Scientific
Research
Publishing

Mapping the Genetic Diversity of *Castanea sativa*: Exploiting Spatial Analysis for Biogeography and Conservation Studies

Francesca Chiocchini[1]*, Claudia Mattioni[1], Paola Pollegioni[1], Ilaria Lusini[1], Maria Angela Martín[2], Marcello Cherubini[1], Marco Lauteri[1], Fiorella Villani[1]

[1]C.N.R. Institute of Agro-environmental and Forest Biology, Terni, Italy
[2]Departamento de Ingeniería del Medio Agronómico y Forestal, Avda. Virgen del Puerto 2, Centro Universitario de Plasencia, Universidad de Extremadura, Caceres, Spain

Email: *francesca.chiocchini@ibaf.cnr.it, claudia.mattioni@ibaf.cnr.it, paola.pollegioni@ibaf.cnr.it, ilaria.lusini@ibaf.cnr.it, ge2macum@uco.es, marcello.cherubini@ibaf.cnr.it, marco.lauteri@ibaf.cnr.it, fiorella.villani@ibaf.cnr.it

Abstract

The current distribution of forest tree species is a result of natural or human mediated historical and contemporary processes. Knowledge of the spatial distribution of the diversity and divergence of populations is crucial for managing and conserving genetic resources in forest tree species. By combining tools from population genetics, landscape ecology and spatial statistics, landscape genetics thus represents a powerful method for evaluating the geographic patterns of genetic resources at the population level. In this study, we explore the possibility of combining genetic diversity data, spatial statistic tools and GIS technologies to map the genetic divergence and diversity of 31 *Castanea sativa* populations collected in Spain, Italy, Greece, and Turkey. The IDW technique was used to interpolate the diversity values and divergence indices as expected heterezygosity (He), allelic richness (Rs), private allelic richness (PRs), and membership values (Q) of each population to different clusters. Genetic diversity maps and a synthetic map of the spatial genetic structure of European chestnut populations were produced. Spatial coincidences between landscape elements and statistically significant genetic discontinuities between populations were investigated. Evidence is provided of the significance of cartographic outputs produced in the study and on their usefulness in managing genetic resources.

Keywords

Landscape Genetics, Microsatellites, Genetic Structure, Spatial Interpolation, Genetic Barriers

---

*Corresponding author.

1. Introduction

The current distribution of species was shaped by complex interactions occurring over time and space between biological, physical and socio-cultural processes. Understanding the biogeographical distribution, the diversity patterns of living species and the underlying evolutionary processes are key to being able to conserve and manage genetic resources [1]. Both natural and anthropic factors have led to habitat fragmentation and a loss of biodiversity, which are considered as major issues in conservation biology [2]. Important ecological and landscape processes occur at different and multiple spatial and temporal scales, influencing the evolutionary biology of plant species [3]. In addition, anthropogenic factors, such as overexploitation, the introduction of invasive species and climate change can modify the geographical distribution of plant species over space and time [4] [5].

In the last decades, interdisciplinary methods have thus been developed in order to understand the complex relationships between natural and human processes. Landscape genetics has emerged as a discipline that combines the concepts and methods of landscape ecology, population genetics, geography and spatial statistics [6] [7], and requires significant interdisciplinary collaboration [8]. It aims to understand how the spatial and temporal heterogeneity of the landscape has shaped the spatial distribution of genetic variations influencing environmental selection and gene flow [9].

Landscape genetics could therefore be crucial for estimating the impact of landscape composition on the spatial genetic structure of natural populations of plants and for new conservation genetic programs. Progress in molecular analysis, spatial analytical methods, GIS technologies and the acquisition of spatially-explicit datasets of environmental variables has contributed to the development and diffusion of landscape genetics [10]. GIS spatial analysis highlights the spatial relationships among biological, physical and human landscape components [11]. In addition, the use of geo-referenced genetic data within a GIS has become a prerequisite for the spatial analysis of ecological processes [12]. Spatial analysis also 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 of plant species and to prioritize areas for conservation [13].

Natural forests are key ecosystems with a high level of biodiversity [14]. Landscape genetics studies are particularly suitable for tree species that are increasingly vulnerable to losses in genetic diversity due to land use change and land degradation [15]. An increasing number of studies has been published to date on the landscape genetics of forest trees [10] [16]-[18]. In particular recent studies have mapped the genetic diversity of *Araucaria araucana* [19], *Juglans regia* [20] and *Annona cherimola* [21].

We focus on *Castanea sativa*, a tree of great economic and ecological importance, and the only species of the genus *Castanea* in Europe. The current patterns of biogeographic distribution of *C. sativa* are known to be a response to the evolutionary history of the species, in relation both to the climatic and geomorphological processes that have occurred over the geological time scale and to the anthropic activities over the centuries.

Although there are many studies on the genetic structure and diversity of chestnut across Europe [22]-[26], only recently the spatial patterns of genetic structure and diversity of chestnut have been taken into consideration. For instance, a landscape genetic approach was applied to the chestnut by Martín [27] and Lusini [28], who assessed the spatial patterns of genetic structure and diversity of Spanish and Bulgarian natural chestnut populations using microsatellite markers. However, these studies referred to populations in restricted geographical areas and provided no information on the species diversity spatial distribution across Europe.

In this study, we explore the possibility of combining spatial analysis techniques with molecular data to map the genetic diversity of a large set of European *C. sativa* populations along the distribution area of the species. We used the landscape genetic overlay technique [17] to analyze a set of geo-referenced genetic data provided in our previous publication [26]. The aims were to: 1) test the potential of spatial analysis to visualize and better understand the intra population genetic diversity and the spatial genetic structure of European natural populations; 2) identify areas as reservoirs of genetic diversity; 3) detect the overlap of genetic and geographic barriers; 4) investigate the effects of the landscape structure on gene flow and genetic variability; 5) make recommendations for conservation planning.

