Genet Resour Crop Evol (2025) 72:6507–6519
https://doi.org/10.1007/s10722-025-02342-x

RESEARCH ARTICLE

# Genetic diversity and population structure of Turkish European chestnut (*Castanea sativa*) genotypes assessed using start codon targeted polymorphism (SCoT) markers

**Erdal Orman · Deniz Çakar · Mehtap Alkan · Göksel Özer · Emrah Güler · Muttalip Gündoğdu**

Received: 19 September 2024 / Accepted: 14 January 2025 / Published online: 27 January 2025
© The Author(s) 2025

**Abstract** The European chestnut (*Castanea sativa*) is an important nut crop that grows naturally in the Black Sea and Aegean regions of Türkiye. This study examined the genetic diversity and population structure of chestnut genotypes from prominent regions in Türkiye using Start Codon Targeted Polymorphism (SCoT) markers. A total of 44 Turkish chestnut genotypes from the Aegean, Marmara, and Black Sea regions, along with a control group of French variety, were analyzed. The SCoT primers underwent tests to select the most suitable ones, producing 8 selected amplified fragments, 65.34% of which were found to be polymorphic. The unweighted pair group method with arithmetic mean (UPGMA) and principal coordinate analysis (PCoA) analyses showed clear discrimination between two populations based on their origins, which was supported by the population structure analysis. The analysis of molecular variance (AMOVA) analysis revealed that 3% of the genetic variation was within populations and 97% was among individuals. The out-group (French variety) showed the furthest genetic similarity, and genetic similarity values decreased with increasing geographic distance. The SCoT primers successfully fingerprinted chestnut genotypes and could be used in future research to analyze the phylogeny of chestnuts using genomic DNA.

**Keywords** SCoT-PCR · DNA marker · Genetic diversity · Chestnut

E. Orman
Van Agricultural Research Institute Directorate, 65000 Van, Turkey

D. Çakar
Central Research Laboratory Application and Research Center, Çankırı Karatekin University, 18100 Çankırı, Turkey

M. Alkan · G. Özer (✉)
Department of Plant Protection, Faculty of Agriculture, Bolu Abant Izzet Baysal University, 14030 Bolu, Turkey
e-mail: gokozer@gmail.com

E. Güler · M. Gündoğdu (✉)
Department of Horticulture, Faculty of Agriculture, Bolu Abant Izzet Baysal University, 14030 Bolu, Turkey
e-mail: gundogdumuttalip@gmail.com

## Introduction

The genus *Castanea*, belonging to the family Fagaceae, consists of self-incompatible, insect-pollinated trees native to the Northern Hemisphere. The *Castanea* genus is categorized into three geographically distinct groups, each consisting of interfertile species. In Asia, there are four recognized species: *C. mollissima*, *C. henryi*, *C. seguinii*, and *C. crenata*. In North America, there are several species, including *C. dentata*, *C. ozarkensis*, and *C. pumila* (Dane et al. 2003), while Europe and Türkiye are home to *C. sativa* (Mellano et al. 2012). This geographic differentiation illustrates

the evolutionary paths and regional adaptations of the species, emphasizing the significance of genetic diversity research for both conservation and breeding programs.The three species—Chinese chestnut (*C. mollissima*), European chestnut (*C. sativa*), and Japanese chestnut (*C. crenata*)—are widely cultivated for their highly nutritious nuts (Pereira-Lorenzo et al. 2010; Barreneche et al. 2019).

The European chestnut (*C. sativa*) is a medium-sized deciduous tree that can grow up to 30–35 m in height. Notably long-lived, some trees have been known to survive for up to 1,000 years, with trunks reaching impressive diameters of up to 12 m. As the only chestnut species indigenous to Europe, *C. sativa* boasts a wide distribution and holds significant economic importance, particularly in rural areas. It is found in 25 countries, covering an area of more than 2.5 million hectares (Conedera et al. 2016). Historically, chestnuts were a key food source for mountain communities throughout Europe (Bellini 2005). Beyond its culinary applications, chestnut wood has been extensively utilized for construction, furniture, and tannin extraction. Anatolia is recognized as one of the regions where chestnuts originated and were first cultivated (Soylu 2004).

In 2022, global chestnut production (in shell) totaled 2,131,240.6 tons, with Türkiye ranking as the fourth largest producer after China, Spain, and Bolivia, contributing 80,200 tons (FAOSTAT 2024). In Türkiye, chestnut cultivation is mainly concentrated in the Marmara and Aegean regions, spanning an area of 81,232 hectares (OGM 2021). The majority of chestnut trees in the country are grown for fruit production, typically using traditional varieties grafted onto seedling rootstocks propagated from seed (Martin et al. 2017). One of the main goals in breeding woody species is the precise identification of genotypes and the early selection of seedlings with desirable traits. Traditionally, genotype identification has been based on morphological traits, which are largely affected by environmental conditions and cultivation methods (Xu 2016). However, these observations can be unreliable due to environmental influences on the expression of genetic traits (Pacheco-Hernández et al. 2021).

Understanding the genetic diversity and relationships among and within plant species is crucial for creating effective conservation and breeding strategies (Bulunuz Palaz et al. 2023). Molecular markers are essential tools for distinguishing and characterizing chestnut species and hybrids, as well as for assessing the status of natural chestnut populations threatened by diseases and genetic contamination. Various molecular markers have been developed and applied to chestnuts, including Amplified Fragment Length Polymorphism (AFLP) (Abdelhamid et al. 2014), isozymes (Pereira-Lorenzo et al. 2010), Random Amplified Polymorphic DNA (RAPD) (Abdelhamid et al. 2014), and chloroplast DNA (Mattioni et al. 2020). Other markers, such as cotyledon storage proteins, inter-simple sequence repeats (ISSR) (Mattioni et al. 2008), Simple Sequence Repeats (SSR) (Abdelhamid et al. 2014; Jiang et al. 2022; Marinoni et al. 2003), chloroplast SSR (cpSSR) (Alcaide et al. 2019; Janfaza et al. 2017; Nie et al. 2021), EST-SSR (Martin et al. 2010), nuclear ribosomal DNA (nrSS) (Zulfiqar et al. 2024), and Single Nucleotide Polymorphisms (SNP) (Nunziata et al. 2020), have also been employed. These markers are widely used (Buck et al. 2003; Marinoni et al. 2003; Laurent et al. 2020) and have been extensively applied (Nishio et al. 2021) to evaluate the genetic diversity of chestnut populations (Mattioni et al. 2008; Martin et al. 2010).

Start Codon Targeted (SCoT) polymorphism markers were developed by Collard and Mackill (2009) based on the conserved region flanking the translation start codon (ATG) in plant genomes. This system employs a single primer, which functions as both forward and reverse primers. In this regard, the SCoT technique is similar to other single primer amplification marker systems, such as Random Amplified Polymorphic DNA (RAPD) and Inter-Simple Sequence Repeat (ISSR). However, several researchers have reported that SCoT markers exhibit greater reproducibility compared to RAPD and ISSR markers (Amom et al. 2020; Gogoi et al. 2020; Tikendra et al. 2021). SCoT markers have been successfully utilized across various plant species and other eukaryotic organisms to assess genetic variation at both interspecies and intraspecies levels (Gorji et al. 2011; Igwe et al. 2017; Yılmaz and Çiftçi 2021; Aydın et al. 2022; Yeken et al. 2022, 2024; Çakar et al. 2023; Palacıoğlu et al. 2023; Abdelhameed et al. 2024; Tekin et al. 2024).

