Theor Appl Genet (2003) 107:730–735
DOI 10.1007/s00122-003-1298-9

J. B. Alvarez · C. Muñoz-Diez · A. Martín-Cuevas · S. Lopez · L. M. Martín

Cotyledon storage proteins as markers of the genetic diversity in *Castanea sativa* Miller

Received: 30 January 2003 / Accepted: 17 March 2003 / Published online: 15 May 2003

Abstract This study has been to analyse the useful nut globulin proteins as a marker of the genetic diversity in *Castanea sativa*. The evaluated populations were highly polymorphic for the globulins, being detected up to 35 polymorphic bands with a wide distribution among all the evaluated populations. Taken together for populations from all the chestnut regions, about 39.3% of total allelic variation was distributed among the populations. The estimates of genetic similarity between populations were clearly associated with the collecting site. This method of analysis of the nut storage proteins (globulins) could be a useful tool for the evaluation of genetic diversity in this and other species of the Fagaceae.

Keywords Cotyledon storage proteins · Genetic resources · Globulins · Sweet chestnut

Introduction

The genus *Castanea* contains up to 13 species and a number of naturally occurring hybrids (Goldsbrough 1990). These belong to the family Fagaceae, which includes the oaks. All members of the genus are obligate out-crossers, and the genus as a whole is noted for a high level of variation. All species have a somatic chromosome number of $2x = 24$ (Jaynes 1963).

The only native species of Europe is the sweet chestnut (*Castanea sativa* Mill.), which was mainly cultivated both for wood and for fruit since ancient times in the Mediterranean region, being expanded as a crop for the Romans (Adua 1999). This species is widely distributed for all the continents, being the principal extension in France, Italy and Spain. In Spain, the main stands appear in the Northern, although important stands have been found in Andalusia, mainly in the provinces Huelva and Malaga (Berrocal et al. 1998).

In the South of Spain (Andalusia), this species is cultivated for fruit in two areas, the Genal river Valley (Malaga) and the Natural Park of Aracena and Aroche Peaks (Huelva). Minor extensions of this species appear in other provinces of Andalusia where the use is mainly for wood. Recently, one study has begun for the maintainance of this species in the region where its forestry and agriculture use can be associated to the landscape values.

The use of sweet chestnut as a fruit tree must be understood as a method of in situ conservation of genetic resources. The trees are the result of two or more individuals, where the radical part came from sexual reproduction while the productive part came from vegetative multiplication. Consequently, the variability present in the patterns may be recuperated at any moment. On the other hand, some varieties used as grafts must be considered as landraces, where the level of variability is still considerable. Unfortunately, scarce studies on the genetic variability of sweet chestnut in Spain have been determined. The varieties are mainly local denominations, but their autochlonal origin cannot be established in most cases.

The storage seed proteins have been used as an important genetic marker in other species, mainly in cereals where their variability is related with the technological properties of flour (Shewry and Miflin 1985; Shewry et al. 1994). These proteins have proved to be a useful tool in the evaluation of the genetic variability of many species (Gepts 1990). The main advantages of these proteins as markers are their high polymorphism level, their simple genetic control, their environmental independency and the economy, easy and fast, of their analysis. The role of these proteins in the forest species has been scarcely explored. Only a few studies have been carried out on the nut storage proteins of the Fagaceae, mainly on the biochemical characteristics of these proteins (Collada et al. 1986, 1991; Fonseca et al.

Communicated by H.F. Linskens

J. B. Alvarez (✉) · C. Muñoz-Diez · A. Martín-Cuevas · S. Lopez · L. M. Martín
Departamento de Genética,
Escuela Técnica Superior de Ingenieros Agrónomos y de Montes,
Universidad de Córdoba, Apdo. 3048, E-14080 Córdoba, Spain,
e-mail: ge2alcaj@uco.es

1997). However, the use of these proteins for evaluation of the structure of populations and their genetic variability has not been careful.