2. Material and Methods

In the last three decades, the CNR Institute of Agro-environmental and Forest Biology (IBAF, Porano, Italy) has been involved in genetic diversity study of European chestnut populations, assembling a large collection of *C. sativa* populations. In this study, we performed spatial analysis of 31 European chestnut populations (779 wild

trees), using a set of genetic data previously obtained by Mattioni [26] by means of microsatellite markers (**Figure 1**; **Table S1**[1]). Geo-referenced genetic data and geographic data were organized in an ESRI ArcGIS 9.3 geodatabase. The global digital elevation model GTOPO30 was acquired (https://lta.cr.usgs.gov/GTOPO30). We also reconstructed the main refugium areas of European chestnut according to their probability level as produced by Krebs [29]. We used ArcGIS 9.3 software for both our consolidated experience in its use and for its popularity and spread in recent studies of landscape genetics. Population genetic data and spatial statistic tools were combined by GIS technologies. We used the overlay technique described by Holderegger [17] that involves: 1) the identification of the main genetic clusters of populations; 2) the detection of genetic discontinuities or barriers; 3) the interpolation of genetic parameters to elaborate genetic surface maps. Subsequently, the resulting genetic surface maps are overlaid on topographic maps to search for spatial coincidences between genetic barriers and discontinuities with landscape elements.

2.1. Spatial Representation of intra Population Genetic Indices

Allelic richness (Rs) and private allelic richness (PRs) were computed by the rarefaction method with HP-RARE software [30], expected (He) heterozygosity [31] was calculated using GeneAlEx 6 software [32]. The Inverse Distance Weighted (IDW) [33] [34] algorithm implemented in ArcGIS 9.3 was used to interpolate values of expected (He) heterozygosity, allelic richness (Rs) and private allelic richness (PRs) of all 31 chestnut populations and to derive maps of genetic diversity. IDW interpolation determines cell values using a linearly weighted combination of a set of sample points, considering that local influence of each single point decrease with distance.

2.2. Spatial Representation of Population Structure

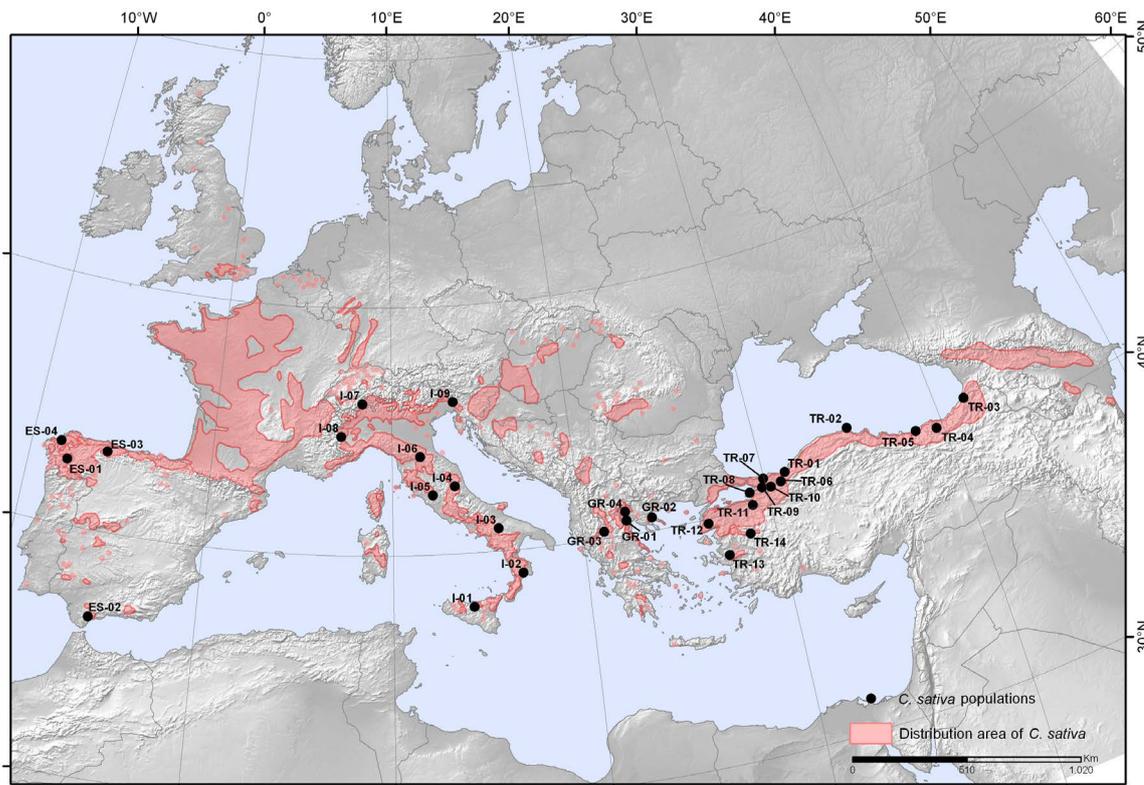A Bayesian clustering approach implemented in STRUCTURE 2.3.3 software [35] was applied to detect the

Figure 1. Location of 31 populations of *Castanea sativa* examined in the study and geographical distribution of the species ([41] http://www.euforgen.org/distribution_maps.html): Spanish (ES), Italian (I), Greek (GR) and Turkish (TR) populations (see **Table S1**).

[1]List of sampling location of 31European chestnut populations.

genetic structure of *C. sativa* populations. A complete description of the method was reported by Mattioni [26]. Population Q-membership values for the K clusters inferred by STRUCTURE were spatially interpolated using IDW function of ArcGIS 9.3 software [20] [36]. K clustering surface maps, representing the spatial patterns of the inferred genetic clusters, were produced. Finally we spatially overlaid the K clustering surfaces, combining multiple raster bands (the K interpolated raster) in a single multiband raster dataset (the resulting synthetic map), by the Composite Bands function of ArcGIS 9.3, to produce a synthetic map showing the spatial genetic structure of the 31 European chestnut populations. The Color Composite tool enabled us to graphically display a composition of three different bands (red-green-blue, corresponding to our three clustering surfaces) at a time. In addition, we investigated the presence of genetic barriers among the chestnut populations, using Monmonier's maximum difference algorithm implemented in BARRIER 2.2 software [37]. The detection of the genetic discontinuities was based on the Delauney triangulation (network connection between the geographical coordinates of each population) and the resulting Voronoi tessellation. Each edge of the Voronoi polygons was associated with the value of the corresponding Nei's [38] genetic distance (D_A) between pairs of populations, calculated with the GENDIST function in the PHYLIP software [39]. The procedure traces the boundary along the Voronoi tessellation starting with the edge to which is associated the highest value of D_A between two populations, and goes on to identify multiple barriers in hierarchical progression. The significance of these genetic barriers was tested using 100 resampled bootstraps matrice of Nei's [38] genetic distances, calculated with SEQBOOT and GENDIST function in the PHYLIP software package [39]. We only considered genetic discontinuities with arbitrary bootstrap support of $P > 0.50$. The genetic barriers detected, corresponding to abrupt change in the patterns of genetic variation among populations were reconstructed using the Voronoi polygons function in ArcGis 9.3 and overlaid on an GTOPO30 digital elevation model to search for spatial coincidences of these genetic boundaries with landscape elements.