Numerous investigations in the literature have evaluated the genetic variation and DNA fingerprinting of chestnut genotypes and cultivars using various established marker systems, as previously discussed. However, SCoT markers have

not yet been employed to investigate the genetic diversity of chestnut populations on a global scale. This study aimed to explore the potential of SCoT markers in analyzing the genetic variation and population structure of different chestnut genotypes and populations.

## Material and methods

### Plant materials and DNA extraction

Chestnut varieties used in this study were sourced from the Atatürk Horticultural Central Research Institute, Republic of Türkiye Ministry of Agriculture and Forestry. Detailed information regarding the locality and origin of the genotypes is provided in Table 1. Additionally, the sampling locations of these genotypes are marked on the map of Türkiye (Fig. 1).

Leaf samples were collected from the scion portion of 44 grafted chestnut individuals during the first leaf emergence in spring. The samples were placed in sterile bags and transported to the laboratory under cold chain conditions, ensuring delivery within 24 hours. Once in the lab, the young leaf samples were treated by washing them in 70% ethanol, followed by three rinses with sterile distilled water. The samples were then frozen in liquid nitrogen and stored at -80°C. For genomic DNA extraction, approximately 100 mg of leaf tissue from each chestnut genotype/variety was used. The extraction followed a CTAB-based method as per the plant DNA extraction protocol for DArT (https://www.diversityarrays.com). DNA concentrations were measured using a DS-11 FX+ spectrophotometer (Denovix Inc., Wilmington, DE, USA) and subsequently diluted to a final concentration of 10 ng/µL using sterile ultrapure water for further PCR analysis.

### SCoT amplification analyses

SCoT analysis was performed using eight primers chosen from a set of 36 primers originally designed by Collard and Mackill (2009). These primers were selected due to their high polymorphism rates observed in preliminary tests. PCR amplification was carried out in a T100 thermal cycler (Bio-Rad, Hercules, CA, USA). Each PCR reaction had a final

**Table 1** List of chestnut varieties used in the current study

| No | Genotypes/varieties name | Location | Origin |
|---|---|---|---|
| 1 | 2430 | İzmir/Beydağ | AR |
| 2 | 2629 | İzmir/Ödemiş | AR |
| 3 | 2643 | İzmir/Ödemiş | AR |
| 4 | 2647 | İzmir/Ödemiş | AR |
| 5 | 2649 | İzmir/Ödemiş | AR |
| 6 | 2650 | İzmir/Ödemiş | AR |
| 7 | 2665 | İzmir/Ödemiş | AR |
| 8 | 2668 | İzmir/Ödemiş | AR |
| 9 | 2669 | İzmir/Ödemiş | AR |
| 10 | 2686 | İzmir/Ödemiş | AR |
| 11 | 2705 | Manisa/Demirci | AR |
| 12 | 2706 | Manisa/Demirci | AR |
| 13 | 2710 | Manisa/Demirci | AR |
| 14 | 2727 | Manisa/Demirci | AR |
| 15 | Nazilli 23–1 | Aydın/Nazilli | AR |
| 16 | Nazilli 2–5 | Aydın/Nazilli | AR |
| 17 | 51,112 | Yalova/Çınarcık | MR |
| 18 | 51,205 | Bursa/İnegöl | MR |
| 19 | 51,206 | Yalova/Çınarcık | MR |
| 20 | 51,209 | Bursa/İnegöl | MR |
| 21 | 51,301 | Bursa/İnegöl | MR |
| 22 | 51,312 | Bursa/İnegöl | MR |
| 23 | 51,314 | Yalova/Çınarcık | MR |
| 24 | 51,315 | Bursa/İnegöl | MR |
| 25 | 51,509 | Yalova/Çınarcık | MR |
| 26 | 52,510 | Yalova/Termal | MR |
| 27 | 52,104 | Bursa/İnegöl | MR |
| 28 | 62,305 | İzmit/Karamürsel | MR |
| 29 | 63,110 | İzmit/Karamürsel | MR |
| 30 | Karamehmet | İzmit/Karamürsel | MR |
| 31 | Firdola | İzmit/Karamürsel | MR |
| 32 | **Hacıömer | Yalova/Çınarcık | MR |
| 33 | **Dursun Kestanesi | Bursa/İnegöl | MR |
| 34 | Alimolla | Bursa/Cumalıkızık | MR |
| 35 | **Osmanoğlu | Bursa/Center | MR |
| 36 | **Mahmutmolla | Bursa/Cumalıkızık | MR |
| 37 | Derekızık | Bursa/Cumalıkızık | MR |
| 38 | Kızılcık | Bursa/Fidyekızık | MR |
| 39 | Gavuraşı | Bursa/Fidyekızık | MR |
| 40 | **Vakit Kestanesi | Yalova/Esenköy | MR |
| 41 | **Sarıaşlama | Bursa/Center | MR |
| 42 | Serdar | Samsun/Terme | BR |
| 43 | **Erfelek | Sinop/Erfelek | BR |
| 44 | **Bouch de betizac | INRAInstitute | France |

*Registered officially by Variety Registration and Seed Certification Center, Republic of Türkiye Ministry of Agriculture and Forestry

