Caryologia
International Journal of Cytology, Cytosystematics and Cytogenetics

Firenze University Press
www.fupress.com/caryologia

OPEN ACCESS

Citation: Tinglu Liu, Shuangshuan Zhang, Yonghe Hao, Xiao Liang, Mohsen Farshadfar (2022) Genome survey of pistachio (*Pistacia vera* L.) accessions revealed by Start Codon Targeted (SCoT) markers. *Caryologia* 75(1): 65-76. doi: 10.36253/caryologia-1310

Received: May 10, 2021

Accepted: August 24, 2021

Published: July 6, 2022

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Genome survey of pistachio (*Pistacia vera* L.) accessions revealed by Start Codon Targeted (SCoT) markers

Tinglu Liu[1], Shuangshuan Zhang[1,*], Yonghe Hao[1], Xiao Liang[2], Mohsen Farshadfar[3]

[1] *Ordos agriculture and animal husbandry technology popularizing center, Ordos, Inner Mongolia 017000*
[2] *Ordos city, this paper flag farming technology promotion center, Ordos, Inner Mongolia 017000*
[3] *Department of Agriculture, Payame Noor University (PNU), Tehran, Iran*
*Corresponding author. E-mail: zhang123123301@163.com

Abstract. Pistachio (*Pistacia vera* L.) is the only cultivated and commercially important species in the genus *Pistacia*, consisting of a deciduous, dioeciously and wind-pollinated at least 11 tree species. *Pistacia vera* is native to north Afghanistan, northeast Iran, and central Asian republics. To investigate the genetic diversity of pistachio (*Pistacia vera*), we genotyped 30 cultivars of this species using 10 Start Codon Targeted (SCoT) markers. The SCoT markers generated 9-25 alleles (155 in total) with an average of 16 per locus. The highest value of percentage polymorphism (61.99%) was observed in Ghafori Rafsanjan (cultivars No.27) which shows high value for gene diversity (0.42) and Shanon· information index (0.39). Genotype Shahpasand (Pust Ghermez) (No.10) has the lowest value for percentage of polymorphism (20%) and the lowest value for Shanon, information index (0.15), and He (0.010). Genetic similarity values obtained from Dice's coefficient ranged from 0.66 (between Akbari (Pust Ghermez) and Badami Dishkalaghi) to 0.88 (between populations Menghar Kalaghi and Kaleghochi (Pust Ghermez). The main objectives of this study were to assess the genetic diversity and genetic relationship of pistachio cultivars in Iran. These results could benefit Irainian pistachio germplasm collection, conservation and future breeding.

Keywords: population structure, gene flow, network, genetic admixture, pistachio (*Pistacia vera* L.).

INTRODUCTION

Genetic variability description specifies differences among individuals or populations of the same species and serves as a very good tool for plant breeding and conservation programmes (Minn *et al.* 2015). Different types of DNA markers have been applied in evaluation of genetic diversity of different plants, considering also the effects of the plant growing environment and developmental stage (Hopla *et al.* 2021; Fikirie *et al.* 2020; Gondal *et al.*

2021). The existing genetic variability of the individual species within and among the populations is connected to this species ability to mirror the short- and long-term specific regimes of their living habitats. The analysis of the distribution of the genetic variability patterns specific for landscape and ecological parameters is valuable for identification of the taxa most vulnerable to the anthropogenic impacts (Brandvain *et al*., 2014).

The genus *Pistacia* is a member of the Anacardiaceae family, which comprises 11 or more species (Zohary 1952). *Pistacia vera* L., is a diploid (2n=30) member of the Anacardiaceae family (Zohary 1952; Whitehouse 1957). *Pistacia vera* is native to north Afghanistan, northeast Iran, and central Asian republics (Browiez 1988; Kafkas 2006). Among the nut tree crops, pistachio tree ranks sixth in world production behind almond, walnut, Cashew, hazelnut and chestnut (Mehlenbacher 2003). Iran is the main world producer with more than 400,000 tons followed by Turkey, USA and Syria (Faostat 2004). The main cultivars grown in Iran are Ohady, Kaleh ghochi, Ahmad Aghai, Badami Zarand, Rezaii and Pust piazi (Esmailpour 2001). Iran is the center of origin for four important *Pistacia* species: *P. vera, P. khinjuk* Stocks, *P. eurycarpa* Yalt. (*P. atlantica* subsp. *Kurdica* Zoh.), and *P. atlantica* Dsef. (Karimi *et al*. 2009). Three essential wild *Pistacia* species, including *P. vera, P. khinjuk,* and *P. atlantica* grow in Iran. Although Wild *P. vera* has spread to a territory of around 75,000 ha, in focal Asia, which envelopes Turkmenistan, Afghanistan, and Northeast Iran, where *P. vera* develops in the Sarakhs region, covering around 17,500 ha (Behboodi 2003). Numerous studies have addressed genetic variability in *Pistacia* that were based on evaluation of morphological, physiological, and biochemical characteristics (Zohary 1952; Barone *et al*. 1993; Dollo 1993; Tayefeh Aliakbarkhany *et al*. 2013).

