Scientific Papers. Series B, Horticulture. Vol. LXX, No. 1, 2026
Print ISSN 2285-5653, CD-ROM ISSN 2285-5661, Online ISSN 2286-1580, ISSN-L 2285-5653

GENETIC DIVERSITY ANALYSIS OF *CORYLUS AVELLANA* L. GENOTYPES USED FOR NEW VARIETY DEVELOPMENT

Monica HÂRŢA[1], Doina CLAPA[1], Claudiu Ioan BUNEA[1], Lehel LUKACS[2], Gabriela Maria ROMAN[2], Eugenia HÂRŞAN[2], Daniel Vasile AVRAM[3]

[1]Faculty of Horticulture and Business in Rural Development, University of Agricultural Sciences and Veterinary Medicine of Cluj-Napoca, 3-5 Calea Mănăştur, 400372, Cluj-Napoca, Romania
[2]Cluj-Napoca Horticultural Research Station, University of Agricultural Sciences and Veterinary Medicine of Cluj-Napoca, 3-5 Calea Mănăştur, 400372, Cluj-Napoca, Romania
[3]Agricultural Cooperative 'Agriculture of Transylvania' SA, 3A Dealului Street, 400243, Cluj-Napoca, Romania

Corresponding author email: doina.clapa@usamvcluj.ro

Abstract

A study compared nine hazelnut genotypes from Transylvania's spontaneous flora with five Corylus avellana cultivars, using 10 microsatellite markers to assess genetic diversity and identify valuable genotypes for developing new varieties. Among the primer sets tested, cavSSR14418 was the most informative, with Na = 5, Ne = 3.88, I = 1.56, PIC = 0.71, and He = 0.74. The observed heterozygosity (Ho) ranged from 0.64 (cavSSR 2135; cavSSR 3909; cavSSR 4217; cavSSR 14875) to 1.0 (cavSSR 2704), with an overall mean of 0.72. The NJ dendrogram grouped the genotypes into two main clusters. The clear separation between clusters grouping the established cultivars and those including the new selections demonstrates the latter's potential to enrich the current genetic pool and confirms that wild genotypes exhibit high genetic variability, making them valuable biological resources for biodiversity conservation and future breeding programs.

Key words*: hazelnuts, wild accessions, cultivars, morphological characteristics, SSR markers.*

INTRODUCTION

Corylus avellana L., The European hazelnut belongs to the Betulaceae family and is a diploid (2n = 22), monoecious, wind-pollinated plant with sporophytic incompatibility that requires cross-pollination (Boccacci et al., 2013; Martins et al., 2015). It can be cultivated in commercially intensive orchards as a shrub or as a single-trunk tree because of its importance in nut production (Brown et al., 2016; Oztolan-Erol et al., 2021).

In Romania, investments in fruit growing have increased the number of hazelnut orchards, with many using organic methods (Stănică et al., 2023). Hazelnut trees also occur in spontaneous or semi-spontaneous stands at the edges of deciduous forests, along watercourses and roads, and in almost all hilly and mountainous areas, reaching altitudes of 900-1200 m (Botu et al., 2018). As a result, many Romanian hazelnut accessions from *in situ* collections maintained by research and development stations have wild ancestry, and natural selection has strongly influenced their genetic structure (Pop et al., 2010).

For example, morphological characterization can reveal wide phenotypic variation in fruit traits and in the environmental adaptation of selected individuals, which can be exploited in breeding programs. However, there is limited information on the relationship between cultivated and wild hazelnut forms in Romania, as assessed using SSR (Simple Sequence Repeats) molecular markers. In this context, exploiting the genetic diversity of wild hazelnut individuals can be crucial for developing new varieties that are more resilient to climate change.

Examining the relationship between wild and cultivated forms within a geographic area using SSR markers can provide valuable insights into possible domestication events, evolutionary links, or gene flow between them (Campa et al. 2011).

Previous research has highlighted the importance of microsatellite (SSR) markers in hazelnuts. Several studies, including those conducted by Bassil et al. (2004), Gökirmak et al. (2009), Gürcan et al. (2010), Leinemann et al. (2013), Tanhuanpää et al. (2019), Kavas et al. (2020), and Farrokhi and Ershadi (2022), have

underscored the value of these DNA-based molecular markers in research related to *C. avellana*.