**AR: Aegean region; MR: Marmara region; BR: Blacksea region

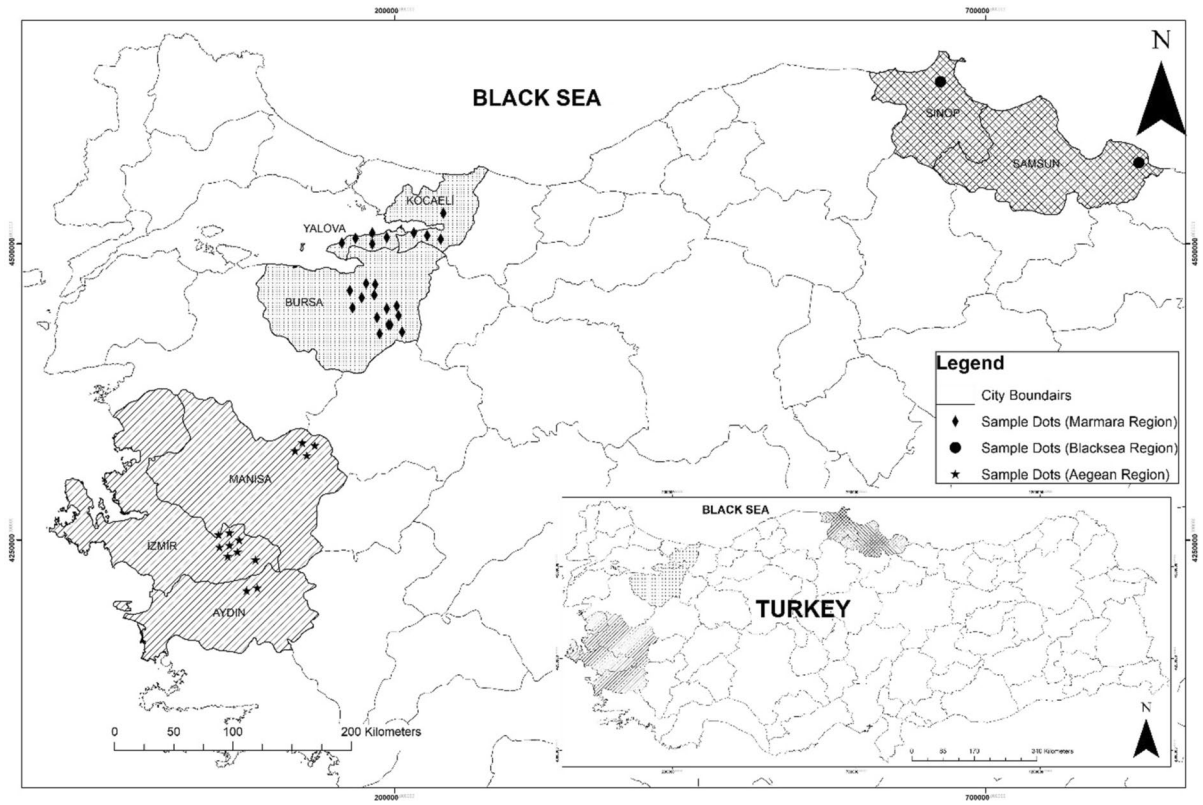



**Fig. 1** Sampling locations of chestnut samples on the Türkiye map

volume of 20 µL, consisting of 1 × reaction buffer containing 2 mM MgCl2, 0.24 mM dNTPs, 0.8 µM primer, 1 unit of Dream Taq polymerase (Thermo Fisher Scientific, Waltham, MA, USA), and 20 ng of template DNA.

The thermal cycling program included an initial denaturation at 95 °C for 3 min, followed by 35 cycles of denaturation at 95 °C for 1 min, annealing at 50 °C for 1 min, and extension at 72 °C for 2 min. The program concluded with a final elongation step at 72 °C for 5 min. After amplification, PCR products were separated on a 1.2% agarose gel using 1 × TAE buffer, stained with ethidium bromide, and visualized under UV light with a gel imaging system(G:BOX F3, Syngene, Cambridge, UK).

## Evaluation of SCoT-PCR data

To ensure reproducibility of the amplification results, each of the eight primers used in the study was evaluated for consistency. The PCR products yielded clear and unambiguous bands, which were manually scored as either present (1) or absent (0), generating a binary data matrix. A 100 bp DNA ladder (Solis BioDyne, Tartu, Estonia) was used to determine the size of the PCR products as a molecular weight marker. To assess the effectiveness of SCoT markers in detecting genetic variation, two key parameters were calculated: resolving power (RP) and polymorphic information content (PIC). The PIC value for each SCoT marker was determined using the formula $PIC = 2f\ (1 - f)$, where "f" represents the frequency of the amplified allele, following the method of Roldán-Ruiz et al. (2000).

Additionally, for dominant markers used in single primer amplification reactions, the band informativeness (Ib) was calculated using the formula $Ib = 1–(2 \times |0.5—p|)$, where “p” is the proportion of individuals with the band. The resolving power (RP) was then computed as the sum of the Ib values for each band produced by each primer, according to

Prevost and Wilkinson (1999). This approach provided a comprehensive evaluation of the discriminating capacity of the SCoT markers used in the study.

The binary data matrix, generated from the SCoT markers, was converted into a genetic similarity matrix using Jaccard's similarity coefficient. This transformation was performed using the NTSYS-pc numerical taxonomy software, version 2.02 (Rohlf, 1992). To represent the genetic relationships among the samples, an unweighted pair-group method with arithmetic averaging (UPGMA) was applied, which resulted in the construction of a dendrogram.

Genetic variation both between and within individuals was further analyzed using GenAlEx 6.5 (Peakall and Smouse 2012). The analysis included several important parameters, such as the observed number of alleles, effective number of alleles, Nei's gene diversity, and Shannon's information index, offering a detailed look into the genetic diversity present in the population. In addition to the dendrogram, a principal coordinate analysis (PCoA) was performed using GenAlEx 6.5 to complement the UPGMA results and provide a clearer visualization of the genetic variation within the studied chestnut individuals.

### Population analyses of chestnut populations

The sampling regions were categorized into three sub-populations for the purpose of population analysis: BR, AR, and MR accessions. To assess genetic variation among these sub-populations, analyses of molecular variance (AMOVA) were performed using GenAlEx 6.5, evaluating contributions of variation both among and within groups. The significance of these variance components was determined through 999 permutations.

For a more detailed evaluation of sub-populations, the binary data matrix was analyzed without prior population origin indication using STRUCTURE v.2.3.4 (Earl 2012). The analysis tested K values ranging from 1 to 10, with ten iterations per K value, to identify the optimal ΔK that represents the most significant subgroups within the populations. The burn-in period and Markov Chain Monte Carlo (MCMC) repetitions were both set to 100,000. The results for different K values were analyzed using STRUCTURE HARVESTER, and the ΔK value was determined using the Evanno method (Evanno et al. 2005).

## Results

### Genetic variation among chestnuts

The primers listed in Table 2 consistently produced reproducible results among chestnut samples, as demonstrated in Fig. 2. Using eight SCoT primers, a total of 101 reproducible and scorable PCR bands were

**Table 2** Primers used in SCoT analysis. The annealing temperature was 50 ℃ for all primers

| SCoT ID | Primer sequences (5′–3′) | GC (%) | TB | PB | PPB (%) | PIC | RP |
|---|---|---|---|---|---|---|---|
| 12 | ACGACATGGCGACCAACG | 61 | 14 | 10 | 71.43 | 0.22 | 4.64 |
| 13 | ACGACATGGCGACCATCG | 61 | 16 | 13 | 81.25 | 0.18 | 4.14 |
| 16 | ACCATGGCTACCACCGAC | 56 | 11 | 5 | 45.45 | 0.10 | 1.45 |
| 18 | ACCATGGCTACCACCGCC | 67 | 6 | 3 | 50.00 | 0.11 | 1.00 |
| 24 | CACCATGGCTACCACCAT | 56 | 13 | 8 | 61.54 | 0.15 | 2.36 |
| 29 | CCATGGCTACCACCGGCC | 72 | 24 | 18 | 75.00 | 0.16 | 5.16 |
| 30 | CCATGGCTACCACCGGCG | 72 | 8 | 5 | 62.50 | 0.06 | 0.59 |
| 32 | CCATGGCTACCACCGCAC | 67 | 9 | 4 | 44.44 | 0.08 | 0.82 |
| | ***Total*** | | ***101*** | ***66*** | | | |
| | ***Avg./primer*** | | ***12.63*** | ***8.25*** | ***65.34*** | ***0.13*** | ***2.52*** |

GC (%), percentage of guanine-cytosine content; TB, total band; PB, polymorphic band; PPB (%), percentage of polymorphic band; PIC, polymorphism information content; RP, resolving power

The underlined regions (ATG) are the "translation start (or initiation) codon"

