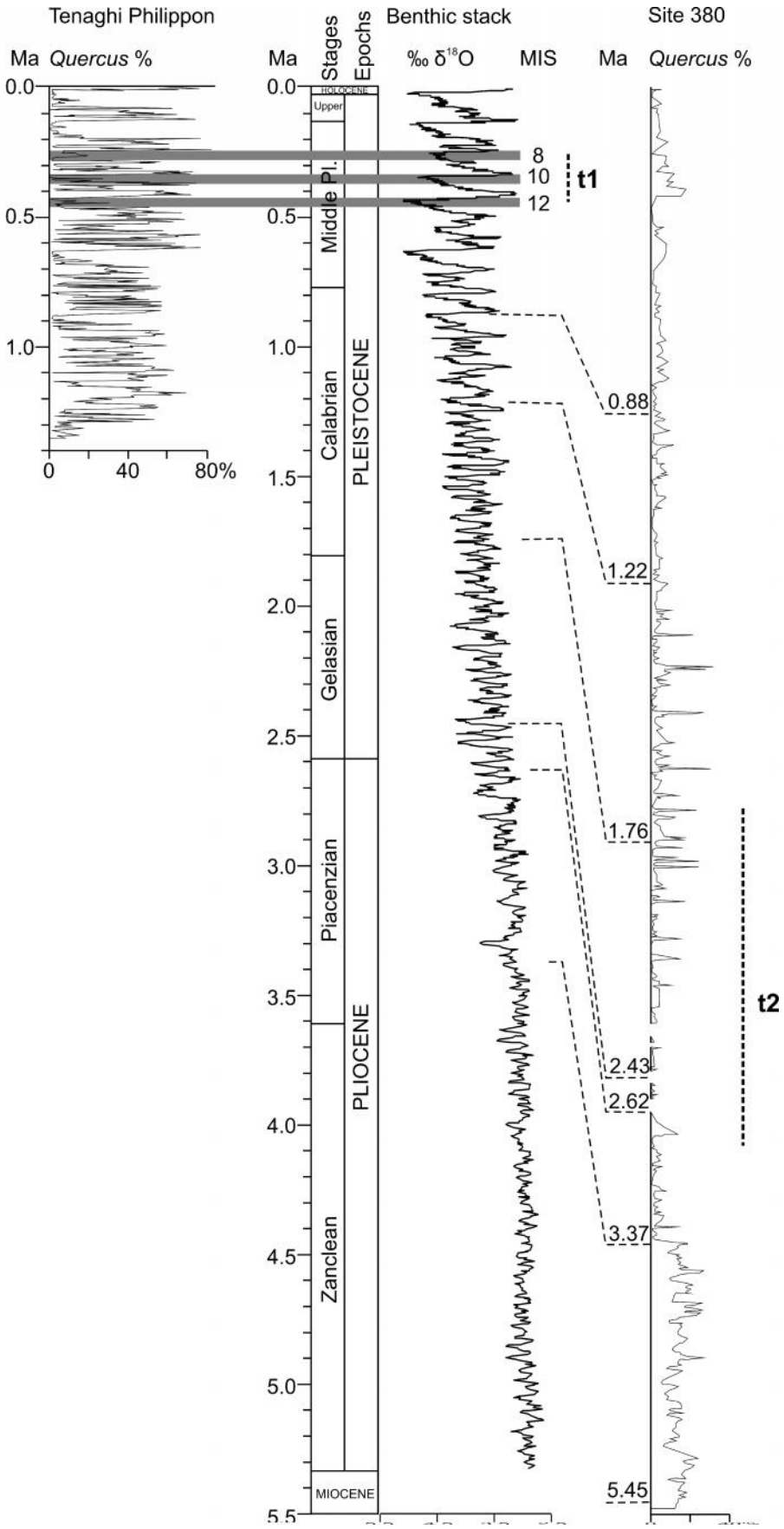


Figure 3



735
736
737 Figure 4