



Functional genetic diversity of chestnut (*Castanea sativa* Mill.) populations from southern Spain

María I. Cuestas¹, Claudia Mattioni², Luis M. Martín¹, Enrique Vargas-Osuna³, Marcello Cherubini² and María A. Martín⁴

¹ Universidad de Córdoba, Dept. Genética, Campus de Excelencia Internacional Agroalimentario (ceiA3), 14071 Córdoba, Spain. ² Consiglio Nazionale delle Ricerche (CNR), Istituto di Biologia Agroambientale e Forestale, Viale Marconi, 2, 05010 Porano (TR), Italy. ³ Universidad de Córdoba, Dept. Ciencias y Recursos Agrícolas y Forestales, Campus de Excelencia Internacional Agroalimentario (ceiA3), 14071 Córdoba, Spain. ⁴ Universidad de Extremadura, Centro Universitario de Plasencia, Dept. Ingeniería del Medio Agronómico y Forestal, Avda. Virgen del Puerto N° 2, 10600 Plasencia (Cáceres), Spain.

Abstract

Aim of the study: To evaluate the adaptive genetic variability of chestnut (*Castanea sativa* Mill.) populations from southern Spain in relation to bud burst and water stress.

Area of study: Andalusia (southern Spain) where many chestnut groves were progressively abandoned and have become ‘naturalized’.

Material and methods: A total of 126 chestnut trees from eight populations were assessed by means of nine genic microsatellite loci (expressed sequence tag simple sequence repeat markers) related to bud burst and water stress.

Main results: Significant differences in genetic diversity were detected within and among populations, not found with neutral microsatellite markers. The structure analysis indicated the presence of two different gene pools.

Research highlights: These results could contribute to the development of conservation strategies for this species in southern areas exposed to the effects of climate change. The genetic diversity of these populations could be useful in minimizing this risk and other predictable factors related to global change.

Additional keywords: functional markers; adaptation; population genetic structure.

Abbreviations used: EST_s (expressed sequence tag); SSR_s (simple sequence repeats).

Authors' contributions: MIC conducted the laboratory analyses and wrote the paper. CM participated in the design of the experiment, data analysis, discussion and redaction of the paper. LMM conducted the collection of data in the field and coordinated the research project. EVO participated in the collection of data in the field. MC led the laboratory analysis. MAM supervised the work and participated in the design of the experiment, discussion and redaction of the paper.

Citation: Cuestas, M. I.; Mattioni, C.; Martín, L. M.; Vargas-Osuna, E.; Cherubini, M.; Martín, M. A. (2017). Short communication: Functional genetic diversity of chestnut (*Castanea sativa* Mill.) populations from southern Spain. Forest Systems, Volume 26, Issue 3, eSC06. <https://doi.org/10.5424/fs/2017263-11547>

Supplementary material (Tables S1, S2) accompanies the paper on FS's website.

Received: 17 Apr 2017. **Accepted:** 21 Nov 2017

Copyright © 2017 INIA. This is an open access article distributed under the terms of the Creative Commons Attribution (CC-by) Spain 3.0 License.

Funding: Spanish Ministry of Economy and Competitiveness (grants AGL2013-48017-C2-1-R and AGL2014-53822-C2-1-R); European Regional Development Fund (FEDER); Secretaría General de Ciencia, Tecnología e Innovación de la Consejería de Economía e Infraestructuras, Regional Government of Extremadura, Spain (financial support to M. A. Martín).

Competing interests: The authors have declared that no competing interests exist.

Correspondence should be addressed to María A. Martín: angelamartin@unex.es

Introduction

Sweet chestnut (*Castanea sativa* Mill.) can be found over a wide range of climatic conditions in Europe, and its current distribution is the result of natural colonization and human intervention (Conedera *et al.*, 2004; Mattioni *et al.*, 2013).

Currently, chestnut forests are ecosystems that, as well as fruit and timber, provide environmental and landscape values; however, they face multiple threats associated with global change as the abandonment of chestnut stands after land-use changes. This has caused

a severe decrease of biodiversity in the affected regions and reduced ecosystem service provision (Conedera *et al.*, 2004).