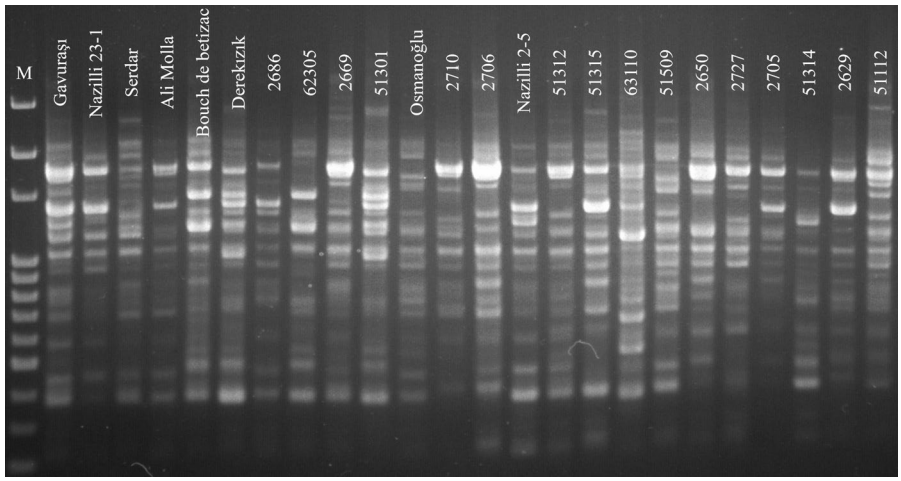



**Fig. 2** Banding patterns of chestnut genotypes obtained by SCoT 29. * M: 100 bp DNA ladder (Solis Bio-Dyne, Tartu, Estonia)

generated to evaluate genetic variation. The DNA fingerprint patterns of chestnuts using primer SCoT 29 are depicted in Fig. 2. Table 2 presents the Polymorphic Information Content (PIC) and Resolving Power (RP) values for each primer, indicating the markers' effectiveness and their ability to differentiate.

Out of the 44 chestnut samples analyzed with the eight SCoT primers, 101 reproducible fragments were identified, with 66 of them being polymorphic (65.3%). The number of amplicons produced by individual primers ranged from 6 (SCoT 18) to 24 (SCoT 29), averaging 12.63 amplicons per marker. The highest PIC value recorded was 0.22 from SCoT 12, while the lowest was 0.06 from SCoT 30. SCoT 29 exhibited the highest RP value of 5.16, whereas SCoT 30 had the lowest RP value of 0.59. The mean PIC and RP values across all primers were calculated to be 0.13 and 2.52, respectively.

The chestnut samples were divided into two main clusters based on the dendrogram analysis. Cluster I included the French chestnut cultivar Bouche de Bétizac, which is a hybrid of *C. sativa* and *C. crenata* (female Bouche rouge × male *Castanea crenata* CA04). On the other hand, Cluster II consisted of all *C. sativa* chestnuts. Within Cluster II, further sub-clustering was observed, primarily based on the geographical origins of the chestnut samples. The Serdar and Erfelek varieties were classified as a subgroup within the main cluster of *C. sativa*. Morover, the Serdar and Erfelek varieties were separated as a subgroup within the *C. sativa* main cluster. The other subcluster consisted of two subclusters, one of which also comprised two genotypes, 2686 and Nazilli23-1 (Fig. 3).

The Principal Coordinate Analysis (PCoA) results indicated clear differentiation among the chestnuts (Fig. 4). The first two components explained 25.84% of the total variation, with PC1 at 13.69% and PC2 at 12.15%. The PCoA plot displayed two primary clusters corresponding to their origins, consistent with the UPGMA clustering results: the first cluster included MR and BR chestnuts, while the second comprised only AR chestnuts (Fig. 4).

## Population analyses of chestnuts

Nei's genetic diversity and Shannon's information index analyses indicated significant genetic variation among chestnuts, demonstrating that SCoT markers provide valuable insights at the intraspecies level. The distribution of chestnuts across various populations enabled an evaluation of population structure. For each population, we analyzed the number of alleles, effective alleles, Nei's gene diversity, and Shannon's information index. The AR and MR populations showed greater genetic diversity compared to the other groups, whereas the BR population exhibited the lowest genetic diversity. The key genetic variation parameters, as identified by GenAlEx, were: observed number of alleles ($1.08 \pm 0.04$), effective number of alleles ($1.15 \pm 0.01$), Nei's gene diversity ($0.13 \pm 0.01$), and Shannon's information index ($0.09 \pm 0.01$). The percentage of polymorphic loci across the populations ranged from 21.78% in the BR

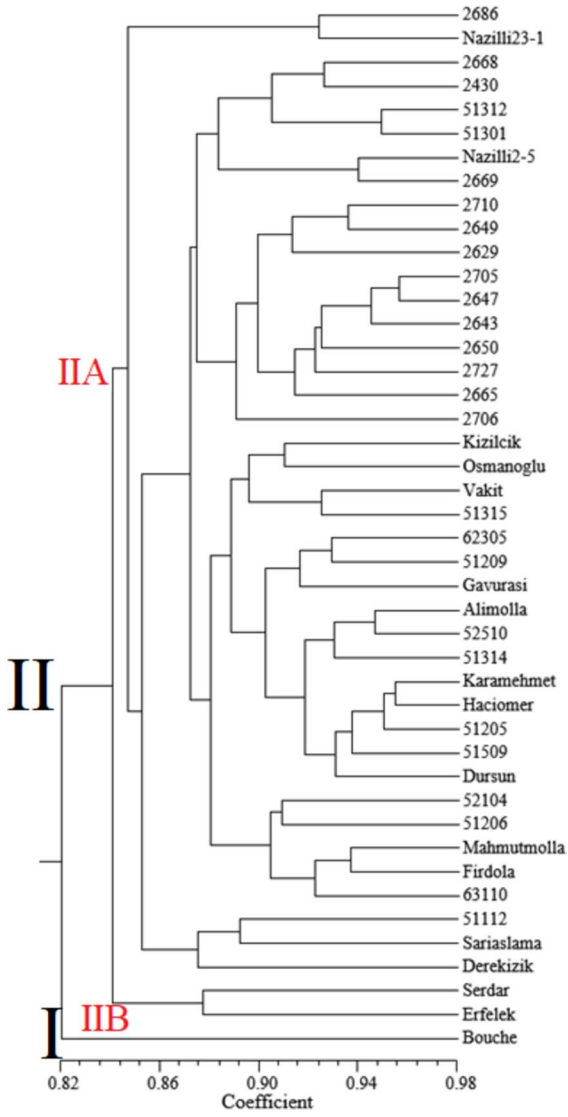



**Fig. 3** Genetic similarity from Start Codon Targeted (SCoT) profiles of chestnut genotypes through cluster analysis using the unweighted pair group method with arithmetic mean (UPGMA) method based on Jaccard's coefficient

population to 46.53% in the AR population, with an average of 30.94% across all groups (Table 3).

The genetic similarity matrix developed by Nei was utilized to evaluate the relationships among populations, with the similarity values detailed in Table 4. The greatest genetic similarity, noted at 0.974, was identified between the AR and MR populations. In contrast, the lowest similarity value of 0.874 was observed between the BR and French populations. The eight SCoT primers employed in the study showed a strong capacity for distinguishing between the chestnut populations.