Among them, RAPD (Williams *et al*. 1990) has been the most commonly used method in pistachio cultivars characterization (Hormaza *et al*. 1994, 1998; Kafkas *et al*. 2002; Katsiotis *et al*. 2003; Golan-Gpldhirsh *et al*. 2004; Mirzaei *et al*. 2005). AFLP and SSR techniques have been also used in pistachio to study genetic relationship among *Pistacia* species and cultivars (Golan-Goldhirsh *et al*. 2004; Katsiotis *et al*. 2003; Ibrahim Basha *et al*. 2007; Ahmad *et al*. 2003; Ahmad *et al*. 2005; Ahmadi Afzadi *et al*. 2007).

Although previous studies have partially characterized pistachio diversity in Iran, they did not conduct a full analysis regarding discrimination of wild *Pistacia* and its potential breeding and implication of its conservation. Induction of diversity in *Pistacia* species are based on morphological characteristics which usually can be achieved by budding or grafting selected scions onto seedling rootstocks of the same species or other *Pistacia* species. *Pistacia* species have a high genetic diversity due to their dioecious character, pollination mechanism. Because of these factors high selectivity in rootstocks breeding is required, and therefore knowledge of the genetic relationships among *Pistacia* species would be very useful in pistachio rootstock breeding.

With the progress in plant molecular biology, numerous molecular marker techniques have been developed and used widely in evaluating genetic diversity, population structure and phylogenetic relationships. In recent years, advances in genomic tools provide a wide range of new marker techniques such as, functional and genetargeted markers as well as develop many novel DNAbased marker systems (Collard and Mackill 2009). Start codon targeted (SCoT) polymorphism is one of the novel, simple and reliable gene-targeted marker systems. This molecular marker offers a simple DNA-based marker alternative and reproducible technique which is based on the short conserved region in the plant genes surrounding the ATG (Collard and Mackill 2009) translation start codon. This technique involves a polymerase chain reaction (PCR) based DNA marker with many advantages such as low-cost, high polymorphism and extensive genetic information (Collard and Mackill 2009; Wu *et al*. 2013; Luo *et al*. 2011). The SCoT system has been successfully used to assess genetic diversity, carry out structure analysis, identify cultivars, map quantitative trait loci (QTL), as well as perform DNA fingerprinting and diagnosis in different species (Elshibli and Korpelainen 2008; Rhouma *et al*. 2009).

The present study is the first attempt to use SCoT markers to assess the level of genetic diversity of Irainian pistachio cultivars which were collected from the wild populations. The main objectives of this study were to assess the genetic diversity and genetic relationship of pistachio cultivars in Iran. These results could benefit Irainian pistachio germplasm collection, conservation and future breeding.

MATERIALS AND METHODS

Plant materials

Thirty specimens belonging to three geographical populations of *Pistacia vera* were collected from different localities that were placed between three provinces Semnan, Damghan, Khorasan, Mashhad and Kerman, Rafsanjan. Details of geographical populations are given in Table 1, Fig. 1. Different references were used for the correct identification of species *Pistacia vera* (Zohary 1952; Barone *et al*. 1993; Dollo 1993). Vouchers were deposited

Table 1. List of pistachio cultivars examined for genetic relatedness using SCoT marker system in this study.by Majid Khayatnezhad.

No	Genotypes	Locality	Latitude	Longitude
1	Sarakhs	Khorasan, Mashhad	36.321247	59.532639
2	Ebrahimi	Khorasan, Mashhad	36.321247	59.532639
3	Karimi	Khorasan, Mashhad	36.321247	59.532639
4	Aliabadi	Khorasan, Mashhad	36.321247	59.532639
5	Kaleghochi (Pust Sefid)	Semnan, Damghan	36°9'52.6824'	54°21'27.52
6	Shahpasand (Pust Sefid)	Semnan, Damghan	36°9'52.6824'	54°21'27.52
7	Akbari (Pust Ghermez)	Semnan, Damghan	36°9'52.6824'	54°21'27.52
8	Khanjari Damghan	Semnan, Damghan	36°9'52.6824'	54°21'27.52
9	Kaleghochi (Pust Ghermez)	Semnan, Damghan	36°9'52.6824'	54°21'27.52
10	Shahpasand (Pust Ghermez)	Semnan, Damghan	36°9'52.6824'	54°21'27.52
11	Fakhri	Semnan, Damghan	36°9'52.6824'	54°21'27.52
12	Akbari (Pust Sefid)	Semnan, Damghan	36°9'52.6824'	54°21'27.52
13	Abbas-Ali	Semnan, Damghan	36°9'52.6824'	54°21'27.52
14	Ahmad Agaei	Semnan, Damghan	36°9'52.6824'	54°21'27.52
15	Menghar Kalaghi	Semnan, Damghan	36°9'52.6824'	54°21'27.52
16	Pust Khormaei	Kerman, Rafsanjan	30.3548893	56.002705
17	Ghazvini	Kerman, Rafsanjan	30.3548893	56.002705
18	Fandoghi	Kerman, Rafsanjan	30.3548893	56.002705
19	Javad Aghaei	Kerman, Rafsanjan	30.3548893	56.002705
20	Badami Dishkalaghi	Kerman, Rafsanjan	30.3548893	56.002705
21	Vahedi		30.3548893	56.002705
22	Behesht Abadi	Kerman, Rafsanjan	30.3548893	56.002705
23	Hasan Zadeh	Kerman, Rafsanjan	30.3548893	56.002705
24	Gholamrezaei	Kerman, Rafsanjan	30.3548893	56.002705
25	Ohadi	Kerman, Rafsanjan	30.3548893	56.002705
26	Saiffodini	Kerman, Rafsanjan	30.3548893	56.002705
27	Ghafori Rafsanjan	Kerman, Rafsanjan	30.3548893	56.002705
28	Ravare	Kerman, Rafsanjan	30.3548893	56.002705
29	Italiaei	Kerman, Rafsanjan	30.3548893	56.002705
30	Shasti	Kerman, Rafsanjan	30.3548893	56.002705

at the herbarium of Islamic Azad University, Science and Research Branch, Tehran, Iran (IAUH).