From this perspective, the present study aims to evaluate the genetic relationships between wild hazelnuts and cultivated varieties to identify locally valuable germplasm for developing new hazelnut varieties.

MATERIALS AND METHODS

The plant material used in this study for SSR analysis consisted of young leaves provided by a private commercial orchard for the varieties: 'Tonda di Giffoni', 'Tonda Gentile delle Langhe', and 'Tonda Gentile Romana'. The leaves for 'Uriaşe de Vâlcea' and 'Cozia' used for DNA isolation were provided by SCDP Vâlcea. Additionally, leaves from nine wild hazelnut genotypes were collected from the experimental field of the Cluj-Napoca Horticultural Research Station of UASVMCN (see Table 1).

Table 1. The codification and provenance of local wild hazelnut genotypes used for SSR analysis

No.	Sample Code	Provenance
1	al_2	Făget forest, Cluj county
2	al_3	Plăieşti forest, Cluj county
3	al_4	Gârbova forest, Alba county
4	al_5	Făget forest, Cluj county
5	al_6	Făget forest, Cluj county
6	al_9	Gheorgheni forest, Cluj county
7	al_10	Făget forest, Cluj county
8	a1_14	Făget forest, Cluj county
9	al_20	Făget forest, Cluj county

Genomic DNA was extracted using a DNeasy Plant Pro kit (Qiagen, Germany) according to the manufacturer's protocol. The concentration and purity of the extracted DNA were measured using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, USA).

A set of 10 SSR primers was used in this study. The characteristics of the primers are presented in Table 2.

The PCR amplification profile was set on a thermocycler (SuperCycler Trinity, Kyratec, Australia), as described by Öztürk et al. (2018). Thermal cycling conditions included an initial denaturation step of 10 minutes at 94°C, followed by 30 cycles each consisting of 30 seconds at 94°C, 30 seconds at 55°C, and 30 seconds at 72°C, ending with a final extension of 10 minutes at 72°C.

Table 2. SSR primer names and sequences

Primer name	Forward primer sequences (5'-3')	Reverse primer sequences (5'-3')	SSR motifs
cavSSR 1361	GATATCACTCACG TCTACCG	GGTCTCTTGGTC TTGATGTT	(AG/TC) 14
cavSSR 2135	ATGTAGCGAGCCT TGATAAG	GTTGTCAGGTA GCTTGAAGA	(TTAA/AATT) 6
cavSSR 2704	GCGGAGTTGGTA GTGATAAT	ATATAGGTATA AGGGGGCCC	(TAG/ATC) 10
cavSSR 3909	AGATGAAGCTGA AGAAAGGG	TATCGCCATCA CACCATTAG	(GGA/CCT) 8
cavSSR 4217	GACAGTTGGCAT GAAAGATG	GCACTCATCAG AGAGTCAAA	(ATT/TAA) 10
cavSSR 12862	TAAAATGGGCCT ACACTTCC	CCAGTACAGGA AGATACGAA	(CT/GA) 13
cavSSR 14219	TATATGGACAGCT GACTCCA	GAGGGAGTTTG TCTGTCTTT	(AAT/TTA) 9
cavSSR 14418	GACTGCAAGAAT GACAACAG	GTCCTCCTCCTT TTTCGTAG	(TTGG/AACC) 6
cavSSR 14875	CACAAGATGATA CCCATGCT	TATCAGCTCCT AAAACGACG	(TACA/ATGT) 6
cavSSR 14937	TGAGCTCTCTGGT TTCTTTC	ACTGGATCTGC TTTTATGGG	(CT/GA) 12

