Tree Genetics & Genomes

Landscape genetics structure of European sweet chestnut (Castanea sativa Mill): indications for conservation priorities

--Manuscript Draft--

Manuscript Number:	TGGE-D-16-00139R2
Full Title:	Landscape genetics structure of European sweet chestnut (Castanea sativa Mill): indications for conservation priorities
Article Type:	Original Article
Funding Information:	
Abstract:	Sweet chestnut is a tree of great economic (fruit and wood production), ecological and cultural importance in Europe. A large-scale landscape genetic analysis of natural populations of sweet chestnut across Europe is applied to 1) evaluate the geographic patterns of genetic diversity 2) identify spatial coincidences between genetic discontinuities and geographic barriers 3) propose certain chestnut populations as reservoirs of genetic diversity for conservation and breeding programmes. Six polymorphic microsatellite markers were used for genotyping 1608 wild trees sampled in 73 European sites. The Geostatistical IDW technique (ArcGIS 9.3) was used to produce maps of genetic diversity parameters (He, Ar, PAr) and a synthetic map of the population membership (Q value) to the different gene pools. Genetic barriers were investigated using BARRIER 2.2 software and their locations were overlaid on a Digital Elevation Model (GTOPO30). The DIVA-GIS software was used to propose priority areas for conservation. High values of genetic diversity (He) and allelic richness (Ar) were observed in the central area of C. sativa's European distribution range. The highest values of private allelic richness (PAr) were found in the eastern area. Three main gene pools and a significant genetic barrier separating the eastern from the central and western populations were identified. Areas with high priority for genetic conservation were indicated in Georgia, eastern Turkey and Italy. Our results increase knowledge of the biogeographic history of C. sativa in Europe, indicate the geographic location of different gene pools and identify potential priority reservoirs of genetic diversity.
Corresponding Author:	Claudia Mattioni, PhD Consiglio Nazionale delle Ricerche Porano, Terni ITALY
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Consiglio Nazionale delle Ricerche
Corresponding Author's Secondary Institution:	
First Author:	Claudia Mattioni, PhD
First Author Secondary Information:	
Order of Authors:	Claudia Mattioni, PhD
	Angela Martin
	Francesca Chiocchini
	Marcello Cherubini
	Muriel Gaudet
	Paola Pollegioni
	Ivaylo Velichkov
	Rob Jarman
	Frank Chambers

	Ladislave Paule
	Vasilica Damian
	Ghita Crainic
	Fiorella Villani
Order of Authors Secondary Information:	
Author Comments:	

ORIGINAL ARTICLE

Landscape genetics structure of European sweet chestnut (*Castanea sativa* Mill): indications for conservation priorities

Claudia Mattioni*[1], M. Angela Martin[2], Francesca Chiocchini[1], Marcello Cherubini[1], Muriel Gaudet[1], Paola Pollegioni[1-8], Ivaylo Velichkov[3], Rob Jarman[4], Frank M. Chambers[4], Ladislave Paule[5], Vasilica L. Damian [6], Ghiţă C. Crainic[7] and Fiorella Villani[1].*1 Istituto di Biologia Agroambientale e Forestale (IBAF), Consiglio Nazionale delle Ricerche (CNR), Italy*

2 Departamento de Ingeniería del Medio Agronómico y Forestal, Centro Universitario de Plasencia Universidad de Extremadura, Spain

3 Forest Research Institute, Sofia 1756, Bulgaria

4 Centre for Environmental Change and Quaternary Research, School of Natural & Social Sciences, University of Gloucestershire, Cheltenham, UK

5 Faculty of Forestry, Technical University, 96053 Zvolen, Slovakia

6 University of Agricultural Sciences and Veterinary Medicine, Cluj-Napoca, Romania

7 University of Oradea, Faculty of Environmental Protection, Oradea, Romania,

8 Genomics, Genetics and Biology innovation Pole, Perugia, Italy

Running title: Genetic diversity of European sweet chestnut populations.

For correspondence: claudia.mattioni@ibaf.cnr.it

ABSTRACT

Sweet chestnut is a tree of great economic (fruit and wood production), ecological and cultural importance in Europe. A large-scale landscape genetic analysis of natural populations of sweet chestnut across Europe is applied to 1) evaluate the geographic patterns of genetic diversity 2) identify spatial coincidences between genetic discontinuities and geographic barriers 3) propose certain chestnut populations as reservoirs of genetic diversity for conservation and breeding programmes. Six polymorphic microsatellite markers were used for genotyping 1608 wild trees sampled in 73 European sites. The Geostatistical IDW technique (ArcGIS 9.3) was used to produce maps of genetic diversity parameters (He, Ar, PAr) and a synthetic map of the population membership (Q value) to the different gene pools. Genetic barriers were investigated using BARRIER 2.2 software and their locations were overlaid on a Digital Elevation Model (GTOPO30). The DIVA-GIS software was used to propose priority areas for conservation. High values of genetic diversity (He) and allelic richness (Ar) were observed in the central area of *C. sativa*'s European distribution range. The highest values of private allelic richness (PAr) were found in the eastern area. Three main gene pools and a significant genetic barrier separating the eastern from the central and western populations were identified. Areas with high priority for genetic conservation were indicated in Georgia, eastern Turkey and Italy. Our results increase knowledge of the biogeographic history of *C. sativa* in Europe, indicate the geographic location of different gene pools and identify potential priority reservoirs of genetic diversity.

Key words

Castanea sativa, sweet chestnut, microsatellite markers, population genetics, landscape genetics, GIS, conservation.

INTRODUCTION

The current distribution of plant species is shaped by complex interactions between biological, physical and socio-cultural processes. Important ecological and landscape processes take place at different multiple spatial and temporal scales, influencing the evolutionary biology of plants (Anderson et al. 2010). Substantial environmental changes have occurred since the Last Glacial Maximum (LGM) and most areas of temperate forest vegetation are believed to have been established since the onset of favorable climates through expansion from refugial populations (Petit et al. 2003). In addition, anthropogenic factors such as exploitation, land-use change and introduction of non-indigenous and/or invasive species have modified the geographical distribution of plants (Butchart et al. 2010; Rands et al. 2010). As a result, there is experimental evidence that both natural and anthropic factors lead to habitat fragmentation and a loss of biodiversity (Fischer and Lindenmayer 2007). Therefore, understanding the biogeographical distribution and the genetic diversity patterns of living species and the underlying evolutionary processes are key to being able to conserve and manage genetic resources (Cox and Moore 2005; Latta 2006).

A new field of research, called landscape genetics, represents a powerful tool for evaluating the geographic patterns of genetic resources at the population level. This discipline combines population genetics and landscape studies, using molecular and Geographic Information System (GIS) tools (Manel et al. 2003; Holderegger et al. 2010). Studies on landscape genetics aim to quantify genetic variation in space and time and its relationships to environmental conditions, providing information about the interaction between landscape features and evolutionary processes such as gene flow or local adaptation (Manel et al. 2013; Storfer et al. 2007). Moreover, the use of geo-referenced genetic data within a GIS can provide significant information on the diversity of species within a specific geographical area, so as to be able to evaluate the current conservation status, optimize the use of genetic resources, find interesting material for breeding and prioritize areas for conservation (Scheldeman and van Zonneveld 2010). In this context, geo-referenced genetic divergent gene pools and measures of genetic diversity, such as heterozygosity and/or allelic richness, can be used for identifying populations for conservation (Petit et al. 1998), since a high level of genetic variation is expected to increase the potential of populations to respond to the natural selection and healthy maintenance of individuals (Kalinowski 2004).

Landscape genetics studies are appropriate for tree species that are increasingly vulnerable to habitat fragmentation and losses of genetic diversity as a result of land-use change and land degradation. Even though several studies have reviewed the role of landscape genetics on different organisms, relatively few studies have focused on

landscape genetics of forest trees (see Lusini et al. 2014; Martin et al. 2012, 2014; Mattioni et al. 2013; Pollegioni et al. 2013, 2015; Sork et al. 2013; Storfer et al. 2010; van Zonneveld et al. 2012).

The landscape genetics approach could be very useful to indicate measures of management and conservation of genetic resources in an economically, ecologically and culturally important tree such as sweet chestnut (*Castanea sativa*), which is widely distributed in Europe. Previous studies have given an important overview on the genetic diversity and structure of chestnut populations from Spain (Martin et al. 2012), Italy, Greece, Turkey (Villani et al. 1999; Mattioni et al. 2008), Bulgaria (Lusini et al. 2014). Mattioni et al. (2013) studied the relationship between 31 chestnut populations from Greece, Turkey and Western Europe, highlighting the genetic divergence between the eastern and western European chestnut populations, confirming the probable localization of refugia as postulated by Huntley and Birks (1983) and Krebs et al. (2004) and assessing the roles of biogeographical history and human activity in determining the current distribution of the species as postulated by Conedera et al. (2004). For this new article we expanded the sampling area from Western to Eastern Europe in order to get a more representative coverage of the species distribution, particularly from central eastern Europe. We also included populations from England for the first time. We applied innovative spatial analysis to explore the power of georeferenced molecular data for understanding the spatial pattern of genetic diversity and divergence of natural populations to provide an exhaustive indication of migration, genetic structure and to locate populations or areas that deserve major attention in terms of conservation priority. This information is essential for optimizing the use of chestnut genetic resources and increasing the likelihood of finding chestnut material resistant to climate changes and biotic stresses. Chestnut natural forests and regional cultivars may harbour high levels of genetic diversity, which could provide the majority of adaptive genetic diversity for future management in response to abiotic stresses and attacks by non-indigenous species. This could be particularly relevant considering that fungal pathogens causing ink and blight disease, plus the newly introduced Asian chestnut gallwasp, together account for a significant recent reduction of sweet chestnut's range and productivity (Vannini et al. 2010; Ježić et. al. 2014). Thus, we structure this new work, combining the data of 47 populations already obtained in previous researches (Martin et al.2012; Mattioni et al. 2013; Lusini et al. 2014) (data from Turkish, Greek, Italian, Spanish and Bulgarian populations) with new analysis of additional 26 populations from Spain, Portugal, France, England, Slovakia, Hungary, Romania, Georgia, Russia and Azerbaijan. Our objectives were to apply an innovative spatial analysis to 1) improve understanding of the geographic distribution of sweet chestnut genetic diversity; 2) identify spatial coincidences between genetic discontinuities and geographic barriers; 3) give indication for prioritizing areas for sweet chestnut conservation.