The main goal of the present study has been to analyse the useful nut globulin proteins as marker of the genetic diversity in *C. sativa*, for a collection of chestnut trees collected in Andalusia (South of Spain).

Materials and methods

Seed samples

Samples of fruits from 146 sweet chestnut (*C. sativa* Mill.) trees collected in the principal cultivated regions of this species in Andalusia (South of Spain) have been used. These materials were grouped into 20 populations by geographic criteria (Table 1).

Protein extraction and electrophoresis analysis

Twenty milligrams of cotyledons are crushed and mixed with 250 μl of 50 mM Tris-HCl pH 8.5, + 500 mM of NaCl. The samples were incubated at room temperature (about 20 °C) for 1 h. After centrifugation, the supernatant was transferred to a new tube, and the proteins were precipitated by the addition of cold-acetone. This dried pellet was resuspended in buffer containing 625 mM of Tris-HCl pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 0.02% (w/v) bromophenol blue and 2% (w/v) dithiothreitol in a 1:5 ratio (mg/μl).

Reduced proteins were fractionated by electrophoresis in vertical SDS-PAGE slabs in a discontinuous Tris-HCI-SDS buffer system (pH 6.8/8.8) at a polyacrylamide concentration of 8%, 10% and 12%, and C = 2.67%. The Tris-HCl/glycine buffer system of Laemmli (1970) was used. Electrophoresis was performed at a constant current of 30 mA/gel at 18 °C. Gels were stained overnight with 12% (w/v) trichloroacetic-acid solution containing 5% (v/v) ethanol and 0.05% (w/v) Coomassie Brilliant Blue R-250. De-staining was carried out with tap water.

Statistical analysis

The following genetic variability parameters were calculated in all the populations: namely, the number of alleles per locus (A), the percentage of polymorphic loci (P), the effective number of alleles per locus (Ne) and the genetic diversity (He).

The gene diversity over all populations (Ht), together with the average gene diversities within (Hs) and among (Dst) populations, were calculated according to Nei (1973). The relative magnitude of genic differentiation among populations, Gst, was estimated as Dst/Ht.

The genetic identify (I) values were calculated between populations (Nei 1972). The Nei's genetic distances (D) were used to generate an unweighted pair-group clustering based on the arithmetic averages (UPGMA) phenogram (Sneath and Sokal 1973). This was tested by the co-phenetic matrix correlation during the re-construction of a co-phenetic matrix based on the tree matrix.

Results

Globulin differentiation within and among populations

The analysis of the evaluated trees for globulins has shown up to 35 polymorphic bands. Three zones were determinated in the gel, named as zone A, zone B and zone C, which were determined by the range of molecular weights (Fig. 1.1). Zone A contains the bands with a molecular weight between 45 and 67 kDa, zone B between 30 and 45 kDa, and zone C between 20 and 30 kDa. Eleven, twelve and twelve bands were detected in each zone, respectively (Fig. 1.2, 1.3 and 1.4). The differences between bands were mainly by mobility, although in some cases differences can be observed due to intensity. The bands were named with an increased number for mobility and a capital letter for the gel zone (p.e. 1A is band 1 of zone A). In two cases the differences in mobility were low, and a lower letter was added (p. e. 1Aa and 1Ab).

The bands that showed a lower frequency were 1Ab and 10A for zone A, 5B, 6B, 7B and 12B for zone B, and 5C, 6Ca and 9C for zone C, which appear in 20% of the chestnut trees evaluated (Table 2). By contrast, the highest frequency bands were 1Aa and 6A for zone A that appear in more than 90% of the evaluated trees.

Table 1 Location of the chestnut populations evaluated in the present work.