In this study, we adjusted the master mix reaction volume to 12 µL per sample, and its composition: 5x GoTaq Flexi Green buffer, nuclease-free water, 1.5 mM $MgCl_2$, 0.2 mM dNTP mix (Promega, USA), 1 µM Cy5-labeled forward SSR primer, 1 µM reverse SSR primer (GeneriBiotech, Czech Republic), and 0.8 U GoTaq polymerase (Promega, USA). A volume of 3 µL of each DNA sample (20 ng/µL) was used for each PCR reaction. Amplification products were verified by running them on 1.6% agarose gels (Promega, USA) stained with RedSafeTM nucleic acid stain (iNtRON Biotech, South Korea) in 1x TBE buffer at 0.29 V/cm for 2.5 hours. The PCR fragments were visualized using the UVP Biospectrum AC Imaging System (UVP BioImaging Systems, USA). Prior to migration and separation within the CEQ 8800TM capillary DNA analysis system (Beckman Coulter, USA), each PCR sample was mixed with 30 µL of sample-loading solution and 0.25 µL of the Genome DNA Standard Kit-400 (Beckman Coulter, USA). Then, allele sizes for each SSR locus were determined using the internal CEQ fragment analysis software.

Genetic diversity among hazelnut genotypes was assessed using GenAlEx version 6.503 (Peakall & Smouse, 2012). The analysis included parameters such as the number of alleles per locus (Na), effective number of alleles (Ne), observed heterozygosity (Ho), expected heterozygosity (He), unbiased expected heterozygosity (uHe) and Shannon

information index (I). Allele frequencies for each SSR locus were calculated based on allele size (bp). The polymorphism information content (PIC) and the null allele frequency (F null) were estimated with Cervus 3.0.7 (Kalinowski et al., 2007). Multivariate analysis was also performed. Clustering was carried out using MEGA7 (Kumar et al., 2016) which created a pairwise distance-based tree via the Neighbor-joining (N-J) method in Tree Explorer, with scaling based on codominant genotypic distances computed in GenAlEx 6.503.

RESULTS AND DISCUSSIONS

Assessment of genetic diversity among wild and cultivated hazelnut genotypes shows that the spontaneous flora of the Transylvanian region contains hazelnut genetic resources that can be used for selecting and developing new genotypes. The results of molecular descriptive statistics are presented in Table 3.

Table 3. Descriptive statistics and genetic diversity of the 14 hazelnut genotypes at 10 microsatellite loci

Locus	Na	Ne	Ho	He	uHe	I	PIC	F (null)
cavSSR 1361	5	2.72	0.71	0.63	0.66	1.21	0.58	-0.049
cavSSR 2135	3	2.65	0.64	0.62	0.65	1.03	0.55	-0.036
cavSSR 2704	5	3.47	1.00	0.71	0.74	1.39	0.67	-0.194
cavSSR 3909	3	2.21	0.64	0.55	0.57	0.93	0.49	-0.063
cavSSR 4217	3	2.65	0.64	0.62	0.65	1.03	0.55	-0.002
cavSSR 12862	5	3.47	0.71	0.71	0.74	1.39	0.67	0.007
cavSSR 14219	4	3.81	0.71	0.74	0.76	1.36	0.69	0.008
cavSSR 14418	6	3.88	0.79	0.74	0.77	1.56	0.71	-0.037
cavSSR 14875	3	2.80	0.64	0.64	0.67	1.06	0.57	-0.009
cavSSR 14937	5	3.84	0.71	0.74	0.77	1.47	0.70	0.011
Mean	4.2	3.15	0.72	0.67	0.69	1.24	0.61	-

All ten SSR primer pairs used in this study were polymorphic, generating 42 distinct alleles across 14 genotypes analysed (n = 14); the number of alleles per locus ranged from 3 to 6, with an average of 4.2 alleles per locus. This mean value indicates moderate allelic richness as compared with others previously reported studies (Bocacci et al., 2008; Fiore et al., 2022). It is important to note that comparisons with other studies are only approximate, since factors like the number of SSRs used and the sample sizes analysed can significantly affect the research outcomes. The effective number of alleles, accounting for alleles at relevant frequencies in the hazelnut samples, ranged from 2.21 at locus cavSSR 3909 to 3.88 at locus cavSSR 14418, with a mean of 3.15. This result aligns with those reported by Erfatpour et al. (2011) in a study of genetic diversity among some Iranian 90 hazelnut genotypes.