DNA extraction and SCoT-PCR amplification

Fresh leaves were used randomly from four to eleven plants in each of the studied populations. These were dried by silica gel powder. CTAB activated charcoal protocol was used to extract genomic DNA (Esfandani-Bozchaloyi *et al.* 2019). The quality of extracted DNA was examined by running on 0.8% agarose gel. A total of 25 SCoT primers developed by Collard and Mackill (2009), 10 primers with clear, enlarged, and rich polymorphism bands were chosen (Table 2). PCR reactions were carried in a 25µl volume containing 10 mM Tris-HCl buffer at pH 8; 50 mM KCl; 1.5 mM MgCl2; 0.2 mM of each dNTP (Bioron, Germany); 0.2 µM of a single primer; 20 ng genomic DNA and 3 U of *Taq* DNA polymerase (Bioron, Germany). The thermal program was carried out with an initial denaturation for 1 min at 94°C, followed by 40 cycles in three segments: 35 s at 95°C, 40s at 55°C and 55s at 72°C. Final extension was performed at 72°C for 5 min. The amplification products were observed by running on 1% agarose gel, followed by the ethidium bromide staining. The fragment size was estimated by using a 100 bp molecular size ladder (Fermentas, Germany).

DATA ANALYSES

Morphological studies

In total nineteen morphological (nineteen quantitative) characters were studied. Four to twelve samples

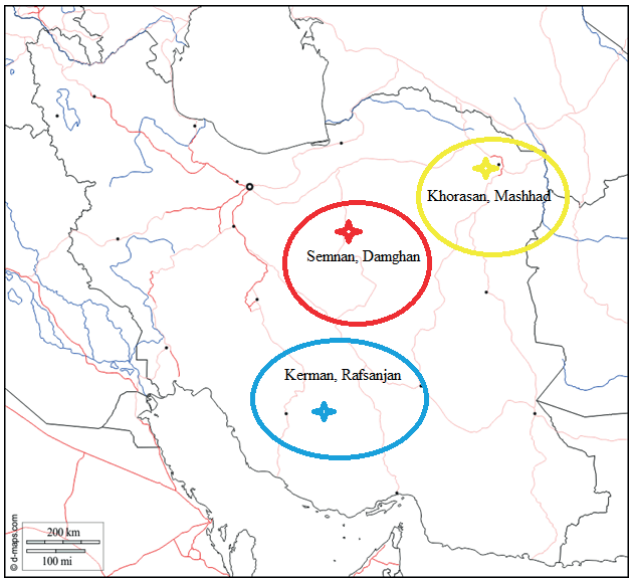

Figure 1. Map of Iran shows the collection sites and provinces where of *Pistacia vera* species were obtained for this study.

from each population were randomly studied for morphological analyses (Appendix 1). Morphological characters were first standardized (Mean = 0, Variance = 1) and used to establish Euclidean distance among pairs of taxa (Podani 2000). For grouping of the plant specimens, The UPGMA (Unweighted paired group using average) and Ward (Minimum spherical characters) as well as ordination methods of MDS (Multidimensional scaling) were used (Podani 2000). PAST version 2.17 (Hammer *et al.* 2012) was used for multivariate statistical analyses of morphological data.

Molecular analyses

Excel 2013 was used to calculate the total number of bands (TNB), the number of polymorphic bands (NPB), and the percentage of polymorphic bands (PPB). The polymorphism information content (PIC) of SCoT primers was determined using POWERMARKER v3.25. Binary characters (presence = 1, absence = 0) were used to encode SCoT bands and used for further analyses. Parameter like Nei's gene diversity (H), Shannon information index (I), number of effective alleles, and percentage of polymorphism (P% =number of polymorphic loci/number of total loci) were determined (Weising *et al.* 2005; Freeland *et al.* 2011).

Shannon's index was calculated by the formula: H' = -Σpiln pi. Rp is defined per primer as: Rp = Σ Ib, were "Ib" is the band informativeness, that takes the values of 1-(2x [0.5-p]), being "p" the proportion of each genotype containing the band. The percentage of polymorphic loci, the mean loci by accession and by population, UHe, H' and PCA were calculated by GenAlEx 6.4 software (Peakall and Smouse 2006)

Nei's genetic distance among populations was used for Neighbor Joining (NJ) clustering and Neighbor-Net networking (Freeland *et al.* 2011; Huson and Bryant 2006). The comparison of genetic divergence or genetic distances, estimated by pairwise F_{ST} and related statistics, with geographical distances by Mantel test is one of the most popular approaches to evaluate spatial processes driving population structure. The Mantel test was performed as implemented in PAST ver. 2.17 (Hammer *et al.* 2012). For this, Nei genetic distance was determined for scot data, while Geographic distance of PAST was determined for geographical data. It is calculated based on the sum of the paired differences among both longitude as well as latitude coordinates of the studied populations. The Mantel test, as originally formulated in 1967, is given by $Z_m = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} \times d_{ij}$ where g_{ij} and d_{ij} are, respectively, the genetic and geographic distances between populations *i* and *j*, considering *n* populations. Because Z_m is given by the sum of products of distances its value depends on how many populations are studied, as well as the magnitude of their distances.