Population	Size	Latitude N	Longitude W	Altitude mean (s.l.s)	Locality
PCO-1	8	37°55′10″	4°50′48″	350	Córdoba (Córdoba)
PGR-1	12	36°57′59″	3°31′11″	750	Mondujar (Granada)
PHU-1	11	35°53′20″	6°46′41″	850	Almonaster la Real (Huelva)
PHU-2	19	37°53′01″	6°32′01″	578	Aracena (Huelva)
PHU-3	3	37°58′17″	6°48′38″	721	Aroche (Huelva)
PHU-4	7	37°53′45″	6°48′33″	705	Cortegana (Huelva)
PHU-5	9	37°53′47″	6°40′11″	822	Fuenteheridos (Huelva)
PHU-6	6	37°53′01″	6°36′17″	711	Los Marines (Huelva)
PHU-7	5	37°57′00″	6°36′19″	717	Los Marines (Huelva)
PMA-1	7	36°36′07″	5°14′37″	483	Benadalid (Malaga)
PMA-2	3	36°36′42″	5°11′56″	640	Faraján (Malaga)
PMA-3	11	36°33′15″	5°13′20″	634	Jubrique (Malaga)
PMA-4	6	36°38′58″	5°07′47″	793	Parauta (Malaga)
PMA-5	10	36°36′19″	5°08′47″	867	Pujerra (Malaga)
PMA-6	5	36°36′58″	5°09′02″	769	Pujerra (Malaga)
PMA-7	5	36°36′17″	5°07′56″	750	Pujerra (Malaga)
PMA-8	5	36°36′10″	5°07′26″	1,006	Pujerra (Malaga)
PMA-9	7	36°36′37″	5°09′09″	855	Pujerra (Malaga)
PSE-1	3	37°55′34″	5°35′34″	550	Constantina (Sevilla)
PSE-2	4	37°53′34″	5°36′48″	556	Constantina (Sevilla)

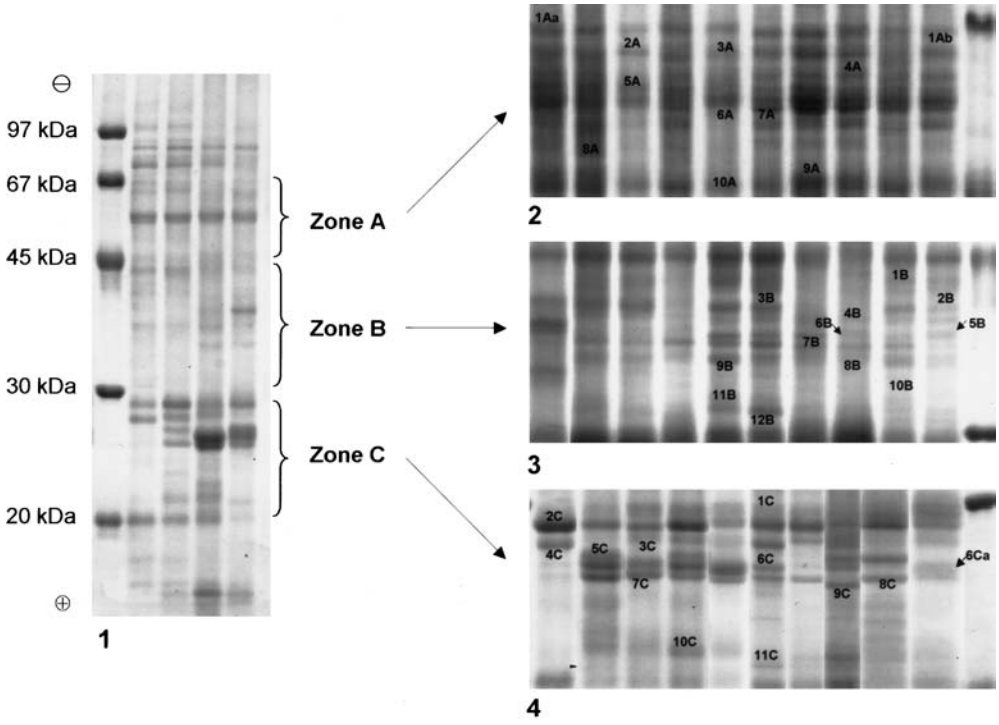

Fig. 1 SDS-PAGE of globulins from cotyledons of chestnut. The different zones of the gel were amplified with increased polyacrylamide concentration. **2** *Zone A* = 8% concentration; **3** *zone B* = 10% concentration; and **4** *zone C* = 12% concentration