MATERIALS AND METHODS

Plant material and DNA extraction

A total of 1608 wild chestnut trees were included in the analysis; 416 of these were newly sampled in 26 sites located in 10 countries (Spain, Portugal, France, England, Slovakia, Hungary, Romania, Georgia, Russia and Azerbaijan) and analyzed by means of molecular markers. Furthermore, the raw data of 1192 samples, collected in 5 different European countries (Italy, Spain, Greece, Turkey, Bulgaria) and previously analyzed in different works (Mattioni et al 2013; Martin et al 2012; Lusini et al. 2014) were added to the statistical analysis. (Table 1 and Figure1). All the sampling sites refer to wild populations, natural or naturalized, excluding orchard or recent forest plantations. Dormant buds or leaves from each plant were sampled from 10 to 20 randomly selected individual trees per site. Individual trees were selected at least 20 metres apart, so as to minimize the sampling of close relatives. Total genomic DNA was isolated by grinding 50-60 mg of fresh tissue and using the DNeasy 96 Plant Kit (Qiagen) according to the manufacturer's instructions.

Microsatellite analysis

A set of six polymorphic microsatellite markers (CsCAT1, CsCAT3, CsCAT6, CsCAT16, EMCs25, EMCs38) developed for *C. sativa* (Buck et al. 2003; Marinoni et al. 2003) was selected and used for the analysis.
These markers were mapped on F_1 intraspecific cross (*C. sativa* x *C. sativa*) and each selected locus has been shown to belong to a different linkage group (Barreneche et al. 2004). The unbiased probability of identity (PIunb) (Paetkau et al. 1998) computed for the combination of the six markers was PI=0.00. This value indicates the probability that two unrelated trees selected at random from a population would have identical genotypes at multiple loci: the lower this value, the higher is the capacity of the markers used to capture the variability present in the data set. Polymerase chain reaction (PCR) was carried out on a GeneAmp 2700 Thermal Cycler (Applied Biosystems, Foster City, USA). The PCR reactions were performed in 20 µl total volume containing 20 ng of genomic DNA following the Qiagen multiplex PCR kit protocol. Cycling parameters were as follows: 15 mins at 95 °C; 30 cycles for 30 secs at 94 °C, 90 secs at 57 °C and 1 min at 72 °C; and a final step of 30 mins at 72 °C. Amplification products (1 µL) were added to 20µL formamide and 0.3 µL LIZ and denaturated at 95 °C for 5 mins. The samples were run on ABI Prism 3130 Avant DNA sequencer. The resulting raw data were collected applying GeneMapper software (Life Technologies). The alleles were determined by automated binning and checked by visual inspection.

Data analysis

Genetic diversity and population genetic structure.

A set of measures of intra- and inter-population genetic statistics were calculated using the GeneAlEx 6.5 software (Peakall and Smouse 2005): observed (A) and effective (Ae) number of alleles, observed (Ho) and expected (He) heterozygosity (Nei 1973) and unbiased estimate of mean expected heterozygosity (UHe) (Nei and Roychoudhury 1974). The estimation of mean number of alleles per locus as a measure of allelic richness can be affected by differences in sample size (Kalinowski 2004). For this reason, allelic richness (Ar) and private allelic richness (PAr), which are independent of sample size (El Mousadik and Petit 1996), were computed by the rarefaction method with the HP-rare software (Kalinowski 2005). This approach uses the frequency of alleles at a locus to estimate the expected number of alleles and/or private alleles in a sub-sample of n individuals selected at random from a sample of N individuals in each population. The n value corresponds to the smallest number of individuals typed for a population. In our data set the smallest number of individuals was nine. The estimated null allele frequency for each locus was calculated using the FREENA software (Chapuis and Estoup 2007). The inbreeding coefficient (Fis) for each population over all loci was computed using hierarchical AMOVA as implemented in Arlequin 3.1.1 software (Excoffier 2005). The statistical significance of Fis was tested using a non-parametric approach, described by Excoffier et al. (1992), using 1000 permutations. The Inverse Distance Weighted (IDW) (Shepard 1968; Hengl, 2009) algorithm implemented in ArcGIS 9.3 (ESRI, Redlands, Calif. USA) was used to interpolate values of expected (He) heterozygosity, allelic richness (Ar) and private allelic richness (PAr) of all 73 chestnut populations and to derive maps of genetic diversity. The IDW algorithm is a method of interpolation that estimates the values of target variables at a new location using a linearly weighted combination of a set of samples points. IDW is based on the assumption that things that are close to each other are more related than things that are a long away from each other. It weights the points closer to the prediction location greater than those farther away.

To delineate genetic repartition of chestnut populations, an UPGMA phylogenetic tree (Sneath and Sokal 1973) was constructed based on Nei genetic distance (Nei 1972) using the software POPTREE2 (Takezaki et al. 2010) and visualized with the software Fig Tree 1.4.2 (Rambaut 2009).

Subsequently, a Bayesian approach was performed using the software STRUCTURE v.2.3.4 (Pritchard and Wen 2004). This method attempts to reveal the population structure by placing individuals in most likely *K* number of clusters. STRUCTURE was run with the option of including prior information on the spatial location of populations and using the admixture model on the whole dataset and the correlated allele frequencies (Falush et al. 2007;

Hubisz et al. 2009). Based on the initial results, six independent runs (from 1 to 6) were performed for each *K* value, with a burn-in period of 10000 steps followed by 10^5 MCMC replicates. To identify the number of clusters (*K*) that best explained the data, the rate of change on L(*K*) (*ΔK*) between successive *K* values was calculated according to Evanno et al. (2005) using STRUCTURE HARVESTER software (Earl and von Holdt 2012). The six runs for each simulation were averaged using CLUMPP 1.1.2 software (Jakobsson and Rosenberg 2007) and represented graphically with DISTRUCT (Rosenberg 2004). The groups indicated by the STRUCTURE analysis were subsequently analysed separately in order to identify subgroups (K´) within each cluster. A spatial interpolation of population membership values (Qi) in the inferred K=3 and K'=3 clusters estimated by STRUCTURE and CLUMPP was calculated using the IDW algorithm; and K clustering surface maps were produced to display the spatial patterns of the inferred genetic clusters.

The presence of genetic barriers, corresponding to geographic zones of sharp genetic variation among populations, was investigated using Monmonier's maximum difference algorithm as implemented in BARRIER software 2.2 (Manni et al. 2004). The locations of genetic barriers were overlaid on a Digital Elevation Model (GTOPO30) in order to reveal overlaps between geographical and genetic discontinuities. The geographic coordinates of each sampling site were connected by Delauney triangulation and the corresponding Voronoï tessellation was derived. Once the network was obtained, each edge of the polygons was associated to its pairwise Nei's (1973) genetic distance (D_A) calculated using GENDIST in the PHYLIP software package (Felsenstein 2005). We tested from 1 to 7 genetic barriers and their significance was evaluated by means of 100 resampled bootstrap matrices of Nei's (1973) genetic distances. Genetic barriers with arbitrary bootstrap support of $P > 0.50$ were considered.

Prediction of the priority areas for conservation

The DIVA-GIS software (www.diva-gis.org) was used to predict the areas with priority for conservation. We used the Reserve Selection analysis that is based on an optimization algorithm developed to minimize the area needed to conserve flowering plants in South Africa (Rebelo and Siegfried 1992). This analysis identifies the minimum number of geographic units needed to conserve all genetic diversity and, in sequence of importance, the geographic units that should be prioritized for conservation. The first population chosen has the highest allelic richness, then each successive population is selected to best complement the intraspecific diversity already represented within the previously selected populations. That means that the second priority cell is not the one with the second highest level of diversity, but that with a number of new alleles not found in the previous population.

RESULTS

Populations' genetic diversity

Genetic diversity estimates for each population are reported in Table 2. The number of alleles per locus (A) ranged between 4.20 (RO02, RO03) and 9.80 (TR08, BU01). The highest number of effective alleles (Ae) was observed in some populations of Italy and central and western Turkey (IT06=5.57, TR08= 5.51 and TR13= 5.41). A positive and significant Fis was found in all populations from eastern Turkey and in some populations from western and central Turkey (TR01, TR06, TR08, TR11, TR12), Spain (SP09, SP16, SP17), Bulgaria (BU05), Greece (GR01), Georgia (GE01, GE04), France (FR02), and Russia (RU01) (Table 2). Figure 2 displays geographically the values of expected heterozygosity (He), allelic richness (Ar) and richness of private alleles (PAr) reported in Table 2. The geospatial interpolation of those indices enabled us to produce new spatial data representative of the genetic diversity of the European sweet chestnut populations (Fig.2). High values of He were observed in populations located in the central area of the species' distribution. Some populations from western Turkey, Greece, Bulgaria and Italy, the population from England and some populations from Spain showed high values of expected heterozygosity, while lower values of He were shown in the eastern Turkey populations, in two Italian populations and in some Spanish and French populations (Fig. 2a). The allelic richness (Ar) values ranged between 3.68 (SK02) and 7.00 (TR13) (Table 2). Higher values of allelic richness were observed in populations from the central geographic area of the species distribution and from some populations in eastern Turkey and Georgia (Fig. 2b). Evident differences between populations were observed for the values of private allelic richness (PAr) (Fig. 2c): the highest values were observed in the eastern area of the distribution and decreased to the west. The populations from eastern Turkey (TR03 and TR04) and Georgia (GE02) had the highest number of private allelic richness, while lower values were observed in the other regions. Spots of high values of PAr are found in some Spanish, Slovakian, Romanian and Bulgarian populations.