AMOVA (Analysis of molecular variance) test (with 1000 permutations) as implemented in GenAlex 6.4 (Peakall and Smouse 2006), and Nei's Gst analysis as implemented in GenoDive ver.2 (2013) (Meirmans and Van Tienderen 2004) were used to show genetic difference of the populations. Moreover, populations' genetic differentiation was studied by G'ST est = standardized measure of genetic differentiation (Hedrick 2005), and D_est = Jost measure of differentiation (Jost 2008).

To assess the population structure of the pistachio genotypes, a heuristic method based on Bayesian clustering algorithms were utilized. The clustering method based on the Bayesian-model implemented in the software program STRUCTURE (Pritchard *et al.* 2000; Falush *et al.* 2007) was used on the same data set to better detect population substructures. This clustering method is based on an algorithm that assigns genotypes to homogeneous groups, given a number of clusters (K) and assuming Hardy-Weinberg and linkage equilibrium within clusters, the software estimates allele frequencies in each cluster and population memberships for every individual (Pritchard *et al.* 2000). The number of potential subpopulations varied from two to ten, and their contribution to the genotypes of the accessions was calculated based on 50,000 iteration burn-ins and 100,000 iteration sampling periods. The most probable number

(K) of subpopulations was identified following Evanno *et al.* (2005). In K-Means clustering, two summary statistics, pseudo-F, and Bayesian Information Criterion (BIC), provide the best fit for k (Meirmans 2012). Gene flow (Nm) which were calculated using POPGENE (version 1.31) program (Yeh *et al.* 1999). Gene flow was estimated indirectly using the formula: Nm = 0.25(1 - FST)/FST. In order to test for a correlation between pair-wise genetic distances (FST) and geographical distances (in km) between populations, a Mantel test was performed using Tools for Population Genetic Analysis (TFPGA; Miller 1997) (computing 999 permutations). This approach considers equal amount of gene flow among all populations.

RESULTS

SCoT polymorphisms

Twenty-five SCoT primers were tested with four of *Pistacia vera* cultivars as DNA templates; all primers produced amplification products, and only primers showing clear and reproducible band patterns were selected for further analysis. The size of the amplified fragments ranged from 100 to 2500 bp (Fig. 2). Ten primers were then chosen for the genotypes identification and phylogenetic analysis. As shown in Table 2, all 10 primers used for SCoT analysis. A total of 155 fragments were obtained, and 143 of the fragments were polymorphic. The number of polymorphic fragments for each SCoT primer ranged from 8 (ST3) to 25 (ST14), with an average of 12. The percentage of polymorphic fragments was from 84.57% to 100.00%, with an average of 94.55% polymorphism. Polymorphism information content (PIC) values were 0.22 to 0.59, with an average of 0.41. The number of different alleles was 0.43 at the species (Table 3). These results indicated that a high level of polymorphism could be detected among *Pistacia vera* cultivars using SCoT markers.

Populations' genetic diversity

Genetic diversity parameters determined in three geographical populations of *Pistacia vera* are presented in Table 3. The percentage of polymorphic loci (*P*) and Nei's gene diversity (*H*) were important parameters for measuring the level of genetic diversity. In Table 3, the genetic diversity parameters of the 30 *Pistacia vera* cultivars are shown. The highest value of percentage polymorphism (61.99%) was observed in Ghafori Rafsanjan (cultivars No.27) which shows high value for gene diversity (0.42) and Shanon' information index (0.39). Genotype Shahpasand (Pust Ghermez) (No.10) has the lowest value for percentage of polymorphism (20%) and the lowest value for Shanon, information index (0.15), and He (0.010).

Population genetic differentiation

AMOVA (PhiPT = 0.29, P = 0.010), revealed significant difference among the studied genotypes (Table 4, Fig. 3). It also revealed that, 23% of total genetic variability was due to within genotypes diversity and 55% was due to among genotypes genetic differentiation.

Moreover, pair-wise AMOVA revealed significant genetic difference almost among all the studied genotypes. These results indicate that of pistachio genotypes are genetically differentiated and we can use such genetic difference in future breeding programs of this