3. Results and Discussion

3.1. Spatial Representation of intra Population Genetic Diversity

The geospatial interpolation of genetic diversity indices enabled us to produce new spatial data representative of the genetic diversity of the European chestnut. Expected heterozygosity (He), allelic richness (Rs) and private allelic richness (PRs) values (data reported in [26]) were interpolated using the IDW function and the respective derived maps of genetic diversity are shown in **Figure 2**. The spatially explicit representation of diversity indices facilitates the interpretation of the results, in order to understand the geographical distribution of the genetic diversity of the chestnut.

Eastern Turkish populations showed intermediate values of expected heterozygosity (He), as well as the highest values of allelic richness (Rs) and private allelic richness (PRs). Various western Turkish populations and two populations from central and northern Italy also showed high values of expected heterozygosity (He) and allelic richness (Rs) (**Figure 2**). Spanish populations presented low values of both expected heterozygosity (He) and allelic richness (Rs). Only one population from Galicia (northern Spain) showed high values of private allelic richness (PRs). Greek populations showed low values of expected heterozygosity (He), and low to intermediate allelic richness (Rs) values, except for one population from central Macedonia. Null values of private allelic richness (PRs) were shown for all Greek populations.

The spatial mapping of genetic diversity also contributed to the identification of priority units for the conservation of genetic resources. Thus outputs from the spatial overlay of the three maps of expected heterozygosity (He), allelic richness (Rs) and private allelic richness (PRs) provided critical information on the diversity of chestnut populations across Europe and can be assumed as basic criteria for identifying areas for conservation. Given that allelic richness is considered a key measurement in analyzing the conservation of the genetic diversity of species [40], we identified those areas with the highest level of diversity as the most suitable for conservation. In addition, gene pools with rare or private alleles, which can be considered in any case as important for conservation, were geographically identified.

Our results indicated that the main centers of genetic diversity for the European chestnut are located in the easternmost areas of Turkey, in western Turkey, and in central and northern Italy. We therefore observed a spatial matching between these sites and the geographical areas identified by Krebs [29] as Quaternary refugia of *C. sativa*.
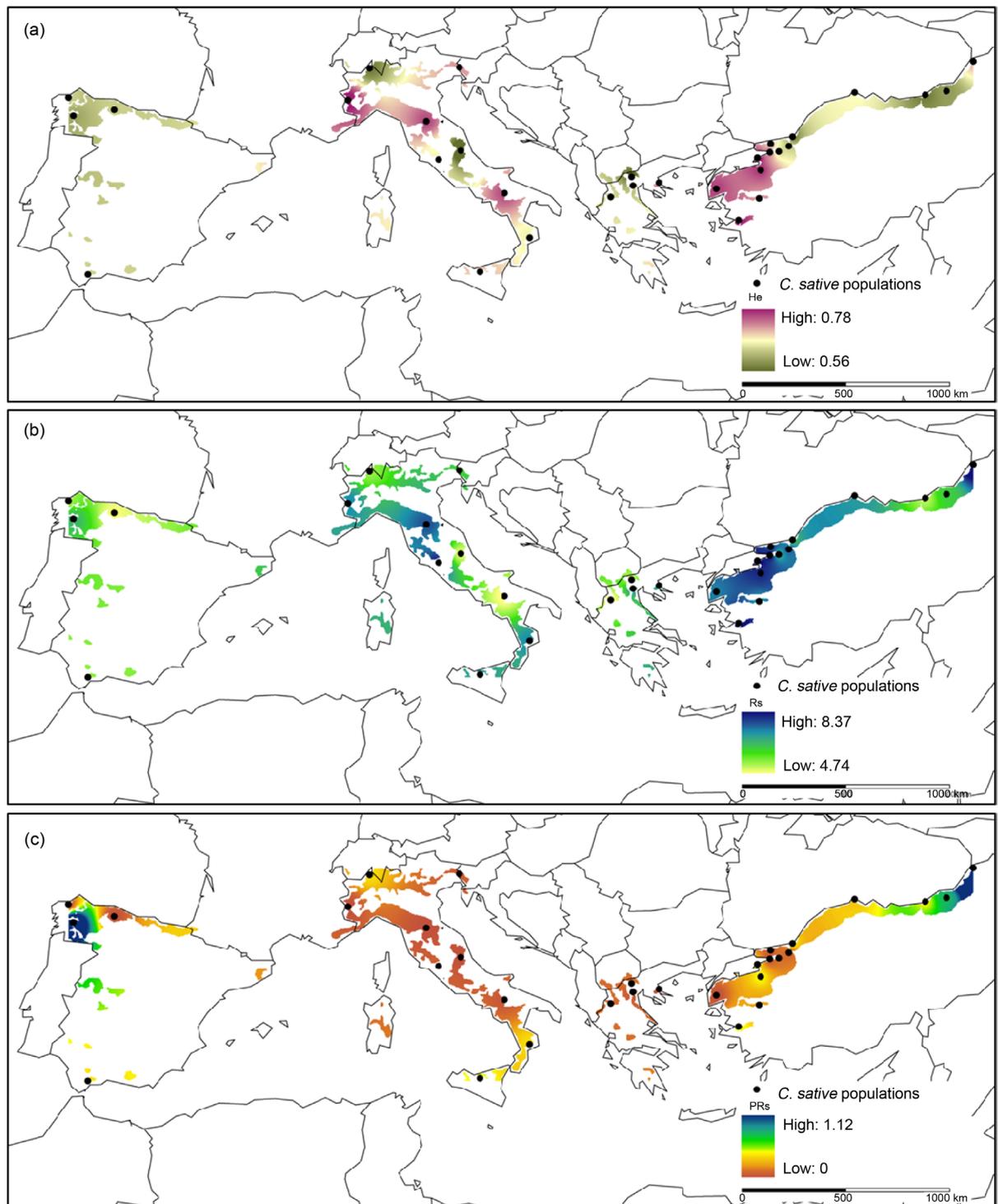
Figure 2. Genetic diversity maps of 31 European chestnut populations: IDW interpolation of (a) expected heterozygosity values (He), (b) allelic richness (Rs) values, and (c) private allelic richness (PRs) values.