Based on the AMOVA results, no significant genetic differences were observed within and among the populations (see Table 5). The total genetic variation showed that 3% of the variation existed among populations, while 97% was within populations, consistent with the low FST value of 0.028. However, the genetic variation within the populations was found to be statistically significant ($P < 0.0001$).

To investigate the population structure arising from ancestral groups, we conducted an analysis using STRUCTURE (Pritchard et al. 2000). The analysis indicated that the highest probability of the data was achieved when the individuals were divided into two populations ($\Delta K = 2$) (Fig. 5a). This result suggests that two subgroups optimally represent the chestnut populations, providing insight into their structure and allowing for the estimation of the membership matrix for each individual cluster. The bar plots illustrating these subgroups are presented in Fig. 5b. The Bayesian clustering results align with the PCoA plot (Fig. 4), which similarly classified the individuals into two major clusters.

## Discussion

The diverse array of molecular markers available for assessing genetic diversity facilitates the comparison of different techniques to determine their suitability for specific species (Biswas et al. 2012). This study explores the effectiveness of SCoT markers in distinguishing species and analyzing the population structure of chestnut genotypes. The eight SCoT primers used in this study yielded a polymorphism rate of 65.34% among the *C. sativa* varieties examined. These primers generated a total of 66 polymorphic bands, which is higher than the number of polymorphic bands reported with iPBS markers (Coutinho et al. 2018). Notably, SCoT 29 produced the highest number of polymorphic bands. For comparison, Ho et al. (2024) reported polymorphic band ratios of 82.09% and 26.87% for RAPD and SRAP markers in *C. crenata* and *C. mollissima*, respectively. The results of this study demonstrate a level of discriminatory power comparable to that observed with iPBS

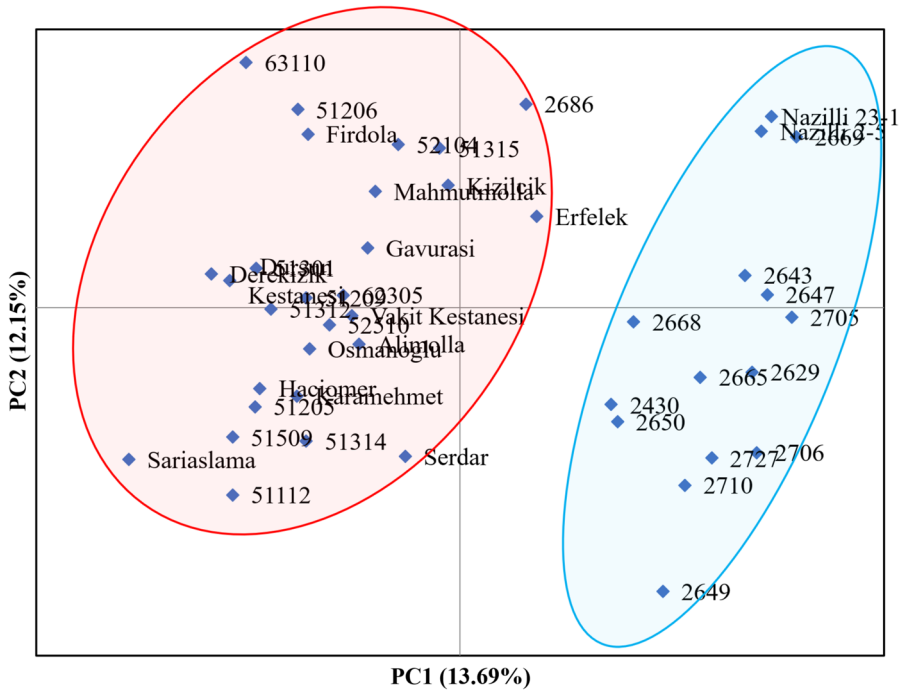



**Fig. 4** The distribution of chestnut genotypes, based on principal coordinate analysis (PCoA)

**Table 3** Genetic diversity population statistics of chestnut populations

| Population* | Na** | Ne | I | He | %P |
|---|---|---|---|---|---|
| AR | 1.34 ± 0.07 | 1.21 ± 0.03 | 0.20 ± 0.03 | 0.13 ± 0.02 | 46.53 |
| MR | 1.51 ± 0.06 | 1.22 ± 0.03 | 0.20 ± 0.03 | 0.13 ± 0.02 | 55.45 |
| BR | 0.53 ± 0.05 | 1.00 ± 0.00 | 0 ± 0 | 0 ± 0 | 21.78 |
| French | 0.92 ± 0.07 | 1.15 ± 0.03 | 0.13 ± 0.02 | 0.09 ± 0.02 | |
| Total | 1.08 ± 0.04 | 1.15 ± 0.01 | 0.13 ± 0.01 | 0.09 ± 0.01 | 30.94*** |

*AR, Aegean region; MR, Marmara region; BR, Blacksea region

**Na, The number of alleles; Ne, the effective number of alleles; I, Shannon's information index; He, Nei's (1973) gene diversity; %P, percentage of polymorphic loci

***: Mean value for %P

**Table 4** Nei's genetic similarity matrix of chestnut populations

| Population* | AR | MR | BR | French |
|---|---|---|---|---|
| AR | 1 | | | |
| MR | 0.974 | 1 | | |
| BR | 0.951 | 0.948 | 1 | |
| French | 0.924 | 0.904 | 0.874 | 1 |

*AR, Aegean region; MR, Marmara region; BR, Blacksea region

**Table 5** The analysis of molecular variance (AMOVA) results for studied chestnuts

| Source | Degrees of freedom | Sun of squares | Mean squares | Est. Var | % |
|---|---|---|---|---|---|
| Among pops | 2 | 1.202 | 0.601 | 0.013 | 3 |
| Within pops | 40 | 18.170 | 0.454 | 0.454 | 97 |
| Total | 42 | 19.372 | | 0.467 | 100 |