In order to assess the distribution of alleles in different populations, the classification of Marshall and Brown (1975) was used. In general, the bands presented a wide distribution among all the evaluated populations. The bands that showed low frequency appear with two types of distribution. For zone A, the band 1Ab was detected in three populations (PGR-1, PSE-1 and PSE-2) from two chestnut regions separated geographically by a large distance; the band 10A appears exclusively in populations of the Huelva province (PHU-1 to PHU-7). In both cases, although the percentage was higher than 5%, these bands may be considered as a local distribution. Only one band of zone B can be considered, while the rare, band 11B only appeared in the Granada population (PGR-1). The band 4B appeared in three out of five collected regions (Córdoba, Granada and Sevilla), while band 5B appeared in the other two (Huelva and Malaga). On the other hand, band 6B was detected in all provinces with the exception of Huelva. For zone C, band 5C was not detected in the Granada population (PGR-1). Band 6Ca was only detected in one individual of the Córdoba population (PCO-1), while band 9C appeared in two populations of Huelva (PHU-1 and PHU-2) and one of Sevilla (PSE-2).

A summary of the genetic data of these 35 globulin genes for the 20 populations is given in Table 3. The PHU-3 population was the lowest polymorphic one, due to only seven out of the 35 globulin bands that showed variation. The highest polymorphic populations were PHU-2 and PMA-5 that presented variation in 26 out of them. The expected heterozygosity (He) showed a mean value of 0.214, ranging between 0.089 for the PHU-3 population and 0.278 for the PMA-5 population.

Additional characterization of the diversity in chestnut for the globulin proteins is presented in Table 4. The

Table 2 Frequencies of each band in the 146 evaluated chestnut trees and 20 populations

Zone	Band	Accessions (n = 146)		Population (n = 20)	
		N	%	N	%
Zone A	1Aa	134	91.78	20	100.00
	1Ab	12	8.22	3	15.00
	2A	70	47.95	18	90.00
	3A	80	54.79	19	95.00
	4A	77	52.74	16	80.00
	5A	126	86.30	20	100.00
	6A	133	91.10	20	100.00
	7A	91	62.33	15	75.00
	8A	49	33.56	16	80.00
	9A	39	26.71	11	55.00
	10A	29	19.86	7	35.00
Zone B	1B	80	54.79	19	95.00
	2B	38	26.03	10	50.00
	3B	105	71.92	19	95.00
	4B	26	17.81	4	20.00
	5B	18	12.33	10	50.00
	6B	24	16.44	9	45.00
	7B	47	32.19	11	55.00
	8B	126	86.30	20	100.00
	9B	80	54.79	19	95.00
	10B	44	30.14	15	75.00
	11B	4	2.74	1	5.00
	12B	91	62.33	20	100.00
Zone C	1C	69	47.26	15	75.00
	2C	92	63.01	20	100.00
	3C	36	24.66	16	80.00
	4C	85	58.22	19	95.00
	5C	28	19.18	14	70.00
	6C	92	63.01	18	90.00
	6Ca	1	0.68	1	5.00
	7C	32	21.92	14	70.00
	8C	72	49.32	18	90.00
	9C	3	2.05	3	15.00
	10C	31	21.23	13	65.00
	11C	66	45.21	17	85.00

Table 3 Storage-protein diversity based on the 35 bands in the 20 populations of chestnut