3.2. Spatial Representation of Population Structure

In a previous work [26], we highlighted two main genetic clusters: K = 2 and K = 3. However, for the analysis carried out in the present work we considered only the data obtained for K = 3 as the best representation of the

hierarchical structure of European chestnut populations. **Figure 3** shows the three clustering surface maps, as a result of the spatial interpolation of the estimated population membership values (Q) in the K = 3 clusters inferred by STRUCTURE. In each map (**Figures 3(a)-(c)**) showing a different cluster of populations, the more intensely colored (saturated) areas indicate the strongest genetic similarity between populations belonging to the same cluster, and the gradual change to lighter colors indicates a gradual decrease in genetic similarity.
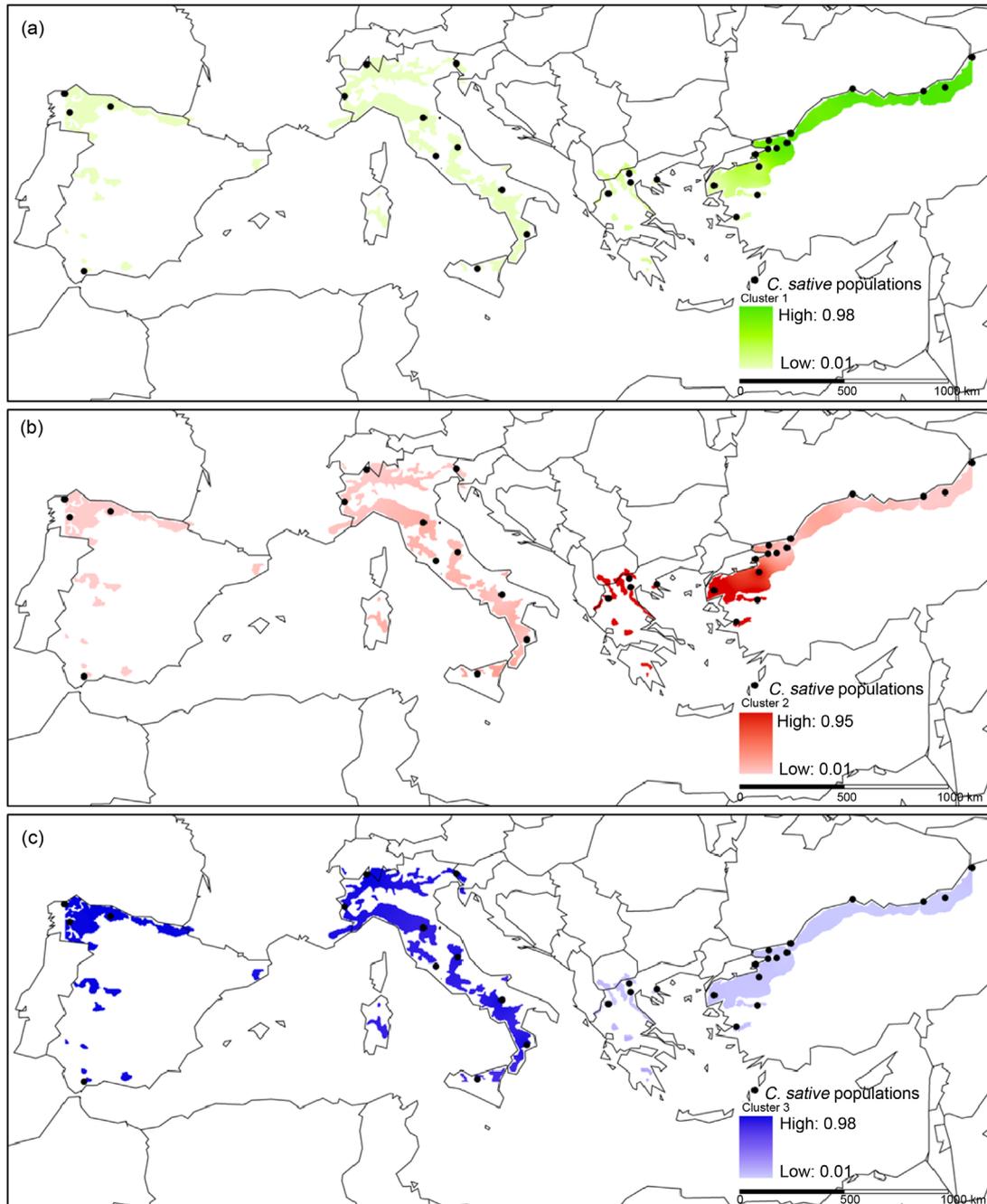

Figure 3. Clustering surface maps for 31 European chestnut populations from the IDW interpolation of the estimated population membership values (Q) in the K = 3 clusters inferred by STRUCTURE. In each map, the more intensely colored area indicates the strongest genetic similarity between populations belonging to the same cluster (cluster 1 (a) in green, cluster 2 (b) in red and cluster 3 (c) in blue), and the gradual change to light colors indicates a gradual decrease in genetic similarity.

The first cluster, in green in **Figure 3(a)**, includes populations from eastern and central Turkey. Populations from western Turkey and Greece were assigned to the second cluster, in red (**Figure 3(b)**). The third cluster, in blue in **Figure 3(c)** includes all the populations from Italy and Spain. By combining the previous three clustering surfaces, **Figure 4** shows the synthetic map of the spatial genetic structure of chestnut populations, as well as the bar diagram of the genetic relationship among populations estimated by STRUCTURE analysis for K = 3. The synthetic map corresponds to the spatial representation of the bar diagram of STRUCTURE: different hues correspond to different combinations of the three components, which are the three clustering surface maps. In the synthetic map, three raster cell values correspond to each population, thus indicating the estimated population Q-membership percentage to clusters 1, 2 and 3, and a combination of the three colored components (green = cluster 1, red = cluster 2, blue = cluster 3).