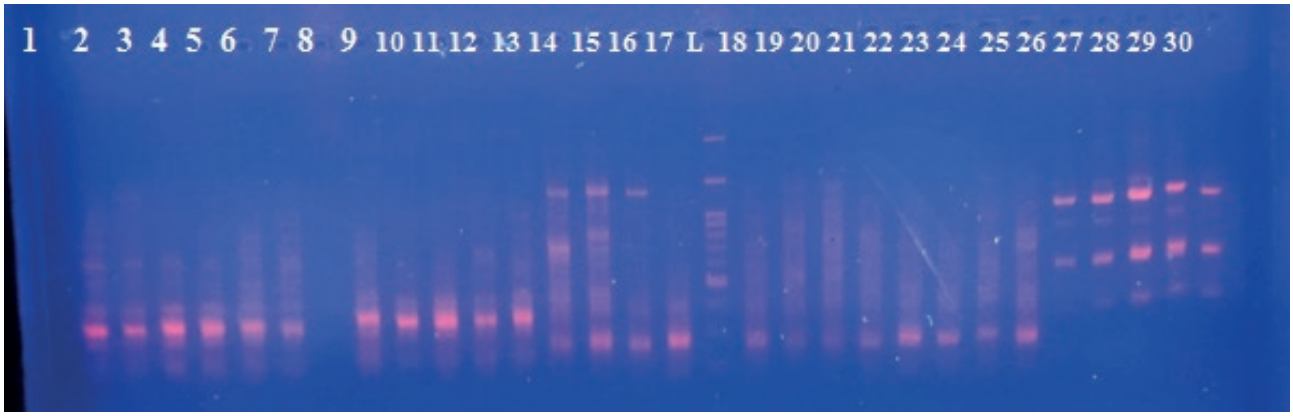

Figure 2. Electrophoresis gel of *Pistacia vera* species from DNA fragments produced by SCoT-11 molecular markers, (Population numbers are according to Table 1).

Table 2. SCoT primers used for this study and the extent of polymorphism. TNP: total number of bands; NPB: number of polymorphic bands; PPB: percentage of polymorphic bands; PIC: polymorphism information content.

Primer name	Primer sequence (5'-3')	TNB	NPB	PPB	PIC
SCoT-1	CAACAATGGCTACCACCA	13	13	100.00%	0.55
SCoT-3	CAACAATGGCTACCACCG	9	8	86.99%	0.43
SCoT-6	CAACAATGGCTACCACGC	19	19	100.00%	0.34
SCoT-11	AAGCAATGGCTACCACCA	17	16	94.33%	0.47
SCoT-14	ACGACATGGCGACCACGC	25	25	100.00%	0.35
SCoT-15	ACGACATGGCGACCGCGA	14	12	94.74%	0.59
SCoT-16	CCATGGCTACCACCGGCC	15	12	92.31%	0.49
SCoT-17	CATGGCTACCACCGGCCC	10	10	100.00%	0.22
SCoT-18	ACCATGGCTACCACCGCG	12	10	84.57%	0.50
SCoT-19	GCAACAATGGCTACCACC	24	24	100.00%	0.37
Mean		16	12	94.55%	0.41
Total		155	143		

Table 3. Genetic diversity parameters in the studied populations of pistachio cultivars (N = number of samples, Na = Number of different alleles, Ne = number of effective alleles, I= Shannon's information index, He = genetic diversity, UHe = unbiased gene diversity, P%= percentage of polymorphism, populations).

Code genotypes	N	Na	Ne	I	He	UHe	%P
Sarakhs	5.000	0.555	1.020	0.22	0.25	0.28	43.53%
Ebrahimi	8.000	0.431	1.088	0.20	0.22	0.25	49.53%
Karimi	8.000	0.255	1.021	0.25	0.28	0.22	37.15%
Aliabadi	5.000	0.261	1.024	0.292	0.23	0.23	53.15%
Kaleghochi (Pust Sefid)	5.000	0.886	1.183	0.184	0.116	0.122	24.29%
Shahpasand (Pust Sefid)	8.000	0.686	1.157	0.30	0.11	0.22	39.43%
Akbari (Pust Ghermez)	4.000	0.344	1.042	0.28	0.23	0.20	33.53%
Khanjari Damghan	5.000	0.455	1.077	0.277	0.24	0.22	53.05%
Kaleghochi (Pust Ghermez)	3.000	0.255	1.021	0.15	0.18	0.19	48.45%
Shahpasand (Pust Ghermez)	3.000	0.643	1.173	0.154	0.010	0.010	20.00%
Fakhri	8.000	0.431	1.088	0.20	0.22	0.25	49.53%
Akbari (Pust Sefid)	9.000	0.255	1.021	0.25	0.28	0.22	37.15%
Abbas-Ali	6.000	0.261	1.024	0.292	0.23	0.23	40.15%
Ahmad Agaei	10.000	0.287	1.253	0.266	0.254	0.28	50.99%
Menghar Kalaghi	5.000	0.358	1.430	0.28	0.20	0.29	23.50%
Pust Khormaei	6.000	0.299	1.029	0.231	0.28	0.23	24.38%
Ghazvini	5.000	0.462	1.095	0.288	0.29	0.22	22.05%
Fandoghi	8.000	0.399	1.167	0.24	0.21	0.213	32.88%
Javad Aghaei	5.000	0.336	1.034	0.23	0.25	0.29	41.83%
Badami Dishkalaghi	4.000	0.344	1.042	0.28	0.23	0.20	57.53%
Vahedi	5.000	0.455	1.077	0.277	0.24	0.22	55.05%
Behesht Abadi	3.000	0.255	1.021	0.15	0.18	0.19	38.45%
Hasan Zadeh	3.000	0.643	1.173	0.154	0.102	0.109	30.00%
Gholamrezaei	8.000	0.431	1.088	0.20	0.32	0.25	41.53%
Ohadi	9.000	0.255	1.021	0.25	0.28	0.22	27.15%
Saiffodini	6.000	0.261	1.024	0.292	0.23	0.23	43.15%
Ghafori Rafsanjan	10.000	0.287	1.253	0.396	0.424	0.44	61.99%
Ravare	3.000	0.567	1.062	0.24	0.224	0.213	34.73%
Italiaei	3.000	0.499	1.067	0.24	0.281	0.24	49.26%
Shasti	9.000	0.352	1.083	0.23	0.22	0.24	45.05%

Table 4. Analysis of molecular variance (AMOVA) of the studied species.