In the present study, the al_4 hazelnut genotype exhibited private alleles at three loci (cavSSR1361, cavSSR14418, cavSSR14937), similar to Tonda Gentile delle Langhe and Tonda Gentile Romana (Table 4).

Table 4. List of the genotypes with one or more private alleles

Genotype name	No. of loci	Loci with private alleles
al_2	1	cavSSR14418
al_3	1	cavSSR2704
al_4	3	cavSSR1361;cavSSR14418; cavSSR14937
al_5	2	cavSSR2704; cavSSR14937
al_6	2	cavSSR14418; cavSSR14937
al_9	1	cavSSR12862
al_10	1	cavSSR14937
a1_14	2	cavSSR14418; cavSSR14937
al_20	2	cavSSR12862; cavSSR14418
Tonda di Giffoni	3	cavSSR1361; cavSSR12862; cavSSR14418
Tonda Gentile delle Langhe	3	cavSSR1286; cavSSR14418; cavSSR14937
Tonda Gentile Romana	2	cavSSR14418; cavSSR14937
Uriaşe de Vâlcea	1	cavSSR14937

This suggests that the al_4 genotype, collected from wild flora, may possess valuable traits that can be exploited in hazelnut breeding and conservation strategies Table 4 lists the hazelnut genotypes with one or more private alleles identified by GenAlEx analysis.

In this study, SSR markers were employed to evaluate genetic variability among the analysed varieties, using observed and expected heterozygosity (Ho and He). The observed heterozygosity (Ho) ranged from 0.64 (cavSSR 2135; cavSSR 3909; cavSSR 4217 and cavSSR 14875) to 1.0 (cavSSR 2704), with an overall mean of 0.72. The expected heterozygosity (He) ranged from 0.55 (cavSSR 3909) to 0.74 (cavSSR 14219; cavSSR 14418 and cavSSR 14937), with an average of 0.67. The overall mean observed heterozygosity was higher than expected (0.72 versus 0.67) due to presumed admixture of individuals from distant populations and different origins, indicating that

the analysed hazelnut samples do not originate from a single isolated population but are a mix of genetically distinct groups (for example, due to recent migration or gene flow) (Allendorf et al., 2013). Furthermore, this excess of heterozygosity signifies a high level of outbreeding, where reproduction occurs between unrelated individuals from different source populations. This pattern aligns with recent genomic studies on hazelnut diversity, which show that human-mediated transport and the absence of strict reproductive barriers often lead to increased genetic variation and the preservation of high heterozygosity levels (Öztürk et al., 2021; Wrigley et al., 2020). The cavSSR 3909 primer pair produced the lowest PIC (0.49), He (0.55), and I (0.93) values. The highest values for PIC (0.71), He (0.74), and I (1.56) were recorded with the cavSSR 14418 primer pair, as shown in Table 3. The recorded values for the null allele frequencies (F null) were negative for almost all of the user SSR primer pairs, except cavSSR 12862, cavSSR 14219, and cavSSR 14937, where F (null) values were positive but less than 0.02. A null allele frequency below 0.02 (2%) in SSR (Simple Sequence Repeat) markers is generally considered very low and acceptable in most genetic studies. Although null alleles (alleles that fail to amplify in PCR due to mutations at primer-binding sites) are common (Jahnke et al., 2022), the low frequency observed in this study indicates that the SSR markers successfully amplified all DNA samples.

The NJ dendrogram grouped the analyzed genotypes into two main clusters, designated I and II. The first main cluster comprises two subclusters (A and B) that group all wild hazelnut genotypes, as well as one international reference genotype (Tonda di Giffoni). Subcluster B includes two genotypes: al_4 and al_5, both wild samples, which are separated from the other genotypes in the first main cluster, suggesting greater genetic diversity than the genotypes grouped in the A subcluster. The second main cluster (II) comprises two subclusters, each containing only the cultivars considered reference varieties, namely Tonda Gentile delle Langhe and Tonda Gentile Romana, and Cozia and Uriaşe de Vâlcea, respectively, as shown in Figure 1.