The presence of populations belonging to three different gene pools is highlighted in **Figure 4** (as for **Figures 3(a)-(c)**, as an introgression zone. This zone is made up of individuals from central Turkey belonging to two different gene pools (cluster 1 in green, and cluster 2 in red), and is shown in shades from green to red in the synthetic map and with two colors (green and red) in the same bar of the STRUCTURE's diagram.

Although the synthetic map of genetic structure provides an exhaustive geographical representation of the main gene pools in a single map, it has limitations. In fact, if the K number of clusters inferred by STRUCTURE analysis is greater than three, it cannot be graphically displayed and it is necessary to map K different clustering surface maps. These limits are derived from the combination of up to three color bands (red-green-blue) at a time, to graphically display the multiband raster dataset representing the genetic structure. However, it is still possible to represent the K clustering surface map, regardless of the K number of clusters inferred by STRUCTURE.

Mapping the spatial clustering of European chestnut populations highlighted three distinct gene pools, in agreement with those we had detected previously [26]: the first corresponding to populations from eastern and central Turkey, the second to populations from western Turkey and Greece, and the third to Italian and Spanish populations. In addition, the spatial patterns of genetic structure confirm the geographical correspondence that we had already observed between the detected genetic clusters and the glacial refugia of European chestnut hypothesized by Krebs [29], based on an extended palynological approach.

Three significant genetic barriers were identified using Monmonier's maximum difference algorithm, with bootstrap support >50%. Populations from Italy and Spain were separated from Greek and Turkish populations by the main genetic barrier ("a" in **Figure 5**). A sampling gap between western and eastern populations probably
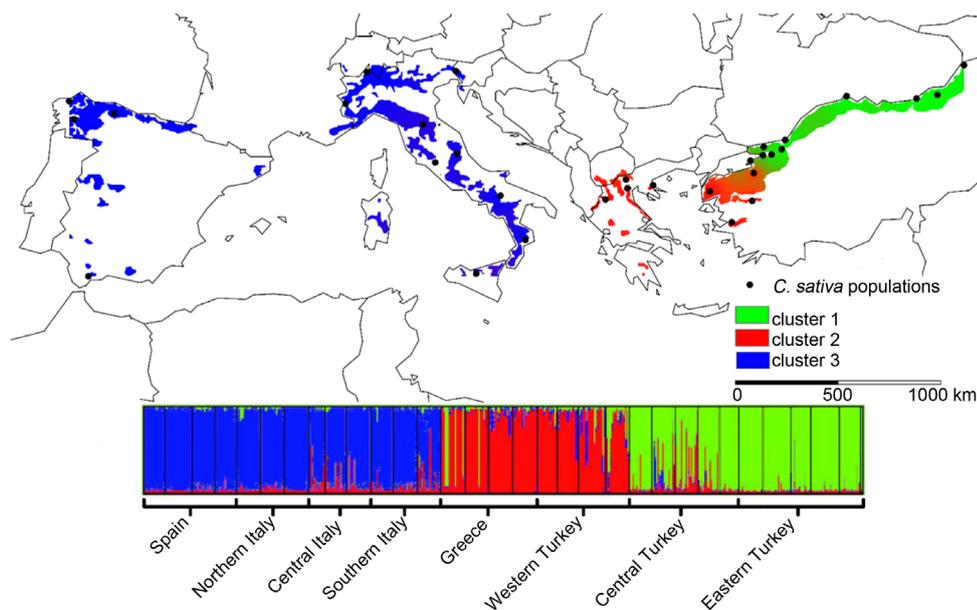

Figure 4. Synthetic map of spatial genetic structure and corresponding population structure inference for 779 chestnut tree samples estimated by Bayesian assignment using STRUCTURE for K = 3 clusters. Cluster 1 is displayed in green, cluster 2 in red and cluster 3 in blue.
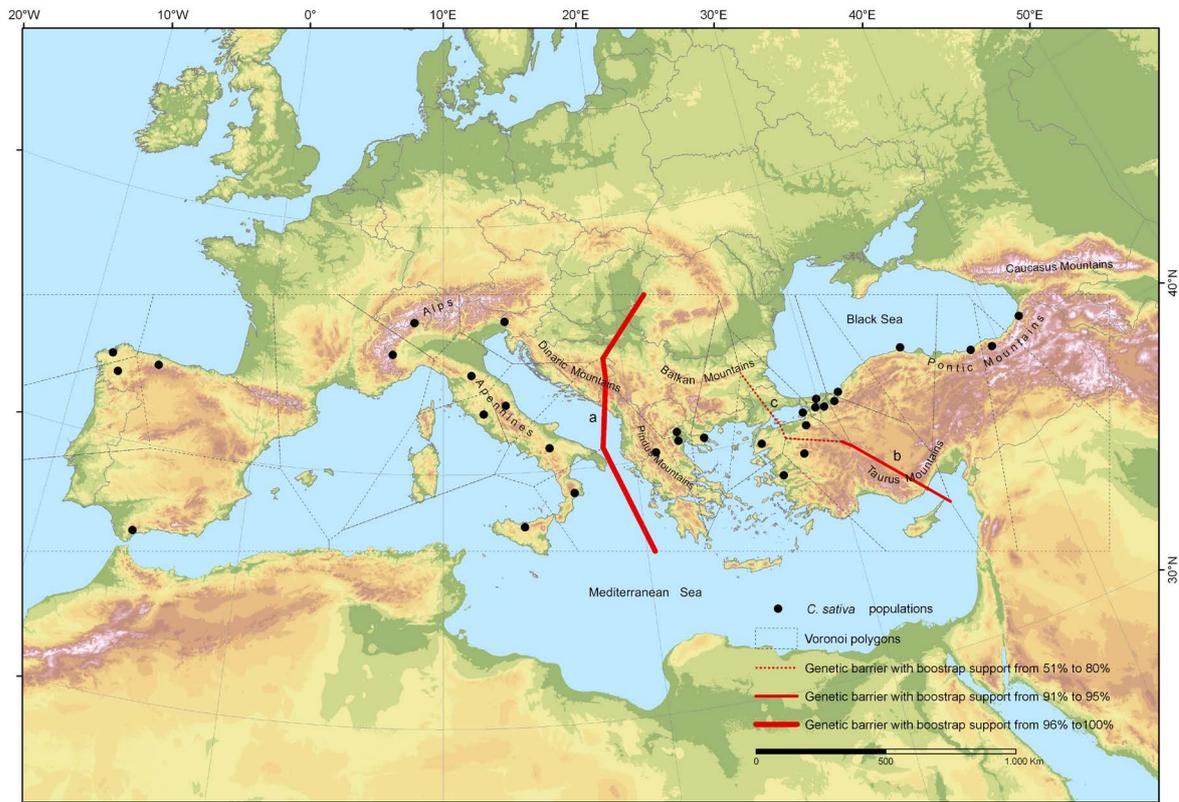
Figure 5. Three significant genetic barriers among 31 European chestnut populations identified using BARRIER software [37] and overlaid on a GTOPO30 global digital elevation model (https://lta.cr.usgs.gov/GTOPO30). The significant genetic barriers (red lines) were reconstructed basing on the Voronoi tessellation (light gray dotted lines).