Genetic structure of populations

The UPGMA (Fig. 3) analysis, based on Nei genetic distance, indicates the presence of three main clusters. Clusters 1 and 2 (represented with the blue and green color) are more genetically similar than cluster 3 (represented with the red color). Cluster 1 includes the populations from eastern Turkey and from Azerbaijan, Georgia and Russia and only one population from Romania. Cluster 2 includes populations from western Turkey, Greece, and Bulgaria. All the western European populations are included in Cluster 3.

The Bayesian analysis is in accordance with the UPGMA results. The STRUCTURE analysis indicated how many genetic groups were contained in the data set. The highest ΔK was observed at K= 2 and a second ΔK peak was detected at K= 3 (Online Resources 1). The estimated population structure inferred from this analysis is shown in Figures 4a and 4b. At K=2, the populations from eastern Mediterranean, eastern Europe and Turkey were clearly separated from the western European populations (Fig.4a). Considering K=3, the Italian, Spanish, Slovakian and Hungarian and only two populations from Romania were included in the same cluster (red). The populations from eastern Europe appeared separated in two main groups; the cluster indicated with the green color included one population from Romania (RO03) and the populations from Bulgaria, Greece and western Turkey, whereas the populations from Russia, Azerbaijan, Georgia and eastern Turkey were grouped in the second blue cluster (Fig. 4a). In order to check the possibility of substructure, the populations belonging to the red cluster (western and central Europe populations) were analyzed separately. The most likely substructure was detected at K'=3, where three main gene pools were detected. The populations from Spain belong to two different gene pools: the southern, central and north-eastern populations belong to sub-cluster 1 (red), while the north-western Spanish populations belong to sub-cluster 3 (blue) together with the Slovakian and Hungarian populations. The Italian, English and French populations belong to the sub-cluster 2 (green) with a degree of admixture with the blue sub-cluster; note also the presence of individuals belonging to the red sub-cluster in the IT01 population, geographically located in south Italy. In Figure 4b and 4c synthetic maps show the spatial representation of the estimated population membership values (Q) in the K=3 clusters and in the substructure K'=3 sub-clusters inferred by STRUCTURE. These maps were produced by combining the three single clustering surface maps (not represented) for each cluster obtained by interpolating the population's Q values using the GIS software. In the synthetic maps (Fig. 4b, 4c), for each population three raster cell values indicate the estimated population Q-membership percentage in cluster 1, cluster 2 and cluster 3, plus a combination of the three colored components (green = cluster 1, red = cluster 2, blue = cluster 3). In this figure, the more intensely colored area indicates the strongest genetic similarity between populations belonging to the same cluster and the gradual change indicates a gradual decrease in the genetic similarity; different hues correspond to different combinations of three components (Lusini *et al.*, 2014).

In Figure 4b, the genetic barriers, calculated with BARRIER software, are indicated with the black line. Using Monmonier's maximum difference algorithm, two main statistically significant genetic barriers (bootstrap support ≥ 75%) are identified. The main genetic barrier divides the eastern European populations from the central and western populations. Interestingly, a second barrier separates the Bulgarian, Greek and western Turkish populations from the central and eastern Turkish and Caucasian populations. The hierarchical AMOVA was

carried out based on the grouping for K=3. The inferred molecular variance among the three clusters was 12.18% (P<0.001), while the molecular variance among populations within groups was 11.56% (P<0.001) and the highest variance was observed within populations (76.26% P<0.001). (Online Resources 2)

Conservation priority

Figure 5 reports the results obtained using the Reserve selection procedure of DIVA-GIS software. This analysis defines the minimum number of geographic units needed to conserve all genetic diversity. The areas indicated with the red dots are those with priority for genetic conservation. According to the size of the red points the geographic units that should be prioritized for conservation were identified, in sequence of importance: the larger the red point, the higher is the priority for genetic conservation. Areas with higher priority are indicated in Georgia, eastern Turkey and Italy. Some areas of western Turkey, Bulgaria, Greece, Slovakia and Romania are also considered but with lower priority for conservation.

DISCUSSION

Owing to extensive forest fragmentation, world-wide *in situ* and *ex situ* conservation plans must consider all levels of biodiversity including the genetic component (Suoto et al. 2015). Intraspecific genetic variation is considered a key factor for the ability of a species to survive under changing environmental conditions and it is a fundamental criterion for developing effective conservation strategies (Eckert et al. 2008). It is essential, therefore, to evaluate the geographical patterns of genetic diversity and identify the populations and areas that show high values of genetic diversity and divergence so as to identify which populations merit the most attention in terms of conservation priority. However, until now, practitioners have seldom taken into account the relevance of genetics research (Bowman et al. 2016) and few studies on landscape genetics have been applied in practical management (Keller et al 2015). From our point of view landscape genetics can be considered an easy and self-explaining tool to transfer information about spatial distribution of genetic variation into practice In the present work, the application of the innovative spatial analysis to a more representative dataset of chestnut European distribution validates the results previously published (Mattioni et al. 2013) and provides more exhaustive and critical information on the genetic diversity and biogeographical history of sweet chestnut, highlighting priority conservation areas. We georeferenced genetic diversity values of 73 European sweet chestnut populations, which are considered broadly representative of the European distribution range (based on present knowledge) and we provided clear self-explanatory maps, easily usable even by the non-scientific community. In this respect, the intra-populations genetic diversity and population structure analysis, together with the outputs from the spatial overlay

of the three maps of expected heterozygosity (He), allelic richness (Ar), private allelic richness (PAr), and the spatial localization of the different gene pools, can be considered valuable information to be used in programs for conservation of genetic resources of chestnut in Europe. Different approaches are often used for making conservation choices including measures of diversity, but rare are the cases in which morphological and demographic parameters are integrated with genetic data (Vinceti et al 2013). In situ conservation should consider populations in which inbreeding and loss of alleles are minimal. In this context, allelic richness and He are very informative measures of genetic diversity considered ideal to determine priorities for conservation (Hollingsworth et al. 2005; van Zonneveld 2012). The DIVA-GIS software can be considered as an informative tool to highlight priority sites based on species richness (Cadima et al. 2014; Phillips et al. 2014) or genetic diversity data as reported in *Prunus africana* and *Annona cherimola* (Vinceti et al 2013; van Zonneveld et al. 2012).

In our study, we applied the Reserve Selection analysis (DIVA-GIS software) to indicate areas as reservoirs of genetic diversity for sweet chestnut. Some populations located in areas from central and eastern Europe, which have different environmental conditions, are indicated with priority for conservation. These high-priority areas are characterized by high allelic richness and they are selected, by the DIVA-GIS procedure, to capture the maximum amount of diversity in the data set. Interestingly the populations with high priority for conservation are representative of the three main observed gene pools. Of course, this is a large-scale indication that gives information disregarding genetic and conservation resources of particular interest at a local scale, such as the varietal diversity found in ancient sweet chestnut orchards or the historical and natural interest of ancient sweet chestnut trees and coppices. Comparing the spatial representation of He, Ar and PAr there is now the possibility to speculate on the biogeographic history of European sweet chestnut. Comps et al. (2000) report in beech a reduced level of allelic richness and an increase of He in newly colonized areas. These findings indicate a quite different dynamic for these two genetic diversity measures: allelic richness may be more useful than He to identify historical processes such as bottlenecks and populations admixture (Widmer 2001). In this context we decided to represent spatially both the He and Ar to give a clearer view of the biogeographical history of sweet chestnut.

Interesting information was obtained considering the marginal areas of distribution. Higher values of He and Ar were observed in the central area of the species' distribution, while values of PAr were higher in the eastern and in some isolated western populations. This can be explained by considering that fragmentation could increase towards the limits of the species' distribution and that fragmented populations are often characterized by a decrease of genetic diversity (Eckert et al. 2008). We can also argue that the high intra-population diversity (He) and allelic richness (Ar) observed in the central area of distribution could be a consequence of mixing colonization routes

from different refugia, as observed in other species (Petit et al. 2003), or could be a result of human introduction of material. On the contrary, the high value of PAr and the positive Fis values and low He observed in the eastern European populations could be an indication of long-term geoclimatic separation of this area from the west. These results are congruent with those obtained in the previous work of Mattioni et al. (2013), in which a higher level of PAr and positive Fis values were reported in the presumed refugia in eastern Turkey and central Italy, possibly suggesting that no human-mediated gene flow has erased ancient signs of refugia.

The structure analysis confirms the results previously obtained with the restricted number of populations but highlights new interesting findings that enable us to speculate on the biogeographical history of chestnut. The populations structure inferred for K=3 shows the presence of three distinct gene pools. The percentage of variation among groups calculated with the AMOVA was 12%; this value is comparable with those obtained in other forest tree species using microsatellite markers (Bagnoli et al 2009, 2016; Mayol et al 2015; Pollegioni et al. 2014).

A significant genetic barrier divides the western and central European populations (including populations in Slovakia, Hungary and northern Romania) from eastern European populations. Considering the structural and geomorphological complexity of the Mediterranean regions, it is likely that geographical barriers have interfered with the gene flow between chestnut populations. The southern Carpathians, the Dinaric Alps and the Adriatic Sea could act as obstacles to the gene flow between eastern and western European populations.

