

Molecular Characterization of Persian Walnut Populations in Iran with Microsatellite Markers

Rohollah Karimi and Ahmad Ershadi¹

Department of Horticulture, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran

Kourosh Vahdati

Department of Horticulture, College of Abouraihan, University of Tehran, Tehran, Iran

Keith Woeste

USDA Forest Service, Hardwood Tree Improvement and Regeneration Center, Department of Forestry and Natural Resources, Pfendler Hall, Purdue University, 715 West State Street, West Lafayette, IN 47907-2061

Additional index words. *Juglans regia*, simple sequence repeat (SSR), genetic diversity

Abstract. The genetic structure and diversity of natural populations of *Juglans regia* L. in Iran were characterized using 11 microsatellite loci. A total of 105 individuals from seven populations were sampled. A high level of genetic diversity was observed within populations with the number of alleles per locus (A) ranging from three to 11 (average = 5.73), the proportion of polymorphic loci was 100%, and the expected heterozygosity ranged from 0.598 to 0.848 (average = 0.707). The proportion of genetic differentiation present among populations accounted for 12% of the total variation. Such considerable interpopulation differentiation detected in *J. regia* L. could have resulted from several factors, including restricted gene flow between populations. Significant departures from Hardy-Weinberg equilibrium were observed for WGA276, WGA32, and WGA321 loci. The deviations were primarily the result of the surplus of heterozygotes. The unweighted pair group method with arithmetic mean cluster analyses based on Nei's unbiased genetic distances separated the seven populations into two main groups.

The family Juglandaceae consists of seven genera, comprising ≈60 monoecious tree species. The genus *Juglans* contains 20 species, all producing edible nuts. Among these, the English or Persian walnut (*Juglans regia*) is the most widely cultivated species (McGranahan and Leslie, 1990). *Juglans* evolved in eastern Europe to Himalayas and parts of China. Walnut stands are found in Kopet Dag, the low land along the southern shore of the Caspian Sea, and in sporadic locations through the ranges of western and southern Iran (Leslie and McGranahan, 1988). Walnut was probably domesticated in Iran and Afghanistan and subsequently introduced to other countries (Bayazit et al., 2007). However, most of the walnut populations in Iran may not be native because trees from open-pollinated seedlings have been used by local people for many years. Walnut is planted in Iran for its fruit, wood, and pharmaceutical properties as well as for forest renovation. According to the FAO (2007), Persian walnut grows on 70,000 ha in Iran, producing 170,000 tons of nuts in shell, ranking third globally.

It was our goal to analyze the current populations of walnut in Iran to detect, select from, and maintain their genetic resources and to improve their genetic potential. Several techniques have been used to examine genetic diversity and relationships among cultivars of Persian walnut, including allozymes (Fornari et al., 2001; Malvolti et al., 1993), restriction fragment length polymorphisms (Fjellstrom and Parfitt, 1995), random amplified polymorphic DNA (RAPD) markers (Nicese et al., 1998), and intersimple sequence repeat (ISSR) markers (Potter et al., 2002). To investigate population structure as well as to facilitate various programs in genetic improvement, forest restoration, conservation, and sustainable management of *J. regia*, suitable molecular markers are needed to assess genetic variability. Because they are hypervariable, codominant, and highly informative, simple sequence repeat (SSR) or microsatellites are ideal for studies in population and conservation genetics (Streiff et al., 1998), the construction of genetic maps (Brondani et al., 1998), analysis of paternity and kinship, and analysis of variation between species and populations (Queller et al., 1993). In the present study, we investigated the level and distribution of genetic diversity in seven Iranian populations of *J. regia* L. using 11 SSR primer pairs originally developed by Woeste et al. (2002) for *J. nigra* and later applied to

breeding populations of *J. regia* (Dangle et al., 2005). We focused our study on Hamedan, Lorestan, Kermanshah, and Kurdistan. These provinces are in western Iran that have the largest populations of walnut in Iran and represent a heterogeneous and interesting genetic resource for breeders. This work will provide insight into the genetic diversity and population genetic structure of *J. regia* and provide valuable information for further conservation and breeding programs for *J. regia* in Iran, a center of diversity for this important crop.