contributed to the detection of this main barrier. The second and third discontinuities ("b" and "c" in **Figure 5**) divided Greek and western Turkish populations from central and eastern Turkish populations. The genetic discontinuities detected, were reconstructed based on the Voronoi tessellation and overlaid on an GTOPO30 digital elevation model (**Figure 5**). The three significant genetic barriers, identified by Monmonier's maximum difference algorithm, divide the chestnut populations into three main groups, thus matching the results of the Bayesian clustering analysis.

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. There is a geographical correspondence between the main genetic barrier, detected in this study, and the Dinaric Alps, which could act as an obstacle to the gene flow between eastern and western European populations of *C. sativa*, and the Adriatic Sea could also have impeded the exchange of germplasm between the Balkan and Italian peninsulas. In addition, genetic discontinuities between western and central-eastern Turkish populations may indicate the Taurus Mountains as a putative physical barrier to the gene flow between Turkish populations.

4. Conclusions

In this study, the spatial patterns of genetic diversity and structure of natural *C. sativa* populations were analyzed at the European scale. This interdisciplinary approach enabled us to combine data and methods from population genetics, landscape ecology and spatial analysis to explicitly quantify the effect of landscape configuration on genetic variation. The use of GIS contributed significantly, due both to its ability to store, manage and integrate molecular and spatial data, and to extract more in depth information from existing data. GIS is also useful for spatial analysis, modeling, data visualization and mapping.

The spatial analyses, such as the interpolation of diversity indices and the Q-membership coefficient value of each population to K clusters inferred by STRUCTURE, produce more visually clear and intuitive representa-

tions of the intra population genetic diversity and structure of European chestnut populations. The landscape genetic overlay technique enabled us to better display areas with a different genetic diversity, to highlight the geographical distribution of different gene pools and to show the overlap of genetic and geographic barriers.

The results also enabled us to indicate areas as reservoirs of genetic diversity, to speculate on the effects of the landscape structure on gene flow and genetic variability, and to provide suggestions for conservation planning. All these spatial outputs support the results of our previous publications, indicating the divergence between eastern and western populations of European chestnut, and thus confirming the presence of two different gene pools and an introgression zone in Turkey [22] [26]. High levels of genetic diversity were found in the majority of sweet chestnut European populations, overlapping the hypothetical chestnut glacial refugia identified by Krebs [29]. The overlay approach indicates that the potential barriers to gene flow, such as the Dinaric Alps and the Adriatic Sea, may have led to a separation between the eastern and western populations. A landscape genetic study including chestnut populations from Spain up to the Transcaucasian region would contribute to clarify the migration route of *C. sativa* across the Mediterranean range.

In conclusion, we have demonstrated how landscape genetics, combining spatial analysis techniques with molecular data in studying the genetic diversity of natural chestnut populations, can contribute to increasing the knowledge and understanding of the biogeographic history and distribution of *C. sativa* in Europe. The significant outputs produced are easily understandable and usable in the inventory, conservation, and management of genetic resources.

Acknowledgements

References

[1] Cox, B. and Moore P. (2005) Biogeography: An Ecological and Evolutionary Approach. John Wiley & Sons, Chichester, 439-472.

[2] Fischer, J. and Lindenmayer, D.B. (2007) Landscape Modification and Habitat Fragmentation: A Synthesis. *Global Ecology and Biogeography*, **16**, 265-280.

[3] Anderson, C.D., Epperson, B.K., Fortin, M.J., Holderegger, R., James, P.M.A., Rosenberg, M.S., Scribner, K.T. and Spear, S.F. (2010) Considering Spatial and Temporal Scale in Landscape-Genetic Studies of Gene Flow. *Molecular Ecology*, **19**, 3565-3575. http://dx.doi.org/10.1111/j.1365-294X.2010.04757.x

[4] Butchart, S.H.M., Walpole, M., Collen, B., van Strien, A., *et al.* (2010) Global Biodiversity: Indicators of Recent Declines. *Science*, **328**, 1164-1168. http://dx.doi.org/10.1126/science.1187512

[5] Rands, M.R.W., Adams, W.M., Bennun, L., Butchart, S.H.M., Clements, A., Coomes, D., Entwistle, A., Hodge, I., Kapos, V., Scharlemann, J.P.W., Sutherland, W.J. and Vira, B. (2010) Biodiversity Conservation: Challenges beyond 2010. *Science*, **329**, 1298-1303. http://dx.doi.org/10.1126/science.1189138

[6] Manel, S., Schwartz, M.K., Luikart, G. and Taberlet, P. (2003) Landscape Genetics: Combining Landscape Ecology and Population Genetics. *Trends in Ecology & Evolution*, **18**, 189-197. http://dx.doi.org/10.1016/S0169-5347(03)00008-9

[7] Holderegger, R. and Wagner, H.H. (2008) Landscape Genetics. *BioScience*, **58**, 199-207. http://dx.doi.org/10.1641/B580306

[8] Holderegger, R. and Wagner, H.H. (2006) A Brief Guide to Landscape Genetics. *Landscape Ecology*, **21**, 793-796. http://dx.doi.org/10.1007/s10980-005-6058-6

[9] Wagner, H.H. and Fortin, M.J. (2013) A Conceptual Framework for the Spatial Analysis of Landscape Genetic Data. *Conservation Genetics*, **14**, 253-261. http://dx.doi.org/10.1007/s10592-012-0391-5

[10] Storfer, A., Murphy, M.A., Spear, S.F., Holderegger, R. and Waits, L.P. (2010) Landscape Genetics: Where Are We Now? *Molecular Ecology*, **19**, 3496-3514. http://dx.doi.org/10.1111/j.1365-294X.2010.04691.x

[11] Fortin, M.J. and Dale, M.R.T. (2005) Spatial Analysis: A Guide for Ecologists. Cambridge University Press, Cam-

bridge, 1-30.