Analyzing the structure of western and central European populations (Portuguese, Spanish, Italian, English, French, Slovakian, Hungarian and northern Romanian populations), has revealed three distinct gene pools: the first includes the north-eastern, central and southern Spanish populations; a second group includes the north-western Spanish populations; the third the Italian, French and English populations. These results are congruent with the existence of glacial refugia in those areas (Martin et al, 2012; Gomez-Sanz et al. 2002; Krebs et al. 2004). Interestingly, the north-western Spanish populations are genetically similar to the Slovakian, Hungarian and northern Romanian populations. A similar structure of populations has been observed for other tree species (Petit et al. 2003) and can be explained by considering that the Pyrenees may not have formed a barrier to colonization (Hewitt 2000): northern Spain, western France, Britain and Ireland were linked during the early postglacial, owing to lower sea level across Biscay and the English Channel (Provan and Bennett 2009; Beatty and Provan 2012). Moreover, the pattern of present divergence and distribution, combined with evidence from fossil pollen of several tree species, could suggest that populations originating from Italy or the Balkans spread into the Iberian Peninsula from the north (Petit et al. 2003; Grivet and Petit 2003). In addition to the hypothesis of natural colonization, the present genetic structure of western European populations of sweet chestnut can also be a result of human-mediated

transport of plant material, because this species had great economic and social importance in many civilizations. The use of sweet chestnut (for food and wood) is known to have become widespread throughout the Roman Empire, while during the Medieval period chestnut cultivation and utilization became much commoner (Conedera et al. 2004; Squatriti 2013). For these reasons, we cannot exclude the possibility that the genetic similarity among the populations from Portugal, Italy, France and England is a consequence of human transplantation of chestnut material (nuts or living saplings/branches). This hypothesis could be true also for the population from Sicily (IT01) that is genetically similar to the eastern Spanish populations. In conclusion, this work integrating the existing and new genetic data with spatial analysis techniques, provides valuable large scale information on genetic diversity and biogeographical history of *C. sativa* in Europe. The easily understandable outputs are usable in the inventory, conservation, and management of chestnut genetic resources.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. I. Beritognolo and Dr. D. Postolache for their critical review of the manuscript. Thanks are due to all the people who collected chestnut samples, especially Dr. P. Aravanopoulos, Dr. R. Botta, M. Akhalkatsi, Z. Manvelidze, V.D. Leiba, M.D. Pinkovskij, A.V. Romanishin and O. Ibadlo for providing us the leaf samples from northern Italy, Greece, Georgia, Russia and Azerbaijan, respectively. M.A. Martín is grateful to the Secretaría General de Ciencia, Tecnología e Innovación de la Consejería de Economía e Infraestructuras from the Regional Government of Extremadura (Spain) for her the financial support.

DATA ARCHIVING STATEMENT

The data are currently being submitted to the TreeGenes Database: the accession number will be supplied when available.

REFERENCES

Anderson CD, Epperson BK, Fortin MJ (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. Mol Ecol 19: 3565-3575

Bagnoli F, Vendramin G G, Buonamici A et al. (2009) Is *Cupressus sempervirens* native in Italy? An answer from genetic and paleobotanical data. Mol Ecol 18: 2276-2286

Bagnoli F, Tsuda Y, Fineschi S et al. (2016) combining molecular and fossil data to infer demographic history of Quercus cerris: insights on European eastern glacial refugia. J Biogeogr 43: 679-690

Barreneche T, Casasoli M, Roussel G, et al (2004) Comparative mapping between *Quercus* and *Castanea* using simple sequence repeats (SSRs). Theor Appl Genet 108: 558-556

Beatty GE, Provan J (2012**)** Post-glacial dispersal, rather than in site glacial survival, best explains the disjunct distribution of the Lusitanian plant species *Daboeacia cantabrica* (Ericaceae). Jof Biogeogr 40: 335-344

Bowman J, Greenhorn JE, Marrotte R.R, McKay MM, Morris KY, Mrentice MB, Wehtje M (2016) On application of landscape genetics. Conserv Genet18 March on line

Buck EJ, Russel K, Hadonou M, James CJ, Blakesley D, Russell K (2003) Isolation and characterization of polymorphic microsatellite in European chestnut (*Castanea sativa* Mill.). Mol Ecol Notes 3: 239-241

Butchart S, Walpole M, Collen B, Strien VA (2010**)** Global biodiversity: indicators of recent declines. Science 328:164-1168

Chapius MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol 24:621-631

Cadima X, van Zonneveld M, Scheldeman X et al (2014) Endemic wild potato (*Solanum* spp) biodiversity status in Bolivia: Reasons for conservation concerns. J Nat Conserv 22:113-131

Conedera M, Krebs P, Tinner W, Prandella M, Torrani T (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

Cox B, Moore P (2005) Biogeography: an ecological and evolutionary approach. John Wiley & Sons, Chichester

Earl DA, von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing structure output and implementing the Evanno method. ConsGenet Res 4: 359-361

Eckert CG, Samis E, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. Mol Ecol 17**:**1170 -1188

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of argan tree [*Argania spinose* (L.) Skeels] endemic to Morocco. Theor Appl Genet 92: 832-839

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611– 2620

Excoffier L, Smouese PE, Quattro JM (1992) Analysis of molecular variance from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479-491

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7:574-578

Fischer J, Lindenmayer, DB (2007) Landscape modification and habitat fragmentation: a synthesis Global Ecol Biogeogr 16: 265–280

Felsenstein J (2005) PHYLIP (phylogeny inference package). Department of Genome Sciences, University of Washington, Seattle.

Gomez-Sanz V, Blanco-Andray A, Sánchez-Palomares O, et al (2002) Autoecología de los castañares andaluces. Investigation Agraria. Sistemas y Recursos forestales 11 :205–226

Grivet D, Petit RJ (2003) Chloroplast DNA phylogeography of hornbeam in Europe: evidence for a bottleneck at the outset of glacial colonization. Conserv Genet 4: 47-56

Hengl T (2009) A practical guide to geostatistical mapping. University of Amsterdam, Amsterdam

Hewitt GM (2000) The genetic legacy of the Quaternary ice age. Nature 405: 907-913.

Holderegger R, Buehler D, Gugerli F, Manel S (2010) Landscape genetics of plants. Trends Plant Sci15:675–683

Hubisz MJ, Falush D, Stephen M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Res 9: 1322-1332

Huntley B, Birks HJB (1983) An Atlas of Past and Present Pollen Maps for Europe: 0-13,000 years ago. CUP, Cambridge

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801-1806

Jezic M, Krstin L, Poljak I, et al. (2014**)** Castanea sativa: genotype-dependent recovery from chestnut blight. Tree Genet Genomes 10: 101-110

Kalinowski ST (2004) Counting alleles with rarefaction: private alleles and hierarchical sampling design. Conserv Genet 5: 539-543

Kalinowski ST (2005) HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. Mol Ecol Notes 5:187-189

Keller D, Holderegger R, van Strien MJ, Bolliger J (2015) How to make landscape genetics beneficial for conservation management? Conserv Genet 16:503–512

Krebs P, Conedera M, Prandella D, Torrioni D, Felber M, Tinner W (2004) Quaternary refugia of sweet chestnut (*Castanea sativa* Mill.): an extended palynological approach. Veget Hist and Archaeob 13: 145-160

Lusini I, Velichkov I, Pollegioni P, et al (2014**)** Estimating the genetic diversity and spatial structure of Bulgarian *Castanea sativa* populations by SSRs: implication for conservation. Conserv Genet 15: 283-293

Manel S, Schwartz MK, Luikhart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. Trends Ecol and Evol 18: 189–197

Manel S, Holderegger R (2013) Ten years of landscape genetics. Trends Ecol Evol 10: 614-621

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, and linguistic) variation: how barriers can be detected by using Monmonier's algorithm. Hum Biol 76: 173-190

Martin MA, Mattioni C, Molina JR, et al (2012) Landscape genetic structure of chestnut (*Castanea sativa* Mill.) in Spain. Tree Genet Genomes 8: 127-136

Martin MA, Mattioni C, Lusini I, et al (2014) New insights into the genetic structure of Araucaria araucana forests based on molecular and historic evidences. Tree Genet Genomes 10: 839-851

Marinoni D, Akkak A, Bonus G, Edwards KJ, Botta R (2003) Development and characterization of microsatellite markers in *Castanea sativa* (Mill.) Mol Breeding 11:127-136

Mattioni C, Martin MA, Pollegioni P, Cherubini M, Villani F (2013) Microsatellite markers reveals a strong geographical structure in European populations of *Castanea sativa* (Fagaceae): Evidence for multiple glacial refugia. Am J Bot 100: 1-11

Mayol M, Riba M, Gonzalez-Martinez S C et al. (2015) Adapting through glacial cycles : insights from a long-lived tree (*Taxus baccata*) New Phytologist 208: 973-986

Nei M (1973) Analysis of gene diversity in subdivided populations. P Natl Acad Sci USA. 70: 3321-3323

Nei M, Roychoudhry AK (1974) Sampling variance of heterozygosity and genetic distance. Genetics 76:379-390

Peakall R, Smouse PE (2005) GeneAlex6: genetic analysis in excel. Population genetic software for teaching and research. Australian National University, Canberra, Australia. http://anu.edu.au/BoZo/GenAl Ex/

Petit RJ, Mousadik EA, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. Conserv Biol 12: 844-855

Petit RJ, Aguinagalde I, de Beaulieu JL, et al (2003) Glacial refugia: hotspotsbut not melting pots of genetic diversity. Science 300: 1563–1565

Phillips J, Kyratzis A, Christoudoulou C, Kell S, Maxted N (2014) Development of a national crop wild relative conservation strategy for Cyprus. Genet Resour Crop Evol 61:817-827

Pollegioni P, Woeste K, Chiocchini F, et al (2014) Landscape genetics of Persian walnut (*Juglans regia* L.) across its Asian range. Tree Genet Genomes 10: 1027–1043

Pollegioni P, Woeste KE, Chiocchini F et al (2015) Ancient humans influenced the current spatial genetic structure of common walnut populations in Asia. *PLoS ONE* **10**: e0135980 doi:10.1371/journal.pone.0135980

Pritchard JK, Wen W (2004) Documentation for STRUCTURE software version 2.3.4. Department of Human Genetics. University of Chicago, Chicago IIlinois USA. Available at http://pritch.bsd.uchicago.edu.