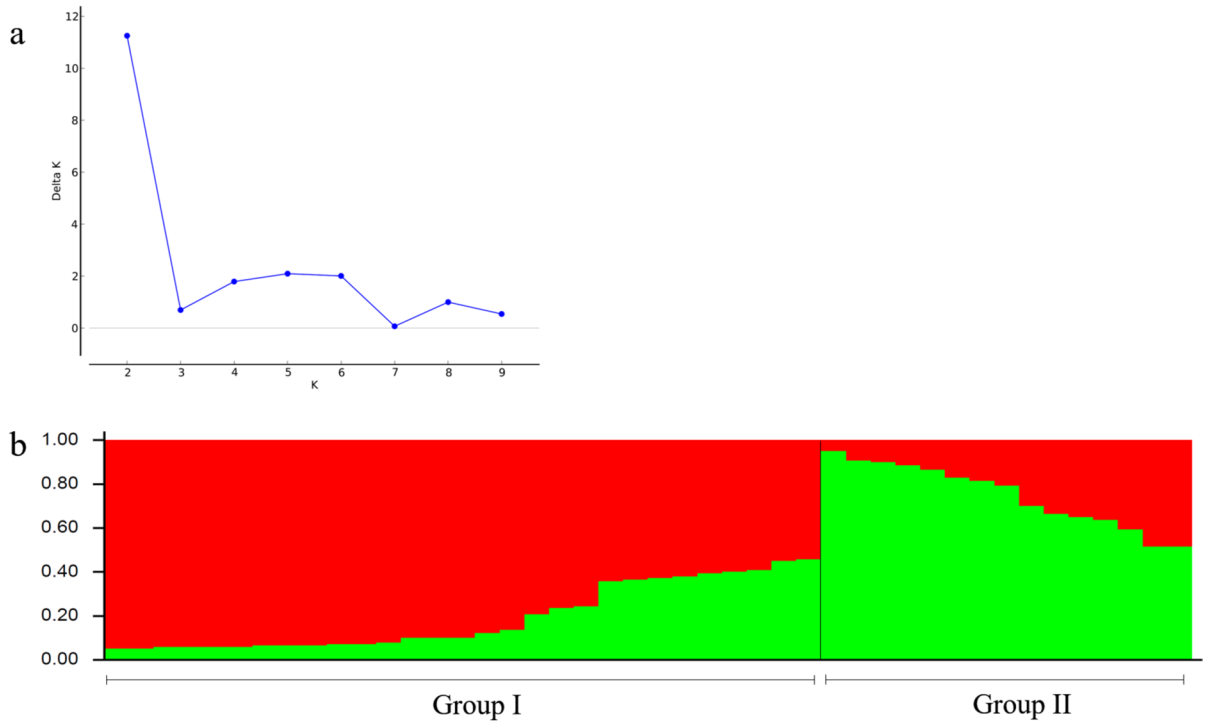



**Fig. 5** Estimated population structure of chestnuts. **a** Bayesian analysis to determine the number of clusters among four populations using ΔK. **b** A plot displaying the assignment probability (Q-values) of individuals to two genetic groups identified in the model for K = 2, with each group represented in a distinct color

markers in previous chestnut research (Coutinho et al. 2018; Kara and Orhan 2023). These findings highlight the efficacy of the SCoT marker system in identifying polymorphism among chestnut varieties. The study emphasizes the importance of preliminary evaluations to select markers with optimal discriminatory power before initiating genetic research. The SCoT29 marker, in particular, exhibited significant polymorphism, underscoring its potential for elucidating intraspecies genetic variation.

In our study, the discriminatory power of SCoT primers is reinforced by the optimal polymorphic information content (PIC) values, which are particularly notable for a dominant marker system. The average PIC values obtained were similar to those reported by iPBS markers in the work of Kara and Orhan (2023). However, Nie et al. (2021) reported higher PIC values using SSR markers, which can be attributed to the fact that SSR is a co-dominant marker system, allowing for the detection of both alleles at a locus, and thus resulting in approximately twice the PIC values compared to dominant markers. Although iPBS markers are robust and effective at targeting retrotransposons, they are universal markers not specific to a particular taxon (Güler et al. 2024). In contrast, SCoT markers were specifically designed based on the conserved regions of the plant genome flanking the start codon, making them inherently more targeted for plants. The relatively high PIC values observed with SCoT markers in this study suggest that they can capture greater genetic variability than previously thought, further underscoring their utility in detecting genetic diversity within and among chestnut genotypes. This finding highlights the importance of selecting markers that are tailored to the genome of the organism being studied to maximize genetic variability detection.

The genotyping of chestnut varieties from different geographical origins demonstrated clear groupings, as observed in both UPGMA and PCoA analyses. This pattern is consistent with findings from previous research using iPBS, ISSR, and RAPD marker techniques in chestnut, as reported by Goulão et al. (2001), Mattioni et al. (2008), and Kara and Orhan

(2023). These results highlight the effectiveness of different marker systems in distinguishing between chestnut genotypes based on geographical origins. Moreover, the intraspecific genetic variation indices, such as Nei's genetic diversity and Shannon's information index, further confirmed that SCoT markers can effectively capture diverse genetic variation at the intraspecific level. Although Shannon's information index in this study was relatively lower than in previous research on *C. sativa* using ISSR markers (Beccaro et al. 2012; Janfaza et al. 2017), the level of polymorphism detected by SCoT markers was still notable. This suggests that SCoT markers can provide valuable information for species discrimination and reveal genetic variation within populations. Therefore, SCoT markers can be used effectively on their own or in combination with other existing marker techniques (such as ISSR or iPBS) to gain a comprehensive understanding of genetic diversity in chestnuts at both the species and intraspecies levels. This combination can help enhance the precision of chestnut genotyping, revealing a more detailed picture of their genetic structure and diversity.

A significant level of genetic differentiation was found among chestnut varieties, based on $F_{ST}$ (0.028) is in the lower part of the range already observed for other *Fagaceae* (Beccaro et al. 2012; Muir and Schloetterer 2005). Similar results indicating a high genetic variation within the individual chestnut varieties using iPBS and SSR markers were also reported by other researchers (Coutibho et al. 2014; Janfaza et al. 2017). Moreover, Martin et al. (2010) noted greater $F_{ST}$ values for SSRs (0.167) and EST-SSRs (0.175) among the European chestnuts revealed by genomic and genic microsatellite markers. Although two sub-groups were obtained under the chestnut varieties cluster on UPGMA, the STRUCTURE addressed two groups ($\Delta K = 2$). Janfaza et al. (2017) were found to belong to two clusters using structure analysis in four *C. sativa* population. Nie et al. (2021) fingerprinted phylogeny of 5 chinese chestnut populations and reported a clustering differantiton according to the regions, which north regions' accessions grouped together and south regions' individiuals grouped together. Our results also indicates a greater dissimilarity when the geographical distance increased, supporting to the previous reports.

Intra-population genetic diversity is crucial for the long-term survival and adaptability of species, especially in the face of environmental changes and pressures (Kahilainen et al. 2014). High genetic diversity within populations enhances their ability to adapt to changing environmental conditions, reduces the likelihood of inbreeding, and decreases extinction risk. In this context, understanding the genetic variation within and between populations of chestnut is vital for both conservation and breeding programs.

It should also be highlighted that dominant molecular markers, such as SCoT, RAPD, ISSR, and iPBS, provide valuable insights into the genetic variation among chestnut varieties. These markers, although dominant (i.e., they do not differentiate between homozygous and heterozygous loci), still offer significant information regarding the genetic structure and diversity of populations. In particular, they are useful for assessing population structure, genetic diversity, and geographic differentiation in chestnut populations.

## Conclusion

This study demonstrated the utility of SCoT markers in revealing the genetic diversity and population structure of *C. sativa* genotypes. The observed polymorphism rate of 65.34% indicated a considerable level of genetic variation among the chestnut populations. The results from UPGMA and PCoA analyses, along with the STRUCTURE analysis, confirmed the geographic-based differentiation of chestnut genotypes. The findings underscore the effectiveness of SCoT markers as a dominant marker system in chestnut genotyping and their potential to complement other marker systems such as iPBS and ISSR. This research provides valuable insights into the genetic variability of chestnut, contributing to future breeding and conservation efforts. The supply chain should be improved by utilizing the proposed divergence of chestnuts and their fruit morphological characteristics in future research.

**Acknowledgements** The genetic material utilized in this study was provided from Atatürk Horticultural Central Research Institute, Türkiye. Authors thanks the instate for providing genetic materials.