Materials and Methods

Plant materials. Samples from seven populations of *J. regia* L. collected in western Iran were included in the study: four from different parts of Hamedan province including Malayer (MR), Hamedan (HN), Tuyserkhan (TN), and Serkan (SN) and one each from Lorestan (LN), Kermanshah (KH), and Kurdistan (KN) provinces (Fig. 1). Populations consisted of old walnut trees from open-pollinated seedlings. The sampled populations inhabit disjunctive mountainous areas with a narrow geographic range of latitude and longitude (32 to 35° N and 46 to 50° E, respectively). The distribution of the sampled populations was from 1385- to 1986-m altitude. Plants within 10 km of each other were considered as belonging to the same deme. To keep relatively uniform sample sizes, we limited our sampling to 15 mother trees per population, each tree separated by a distance of 100 m. Leaves were sampled in early morning hours and stored at -80 °C until their use.

DNA extraction. For each mother tree, at least six young leaves were collected in summer and used for DNA extractions. A 150- to 200-mg sample was ground in 2-mL Eppendorf tubes with 1800 mL of extraction buffer (2% CTAB, 100 mM Tris, 1.4 M NaCl, and 20 mM EDTA, pH 8.3). Then DNA was extracted following the CTAB method (Doyle and Doyle, 1987).

Polymerase chain reaction amplification and electrophoresis. Eleven primers designed from the sequences of clones isolated from an enriched (GA/CT)_n library of *J. nigra* (Woeste et al., 2002) were used to amplify genomic DNA of *J. regia* genotypes (Table 1). The SSR reaction protocol was based on the protocol of Dangle et al. (2005). Amplification reactions were performed in a volume of 15 µL containing 1 × polymerase chain reaction (PCR) buffer, 25 ng genomic DNA, 200 µM dNTPs, 0.2 µM of each primer, 2 mM MgCl₂, and 0.5 unit of Taq DNA polymerase (Sinagen, Tehran, Iran). For DNA amplifications, an Eppendorf thermocycler was programmed according to the following profile: 1) 94 °C for 5 min, one cycle; 2) 30 cycles at 94 °C for 1 min, the appropriate primer annealing temperature (Table 2) for 1 min and 72 °C for 40 s; and 3) a final elongation at 72 °C for 2 min, one cycle. After amplification, 5 µL of each sample was loaded and electrophoresed on a 2% horizontal agarose gel to control for positive amplification and to determine the approximate amount of product. Then, 3 µL of

Received for publication 10 Feb. 2010. Accepted for publication 6 July 2010.
We acknowledge Dr. Sarikhani and Dr. Esna-Ashari for their kindly prereview of the manuscript.
¹To whom reprint requests should be addressed; e-mail Ershadi@basu.ac.ir.



Fig. 1. Sampled populations were located in provinces Hamedan, Kurdistan, Kermanshah, and Lorestan, leading areas for walnut genetic diversity and production in Iran.

Table 1. Properties of the microsatellite loci used to characterize seven walnut populations in Iran.^z

Locus name	Na ^y	Ne ^x	Locus name	Na	Ne
WGA1	4	2.88	WGA321	5	3.40
WGA9	6	4.85	WGA332	5	2.45
WGA89	4	2.58	WGA349	3	2.74
WGA202	5	2.50	WGA32	8	5.25
WGA225	4	3.78	WGA27	8	3.59
WGA276	11	6.21			

^zPrimers developed by Woeste et al. (2002).

^yNa = number of alleles.

^xNe = number of effective alleles.

Table 2. Genetic variation within *J. regia* L. populations based on 11 microsatellite loci.^z

Population	N	A	Ae	Ho	He
MR	15	3.73	2.58	0.714	0.645
HN	15	4.27	3.05	0.702	0.706
TN	15	3.82	2.52	0.662	0.614
SN	15	4.18	2.97	0.701	0.673
KH	15	3.64	2.63	0.675	0.650
KN	15	3.73	2.75	0.612	0.660
LN	15	4.36	3.02	0.697	0.709
Mean	15	3.96	2.70	0.682	0.666

^zN = sample size; A = mean number of alleles; Ae = effective number of alleles per locus; Ho = observed heterozygosity; He = expected heterozygosity.

each sample was electrophoresed on a 6% polyacrylamide gel containing 1 × TBE buffer. After electrophoresis, the gel was silver-stained using the procedure of Bassam et al. (1991). In all cases, PCR reactions were performed at least twice to ensure that allele sizes were consistent.

Allele sizing and scoring was done using a 100 bp DNA ladder (MBI-Fermentas, Vilnius, Lithuania) as the length reference.