Genetic diversity plays a vital role in determining the ability for adaptation of forest populations to new conditions (Fady *et al.*, 2016). Molecular markers such as microsatellites (SSRs), have become the most used tool for studying forest genetics, because they are highly polymorphic, codominant and widespread across the entire genome (Powel *et al.*, 1996). However, these markers are neutral, and unsuitable for estimating adaptive genetic diversity, while the microsatellite

markers developed from expressed sequences (EST-SSRs) having putative functions, can be used to examine functional diversity in relation to adaptive variation (Varshney *et al.*, 2005). In recent years, these markers have been widely used due to their high transferability between related species, and have furnished additional information on genetic diversity obtained with other neutral markers in many forest trees (Homolka *et al.*, 2013; Uchiyama *et al.*, 2013).

Genetic variation in chestnut populations has been extensively studied using different molecular neutral markers (Fineschi *et al.*, 1994; Mattioni *et al.*, 2013). These studies showed a high genetic diversity in chestnut populations throughout Europe and indicated the origin of the current distribution of European populations from different glacial refuges (Mattioni *et al.*, 2013). Likewise, SSRs were used to verify the genetic identity of the main cultivars in Europe (Pereira-Lorenzo *et al.*, 2010). Furthermore, genic markers demonstrated a signal of adaptation related to bud burst in chestnut populations from contrasting environmental conditions in Europe (Martín *et al.*, 2010).

In Spain, a study of the genetic diversity in chestnut populations using neutral markers showed a clear geographic pattern with three different groups of populations corresponding to the northwest, northeast and southeast of Spain, respectively (Martín *et al.*, 2012). In particular, populations from southern Spain displayed both a high level of genetic diversity and differentiation (Martín *et al.*, 2012). In this region, chestnut groves were progressively abandoned due to the rural exodus in the 1960s and 1970s, and have become 'naturalized' stands, constituting in some cases relict woodlands (Martín *et al.*, 2007). Until now, adaptive diversity of naturalized stands in southern Spain has not been evaluated using functional markers. Considering southern Europe as an area particularly affected by climate change, it is expected that consequences of the warming effect will be evident in forests from the Mediterranean basin (Coll *et al.*, 2013).

The current study was carried out to evaluate the adaptive genetic variability in a set of chestnut populations from southern Spain by means of EST-SSR markers. The main objectives were to: 1) evaluate the polymorphism of the markers in these populations, and 2) compare the genetic diversity obtained for these populations with EST-SSRs that reported previously using neutral SSRs (Martín *et al.*, 2012).

Material and methods

A total of 126 chestnut trees were sampled from eight different populations in southern Spain (Table S1).

Seven of these populations were previously evaluated using neutral SSRs in order to assess their genetic structure (Martín *et al.*, 2012).

The genomic DNA was isolated by 20 mg of lyophilized leaf tissue using the Qiagen DNeasy™ Plant mini Kit according to the manufacturer's instructions. Nine EST-SSR loci developed from *Quercus* spp. and expressed during drought stress and bud burst were assessed (Durand *et al.*, 2010; Bodènés *et al.*, 2012). A multiple PCR was conducted in a 12.5 µL volume using the Qiagen Type-it Microsatellite PCR Kit with the following program: initial denaturation at 95°C for 5 min; 28 cycles at 95°C for 30s, annealing at 57°C for 90s, and elongation at 72°C for 30s; and a final extension step at 60° for 30min. Amplification products were separated on an ABI PRISM 3100 DNA sequencer and allele scoring was performed using Gen Mapper software.

The number of total alleles for locus (A), mean number of alleles per population (N_a), the observed (H_o), expected (H_e) and unbiased expected heterozygosity (uHe), and the private alleles in each population (PA) were estimated using Arlequin 3.11 (Excofier *et al.*, 2005). Allelic richness (A_r) based on a minimum sample size of 10 individuals was calculated using FSTAT (Goudet, 2001). The inbreeding coefficient F_{IS} (Weir & Cockerham, 1984) was computed using hierarchical locus-by-locus AMOVA as implemented in Arlequin 3.11 and its deviation from zero tested by 10,000 allele permutations. Differentiation among populations was calculated by F_{ST} (Weir & Cockerham, 1984) and R_{ST} (Slatkin, 1995).