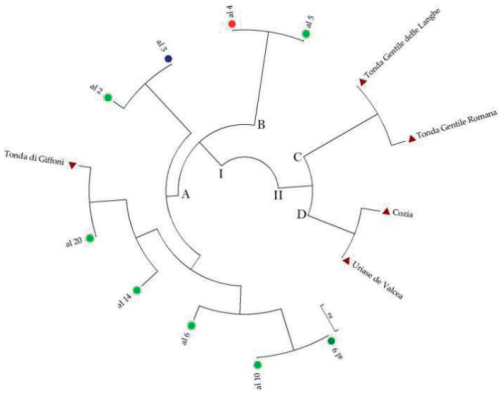

Figure1. NJ dendrogram grouping the wild hazelnut genotypes (marked with red, blue, and green circles) and reference varieties (marked with red triangles); I, II - main clusters; A, B, C, D - subclusters

The Neighbor-Joining (NJ) analysis revealed genetic diversity among the hazelnut genotypes studied, confirming genetic variability within the analysed germplasm that can be used in future breeding programs. The clear separation between clusters grouping established cultivars (C and D) and those including the new selections (A and B) demonstrates the latter's potential to enrich the current genetic pool. In particular, the genetic isolation of selections al_4 and al_5 from the first main genotype group (cluster I) highlights their possible divergent origins and their value as novel genetic resources.

CONCLUSIONS

Based on SSR data analysis, this study confirms that local genotypes exhibit high genetic variability, making them valuable biological resources for biodiversity conservation and for future breeding programs to develop new varieties that enhance genetic resilience.

ACKNOWLEDGEMENTS

This research work was carried out with the support of OG and also was financed from Project C161A0000011861300005/14.05.2021, financed with non-reimbursable European funds through NRDP 2014-2020.

REFERENCES

Allendorf, F. W., Luikart, G., & Aitken, S. N. (2013). *Conservation and the genetics of populations* (2nd ed.). Wiley-Blackwell Publishing House.

Bassil, N. V., Gökirmak, T., & Mehlenbacher, S. A. (2004, June). Investigation of Genetic Diversity among European Hazelnut (*Corylus avellana*) Cultivars Using SSR Markers. In *VI International Congress on Hazelnut*, 686 (pp. 141-149).

Boccacci, P., Botta, R., & Rovira, M. (2008). Genetic diversity of hazelnut (*Corylus avellana* L.) germplasm in northeastern Spain. *HortScience*, *43*(3), 667-672.

Boccacci, P., Aramini, M., Valentini, N., Bacchetta, L., Rovira, M., Drogoudi, P., ... & Botta, R. (2013). Molecular and morphological diversity of on-farm hazelnut (*Corylus avellana* L.) landraces from southern Europe and their role in the origin and diffusion of cultivated germplasm. *Tree Genetics & Genomes*, *9*(6), 1465-1480.

Botu, M., Achim, G., Preda, S., Vicol, A., Scutelnicu, A., Giura, S., & Tarasescu, F. (2018). Assortment of prospective cultivars for hazelnut culture in Romania. *Fruit growing research*, *34*(34):7-13.

Brown, J. A., Beatty, G. E., Montgomery, W. I., & Provan, J. (2016). Broad-scale genetic homogeneity in natural populations of common hazel (Corylus avellana) in Ireland. Tree Genetics & Genomes, 12(6), 122.

Campa, A., Trabanco, N., Pérez-Vega, E., Rovira, M., & Ferreira, J. J. (2011). Genetic relationship between cultivated and wild hazelnuts (*Corylus avellana* L.) collected in northern Spain. *Plant breeding*, *130*(3), 360-366.

Erfatpour, M., Hamidogli, Y., Kaviani, B., Fatahi, R., Falahati, M., Javadi, D., & Hashemabadi, D. (2011). Assessment of genetic diversity among some Iranian hazelnut genotypes using SSR markers. *Australian Journal of Crop Science*, *5*(10), 1286-1291.

Farrokhi, T. J., & Ershadi, A. (2022). Evaluation of genetic diversity of Iranians populations of hazlnut (*Corylus avellana* L.) using SSR Markers. *Journal of Nuts*, *13*(4), 273-288.