Source	Df	SS	MS	Est. Var	%
Among Regions	12	39.211	23.648	0.266	19%
Among Pops	15	96.822	18.802	0.114	55%
Among Indiv	57	64.553	21.130	0.283	20%
Within Indiv	71	15.500	0.284	0.204	8%
Total	141	215.007		1.678	100%

df: degree of freedom; **SS:** sum of squared observations; **MS:** mean of squared observations; **EV:** estimated variance.

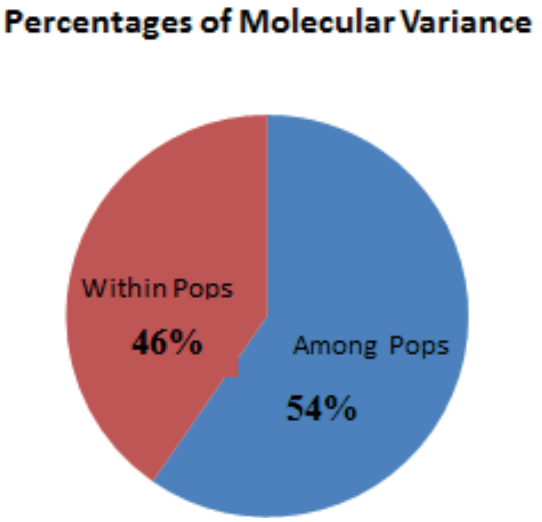

Figure 3. AMOVA test of the studied populations.

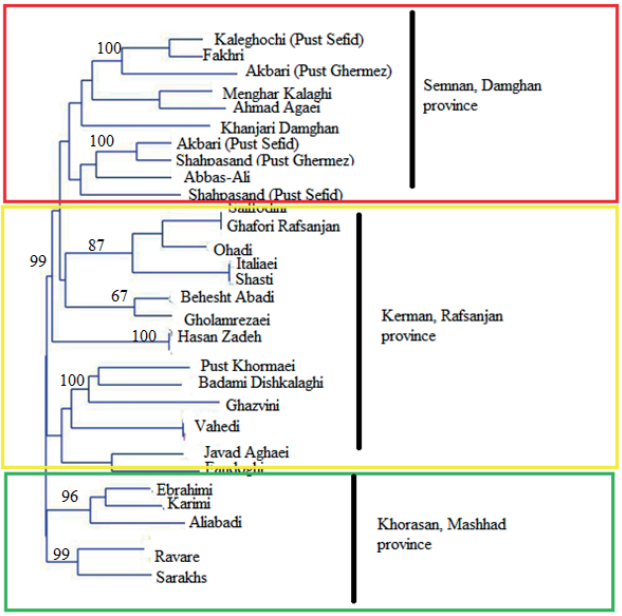

Figure 4. NJ tree of populations in *Pistacia vera* based on SCoT molecular markers.

valuable plant species. The results of this study showed that there is a relatively low level of genetic diversity in the studied samples which are expected in view of the dioecius and outbreeding nature of the cultivated pistachio cultivars and high level of heterozygosity due to the cross-pollinating nature of the plant established during the evolution and domestication processes which have been conserved by the propagation of clones through vegetative reproduction.

The pairwise comparisons of 'Nei genetic identity' among the studied populations of *Pistacia vera* (Table not included) have shown a higher a genetic similarity (0.887) between populations Menghar Kalaghi (province Semnan) and Kaleghochi (Pust Ghermez) (province Semnan), while the lowest genetic similarity value (0.667) occurs between Akbari (Pust Ghermez) (province Semnan) and Badami Dishkalaghi (province Kerman).

Populations genetic affinity

NJ tree and Neighbor-Net network produced similar results therefore only NJ tree is presented and discussed (Fig. 4). This result show that molecular characters studied can delimit *Pistacia vera* genotypes in two different major clusters or groups. In general, two major clusters were formed in NJ tree (Fig. 3), four genotypes of cultivars Sarakhs, Ebrahimi, Karimi and Aliabadi formed a single cluster, and these genotypes were all from Khorasan, Mashhad province. Cluster II contained two sub-clusters, and most of individuals Kaleghochi (Pust Sefid); Shahpasand (Pust Sefid); Akbari (Pust Ghermez); Khanjari Damghan and Kaleghochi (Pust Ghermez), Shahpasand (Pust Ghermez); Fakhri; Akbari (Pust Sefid); Abbas-Ali and Ahmad Agaei (Semnan Province) formed cluster II. There were 26 individuals in this cluster.

Besides, principal coordinate analysis (PCoA) was performed to visualize the association among the genotypes in more detail. The PCoA results showed that the first three principal coordinates account for 64.88% of the total variation (not shown). Based on the results of PCoA analysis, cultivars Sarakhs, Ebrahimi, Karimi and Aliabadi genotype showed the highest dissimilarity with other genotypes. Additionally, the results from Bayesian clustering analysis using STRUCTURE software (Fig. 4) confirmed the groupings we observed in NJ and PCoA clusterings.

The present study indicated that a higher genetic diversity was found in the older genotypes. This fact confirms our speculation that pistachio cultivations have increasingly led to the reduction of their genetic variation due to deployment of improved cultivars and to the availability of private or public grafted seedling nurseries for pistachio, as well as the changing livelihood conditions. Recently, the method of pistachio cultivation is changing leading towards an increased reduction of crop diversity deployed on farm. In the past, pistachio diversity was maintained high in the field through a number of cultivation practices, s. a. use of male varieties derived from seed, use of wild *Pistacia* species to boost pollination and hence the fruit setting, use of natural populations of wild *Pistacia* (*P. atlantica*) as a rootstock due to their well-known resistance to stony and calcareous soils.