The population structure and proportion of membership (Q value) for each predefined population was inferred using STRUCTURE v.2.3.4 software (Pritchard *et al.*, 2000). The analysis was performed using the admixture model on the whole dataset (Falush *et al.*, 2007), considering a minimum threshold of 0.75 in the value of Q (Mattioni *et al.*, 2013). Six independent runs were performed for K with a burn in period of 10,000 steps followed by 10^5 Markov chain Monte Carlo (MCMC) replicates. The ΔK defined by Evanno *et al.* (2005) was used to detect the most likely number of populations.

Results and discussion

A total of 38 different alleles were identified in the 126 individuals with an average of 4.22 alleles per locus. The range of detected alleles was from three (loci *PIE233* and *FIR030*) to seven (locus *PIE227*) (Table 1). Four populations were polymorphic for all loci (Pujerra, Paterna and S.M. Trassierra), whereas the Bubión, S.

Table 1. Genetic diversity assessed for the nine EST-SSR markers

Locus	Size range (bp)	A	Ho	He
<i>FIR030</i>	168-173	3	0.429	0.530
<i>GOT014</i>	108-116	4	0.254	0.698
<i>PIE227</i>	154-179	7	0.579	0.681
<i>PIE228</i>	177-196	5	0.667	0.730
<i>PIE233</i>	162-168	3	0.214	0.365
<i>PIE260</i>	154-168	4	0.746	0.630
<i>POR009</i>	122-140	4	0.635	0.614
<i>POR026</i>	137-148	4	0.714	0.670
<i>WAG004</i>	260-271	4	0.135	0.157
Mean		4.22	0.486	0.566

A, number of alleles; Ho, observed heterozygosity; He, expected heterozygosity.

Norte and Gaucín populations were monomorphic for locus *WAG004* and the S. Elena population for locus *PIE233*. All populations displayed similar values of genetic diversity, although the lowest values were in Gaucín (in terms of Na, Ar and uHe) (Table 2). The Gaucín and S.M. Trassiera populations had private alleles. The inbreeding coefficient (F_{IS}) deviated significantly from zero only in the Bubión population.

These results agree with those obtained in a previous study comparing genomic and genic markers in European chestnut populations (Martín *et al.*, 2010), although the level of polymorphism and genetic diversity was lower than those obtained by Martín *et al.* (2012) using SSR makers in the same populations. Lower levels of polymorphism have also been described in *Quercus* ssp. for EST-SSRs although differences between both

types of markers were not significant, indicating that most of the EST-SSRs chosen for the study reflected neutral variation (Lind & Gailing, 2013).

The coefficients of genetic differentiation among the eight populations were $F_{ST} = 0.150$ and $R_{ST} = 0.099$ (Table 2). Likewise, AMOVA indicated significant differences among populations despite the limited area of the study (Table S2 [suppl]). The results obtained from STRUCTURE indicated that the most probable division with the strongest support in terms of log-likelihood values was for $K = 2$, separating the populations into two groups (Fig. 1). The Paterna, Güejar and Bubión populations clearly belonged to Cluster I with more than 80% of their members grouped in the same gene pool, while 77% of the members of the Pujerra population were in the same gene pool. On

Table 2. Genetic diversity of the eight evaluated chestnut populations

Population	Na	Ar	PA	Ho	He	uHe	F_{IS}	Q_1	Q_2	F_{ST}	R_{ST}
Santa María Trassiera	3.56	3.15	3	0.521	0.466	0.474	-0.204	0.09	0.91		
Sierra Norte	3.22	3.04	0	0.529	0.465	0.482	-0.021	0.47	0.53		
Gaucín	2.78	2.57	1	0.444	0.424	0.437	-0.189	0.33	0.67		
Santa Elena	3.00	3.00	0	0.433	0.454	0.478	-0.226	0.50	0.50		
Pujerra	3.22	3.12	0	0.519	0.443	0.462	-0.206	0.77	0.23		
Güejar	3.22	3.10	0	0.481	0.509	0.527	0.075	0.83	0.17		
Paterna	3.11	2.92	0	0.437	0.508	0.526	0.056	0.85	0.15		
Bubión	3.00	2.94	0	0.489	0.533	0.552	0.251*	0.83	0.17		
										0.150	0.099

Na, mean number of alleles per locus; Ar, allelic richness; PA, private allele; Ho, observed heterozygosity; He, expected heterozygosity; uHe, unbiased expected heterozygosity; F_{IS} , inbreeding coefficient; Q_1 and Q_2 , percentage of membership of predefined population inferred by STRUCTURE (Q values > 0.75 are reported in bold); F_{ST} , coefficient of genetic differentiation among populations according to Weir & Cockerham (1984); R_{ST} , coefficient of genetic differentiation among populations according to Slatkin (1995); * $p < 0.05$.