Data analysis. The data were entered in the form of single-individual genotypes. The following parameters of genetic variation were assessed for each population: the mean number of alleles per locus (A), the effective number of alleles (Ae), percentage of polymorphic loci (P), expected heterozygosity (He) (Nei, 1978), and observed heterozygosity (Ho). Departures from the Hardy-Weinberg (H-W) equilibrium were assessed at each locus for every population and per locus across all populations using *F*-statistics of Wright (1978). The significance of the deviations was evaluated with a chi-square test following the method of Workman and Niswander (1970). Spatial genetic structure was further investigated using Wright's analysis of hierarchical *F*-statistics (Wright, 1978). Nei's (1978) unbiased genetic distances were calculated for all population pairs and used to construct a phylogenetic tree [unweighted pair group method with arithmetic mean (UPGMA)]. All of these calculations were performed using POPGENE Version 1.32 software (Yeh et al., 1997).

Results

Allelic variation at microsatellite loci. All 11 loci assayed possessed a high level of polymorphism with the number of alleles per locus ranging from three at WGA349 to 11 at

WGA276 (Table 2). All 11 microsatellite loci were polymorphic across all seven populations sampled. A total of 63 alleles were detected; the mean number of alleles per locus was 5.73 and they ranged in size from 120 to 282 bp. No private alleles (alleles present in a single population only) were detected. Instead, 25 alleles were common to the majority of the populations (at least five populations), whereas the other alleles were shared among fewer populations. A summary of the loci and alleles for each population is provided in Table 2.

Genetic variation within populations. Genetic diversity parameters based on allelic frequencies are shown in Table 2. In individual populations, the mean number of alleles per locus (A) varied from 3.64 to 4.36 with an average of 3.96, whereas the effective number of alleles per locus (Ae) varied from 2.52 to 3.05 with an average of 2.70. The observed heterozygosities (Ho) ranged from 0.596 to 0.745 with an average of 0.683. The average expected heterozygosity (He) was 0.707 and ranged from 0.598 to 0.848.

A comparison of the genetic diversity of *J. regia* was performed among the seven populations. The highest level of diversity existed in population LN (He = 0.709), whereas the lowest values were in population TN (He = 0.614). The within-population genetic diversity generally decreased from south to north.

Population genetic structure. The genetic analyses revealed high levels of differentiation among populations (Table 2). The coefficient of hierarchical *F*_{ST} (Table 3), estimated according to Wright (1978), ranged from 0.017 for locus WGA27 to 0.202 for locus WGA276 with the average value equaling 0.122. This showed that 12% of the total genetic diversity existed among populations. The loci WGA276 and WGA321 possessed the highest *F*_{ST} value. These values indicated the important role of these two loci in inter-population differentiation. In individual populations, tests for the departure from the H-W equilibrium showed significant deviations for at least one locus in every population. Most of the loci followed the H-W equilibrium in the majority of populations. However, significant departures from the H-W equilibrium were observed in WGA276, WGA32, and WGA321 loci. The deviations were primarily the result of the surplus of heterozygotes.

Genotypic structure and departure from Hardy-Weinberg equilibrium. *F*_{IT} is the overall inbreeding coefficient of an individual relative to the whole set of populations, whereas *F*_{IS} is the inbreeding coefficient of an individual relative to its own population. The mean *F*_{IT} value of 0.026 indicates that, overall, the sampled populations of *J. regia* had a slight deficiency of heterozygotes. Within populations, all loci but two showed a surplus of heterozygotes (*F*_{IS} ranged from -0.026 to -0.310). The two loci that showed a deficiency of heterozygotes were WGA32 and WGA 27 (*F*_{IS} = 0.025 and 0.040, respectively) with the average *F*_{IS} value equaling -0.107, which was significantly different from zero (Table 3). This means that there was a small heterozygote excess among the sampled trees. However, it

is possible to get an excess of heterozygotes if stutter bands are erroneously scored as alleles.

Genetic relationships. Genetic distances were calculated for each pair of populations to estimate the extent of their divergence (Table 4). The average genetic distance among populations equaled 0.297. The lowest genetic distance (0.100) was found between populations TN and SN and the greatest genetic distance (0.582) was found between populations TN and MR. The UPGMA cluster analyses based on Nei's unbiased genetic distances was performed to further show the

genetic relationships among the populations (Fig. 2). The dendrogram separated the seven populations into two main groups. The first group consisted of the populations MR, LN, and KN. The second group consisted of the populations HN, TN, SN, and KH. In the first group, MR and LN were the most similar; TN and SN, in the second group, were the most closely related populations in the study.