**Author contribution** E.O.: Resources, Investigation, Data curation; D.Ç.: Laboratory research; M.A.:Laboratory research, Writing, reviewing and editing; G.Ö.: Data analysis, Sofware, Writing, reviewing, and editing; E.G.: Data analysis, Sofware, Writing, reviewing, and editing; M.G.: Supervision, Writing, reviewing, and editing.

**Funding** Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK). This research did not receive any external funding.

**Data availability** No datasets were generated or analysed during the current study.

**Declarations**

**Conflict of interest** The authors declare that there is no conflict of interest in this research.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

## References

Abdelhameed AA, Ali M, Darwish DBE, AlShaqhaa MA, Selim DAFH, Nagah A, Zayed M (2024) Induced genetic diversity through mutagenesis in wheat gene pool and significant use of SCoT markers to underpin key agronomic traits. BMC Plant Biol 24(1):673

Abdelhamid S, Lê CL, Conedera M, Küpfer P (2014) The assessment of genetic diversity of *Castanea* species by RAPD, AFLP, ISSR, and SSR markers. Turk J Bot 38(5):835–850. https://doi.org/10.3906/bot-1303-30

Alcaide F, Solla A, Mattioni C, Castellana S, Martín MÁ (2019) Adaptive diversity and drought tolerance in *Castanea sativa* assessed through EST-SSR genic markers. For Int J for Res 92(3):287–296

Amom T, Tikendra L, Apana N, Goutam M, Sonia P, Koijam AS, Potshangbam AM, Rahaman H, Nongdam P (2020) Efficiency of RAPD, ISSR, iPBS, SCoT and phytochemical markers in the genetic relationship study of five native and economical important bamboos of North-East India. Phytochemistry 174:112330. https://doi.org/10.1016/j.phytochem.2020.112330

Aydın F, Özer G, Alkan M, Çakır İ (2022) Start codon targeted (SCoT) markers for the assessment of genetic diversity in yeast isolated from Turkish sourdough. Food Microbiol 107:104081. https://doi.org/10.1016/j.fm.2022.104081

Barreneche T, Botta R, Robin C (2019) Advances in breeding of chestnuts. In: Serdar U, Fulbright D (eds) Achieving sustainable cultivation of tree nuts. Burleigh Dodds Science Publishing Limited, Cambridge, pp 317–348

Beccaro GL, Torello-Marinoni D, Binelli G, Donno D, Boccacci P, Botta R, Conedera M (2012) Insights in the chestnut genetic diversity in Canton Ticino (Southern Switzerland). Silvae Genet 61(6):292–300. https://doi.org/10.1515/sg-2012-0037

Bellini E (2005) The chestnut and its resources: images and considerations. Acta Hort 693:85–96

Biswas MK, Chai L, Qiang X, Deng X (2012) Generation, functional analysis and utility of *Citrus grandis* EST from a flowerderived cDNA library. Mol Biol Rep 39:7221–7235. https://doi.org/10.1007/s11033-012-1553-8

Buck EJ, Hadonou M, James CJ, Blakesley D, Russell K (2003) Isolation and characterisation of polymorphic microsatellites in European chestnut (*Castanea sativa* Mill). Mol Eco Notes 3:239–241. https://doi.org/10.1046/j.1471-8286.2003.00410.x

Bulunuz Palaz E, Demirel F, Adali S, Demirel S, Yilmaz A (2023) Genetic relationships of salep orchid species and gene flow among *Serapias vomeracea*× *Anacamptis morio* hybrids. Plant Biotechnology Reports 17(2):315–327

Çakar D, Özer G, Akıllı Şimşek S, Maden S (2023) Determination of vc and mating types of *Cryphonectria parasitica* isolates by multiplex PCR and their genetic diversity in 13 chestnut-growing provinces of Turkey. For Pathol 53(3):e12813

Collard BC, Mackill DJ (2009) Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. Plant Mol Biol Rep 27(1):86–93. https://doi.org/10.1007/s11105-008-0060-5

Conedera M, Manetti MC, Giudici F, Amorini E (2004) Distribution and economic potential of the sweet chestnut (*Castanea sativa* Mill.) in Europe. Ecologia Mediterranea 30:47–61

Conedera M, Tinner W, Krebs P, de Rigo D, Caudullo G (2016) *Castanea sativa* in Europe: distribution, habitat, usage and threats. https://boris.unibe.ch/80790/1/Castanea_sativa.pdf

Coutinho JP, Carvalho A, Martín A, Lima-Brito J (2018) Molecular characterization of *Fagaceae* species using inter-primer binding site (iPBS) markers. Mol Biol Rep 45:133–142

Dane F, Lang P, Huang H, Fu Y (2003) Intercontinental genetic divergence of *Castanea* species in eastern Asia and eastern North America. Heredity 91(3):314–321

Earl DA (2012) Structure harvester: a website and program for visualising STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4(2):359–361. https://doi.org/10.1007/s12686-011-9548-7

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software Structure: a simulation study. Mol Ecol 14(8):2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

FAOSTAT D (2024) Food and agriculture organisation of the United Nations. Statistical database. http://www.fao.org/faostat/en/#data/QC. Accessed 13 Sep 2024.

Gogoi B, Wann SB, Saikia SP (2020) Comparative assessment of ISSR, RAPD, and SCoT markers for genetic diversity in *Clerodendrum* species of North East India. Mol Biol Rep 47(10):7365–7377. https://doi.org/10.1007/s11033-020-05792-x

Gorji AM, Poczai P, Polgar Z, Taller J (2011) Efficiency of arbitrarily amplified dominant markers (SCoT, ISSR and RAPD) for diagnostic fingerprinting in tetraploid potato. Am J Potato Res 88:226–237. https://doi.org/10.1007/s12230-011-9187-2

Goulão L, Valdiviesso T, Santana C, Oliveira CM (2001) Comparison between phenetic characterisation using RAPD and ISSR markers and phenotypic data of cultivated chestnut (*Castanea sativa* Mill.). Genet Resour Crop Evol 48:329–338

Güler E, Karadeniz T, Özer G, Uysal T (2024) Diversity and association mapping assessment of an untouched native grapevine genetic resource by iPBS retrotransposon markers. Genet Resour Crop Evol 71(2):679–690

Ho UH, Kim CH, Kim IJ et al. (2024) Genetic diversity and population structure in Chestnut (*Castanea spp.*) varieties revealed by RAPD and SRAP markers. Agric Res 13:419–428. https://doi.org/10.1007/s40003-024-00722-w

Igwe DO, Afiukwa CA, Ubi BE, Ogbu KI, Ojuederie OB, Ude GN (2017) Assessment of genetic diversity in Vigna unguiculata L. (Walp) accessions using inter-simple sequence repeat (ISSR) and start codon targeted (SCoT) polymorphic markers. BMC Genet 18(1):1–13. https://doi.org/10.1186/s12863-017-0567-6

Janfaza S, Yousefzadeh H, Hosseini Nasr SM, Botta R, Asadi Abkenar A, Torello Marinoni D (2017) Genetic diversity of *Castanea sativa* an endangered species in the Hyrcanian forest. Silva Fenn 51(1):1–15. https://doi.org/10.14214/sf.1705