Provan J, Bennet KD (2008) Phylogeographic insights into criptic glacial refugia. Trends Ecol Evol **23**: 564-571

Rambaut, A. "FigTree, version 1.3. 1." Computer program distributed by the author, website: http://treebioedacuk/software/figtree/[accessed January 4, 2011]

Rands MRW, Adams WM, Bennun L et al (2010) Biodiversity conservation: challenges beyond. Science 329: 1298–1303

Rebelo AG, Siegfried WR (1992) Where should nature reserves be located in the Cape Floristic Region, South Africa? Models for the spatial configuration of a Reserve Network aimed at maximizing the protection of floral diversity. Conserv Biol 6: 243–252

Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. Mol Ecol Not 4:137-138

Sork VL, Aitken SN, Dyer R.J, Eckert AJ, Legendre P, Neale DB (2013) Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. Tree Genet Genomes 9: 901-911

Scheldeman X, van Zonneveld M (2010) Training manual on Spatial Analysis of Plant Diversity and Distribution. Biodiversity International, Rome, Italy

Shepard D (1968) A two-dimensional interpolation function for irregularly-spaced data. In: Blue, R. B. S., Rosenberg, A. M. (Eds.), Proceedings of the 1968 ACM National Conference. ACM Press, New York, pp. 517-524

Squatriti P (2013) Landscape and change in early medieval Italy: chestnut, economy and culture 1st edn. Cambridge: Cambridge University Press

Storfer A, Murphy MA, Evans JS. et al (2007) Putting the 'landscape' in landscape genetics. Heredity 98, 128–142

Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010) Landscape genetics: where are we now? Mol Ecol 19:3496–3514

Souto C, Mathiasen, P, Acosta M. 2015. Identifying Genetic Hotspots by Mapping Molecular Diversity of Widespread Trees: When Commonness Matters. J Hered. 106: 537–545

Takezaki N, Nei M, Tamura K. 2010. POPTREE2: software for constructing population trees from allele frequency data and computing other population statistics with windows interface. Mol Biol Evol 27: 747-752

van Zonneveld M, Scheldeman X, Escribano P (2012) Mapping genetic diversity of Cherimoya (*Annona cherimola* Mill.): application of spatial analysis for conservation and use of plant genetic resources. *PLoS ONE* 7**:** 1-14

Vannini A, Natili G, Anselmi N, Mantaghi A, Vettraini AM (2010) Distribution and gradient analysis of Ink disease in chestnut forests. Forest Pathol 40: 73-86

Vinceti B, Loo J, Gaisberger H et al (2013) Conservation priorities for Prunus africana defined with the aid of spatial analysis of genetic data and climatic variables. PLos ONE 8(3): e 59987.doi:10.1371/journal.pone.0059987

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. Journal Evolution Biol 12: 233-244

Table 1. Number of plants (N), sampling location, Identity code (ID), and geographical coordinates for 73 European chestnut populations

Country	Region	Sampling site	ID	Long.	Lat.	N
Portugal	Bragança District	Bragança	PT01	-7.2465	41.2440	17
Portugal	Vila Real	Vila Pouca	PT02	-7.5995	41.4926	20
Portugal	Guarda District	Guarda	PT03	-7.2890	40.7030	20
Spain	Andalucia	Bubión	SP01	-3.3522	36.9461	14
Spain	Cataluña	Castanyet	SP02	2.6300	41.8900	15
Spain	Galicia	Costa Atlántica	SP03	-8.3697	43.2858	21
Spain	Andalucia	Gaucín	SP04	-5.3094	36.5389	26
Spain	Andalucia	Güejar Sierra	SP05	-3.4306	37.1494	14
Spain	Extremadura	Hervas	SP06	-5.2003	40.2600	14
Spain	Castilla-León	Médulas	SP07	-6.0500	42.4500	15
Spain	Asturias	Mieres	SP08	-5.7639	43.2169	29
Spain	Andalucia	Paterna	SP09	-2.9428	37.0239	14
Spain	Cataluña	Prades	SP10	0.9900	41.3400	15
Spain	Andalucia	Pujerra	SP11	-5.1436	36.5939	12
Spain	Andalucia	Santa Elena	SP12	-3.5603	38.4178	11
Spain	Galicia	Sierra Faro	SP13	-7.8664	42.6081	23
Spain	Andalucia	Sierra Norte	SP14	-5.6292	37.9153	15
Spain	Extremadura	Valverde	SP15	-6.1300	40.2300	15
Spain	Cataluña	Viladrau	SP16	2.4000	41.8467	15
Spain	Andalucia	Trasierra	SP17	-4.8467	37.9194	35
France	Aquitania	Dordogne	FR01	1.0390	44.6871	15
France	Midi-Pyrénées	Aveyron	FR02	2.9175	44.0678	16
France	Aquitania	Pyrénées Atlantiques	FR03	-1.0860	43.2786	9
England	Gloucestershire	Speech House	UK01	-2.5483	51.7952	20
Italy	Sicilia	Madonie	IT01	14.0900	37.8300	26
Italy	Calabria	Sila Piccola	IT02	16.7200	39.0500	26
Italy	Basilicata	Mt. Vulture	IT03	15.6100	40.9200	25
Italy	Marche	Mt. Laga	IT04	13.4100	42.7300	26
Italy	Lazio	Mt. Cimini-	IT05	12.1900	42.4100	23
Italy	Toscana	Mugello	IT06	11.5700	43.9700	19
Italy	Piemonte	Trontano	IT07	8.3300	46.1200	26
Italy	Piemonte	V. Pellice	IT08	7.1400	44.8000	26
Italy	Friuli	V. del Natisone	IT09	13.5600	46.1200	26
Slovakia	Banska Bystrica	Modrý Kameň	SK01	19.3305	48.2410	27
Slovakia	Nitra	Jelenec	SK02	18.2221	48.4116	22
Slovakia	Bratislava	Častá	SK03	17.3582	48.4029	20
Slovakia	Bratislava	Bratislava	SK04	17.1094	48.1733	13
Slovakia	Bratislava	Modra	SK05	17.2888	48.3407	10
Hungary	Pest	Nagymasros	HU01	18.9483	47.7924	30
Bulgaria	Blagoevgrad	Belasitsa	BU01	23.2000	41.3658	50
Bulgaria	Blagoevgrad	Slavyanka	BU02	23.5258	41.4147	21
Bulgaria	Blagoevgrad	North west Pirin	BU03	23.1997	41.8467	20
Bulgaria	Blagoevgrad	South west Pirin	BU04	23.4106	41.5114	21
Bulgaria	Blagoevgrad	Ograjden	BU05	23.0042	41.4611	42
Bulgaria	Montana	Western Stara	BU06	23.1094	43.2133	21
Rumania	Dobresti	Bihor	RO01	22.3140	46.9060	11
Rumania	Baia Sprie	Maramures	RO02	23.6650	47.6640	11
Rumania	Tarnita	Mehedinti	RO03	22.6680	44.9860	12
Greece	S-E-Macedonia	Holomontas	GR01	23.7500	40.5300	26
Greece	C-Macedonia	Hortiatis	GR02	22.3800	40.5900	26
Greece	W- Macedonia	Dafni	GR03	21.1400	40.2800	26
Greece	N- Macedonia	Paiko	GR04	22.3800	40.9500	26
Turkey	Duzce	Akcakoca	TR01	31.1600	41.0700	24
Turkey	Sinop	Sinop	TR02	35.0500	42.0000	31
Turkey	Artvin	Hopa	TR03	41.5700	41.3900	22
Turkey	Trabzon	Meryem Ana	TR04	39.6300	40.7200	30
Turkey	Giresun	Giresun	TR05	38.5228	40.9068	26
Turkey	Sakarya	Karadere	TR06	30.8400	40.7500	21
Turkey	Kocaeli	Sardala	TR07	29.9500	41.0400	23
Turkey	Yalova	Cinarcik	TR08	29.0800	40.6400	26

Country	Region	Sampling site	ID	Long.	Lat.	N
Turkey	Kocaeli	Golcuk	TR09	29.7900	40.7200	25
Turkey	Sakarya	Sapanca	TR10	30.2500	40.6500	24
Turkey	Bursa	Bursa	TR11	29.0800	40.1200	25
Turkey	Canakkale	Bayramic	TR12	26.6100	39.8200	31
Turkey	Izmir	Kemalpasa	TR13	27.3500	38.4000	23
Turkey	Manisa	Demirci	TR14	28.6300	39.0400	22
Georgia	Shida Kartli	Rikoti Pass	GE01	43.4926	42.0524	31
Georgia	Adjara	Uchkhiti	GE02	41.8131	41.5474	30
Georgia	Adjara	Mtirala National Park	GE03	41.8667	41.6833	26
Georgia	Abkhasia	Ochamchira	GE04	41.4363	42.7439	14
Georgia	Abkhasia	Tkwarchelsk	GE05	41.7747	42.8086	15
Azerbaijan	Oguz	Oguz	AZ01	47.4607	41.0770	32
Russia	Krasnodarskiy Kray	Chvigepse	RU01	40.0349	43.7334	28

1

2

3

4

5

6

7

8

Table 2. Genetic diversity of 73 European chestnut populations. Mean number of alleles per locus (A), mean effective number of alleles per locus (Ae), Observed (Ho) and expected (He) unbiased expected (UHe) heterozygosity, allelic richness (Ar), private allelic richness (PAr) and inbreeding coefficient (Fis)