[12] Etherington, T.R. (2011) Python Based GIS Tools for Landscape Genetics: Visualising Genetic Relatedness and Measuring Landscape Connectivity. *Methods in Ecology and Evolution*, **2**, 52-55. http://dx.doi.org/10.1111/j.2041-210X.2010.00048.x

[13] Scheldeman, X. and van Zonneveld, M. (2010) Training manual on Spatial Analysis of Plant Diversity and Distribution. Biodiversity International, Rome, 97-138.

[14] Pautasso, M. (2009) Geographical Genetics and the Conservation of Forest Trees. *Perspectives in Plant Ecology*, *Evolution and Systematics*, **11**, 157-189. http://dx.doi.org/10.1016/j.ppees.2009.01.003

[15] Manel, S. and Holderegger, R. (2013) Ten Years of Landscape Genetics. *Trends in Ecology & Evolution*, **28**, 614-621. http://dx.doi.org/10.1016/j.tree.2013.05.012

[16] Smulders, M.J.M., Cobben, M.M.P., Arens, P. and Verboom, J. (2008) Landscape Genetics of Fragmented Forests: Anticipating Climate Change by Facilitating Migration. *iForest*, **2**, 128-132. http://dx.doi.org/10.3832/ifor0505-002

[17] Holderegger, R., Buehler, D., Gugerli, F. and Manel, S. (2010) Landscape Genetics of Plants. *Trends in Plant Science*, **15**, 675-683. http://dx.doi.org/10.1016/j.tplants.2010.09.002

[18] Ismail, S.A., Ghazoul, J., Ravikanth, G., Kushalappa, C.G., Uma Shaanker, R. and Kettle, C.J. (2014) Forest Trees in Human Modified Landscapes: Ecological and Genetic Drivers of Recruitment Failure in *Dysoxylum malabaricum* (Meliaceae). *PLoS One*, **9**, e89437. http://dx.doi.org/10.1371/journal.pone.0089437

[19] Martín, M.A., Mattioni, C., Lusini, I., Molina, J.K., Cherubini, M., Drake, F., Herrera, M.A., Villani, F. and Martín, L.M. (2014) New Insights into the Genetic Structure of *Araucaria araucana* Forests Based on Molecular and Historic Evidences. *Tree Genetics & Genomes*, **10**, 839-851. http://dx.doi.org/10.1007/s11295-014-0725-1

[20] Pollegioni, P., Woeste, K., Chiocchini, F., Olimpieri, I., Tortolano, V., Clark, J., Hemery, G.E., Mapelli, S. and Malvolti, M.E. (2014) Landscape Genetics of Persian Walnut (*Juglans regia* L.) Across Its Asian Range. *Tree Genetics & Genomes*, **10**, 1027-1043. http://dx.doi.org/10.1007/s11295-014-0740-2

[21] van Zonneveld, M., Scheldeman, X., Escribano, P., Viruel, M.A., Van Damme, P., Garcia, W., Tapia. C., Romero, J., Siguenas, M. and Hormaza, J.I. (2012) Mapping Genetic Diversity of Cherimoya (*Annona cherimola* Mill.): Application of Spatial Analysis for Conservation and Use of Plant Genetic Resources. *PLoS ONE*, **7**, e29845. http://dx.doi.org/10.1371/journal.pone.0029845

[22] Villani, F., Sansotta, A., Cherubini, M., Cesaroni, D. and Sbordoni, V. (1999) Genetic Structure of Natural Populations of *Castanea sativa* in Turkey: Evidence of a Hybrid Zone. *Journal of Evolutionary Biology*, **12**, 233-244. http://dx.doi.org/10.1046/j.1420-9101.1999.00033.x

[23] Fineschi, S., Taurchini, D., Villani, F. and Vendramin, G.G. (2000) Chloroplast DNA Polymorphism Reveals Little Geographical Structure in *Castanea sativa* Mill. (Fagaceae) throughout Southern European Countries. *Molecular Ecology*, **9**, 1495-1503. http://dx.doi.org/10.1046/j.1365-294x.2000.01029.x

[24] Mattioni, C., Cherubini, M., Micheli, E., Villani, F. and Bucci, G. (2008) Role of Domestication in Shaping *Castanea sativa* Genetic Variation in Europe. *Tree Genetics & Genomes*, **4**, 563-574. http://dx.doi.org/10.1007/s11295-008-0132-6

[25] Martín, M.A., Mattioni, C., Cherubini, M., Taurchini, D. and Villani, F. (2010) Genetics Diversity in European Chestnut Populations by Means of Genomic and Genic Microsatellite Markers. *Tree Genetics & Genomes*, **6**, 735-744. http://dx.doi.org/10.1007/s11295-010-0287-9

[26] Mattioni, C., Martín, M.A., Pollegioni, P., Cherubini, M. and Villani, F. (2013) Microsatellite Markers Reveal a Strong Geographical Structure in European Populations of *Castanea sativa* (Fagaceae): Evidence for Multiple Glacial Refugia. *American Journal of Botany*, **100**, 951-961. http://dx.doi.org/10.3732/ajb.1200194

[27] Martín, M.A., Mattioni, C., Molina, J.R., Alvarez, J.B., Cherubini, M., Herrera, M.A., Villani, F. and Martín, L.M. (2012) Landscape Genetic Structure of Chestnut (*Castanea sativa* Mill.) in Spain. *Tree Genetics & Genomes*, **8**, 127-136. http://dx.doi.org/10.1007/s11295-011-0427-x

[28] Lusini, I., Velichkov, I., Pollegioni, P., Chiocchini, F., Hinkov, G., Zlatanov, T., Cherubini, M. and Mattioni, C. (2014) Estimating the Genetic Diversity and Spatial Structure of Bulgarian *Castanea sativa* Populations by SSRs: Implications for Conservation. *Conservation Genetics*, **15**, 283-293. http://dx.doi.org/10.1007/s10592-013-0537-0