Population	A[a]	Ne[a]	P[a]	He[a]
PCO-1	1.37	1.27	37.14	0.154
PGR-1	1.49	1.39	48.57	0.212
PHU-1	1.69	1.49	68.57	0.276
PHU-2	1.74	1.45	74.29	0.264
PHU-3	1.20	1.16	20.00	0.089
PHU-4	1.46	1.34	45.71	0.191
PHU-5	1.63	1.44	62.86	0.246
PHU-6	1.51	1.42	51.43	0.224
PHU-7	1.63	1.47	62.86	0.265
PMA-1	1.63	1.39	62.86	0.229
PMA-2	1.34	1.27	34.29	0.152
PMA-3	1.66	1.44	65.71	0.249
PMA-4	1.51	1.33	51.43	0.195
PMA-5	1.74	1.49	74.29	0.278
PMA-6	1.57	1.45	57.14	0.247
PMA-7	1.51	1.42	51.43	0.229
PMA-8	1.60	1.40	60.00	0.233
PMA-9	1.57	1.38	57.14	0.222
PSE-1	1.40	1.32	40.00	0.178
PSE-2	1.34	1.24	34.29	0.139
Mean	1.53	1.38	53.00	0.214

[a] A = number of alleles per locus. Ne = effective number of alleles. P = percentage polymorphic loci (5%). He = genetic diversity

Table 4 Differentiation of globulin diversity within and among 20 populations of *C. sativa* from the five chestnut regions of Andalusia

Province	N	Ht[a]	Hs[a]	Dst[a]	Gst[a] (%)
Córdoba	1	0.154	0.154	0.000	0.00
Granada	1	0.212	0.212	0.000	0.00
Huelva	7	0.289	0.222	0.067	23.26
Málaga	9	0.291	0.226	0.065	22.23
Sevilla	2	0.185	0.159	0.027	14.42
All	20	0.352	0.214	0.138	39.29

[a] Ht = total gene diversity; Hs = average gene diversity within populations; Dst = average gene diversity among populations; Gst = gene diversity among populations relative to Ht

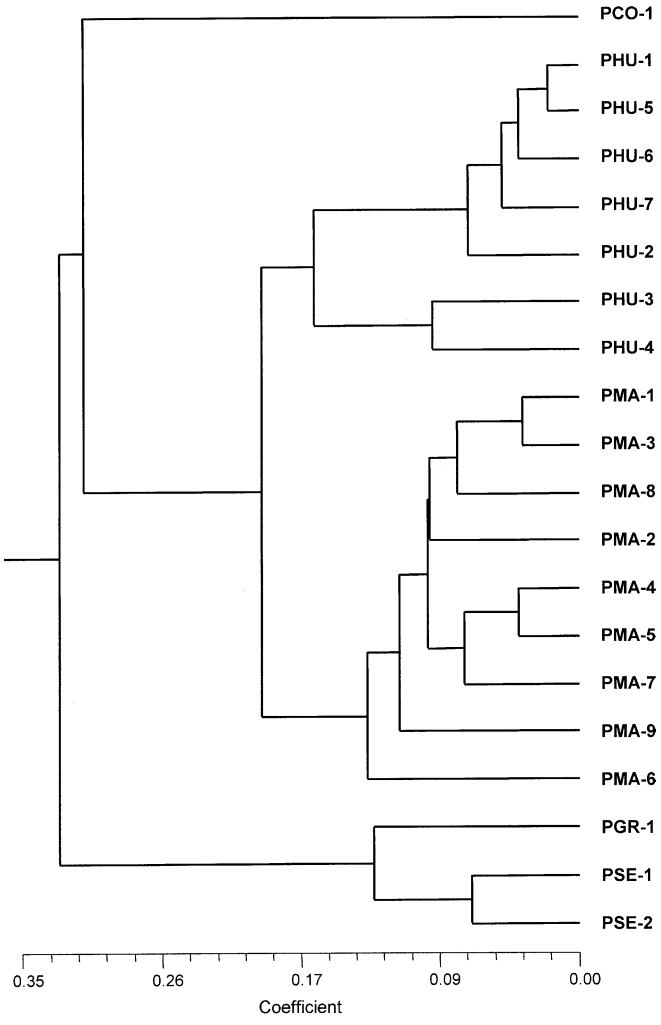

Fig. 2 Dendrogram based on Nei's genetic distance matrix between the 20 chestnut populations (co-phenetic correlation = 0.867***)

genic diversity was similar for the Huelva and Malaga populations, in both cases being about 23.3 and 22.2% of the total genetic diversity (Gst), respectively. The genic diversity of the other chestnut regions was lower, although it is important to emphasise that, due to the little importance of these regions, only one or two populations were collected with a relative small number of individuals (between 7 and 12). Taken together for populations from all chestnut regions, about 39.3% of the total allelic variation was distributed among populations, with 60.7% within populations.