Discussion

In the present study, seven populations of *J. regia* were characterized using 11 highly

polymorphic microsatellite markers. The average number of alleles per locus was 5.73, much higher than 1.3 and 3.9 detected in *J. regia* with RAPDs (Nicese et al., 1998) and ISSRs (Potter et al., 2002), respectively. For all loci, the observed numbers of alleles was lower than that reported in *J. nigra* (Victory et al., 2006), the species from which the markers were originally developed (Woeste et al., 2002). This may be because of the large number of samples and populations analyzed in previous *J. nigra* studies or because SSRs may show a decrease in allele number when applied in related species (Ana et al., 2000). For most of the loci, the length ranges of the amplification products obtained here were wider than those detected in *J. nigra* (Victory et al., 2006), a bias reported in other species as well (Donini et al., 1998).

Although a great number of population genetic studies have been conducted on *Juglans* species (Fornari et al., 2001; Victory et al., 2006), no genetic research based on the use of molecular markers has been published previously on *J. regia* populations in Iran. In this investigation, high levels of polymorphism were detected, the total number of alleles equaling 63. The observed heterozygosity within populations (mean $H_e = 0.68$) was higher than that detected in *J. regia* based on allozymes (mean $H_e = 0.43$) (Fornari et al., 1999), but lower than the heterozygosity of 0.8 reported in wild populations of *Juglans nigra* (Victory et al., 2006), a species that is also highly outcrossed. Within populations, all loci but two showed a surplus of heterozygotes. The negative values of F_{IS} , the intrapopulation genetic diversity, also suggests an excess of heterozygosity present in the populations of walnut, which could be the result of walnut's mating system (Fornari et al., 2001) and could also be caused by selection or disassortative mating. Despite the relatively short geographic distances between populations included in the study, F_{ST} , the interpopulation genetic diversity was considerable ($F_{ST} = 0.122$) and it was clearly higher than the values detected in other *Juglans* species based primarily on isozymes (Fornari et al., 2001). The F_{ST} value detected here in *J. regia* is higher than the differentiation among populations observed in *J. nigra* ($F_{ST} = 0.058$) based on microsatellite markers (Victory et al., 2006) and is much closer to the value $F_{ST} = 0.106$ observed by Fornari et al. (2001) using allozyme on natural populations of *J. regia*. This means that there are some processes that are serving to isolate the populations and that they may have unique genetic qualities that can be used for breeding. It could indicate that there are selection pressures operating in each population that are different from the other populations. This also means that it will be worthwhile to include members of each population in germplasm collections.

The seven natural populations of *J. regia* studied here occupy fragmented habitats, the populations being confined to islands varying in size and degree of isolation within the northwestern Kurdistan province and

Table 3. Relative measurements of genetic differentiation among populations of *J. regia*.^z

Locus	F_{IS}	F_{IT}	F_{ST}	H_o	H_e	PIC
WGA276	-0.077	0.140	0.202	0.723	0.848	0.852
WGA009	-0.026	0.143	0.165	0.681	0.803	0.794
WGA32	0.025	0.170	0.149	0.681	0.818	0.780
WGA089	-0.112	0.029	0.132	0.596	0.617	0.601
WGA321	-0.310	-0.051	0.198	0.745	0.714	0.687
WGA225	-0.093	0.099	0.176	0.660	0.743	0.723
WGA001	-0.031	-0.010	0.021	0.660	0.660	0.623
WGA332	-0.257	-0.156	0.080	0.680	0.598	0.640
WGA349	-0.161	-0.106	0.048	0.702	0.642	0.661
WGA27	0.040	0.056	0.017	0.681	0.729	0.495
WGA202	-0.267	-0.166	0.080	0.702	0.607	0.684
Mean	-0.107	0.026	0.122	0.683	0.707	0.685

^z F_{IS} = fixation index within populations; F_{IT} = fixation index with respect to total populations; F_{ST} = fixation index among populations; PIC = polymorphism information content.

Table 4. Genetic distances among populations of Persian walnut based on Nei's (1978) similarity coefficient and unweighted pair group method with arithmetic mean method.