[29] Krebs, P., Conedera, M., Pradella, M., Torriani, D., Felber, M. and Tinner, W. (2004) Quaternary Refugia of the Sweet Chestnut (*Castanea sativa* Mill): An Extended Palynological Approach. *Vegetation History and Archaeobotany*, **13**, 145-160. http://dx.doi.org/10.1007/s00334-004-0048-5

[30] Kalinowski, S.T. (2005) HP-Rare: A Computer Program for Performing Rarefaction on Measures of Allelic Diversity. *Molecular Ecology Notes*, **5**, 187-189. http://dx.doi.org/10.1111/j.1471-8286.2004.00845.x

[31] Nei, M. (1973) Analysis of Gene Diversity in Subdivided Populations. *Proceedings of the National Academy of Sci-*

ences of the United States of America, **70**, 3321-3323. http://dx.doi.org/10.1073/pnas.70.12.3321

[32] Peakall, R. and Smouse, P.E. (2005) GenAlEx v6: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research. *Molecular Ecology Notes*, **6**, 288-295. http://dx.doi.org/10.1111/j.1471-8286.2005.01155.x

[33] Shepard, D. (1968) A Two-Dimensional Interpolation Function for Irregularly-Spaced Data. *Proceedings of the* 1968 *ACM National Conference*, New York, 27-29 August 1968, 517-524. http://dx.doi.org/10.1145/800186.810616

[34] Hengl, T. (2009) A Practical Guide to Geostatistical Mapping. University of Amsterdam, Amsterdam, 1-26. http://spatial-analyst.net/book/

[35] Pritchard, J.K., Stephens, M. and Donnelly, P. (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, **155**, 945-959.

[36] Murphy, M.A., Evans, J.S., Cushman, S.A. and Storfer, A. (2008) Representing Genetic Variation as Continuous Surfaces: An Approach for Identifying Spatial Dependency in Landscape Genetic Studies. *Ecography*, **31**, 685-697. http://dx.doi.org/10.1111/j.1600-0587.2008.05428.x

[37] Manni, F., Guerard, E. and Heyer, E. (2004) Geographic Patterns of (Genetic, Morphologic, Linguistic) Variation: How Barriers Can Be Detected by Using Monmonier's Algorithm. *Human Biology*, **76**, 173-190. http://dx.doi.org/10.1353/hub.2004.0034

[38] Nei, M. (1972) Genetic Distance between Populations. *The American Naturalist*, **106**, 283-392. http://dx.doi.org/10.1086/282771

[39] Felsenstein, J. (2005) PHYLIP (Phylogeny Inference Package). Department of Genome Sciences, University of Washington, Seattle.

[40] Petit, R.J., El Mousadik, A. and Pons, O. (1998) Identifying Populations for Conservation on the Basis of Genetic Markers. *Conservation Biology*, **12**, 844-855. http://dx.doi.org/10.1046/j.1523-1739.1998.96489.x

[41] EUFORGEN (2009) Distribution Map of Chestnut (*Castanea sativa*). www.euforgen.org.

Supplementary Material 1. List of Sampling Location of 31 European Chestnut Populations

Table S1. Number of plants (N), sampling location and geographic coordinates (latitude, longitude) for 31 European chestnut populations, as reported in [29].

Country	Region	Sampling site	ID[a]	Latitude	Longitude	N
Spain	Galicia	Sierra do Faro	ES_01	42°36'29"N	07°51'59"W	23
	Andalucia	Sierra do Ronda	ES_02	36°32'20"N	0.5°18'34"W	26
	Asturias	Navaliego	ES_03	43°13'01"N	05°45'50"W	29
	Galicia	Costa Atlanticas	ES_04	43°17'09"N	08°22'11"W	21
Italy	Sicilia	Madonie	I_01	37°49'56"N	14°05'12"E	26
	Calabria	SilaPiccola	I_02	39°02'55"N	16°43'03"E	26
	Basilicata	Mt. Vulture	I_03	40°54'56"N	15°36'50"E	25
	Marche	Mt. dellaLaga	I_04	42°44'05"N	13°24'45"E	26
	Lazio	Mt. Cimini-	I_05	42°24'50"N	12°11'38"E	23
	Toscana	Mugello	I_06	43°58'07"N	11°34'03"E	19
	Piemonte	Trontano	I_07	46°07'12"N	08°20'07"E	26
	Piemonte	V. Pellice	I_08	44°48'02"N	07°08'17"E	26
	Friuli	V. del Natisone	I_09	46°07'17"N	13°33'19"E	26
Greece	S-E-Macedonia	Holomontas	GR_01	40°32'01"N	23°44'44"E	26
	C-Macedonia	Hortiatis	GR_02	40°35'40"N	22°22'50"E	26
	W-Macedonia	Dafni	GR_03	40°16'38"N	21°08'37"E	26
	N-Macedonia	Paiko	GR_04	40°57'09"N	22°22'44"E	26
Eastern Turkey	Duzce	Akcakoca	TR_01	41°05'15"N	31°07'25"E	24
	Sinop	Sinop	TR_02	42°01'36"N	35°09'01"E	31
	Artvin	Hopa	TR_03	41°24'01"N	41°29'23"E	22
	Trabzon	Meryem Ana	TR_04	40°44'50"N	39°34'50"E	30
	Giresun	Giresun	TR_05	40°54'24"N	38°31'22"E	26
Central Turkey	Sakarya	Karadere	TR_06	40°43'56"N	30°51'43"E	21
	Kocaeli	Sardala	TR_07	41°02'09"N	29°57'13"E	23
	Yalova	Cinarcik	TR_08	40°38'42"N	29°07'12"E	26
	Kocaeli	Golcuk	TR_09	40°43'20"N	29°48'54"E	25
	Sakarya	Sapanca	TR_10	40°41'29"N	30°16'06"E	24
Western Turkey	Bursa	Bursa	TR_11	40°11'40"N	29°03'35"E	25
	Canakkale	Bayramic	TR_12	39°48'36"N	26°36'37"E	31
	Izmir	Kemalpasa	TR_13	38°26'07"N	27°15'44"E	23
	Manisa	Demirci	TR_14	39°02'41"N	28°39'01"E	22

[a]ID = Identity code of population.