The genetic diversity from the two most-important chestnut regions (Huelva and Malaga regions) was calculated. Data showed that both regions present a similar value of genetic diversity, with most part of the genetic diversity appearing within each region, while the value among regions (Gst) was only 10% of the total genetic diversity.

Genetic distances

Globulin genetic similarity coefficients (I) were calculated for paired comparisons of all 20 populations based on the normalized identity of all loci between each pair of populations (Nei 1972). The mean value of genetic similarity was I = 0.827, ranging between 0.575 and 0.981. The lowest similarity was detected between the populations PCO-1 and PHU-3 (I = 0.575), while the highest similarity was for the pair PHU-1/PHU-5 (I = 0.981).

The estimates of genetic similarity between populations were clearly associated with the collecting site. The mean of the Huelva populations was I = 0.917, the pair PHU-2/PHU-3 showed the lowest value (I = 0.762) while the pair PHU-1/PHU-5 showed the highest one (I = 0.981). On the other hand, the mean genetic similarity for the Malaga populations was I = 0.917, ranging between 0.832 (pair PMA-2/PMA-6) and 0.966 (pair PMA-1/PMA-3).

The dendrogram, based on Nei's genetic distance matrix, was tested for the significance of the clustering method. A coefficient with a co-phenetic correlation of $r = 0.867^{***}$ was observed, that represents a good fit of a cluster to the data (Rohlf and Fisher 1986). In the phenogram showed in Fig. 2, is possible that the above mentioned populations appeared, in general, to be grouped by geographic criteria. All the populations of Huelva appear joined and clearly separated by the populations of Malaga, that also appear joined.

Discussion

The materials used in the present work were collected in the regions of Andalusia where the chestnut crop is maintained. Nevertheless, in some of these regions (province of Córdoba, Granada and Sevilla) this crop could be considerate as a relict, and appear in small wild stands with only an interested landscape or with a limited wood interest. The other two regions, the Natural Park of Aracena and Aroche Peaks in Huelva and Genal river Valley in Malaga, include trees for fruit production. Likewise, in the first one, the Natural Park, the landscape values are very important.

Although all the protein fractions of the nut were analysed, the results were not satisfactory for the prolamin fraction, which are the main seed-storage proteins in the Poaceae family (Shewry and Miflin 1985). Similar results were obtained for the glutelin fraction. However, the analysis of the NaCl-soluble proteins or the globulin fraction was successful, and showed a high degree of polymorphism, up to 35 polymorphic bands in the entire trees selected. For this reason, this fraction was used to evaluate the different trees selected in this study.

The evaluated populations were highly polymorphic for the globulins, and the mean value of polymorphism was $P = 53.0\%$, ranging between 20.0 and 74.3%. These results were similarly obtained by Dane et al. (1999) with natural populations of *Castanea pumila* using isozymes. Thus, the value of He (in our study He = 0.214) was similar to the value (He = 0.227) in the other species (Dane et al. 1999); although they were lower than those found for other authors in Chinese (He = 0.328) and European (He = 0.317) *Castanea* species (Villani et al. 1991; Huang et al. 1994, 1998). Nevertheless, these results should be interpreted cautiously because of limited sampling sizes.