Population	MR	HN	TN	SN	KH	LN	KN
MR	***						
HN	0.325	***					
TN	0.582	0.185	***				
SN	0.457	0.179	0.100	***			
KH	0.337	0.264	0.350	0.167	***		
LN	0.180	0.273	0.409	0.289	0.142	***	
KN	0.251	0.380	0.437	0.362	0.347	0.226	***

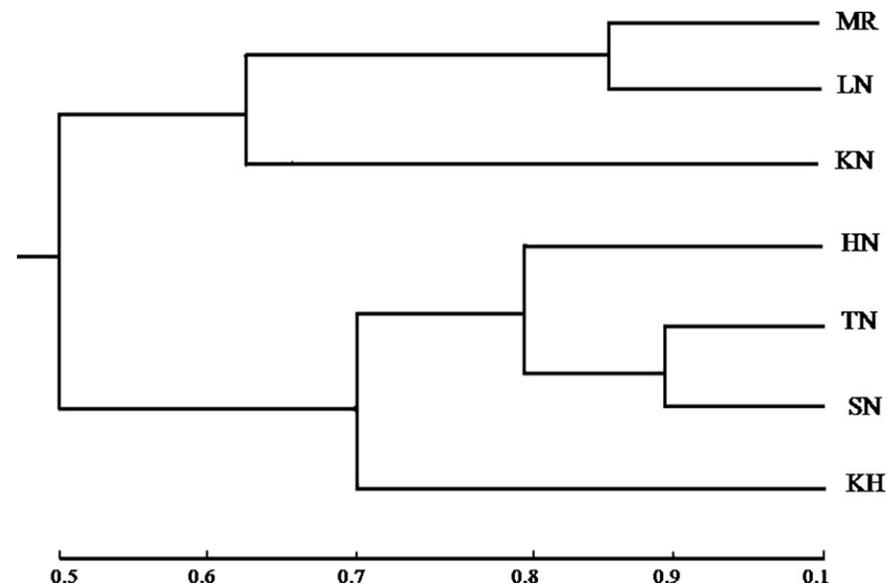


Fig. 2. Dendrogram of genetic diversity among Persian walnut populations based on Nei's (1978) similarity coefficient and unweighted pair group method with arithmetic mean method.

southeastern Lorestan province. The mean effective number of allele per locus was 2.70 (range, 2.52 to 3.05). The highest level of diversity existed in population LN ($H_e = 0.709$) because of high plant density, isolation, and lack of human selection, whereas the lowest value was in population TN ($H_e = 0.614$). The relatively low diversity level in population TN is attributable to the human propagation manner, leading to the selection of trees with the most suitable traits, negatively affecting genetic variability (Fornari et al., 1999). Populations TN and SN are closely located geographically in Hamedan province and are closely related ($D = 0.100$). A decrease in genetic diversity from southern to northern populations was observed. The southern populations (e.g., LN) contained higher levels of expected heterozygosity than did the northern ones. Busov et al. (2002) reported a similar finding when studying the allozyme variation of black walnut (*Juglans nigra*) in the Central Hardwood Region of the United States. However, Lorestan has the highest annual average rainfall among all studied areas and better growing conditions for walnut.

The genetic parameters estimated from the microsatellite data indicated that there are substantial levels of genetic diversity in all sampled populations. This result is in accordance with the common observation of high levels of variation detected in wind-pollinated, long-lived tree species (Streiff et al., 1998; Victory et al., 2006). In summary, we observed: 1) considerable genetic diversity in populations of *J. regia* from western Iran, and the diversity was both within and among populations; 2) the relatively high interpopulation differentiation of *J. regia* could have resulted from several factors, including restricted gene flow between populations and human disturbances; 3) the capability of *J. regia* to maintain a noticeable intrapopulation differentiation, even in the presence of selection pressures, is attributable to its reproductive strategies based on outcrossing; and 4) microsatellite markers are powerful tools for monitoring the genetic structure and diversity of *J. regia* populations. Future research in *J. regia* based on more extensive sampling and more intensive sampling with

microsatellite markers will allow us to describe in greater detail the population genetic structures of Persian walnut in Iran.