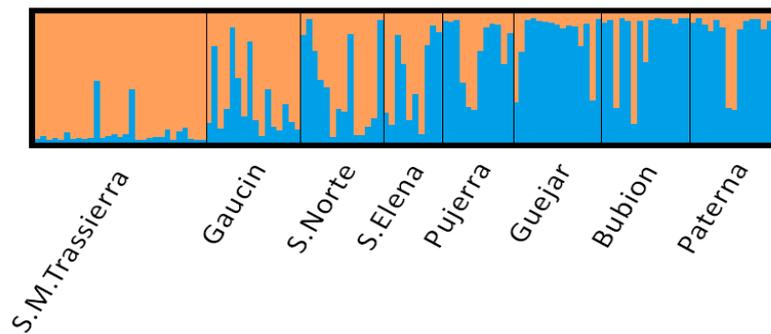


Figure 1. Bayesian clustering of the eight chestnut populations obtained with STRUCTURE software for $K=2$. Each individual is represented by a vertical line and populations are separated by black lines.

the other hand, S.M. Trassierra was the only population in Cluster II. The remaining populations showed an admixture of gene pools, with the Gaucín population more similar to the S.M. Trassierra population ($Q_2=67\%$). Likewise, Gaucín and S.M. Trassierra were the only populations with private alleles, which are usually considered to have potential to respond to selection or to have evolutionary significance (Petit *et al.*, 1998).

The genetic structure obtained with the functional markers was compared with those obtained using neutral markers by Martín *et al.* (2012). The eastern populations (Bubión, Guejar and Paterna) showed the same pattern of clustering with both types of markers. Likewise, the Pujerra, S. Elena and S. Norte populations displayed a high degree of admixture for both types of markers. However, for functional markers, the Gaucín population showed greater similarity with Andalusian populations compared with neutral markers, for which it was associated with northwest populations (Martín *et al.*, 2012). Furthermore, similar coefficients of differentiation were found at both types of markers (0.150 vs. 0.145). This similar clustering between most of populations using both EST-SSRs and neutral SSRs could indicate neutral variation, as also reported by other authors using EST-SSRs (Lind & Gailing, 2013). Nevertheless, the different behaviour of Gaucín could indicate an adaptive signal for this population, although further analyses with a higher number of EST-SSR markers would be necessary.

In conclusion, the EST-SSR markers used revealed a clear genetic structure of Andalusian populations and differences in diversity within populations, although no adaptive signal was detected, may be due to the limited area sampled. Additional sampling and further analysis could provide better insight into adaptation of these populations. Analysis of genetic variation within and between chestnut populations could provide useful information to establish future conservation strategies. Thus, we highlight the importance of the S.M. Trassierra

and Gaucín populations, as representative of the less abundant genetic pool detected with our markers, as well as because of the presence of private alleles.

References

- Bodénès C, Chancere E, Gailing O, Vendramin GG, Bagnoli F, Durand J, Goicoechea PG, Soliani C, Villani F, Mattioni C, *et al.*, 2012. Comparative mapping in the Fagaceae and beyond with EST-SSRs. *BMC Plant Biol* 12: 153. <https://doi.org/10.1186/1471-2229-12-153>
- Coll M, Peñuelas J, Ninyerola M, Pons X, Carnicer J, 2013. Multivariable effect gradients driving forest demographic responses in the Iberian Peninsula. *Forest Ecol Manag* 303: 19-209. <https://doi.org/10.1016/j.foreco.2013.04.010>
- Conedera M, Krebs P, Tinner W, Pradella M, Torriani D, 2004. The cultivation of *Castanea sativa* (Mill.) in Europe, from its origin to its diffusion on a continental scale. *Veg Hist Archaeobot* 13: 161-179. <https://doi.org/10.1007/s00334-004-0038-7>
- Durand J, Bodénès C, Chancerel E, Frigerio JM, Vendramin G, Sebastiani F, Buonamici A, Galiling O, Koelewijn H, Villani F, *et al.*, 2010. A fast and cost-effective approach to develop and map EST-SSR markers: Oak as a case study. *BMC Genomics* 11: 1. <https://doi.org/10.1186/1471-2164-11-570>
- 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. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier L, Laval G, Schneider S, 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinformatics* 1: 47. <https://doi.org/10.1177/117693430500100003>
- Fady B, Cottrell J, Ackzell L, Alía R, Muys B, Prada A, González-Martínez SC, 2016. Forests and global change: what can genetics contribute to the major forest management and policy challenges of the twenty-first