Pop	A	Ae	Ho	He	UHe	Ar	PAr	Fis
PT01	5.50	3.66	0.68	0.69	0.75	5.17	0.06	0.056
PT02	6.17	3.73	0.76	0.69	0.73	5.35	0.00	-0.083
PT03	6.67	3.29	0.70	0.65	0.69	5.49	0.10	-0.072
SP01	5.00	3.40	0.76	0.69	0.72	4.48	0.13	-0.061
SP02	5.20	3.87	0.83	0.72	0.74	4.81	0.00	-0.121
SP03	6.00	3.51	0.66	0.68	0.69	4.70	0.00	0.052
SP04	6.20	3.21	0.61	0.65	0.67	4.84	0.07	0.089
SP05	4.40	3.22	0.64	0.68	0.70	4.12	0.13	0.086
SP06	5.80	3.49	0.67	0.66	0.68	5.07	0.00	0.016
SP07	6.40	3.67	0.75	0.70	0.73	5.45	0.12	-0.028
SP08	5.40	3.13	0.69	0.65	0.66	4.33	0.00	-0.048
SP09	4.40	3.19	0.54	0.65	0.68	4.19	0.00	0.204*
SP10	4.40	3.06	0.77	0.64	0.66	4.07	0.00	-0.173
SP11	5.00	3.51	0.75	0.71	0.74	4.74	0.00	-0.020
SP12	5.80	4.31	0.67	0.72	0.76	5.58	0.00	0.116
SP13	7.40	3.59	0.73	0.63	0.65	5.14	0.08	-0.131
SP14	5.80	4.00	0.73	0.72	0.75	5.31	0.00	0.017
SP15	6.00	3.20	0.73	0.67	0.69	5.11	0.10	-0.067
SP16	5.80	3.87	0.68	0.71	0.74	5.03	0.02	0.079*
SP17	6.50	3.21	0.56	0.63	0.64	4.97	0.05	0.169**
FR01	5.50	3.81	0.67	0.60	0.67	5.27	0.21	-0.070
FR02	6.67	3.82	0.52	0.63	0.73	6.01	0.04	0.162*
FR03	4.33	3.46	0.59	0.66	0.74	4.80	0.00	0.102
UK01	8.40	5.23	0.76	0.73	0.75	6.66	0.00	-0.010
IT01	7.40	4.51	0.67	0.72	0.74	5.91	0.04	0.095
IT02	7.80	3.67	0.66	0.70	0.71	5.51	0.07	0.072
IT03	9.20	4.69	0.84	0.75	0.76	6.36	0.03	-0.102
IT04	6.20	3.24	0.60	0.58	0.60	4.27	0.01	-0.007
IT05	9.00	4.29	0.71	0.72	0.74	6.20	0.03	0.031
IT06	8.20	5.57	0.82	0.80	0.83	6.65	0.00	0.006
IT07	6.40	3.08	0.68	0.65	0.66	4.45	0.07	-0.028
IT08	8.20	4.89	0.84	0.78	0.80	6.00	0.00	-0.054
IT09	6.80	4.00	0.77	0.72	0.73	5.22	0.00	-0.049
SK01	6.20	3.21	0.74	0.67	0.68	4.68	0.12	-0.091
SK02	4.40	2.43	0.64	0.56	0.57	3.68	0.08	-0.116
SK03	5.80	3.04	0.65	0.64	0.66	4.46	0.18	0.014
SK04	6.00	3.75	0.74	0.70	0.73	5.29	0.06	-0.010
SK05	5.20	3.60	0.82	0.72	0.75	5.06	0.00	-0.092
HU01	5.80	3.25	0.73	0.67	0.68	4.54	0.00	-0.075
BU01	9.80	4.63	0.81	0.77	0.78	5.99	0.05	-0.039
BU02	6.40	4.27	0.68	0.76	0.77	5.47	0.20	0.129
BU03	5.40	3.64	0.72	0.72	0.74	4.74	0.00	0.023
BU04	6.40	4.08	0.63	0.75	0.77	5.47	0.00	0.184
BU05	6.60	4.04	0.61	0.73	0.74	5.05	0.00	0.175**
BU06	5.40	3.21	0.67	0.67	0.69	4.76	0.00	0.033
RO01	7.00	4.82	0.76	0.76	0.79	6.50	0.17	-0.007
RO02	4.20	2.92	0.49	0.54	0.56	4.87	0.13	0.049
RO03	4.20	3.09	0.63	0.65	0.69	4.06	0.00	0.040
GR01	8.00	5.38	0.72	0.79	0.80	6.21	0.00	0.108*
GR02	8.20	4.12	0.72	0.73	0.75	6.18	0.04	0.044
GR03	5.60	3.33	0.75	0.69	0.71	4.56	0.00	-0.056

Pop	A	Ae	Ho	He	UHe	Ar	PAr	Fis
GR04	6.00	3.34	0.58	0.68	0.69	4.68	0.09	0.160
TR01	7.00	2.83	0.54	0.59	0.60	4.99	0.13	0.106*
TR02	8.20	3.28	0.56	0.65	0.66	5.11	0.05	0.153**
TR03	9.80	4.68	0.61	0.70	0.71	6.58	0.43	0.149**
TR04	7.20	2.88	0.53	0.60	0.61	4.73	0.20	0.133**
TR05	6.00	2.57	0.48	0.58	0.59	4.21	0.08	0.198**
TR06	8.20	3.40	0.54	0.65	0.67	5.48	0.06	0.189**
TR07	8.60	4.53	0.74	0.72	0.74	6.24	0.02	0.000
TR08	9.80	5.51	0.69	0.81	0.82	6.71	0.04	0.162**
TR09	9.40	4.80	0.70	0.75	0.77	6.62	0.01	0.084
TR10	7.40	3.53	0.66	0.67	0.69	5.28	0.04	0.042
TR11	8.80	4.49	0.69	0.76	0.78	6.41	0.07	0.115*
TR12	8.60	4.80	0.68	0.76	0.78	6.12	0.00	0.121*
TR13	9.60	5.41	0.75	0.80	0.82	7.00	0.01	0.087
TR14	7.80	4.57	0.81	0.77	0.78	6.12	0.02	-0.033
GE01	8.00	3.46	0.55	0.62	0.63	5.32	0.03	0.127**
GE02	9.40	4.42	0.63	0.66	0.67	5.90	0.20	0.052
GE03	5.40	3.68	0.65	0.64	0.66	4.73	0.00	0.015
GE04	5.20	3.50	0.56	0.67	0.70	4.74	0.52	0.205*
GE05	4.60	3.12	0.55	0.64	0.66	4.34	0.03	0.175
AZ01	5.60	3.54	0.68	0.69	0.70	4.37	0.07	0.031
RU01	8.20	3.44	0.59	0.68	0.69	5.43	0.05	0.157**

FIGURES CAPTIONS

Figure 1. Geographical distribution of the 73 *Castanea sativa* populations examined in the study. The distribution area of species is in red. (www.euforgen.org/distribution_maps)

Figure 2. Genetic diversity maps of 73 European chestnut populations: IDW interpolation of a) expected heterozygosity values (He), b) allelic richness (Ar) values, and c) private allelic richness (PAr) values. PT=Portugal, SP=Spain, Fr=France, UK=England, IT=Italy, SK=Slovakia, Hu=Hungary, BU=Bulgaria, RU=Rumania, Gr=Greece, TR=Turkey, GE=Georgia, AZ=Azerbaijan, RU=Russia

Figure 3. UPGMA based on Nei genetic distance. PT=Portugal, SP=Spain, Fr= France, UK = England, IT= Italy, SK=Slovakia, Hu=Hungary, BU=Bulgaria, RU=Rumania, Gr=Greece, TR= Turkey, GE=Georgia, AZ= Azerbaijan, RU=Russia

Figure 4. Population structure inferred for 1608 samples of *Castanea sativa* by Bayesian assignment using structure software STRUCTURE a) Each individual is represented by a vertical line and populations are separated by a vertical black line. Different colors in the same line indicate the individual's estimated membership percentage in K clusters (K=2, K=3 and substructure K'=3). b) Clustering surface map (K=3) for 73 resulting by the IDW interpolation of the estimated population membership values (Q). The more intensely coloured area indicates the strongest genetic similarity between populations belonging to the same cluster, and the gradual change to light colors indicates a gradual decrease in genetic similarity. Black line indicated genetic barriers with boostrap support > 50%. c) clustering surface map of substructure calculated for the population included in cluster 1.

Figure 5. Areas with priority for conservation. The size of red points indicates the priority: the larger the point the highest is the priority.

Online Resources 1. Inference of K, the most probable number of clusters, using the software STRUCTURE. Second order of change of the log-likehood of the data (ΔK) as a function of K, calculated over six replicates

Online Resources 2. The hierarchical AMOVA (Excoffier et al 2005) and F statistical analysis calculated considering STRUCTURE clusters (K=3). Significance of Fvalues was tested using a nonparametric approach described in Excoffier et al (1992) with 1000 permutation (*P<0.01).