Literature Cited

- Ana, C.R., C.A. Paul, C.D. Myriam, M.M. Maria, W.B. Merideth, I. Carlos, and T. Joe. 2000. Cross-species amplification of cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellites: Allelic polymorphism and degree of relationship. *Amer. J. Bot.* 87:1647–1655.
- Bassam, B.J., G. Caetano-Anolles, and P.T. Gresshoff. 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal. Biochem.* 196:80–83.
- Bayazit, S., K. Kazan, S. Golbitti, V. Cevik, H. Ayanoglu, and A. Ergul. 2007. AFLP analysis of genetic diversity in low chill requirement walnut (*Juglans regia* L.) genotypes from Hatay, Turkey. *Sci. Hort.* 111:394–398.
- Brondani, R.P.V., C. Brondani, R. Tarchini, and D. Grattapaglia. 1998. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. *Theor. Appl. Genet.* 97:816–827.
- Busov, V.B., G. Rink, and K.E. Woeste. 2002. Allozyme variation and mating system of black walnut (*Juglans nigra* L.) in the central hardwood region of the United States. *For. Genet.* 9:319–327.
- Dangle, G.S., K.E. Woeste, M.K. Aradhya, A. Koehmstedt, C. Simon, D. Potter, C.A. Leslie, and G.H. McGranahan. 2005. Characterization of fourteen microsatellite markers for genetic analysis and cultivar identification of walnut. *J. Amer. Soc. Hort. Sci.* 130:348–354.
- Donini, P., P. Stephenson, G.J. Bryan, and R.M.D. Koebner. 1998. The potential of microsatellites for high throughput genetic diversity assessment in wheat and barley. *Genet. Resources Crop Evol.* 145:415–421.
- Doyle, J.J. and J.L. Doyle. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochem. Bull.* 19:11–15.
- FAO. 2007. FAOSTAT, Production. 16 Sept. 2007. <<http://faostat.fao.org/>>.
- Fjellstrom, R.G. and D.E. Parfitt. 1995. Phylogenetic analysis and evolution of the genus *Juglans* (Juglandaceae) as determined from nuclear genome RFLPs. *Plant Syst. Evol.* 197:19–32.
- Fornari, B., F. Canata, M. Spada, and M.E. Malvolti. 1999. Allozyme analysis of genetic diversity differentiation in European and Asiatic walnut (*Juglans regia* L.) populations. *For. Genet.* 6: 115–127.
- Fornari, B., M.E. Malvolti, D. Turchini, S. Fineschi, I. Beritognolo, E. McCaglia, and F. Cannata. 2001. Isozyme and organellar DNA analysis of genetic diversity in natural/ naturalized European and Asiatic walnut (*Juglans regia*) populations. *Acta Hort.* 544:167–178.
- Leslie, C.H. and G. McGranahan. 1988. Native populations of *Juglans regia* L. a draft. Intl. Conference on Walnuts; 19–23 Sept.; Yalova, Turkey.
- Malvolti, M.E., M. Paciucci, F. Cannata, and S. Fineschi. 1993. Genetic variation in Italian populations of *Juglans regia* L. *Acta Hort.* 311: 86–94.
- McGranahan, G.H. and C. Leslie. 1990. Walnuts (*Juglans*). *Acta Hort.* 290:905–951.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583–590.
- Nicese, F.P., J.I. Hormaza, and G.H. McGranahan. 1998. Molecular characterization and genetic relatedness among walnut (*Juglans regia* L.) genotypes based on RAPD markers. *Euphytica* 101:199–206.
- Potter, D., F. Gao, G. Aiello, C. Leslie, and G.H. McGranahan. 2002. Inter-simple sequence repeats markers for fingerprinting and determining genetic relationships of walnut (*Juglans regia* L.) cultivars. *J. Amer. Soc. Hort. Sci.* 127: 75–81.
- Queller, D.C., J.E. Stressmann, and C.R. Hughes. 1993. Microsatellites and kinship. *Trees (Berl.)* 8:285–288.
- Streiff, R., T. Labbe, and R. Bacilieri. 1998. Within population genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. *Mol. Ecol.* 7:317–328.
- Victory, E.R., J.C. Glaubitz, O.E. Rhodes, and K.E. Woeste. 2006. Genetic homogeneity in *Juglans nigra* (Juglandaceae) at nuclear microsatellites. *Amer. J. Bot.* 93:118–126.
- Woeste, K., R. Burns, O. Rhodes, and C. Michler. 2002. Thirty polymorphic nuclear microsatellite loci from black walnut. *J. Hered.* 93:58–60.
- Workman, P.L. and J.D. Niswander. 1970. Population studies on southwestern Indian tribes. II. Local genetic differentiation in the Papago. *Am. J. Hum. Genet.* 22:24–49.
- Wright, S. 1978. Evolution and genetics of populations. Vol. 4. University of Chicago Press, Chicago, IL.
- Yeh, F.C., R.C. Yang, and T. Boyle. 1997. POP-GENE. Version 1.32 ed. software. Microsoft Window-based freeware for population genetic analysis. University of Alberta, Edmonton, Canada.