Jiang X, Fang Z, Lai J, Wu Q, Wu J, Gong B, Wang Y (2022) Genetic diversity and population structure of Chinese chestnut (*Castanea mollissima* Blume) cultivars revealed by GBS resequencing. Plants 11(24):3524

Johnson GP (1988) Revision of *Castanea* sect. *Balanocastanon* (Fagaceae). J Arnold Arbor 69(1):25–49

Kahilainen A, Puurtinen M, Kotiaho JS (2014) Conservation implications of species–genetic diversity correlations. Glob Ecol Conserv 2:315–323. https://doi.org/10.1016/j.gecco.2014.10.013

Kara D, Orhan E (2023) Tolerance evaluation and genetic relationship analysis among some economically important chestnut cultivars in Türkiye using drought-associated SSR and EST-SSR markers. Sci Rep 13(1):20950

Laurent B, Larue C, Chancerel E, Guichoux E, Petit RJ, Barreneche T, Lepais O (2020) Microhaplotype genotyping-by-sequencing of 98 highly polymorphic markers in three chestnut tree species. Conserv Genet Resour 12:567–580

Marinoni D, Akkak A, Bounous G, Edwards KJ, Botta R (2003) Development and characterisation of microsatellite markers in *Castanea sativa* (Mill). Mol Breed 11:127–136. https://doi.org/10.1023/A:1022456013692

Martin MA, Mattioni C, Cherubini M, Taurchini 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, Cherubini M, Villani F, Martín LM (2017) A comparative study of European chestnut varieties in relation to adaptive markers. Agrofor Syst 91:97–109

Mattioni C, Cherubini M, Micheli E, Villani F, Bucci G (2008) Role of domestication in shaping *Castanea sativa* genetic variation in Europe. Tree Genet Genomes 4(3):563–574. https://doi.org/10.1007/s11295-008-0132-6

Mattioni C, Ranzino L, Cherubini M, Leonardi L, La Mantia T, Castellana S, Villani F, Simeone MC (2020) Monuments unveiled: Genetic characterization of large old chestnut (*Castanea sativa* Mill.) trees using comparative nuclear and chloroplast DNA analysis. Forests 11(10):1118

Mellano MG, Beccaro GL, Donno D, Marinoni DT, Boccacci P, Canterino S, Bounous G (2012) Castanea spp. biodiversity conservation: collection and characterization of the genetic diversity of an endangered species. Genet Resour Crop Ev 59:1727–1741.

Muir G, Schloetterer C (2005) Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (Quercus spp.). Mol Ecol 14(2):549–561. https://doi.org/10.1111/j.1365-294X.2004.02418.x

Nie XH, Wang ZH, Liu NW, Li SONG, Yan BQ, Yu XING, Cao QQ (2021) Fingerprinting 146 Chinese chestnut (*Castanea mollissima* Blume) accessions and selecting a core collection using SSR markers. J Integr Agric 20(5):1277–1286

Nishio S, Kunihisa M, Taniguchi F, Kajiya-Kanegae H, Moriya S, Takeuchi Y, Sawamura Y (2021) Development of SSR databases available for both NGS and capillary electrophoresis in apple, pear and tea. Plants 10(12):2796

Nunziata A, Ruggieri V, Petriccione M, De Masi L (2020) Single nucleotide polymorphisms as practical molecular tools to support european chestnut agrobiodiversity management. Int J Mol Sci 21(13):4805. https://doi.org/10.3390/ijms21134805

OGM (2021) Türkiye orman varlığı. Orman Genel Müdürlüğü (in Turkish).

Pacheco-Hernández Y, Villa-Ruano N, Lozoya-Gloria E, Barrales-Cortés CA, Jiménez-Montejo FE, Cruz-Lopez MDC (2021) Influence of environmental factors on the genetic and chemical diversity of *Brickellia veronicifolia* populations growing in fragmented shrublands from Mexico. Plants 10(2):325

Palacıoğlu G, Alkan M, Derviş S, Bayraktar H, Özer G (2023) Molecular phylogeny of plant pathogenic fungi based on start codon targeted (SCoT) polymorphism. Mol Biol Rep 50:1–9. https://doi.org/10.1007/s11033-023-08735-4

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Pereira-Lorenzo S, Costa RML, Ramos-Cabrer AM, Ribeiro CAM et al (2010) Variation in grafted European chestnut and hybrids by microsatellites reveals two main origins in

the Iberian Peninsula. Tree Genet Genomes 6(5):701–715. https://doi.org/10.1007/s11295-010-0285-y

Prevost A, Wilkinson MJ (1999) A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. Theor Appl Genet 98(1):107–112. https://doi.org/10.1007/s001220051046

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155(2):945–959. https://doi.org/10.1093/genetics/155.2.945

Rohlf FJ (1992) NTSYS-pc: Numerical taxonomy and multivariate analysis system, version 2.0. — State Univ. New York, Stony Brook, NY.

Roldan-Ruiz I, Dendauw J, Van Bockstaele E, Depicker A, De Loose MA (2000) AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). Mol Breed 6(2):125–134. https://doi.org/10.1023/A:1009680614564

Rutter PA, Miller G, Payne JA (1990) Chestnuts (*Castanea*). Acta Hort 290:761–788

Soylu A (2004) Kestane yetiştiriciliği ve özellikleri. Hasat yayıncılık, pp 45–67.

Tekin F, Altın İ, Aydın F, Alkan M, Orel DC, Ardıç M, Özer G (2024) Unveiling genetic diversity among bacterial ısolates using SCoT markers. Curr Microbiol 81(12):1–8

Tikendra L, Potshangbam AM, Dey A, Devi TR, Sahoo MR, Nongdam P (2021) RAPD, ISSR, and SCoT markers based genetic stability assessment of micropropagated Dendrobium fimbriatum Lindl. var. oculatum Hk. f.-an important endangered orchid. Physiol Mol Biol Plants 27(2):341–357. https://doi.org/10.1007/s12298-021-00939-x

Xu Y (2016) Envirotyping for deciphering environmental impacts on crop plants. Theor Appl Genet 129:653–673

Yeken MZ, Emiralioglu O, Ciftci V, Bayraktar H, Palacioglu G, Ozer G (2022) Analysis of genetic diversity among common bean germplasm by start codon targeted (SCoT) markers. Mol Biol Rep. https://doi.org/10.1007/s11033-022-07229-z

Yeken MZ, Emiralioğlu O, Tekin M, Çelik A, Çiftçi V, Özer G (2024) Genetic diversity and population structure in Turkish barley (*Hordeum vulgare* L.) cultivars using start codon targeted (SCoT) markers. Genetic Resour Crop Evol, pp 1–12 https://doi.org/10.1007/s10722-024-02297-5

Yılmaz A, Çiftci V (2021) Genetic relationships and diversity analysis in Turkish laurel (*Laurus nobilis* L.) germplasm using ISSR and SCoT markers. Mol Biol Rep 48:4537–4547. https://doi.org/10.1007/s11033-021-06474-y

Zulfiqar S, Aslam MM, Ditta A et al (2024) Evaluation of genetic diversity and population structure of the Chinese chestnut (*Castanea mollissima*) by using NR-SSR markers. Genet Resour Crop Evol. https://doi.org/10.1007/s10722-024-02107-y

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.