Figure1

Click here to download Figure figure1.tif

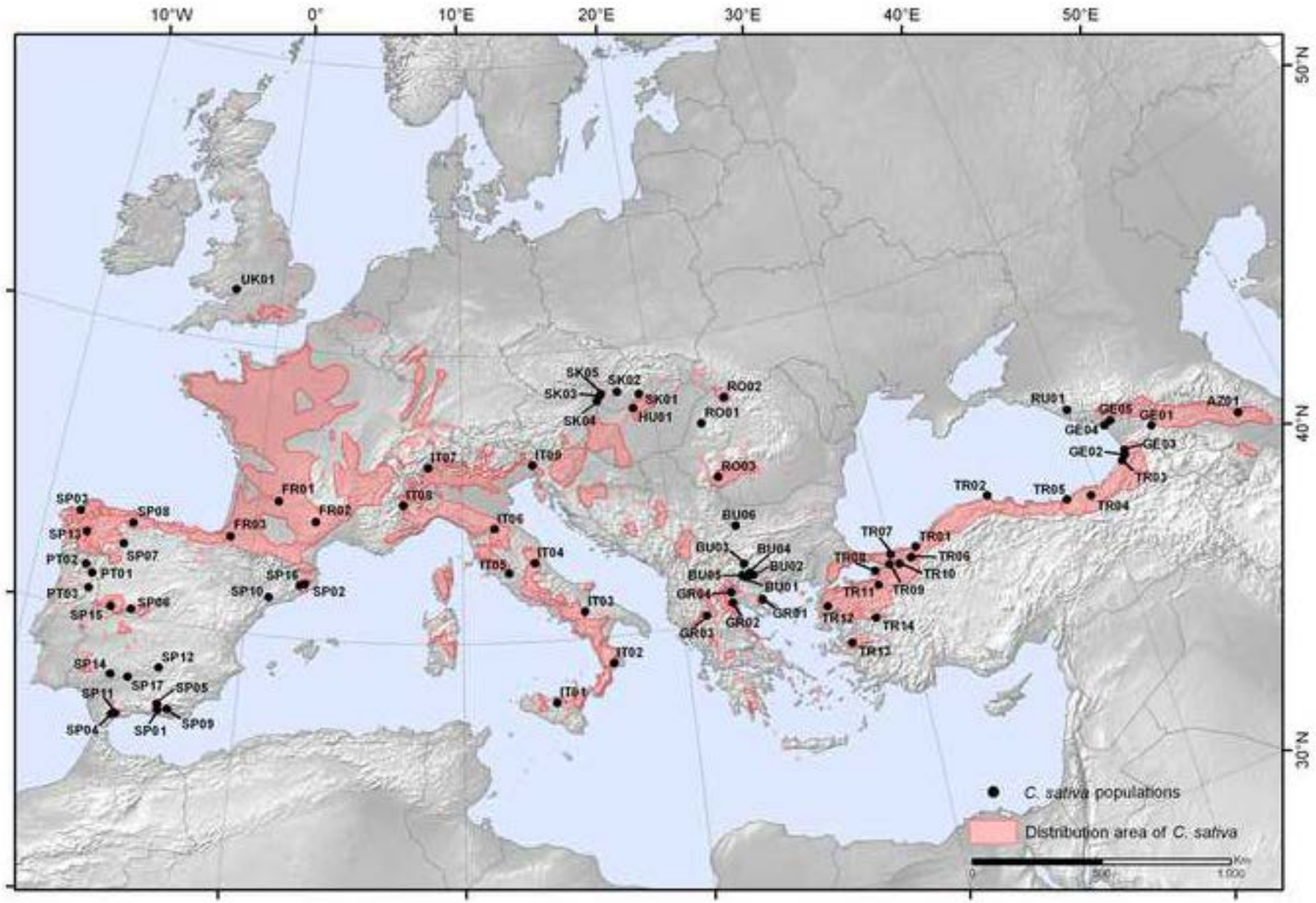


Figure2

Click here to download Figure figure2new.tif

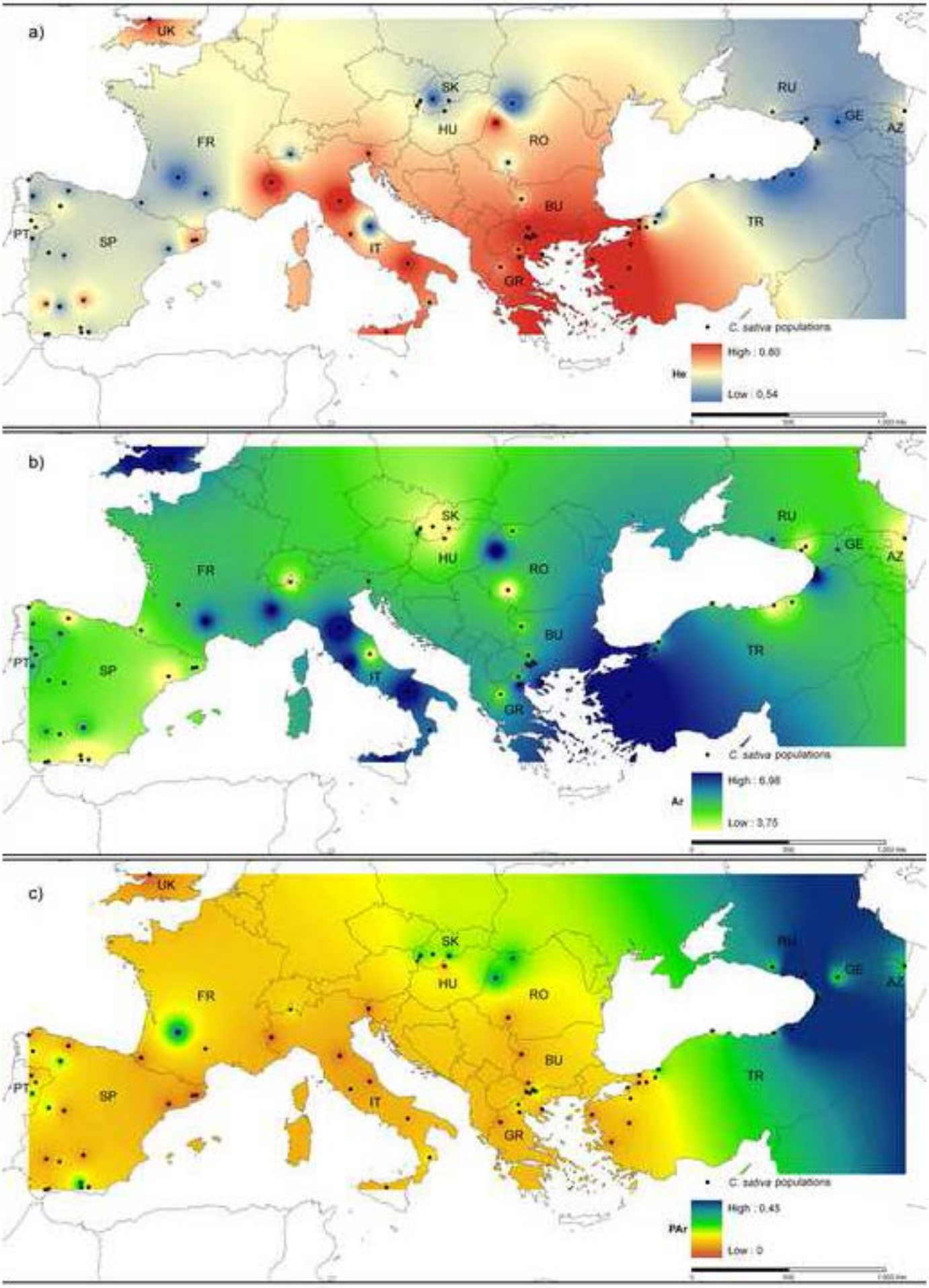


Figure3

Click here to download Figure figure3.tif

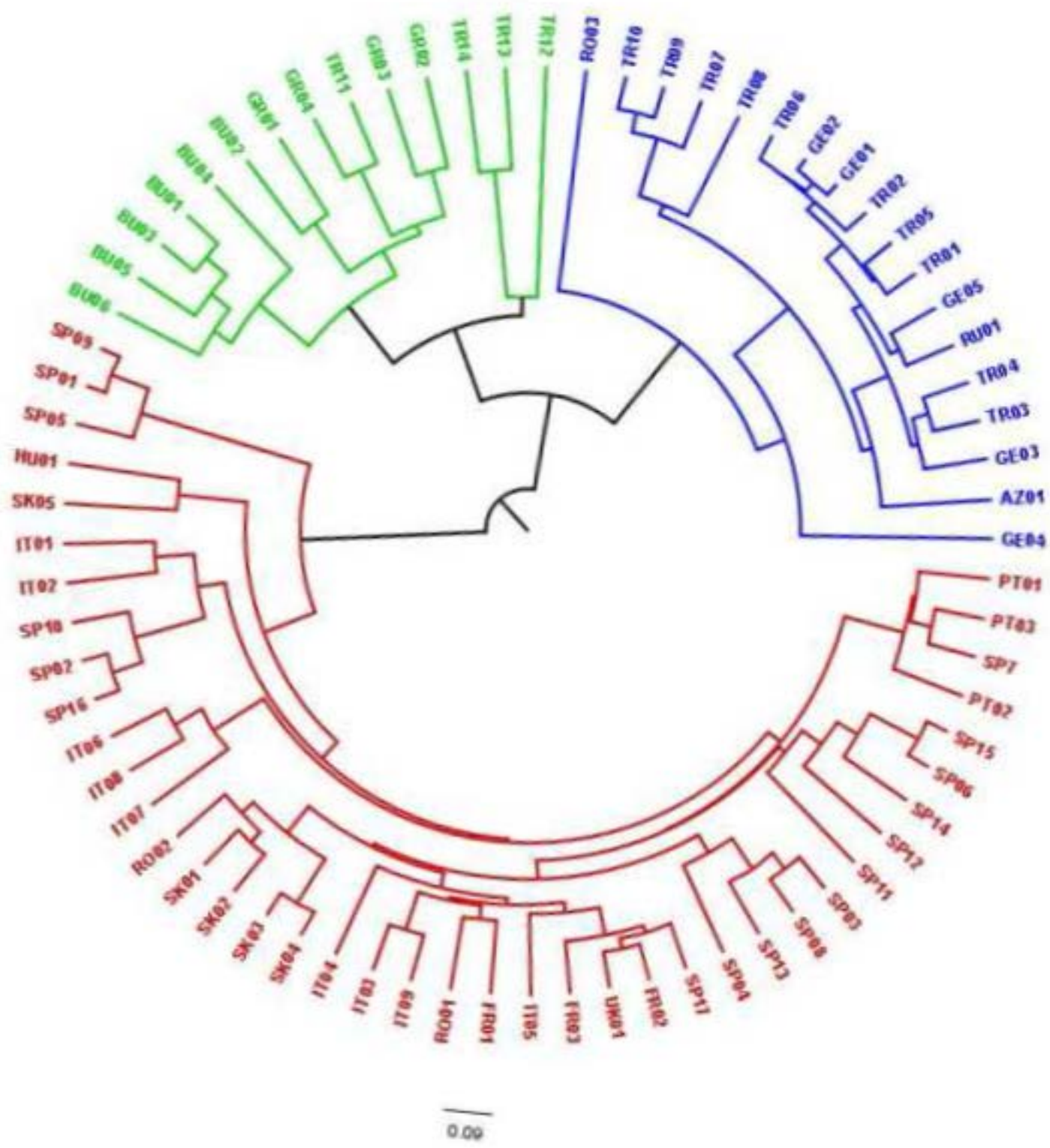


Figure4

Click here to download Figure figure4new.tif

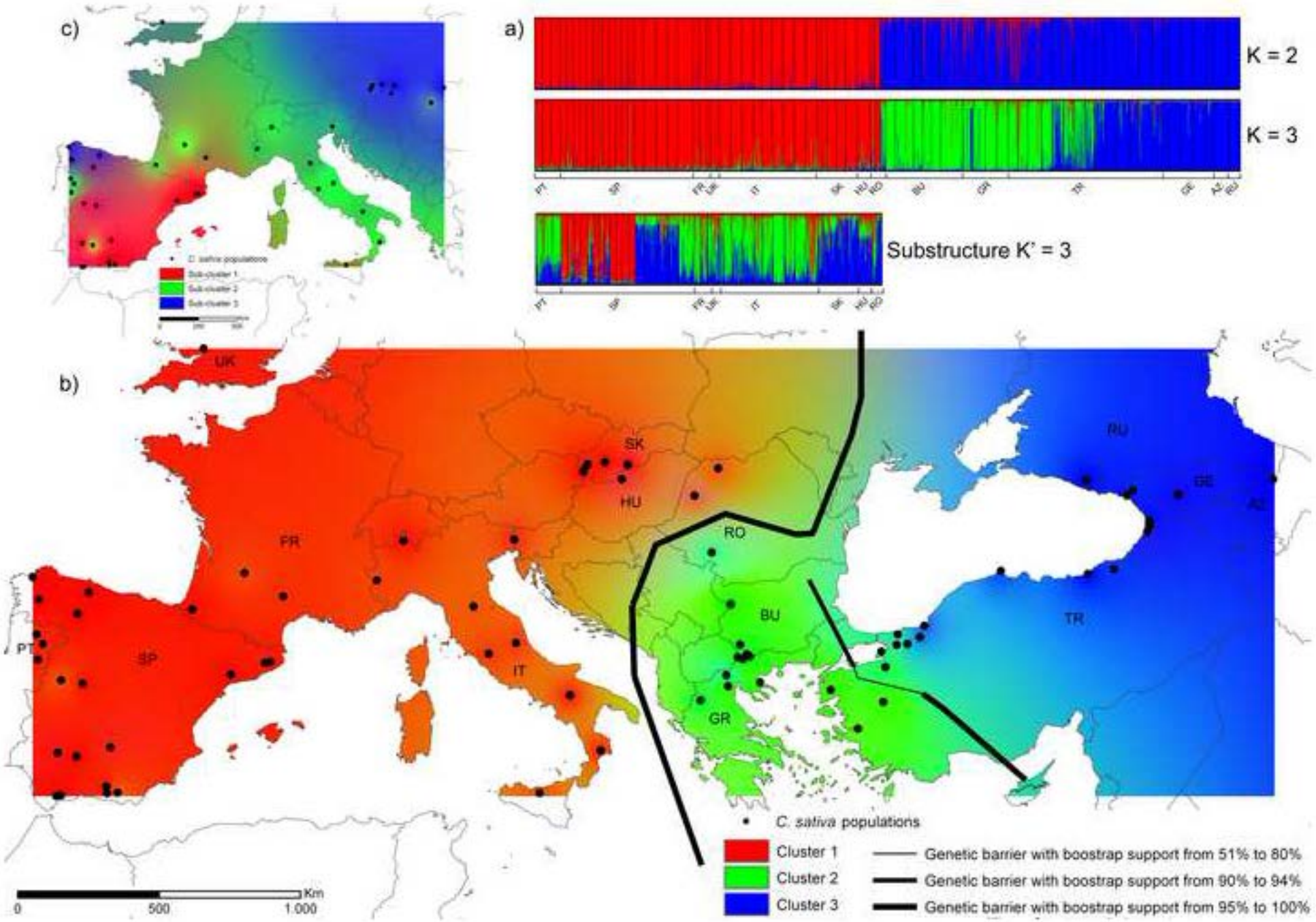
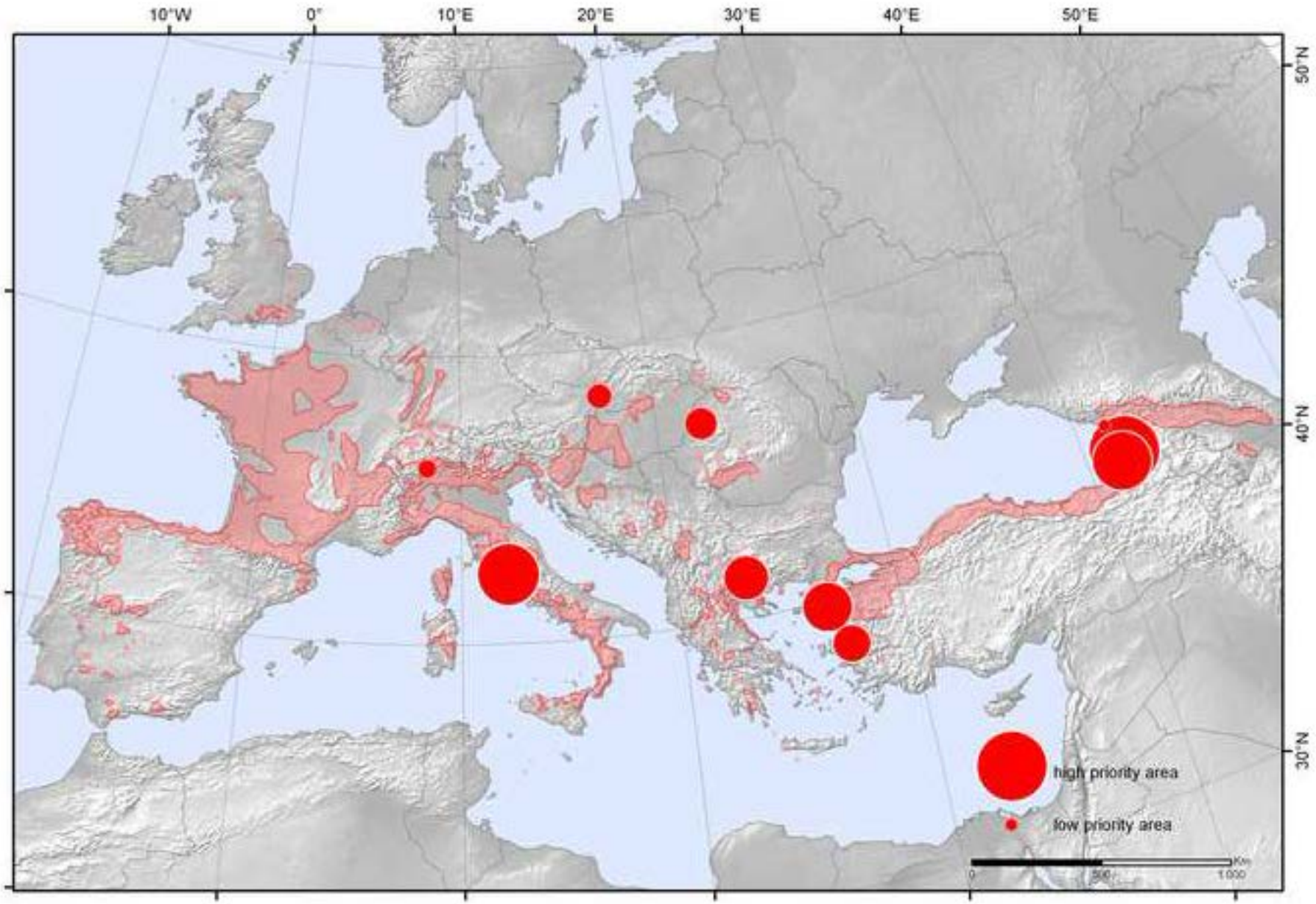


Figure5

Click here to download Figure figure5.tif



supplementary material

Click here to access/download
Supplementary Material
online resources 2.docx

Istituto di Biologia Agroambientale e Forestale

Dr. Claudia Mattioni

Via Marconi 2, 05010 Porano (TR), Italy

Tel. (+39) (0)763 374 940 - 911

Fax: (+39) (0)763 374 980

E-Mail: claudia.mattioni@ibaf.cnr.it

Dear Editor in Chief,
Dr. David Neale

We are sending you the revised version of the manuscript **TGGE-D-16-00139R1 " Landscape genetics structure of European sweet chestnut (*Castanea sativa* Mill): indications for conservation priorities**

by Claudia Mattioni, M. Angela Martin, Francesca Chiocchini, Marcello Cherubini, Muriel Gaudet, Paola Pollegioni, Ivaylo Velichko, Rob Jarman, Frank M. Chambers, Ladislave Paule, Vasilica L. Damian, Ghiţă C. Crainic and Fiorella Villani.

We have completed the revisions required by Reviewer #1:
Our detailed answers (in bold) are as follows:

p. 10 line 8-9. The last sentence in this paragraph is extremely vague. If DIVA-GIS has yielded noteworthy insights, or especially, been used in conservation decisions, examples of those could be cited by way of putting the authors' results in context. Otherwise, this sentence is not helpful.

p. 10 line 8-11. We changed the sentence trying to explain the relevance of results obtained with DIVA-GIS software. We added two references

p. 12 lines 11-14. This sentence tries to make several points but in the end is probably unnecessary. I would suggest shortening it to make a single, clear point, or else eliminating itentirely.

p. 12 line 7-10 we shortened the sentence

Hoping that our paper will be now suitable for publication

Best regards,

Claudia Mattioni