The proportion of genetic diversity found among the 20 *C. sativa* populations evaluated here (Gst = 39.3%) was somewhat higher than that observed for other authors in this same species (*C. sativa* Fst = 10.0%; Villani et al. 1991) or other species of the genus (*Castanea dentata* Gst = 11.0%; Huang et al. 1998). Although, because the diversity was measured with different genetic markers applied here, this could affect the level of the genetic diversity detected. However, this value was similar when the genetic diversity of the two most important chestnut regions (Huelva and Malaga) was exclusively analysed (Gst = 10.0%). Both regions showed high values of internal genetic diversity. It is important to emphasise that both regions represented more than 90% of chestnut stands in Andalusia, and that approximately 80% of the analysed trees in this study were collected in these regions.

Because of an understanding of the genetic diversity present in a species, the distribution of this variation among populations will be critical for the formulation of appropriate management strategies, and the evaluation of this must be considered as a priority for maintainance of genetic resources. This method of analysis of the nut storage proteins (globulins) could be a useful tool for the evaluation of genetic diversity in this and other species of the Fagaceae.

Acknowledgements This research was supported by an agreement (CONV-2000-63) between the Ministry of Environment from the Regional Government of Andalusia (Spain) and the University of Córdoba. The first author is grateful to the 'Ramon y Cajal' Programme from the Spanish Ministry of Science and Technology for financial support. We would like to thank the farmers of the chestnut regions for their kind collaboration of this collection.

References

Adua M (1999) The sweet chestnut throughout history from the Miocene to the third millennium. In: Salesses G (ed) Proc 2nd Int Symposium on Chestnut. ISHS, NL, pp 29–36

Berrocal M, Gallardo JF, Cardeñoso JM (1998) El castaño. MundiPrensa, S.A.

Collada C, Casado R, Barber D, Fernandez de Caleya R, Aragoncillo C (1986) Characterization of seed protein fractions from *Castanea* spp. J Exp Bot 37:1872–1878

Collada C, Caballero RG, Casado R, Aragoncillo C (1991) Seed storage proteins in the Fagaceae: similarity between *Castanea* globulins and *Quercus* glutelins. Plant Sci 75:145–154

Dane F, Hawkins LK, Huang H (1999) Genetic variation and population structure of *Castanea pumila* var. *ozarkensis*. J Am Soc Hortic Sci 124:666–670

Fonseca PA, Ferreira RB, Teixeira AR (1997) Seed proteins from *Quercus suber*. J Agric Food Chem 45:3443–3447

Gepts P (1990) Genetic diversity of seed storage proteins in plants. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding and genetic resources. Sinauer Associates Publishers, Sunderland, Massachusetts, pp 64–82

Goldsbrough G (1990) A beginner's guide to Chestnut growing. Hilton Press

Huang H, Dane F, Norton JD (1994) Allozyme diversity in Chinese, Seguin and American chestnut (*Castanea* sp.). Theor Appl Genet 88:981–985

Huang H, Dane F, Kubisiak TL (1998) Allozyme and RAPD analysis of the genetic diversity and geographic variation in wild populations of the American chestnut *Castanea dentata* (Fagaceae). Am J Bot 85:1013–1021

Jaynes RA (1963) Biparental determination of nut characters in *Castanea*. J Hered 54:84–88

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Marshall DR, Brown AHD (1975) Optimum sampling strategies in genetic conservation. In: Frankel OH, Hawkes JG (eds) Crop genetic resources for today and tomorrow. Cambridge University Press, Cambridge, pp 53–70

Nei M (1972) Genetic distance between populations. Am Natl 106:283–292

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

Rohlf FJ, Fisher DL (1986) Test for hierarchical structure in random data sets. Systematic Zool 17:407–412

Shewry PR, Miflin BJ (1985) Seed storage proteins of economically important cereals. Adv Cereal Sci Tech 7:1–83

Shewry PR, Miles MJ, Tatham AS (1994) The prolamin storage proteins of wheat and related cereals. Prog Biophys Mol Biol 61:37–59

Sneath PHA, Sokal RR (1973) Numerical taxonomy. Freeman, San Francisco

Villani F, Pigliucci M, Benedettelli S, Cherubini M (1991) Genetic differentiation among Turkish chestnut (*Castanea sativa* Mill.) populations. Heredity 66:131–136