This is in agreement with AMOVA and genetic diversity parameters presented before. Mantel test after 5000 permutations produced significant correlation between genetic distance and geographical distance in these populations (r = 0.87, P = 0.001). Therefore, the populations that are geographically more distant have less amount of gene flow, and we have isolation by distance (IBD) in *Pistacia vera* genotypes. The most popular approaches for estimating divergence include calculation of genetic distances and variance partitioning among and within populations using Wright's F_{ST} and other related statistics, such as G_{ST}, A_{ST}, R_{ST}, θ_{ST} and Φ_{ST}. For instance, the F_{ST} gives an estimate of the balance of genetic variability among and within populations, and is an unbiased estimator of divergence between pairs of populations under an island-model in which all populations diverged at the same time and are linked by approximately similar migration rates. However, migration rates usually vary proportionally with geographical distances, so that pairwise F_{ST} estimates between pairs of populations vary. Therefore, the populations that are geographically more distant have less amount of gene flow, and we have isolation by distance (IBD) in *Pistacia vera* genotypes.

Populations genetic structure

The number of genetic groups was determined by two methods of 1—K-Means clustering which is based on the maximum likelihood approach, and 2—Evanno test which is based on STRUCTURE analysis and is a Bayesian approach based method. K-Means clustering based on pseudo-F and BIC (Bayesian Information Criterion) recognized 3 and 5 genetic groups, respectively. This is in agreement with AMOVA result, showing significant genetic difference among date populations of *Pistacia vera* genotypes.

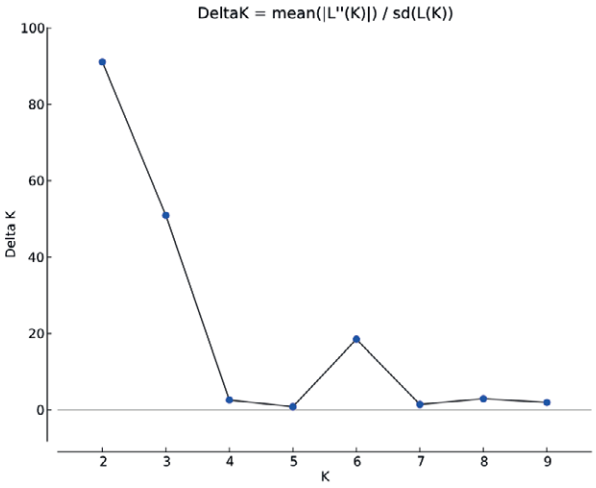

Figure 5. Delta k plot of Evanno's test based on STRUCTURE analysis.

Evan test based on delta k (Fig. 5) identified the optimum number of genetic groups 3. We performed STRUCTURE analysis based on k = 3, to identify the genetic groups (Fig. 6). In the plot of k = 3, the cultivars Sarakhs, Ebrahimi, Karimi and Aliabadi (red colored) are placed in the first genetic group, while the populations of Kaleghochi (Pust Sefid); Shahpasand (Pust Sefid); Akbari (Pust Ghermez); Khanjari Damghan and Kaleghochi (Pust Ghermez), Shahpasand (Pust Ghermez); Fakhri; Akbari (Pust Sefid); Abbas-Ali and Ahmad Agaei (Semnan Province) (blue colored) formed the second genetic group and finally the populations of Kerman province (green colored) formed the third genetic group. These different genetic groups may be used in future breeding and hybridization programs of Iranian date *Pistacia vera* genotypes.

The mean Nm = 0.65 was obtained for all SCoT loci, which indicates low amount of gene flow among the populations and supports genetic stratification as indicated by K-Means and STRUCTURE analyses. This result is in agree with grouping we obtained with PCA plot, as these populations were placed close to each other. As evidenced by STRUCTURE plot based on admixture model, these shared alleles comprise very limited part of the genomes in these populations and all these results are in agreement in showing high degree of genetic stratification within of *Pistacia vera* genotypes.

Morphometric analyses

In present study we used 30 plant accessions (six to fourteen samples from each populations) belonging to four different populations. In order to determine the most