- century? *Reg Environ Change* 16: 927-939. <https://doi.org/10.1007/s10113-015-0843-9>
- 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. <https://doi.org/10.1111/j.1471-8286.2007.01758.x>
- Fineschi S, Malvolti ME, Morgante M, Vendramin GG, 1994. Allozyme variation within and among cultivated varieties of sweet chestnut (*Castanea sativa*). *Can J Forest Res* 24: 1160-1165. <https://doi.org/10.1139/x94-153>
- Goudet J, 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (vers. 2.9.3). Lausanne University, Lausanne.
- Homolka A, Schueler S, Burg K, Fluch S, Kremer A, 2013. Insights into drought adaptation of two European oak species revealed by nucleotide diversity of candidate genes. *Tree Genet Genomes* 9: 1179-1192. <https://doi.org/10.1007/s11295-013-0627-7>
- Lind JF, Gailing O, 2013. Genetic structure of *Quercus rubra* L. and *Quercus ellipsoidal* E. J. Hill populations at gene-based EST-SSR and nuclear SSR markers. *Tree Genet Genomes* 9: 707-722. <https://doi.org/10.1007/s11295-012-0586-4>
- Martín MA, Moral A, Martín LM, Álvarez JB, 2007. The genetic resources of European sweet chestnut (*Castanea sativa* Miller) in Andalusia, Spain. *Genet Resour Crop Ev* 54: 379-387. <https://doi.org/10.1007/s10722-005-5969-z>
- Martín MA, Mattioni C, Cherubini M, Turchini D, Villani F, 2010. Genetic diversity in European chestnut populations by means of genomic and genic microsatellite markers. *Tree Genet Genomes* 6: 735-744. <https://doi.org/10.1007/s11295-010-0287-9>
- Martín MA, Mattioni C, Molina JR, Álvarez JB, Cherubini M, Herrera MA, Villani F, Martín LM, 2012. Landscape genetic structure of chestnut (*Castanea sativa* Mill.) in Spain. *Tree Genet Genomes* 8: 127-136. <https://doi.org/10.1007/s11295-011-0427-x>
- Mattioni C, Martín MA, Pollegioni P, Cherubini M, Villani F, 2013. Microsatellite markers reveal a strong geographical structure in European populations of *Castanea sativa* (Fagaceae): evidence for multiple glacial refugia. *Am J Bot* 100: 951-961. <https://doi.org/10.3732/ajb.1200194>
- Pereira-Lorenzo S, Costa R, Ramos-Cabrera A, Ribeiro C, da Silva M, Manzano G, Barreneche T, 2010. Variation in grafted European chestnut and hybrids microsatellite reveals two main origins in the Iberian Peninsula. *Tree Genet Genomes* 5: 701-715. <https://doi.org/10.1007/s11295-010-0285-y>
- Petit RJ, El Mousadik A, Pons O, 1998. Identifying populations for conservation on the basis of genetic markers. *Conserv Biol* 12: 844-855. <https://doi.org/10.1046/j.1523-1739.1998.96489.x>
- Powell W, Machray GC, Provan J, 1996. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci* 1: 215-222. [https://doi.org/10.1016/S1360-1385\(96\)86898-0](https://doi.org/10.1016/S1360-1385(96)86898-0)
- Pritchard JK, Stephens M, Donnelly P, 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Slatkin M, 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139: 457-462.
- Uchiyama K, Fujii S, Ishizuka W, Goto S, Tsumura Y, 2013. Development of 32 EST-SSR markers for *Abies firma* (Pinaceae) and their transferability to related species. *Appl Plant Sci* 1: 1200464. <https://doi.org/10.3732/apps.1200464>
- Varshney RK, Graner A, Sorrells ME, 2005. Genic microsatellite markers in plants: features and applications. *Trends Biotechnol* 23: 48-55. <https://doi.org/10.1016/j.tibtech.2004.11.005>
- Weir BS, Cockerham CC, 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.