Fiore, M. C., Marchese, A., Mauceri, A., Digangi, I., & Scialabba, A. (2022). Diversity assessment and DNA-based fingerprinting of Sicilian hazelnut (*Corylus avellana* L.) germplasm. *Plants*, *11*(5): 631.

Gökirmak, T., Mehlenbacher, S. A., & Bassil, N. V. (2009). Characterization of European hazelnut (*Corylus avellana*) cultivars using SSR markers. *Genetic Resources and crop evolution*, *56*(2), 147-172.

Gürcan, Kahraman, Mehlenbacher, S. A., & Erdoğan, V. E. L. İ. (2010). Genetic diversity in hazelnut (*Corylus avellana* L.) cultivars from Black Sea countries assessed using SSR markers. *Plant breeding*, *129*(4), 422-434.

Jahnke, G., Smidla, J., Deák, T., Oláh, R., Szőke, B. Á., & Nyitrainé Sárdy, D. Á. (2022). The SSR null allele problem, and its consequences in pedigree reconstruction and population genetic studies in viticulture. *Horticulturae*, *8*(7), 658.

Kalinowski, S.T.; Taper, M.L.; Marshall, T.C. (2007). Revising How the Computer Program CERVUS Accommodates Genotyping Error Increases Success in Paternity Assignment. *Mol. Ecol*., 16, 1099-1106.

Kavas, M., Yıldırım, K., Seçgin, Z., & Gökdemir, G. (2020). Discovery of simple sequence repeat (SSR) markers in hazelnut (*Corylus avellana* L.) by transcriptome sequencing and SSR-based characterization of hazelnut cultivars. *Scandinavian Journal of Forest Research*, *35*(5-6), 227-237.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, *33*(7), 1870-1874.

Leinemann, L., Steiner, W., Hosius, B., Kuchma, O., Arenhövel, W., Fussi, B., ... & Finkeldey, R. (2013). Genetic variation of chloroplast and nuclear markers in natural populations of hazelnut (*Corylus avellana* L.) in Germany. *Plant Systematics and Evolution*, *299*(2), 369-378.

Martins, S., Simões, F., Mendonça, D., Matos, J., Silva, A. P., & Carnide, V. (2015). Western European wild and landraces hazelnuts evaluated by SSR markers. *Plant molecular biology reporter*, *33*(6), 1712-1720.

Öztürk, G., Revord, R. S., Gül, S., & Lucas, S. J. (2021). Unraveling genetic diversity amongst European hazelnut (*Corylus avellana* L.) varieties in Turkey using ddRAD-seq genomic data. *Frontiers in Plant Science*, 12, Article 661274.

Oztolan-Erol, N., Helmstetter, A. J., İnan, A., Buggs, R. J., & Lucas, S. J. (2021). Unraveling genetic diversity amongst European hazelnut (*Corylus avellana* L.) varieties in Turkey. *Frontiers in plant science*, 12, 661274.

Pop, I. F., Pamfil, D., Raica, P. A., Petricele, I. V., Botu, M., Vicol, A. C., ... & Sisea, C. R. (2010). Evaluation of the genetic diversity of several *Corylus avellana* accessions from the Romanian national hazelnut collection. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, *38*(2), 61-67.

Peakall R., Smouse P.E. (2012). GenALEx6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, 28: 2537-2539.

Stănică, F., Mihai, C. A., Prodan, E. C., & Butcaru, A. C. (2023). The behavior of hazelnut cultivars in the eco-pedoclimatic conditions of the Bucharest area. *Romanian Journal of Horticulture*, 95.

Tanhuanpää, P., Heinonen, M., Bitz, L., & Rokka, V. M. (2019). Genetic diversity and structure in the northern populations of European hazelnut (*Corylus avellana* L.). *Genome*, *62*(8), 537-548.

Wrigley, J., Lucas, S. J., & Öztürk, G. (2020). Genetic diversity and domestication of hazelnut (*Corylus avellana* L.) in Turkey. *Plants, People, Planet*, *2*(2), 154-166.