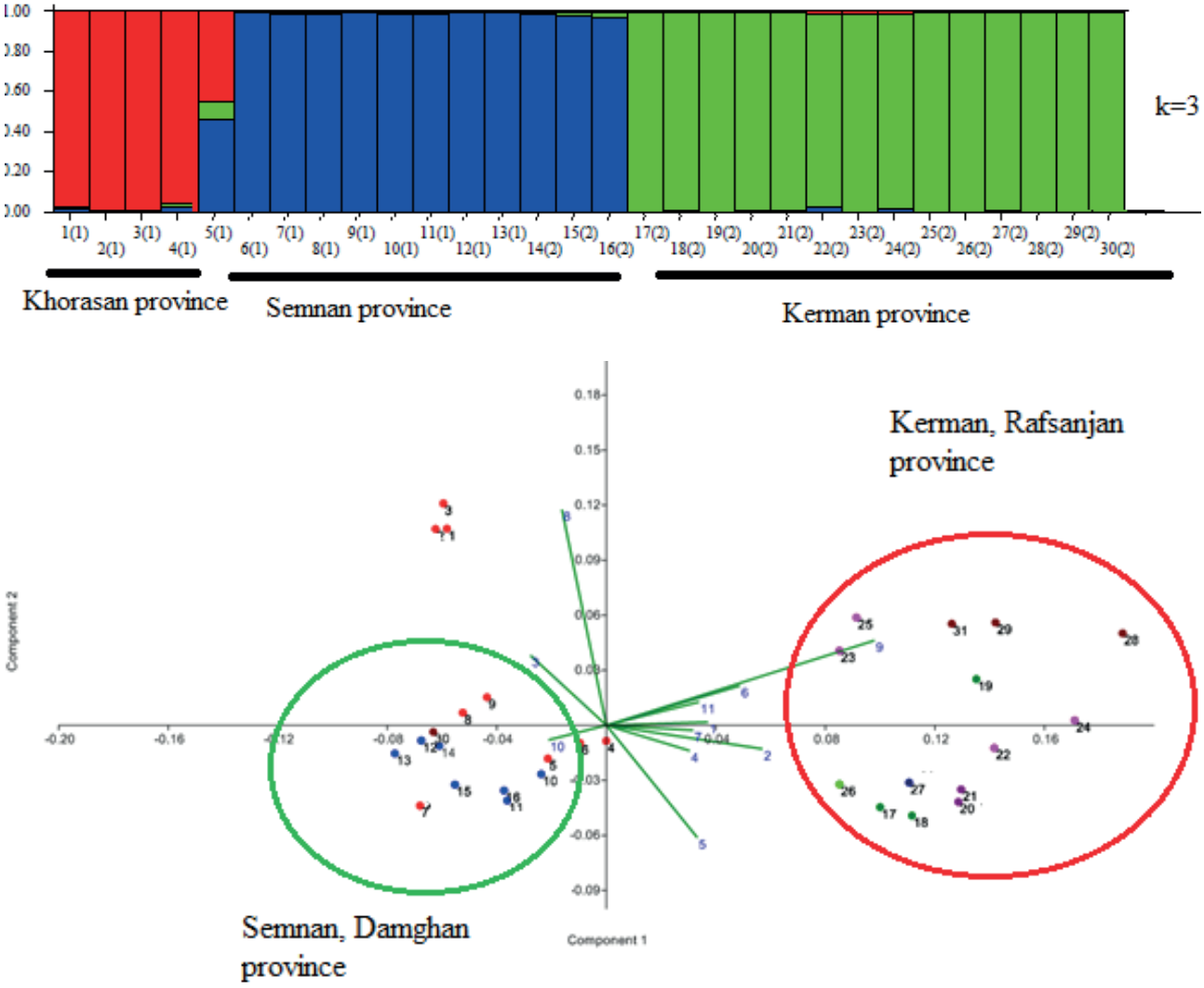

Fig. 6. Top: STRUCTURE plot of *Pistacia vera* populations based on k = 3, Numbers are according to Table 1.Bottom: PCA plot of *Pistacia vera* populations based on morphological characters. Numbers are according to Table 1.

variable characters among the taxa studied, PCA analysis has been performed (Fig. 6). It revealed that the first three factors comprised over 73% of the total variation. In the first PCA axis with 40% of total variation, such characters as length of leaves; width of leaves; length of petioles; length of the terminal leaf; width of the terminal leaf; length of inflorescence have shown the highest correlation (> 0.7), fruit length; fruit width; fruit thickness; number of fruit per inflorescence; kernel infestation were characters influencing PCA axis 2 and 3, respectively.

Different clustering and ordination methods produced similar results therefore, PCA plot of morphological characters are presented here (Fig. 6). The result showed morphological difference/ divergence among most of the studied populations. This morphological difference was due to quantitative characters only.

DISCUSSION

The coupling of ecological and genetic data will provide the most suitable background for preserving the ability of the biota to respond the rapid environmental changes ((Sawadogo *et al.* 2021; Paul *et al.* 2021)). The literature reports the following basic factors influencing the distribution of genetic variation: habitat specify, plant-insect interactions, connectivity and disturbance, dispersal ability, species lifespan, reproductive rates

and existing genetic diversity (Esfandani–Bozchaloyi, *et al*. 2018a, 2018b, 2018c, 2018d). Genetic diversity when analysed by neutral markers does not correspond to the adaptive ability of plant populations, but these types of markers are very useful for the interpretation of the past landscapes, refugia and gene flow (Wankiti *et al.* 2021; Lucena *et al.* 2021). That is, why the selected genes or markers of active parts of plant genomes are used to interprete the plant genome response to the changes to the local climate and environment. Molecular-based population genetic data are very useful for determining the ecological and habitat events in the past and for detection of patterns of the recent genetic divergence. This can be achieved using different types DNA markers. SCoT markers are novel molecular markers that target the translation initiation site and preferentially bind to genes that are actively transcribed. These primers have been shown to exhibit relatively high levels of polymorphism (Collard and Mackill 2009). It was more informative than IRAP and ISSR for the assessment of diversity of plants (Collard and Mackill 2009).

Pistachio has important socio-economic and ecological impacts in the arid and semi-arid agricultural regions of Iran (Kafkas *et al.* 2006). In addition, Iran hosts a wide genetic diversity of *Pistacia* spp. and more than 300 pistachio genotypes have been collected across the country. Iran therefore possesses valuable germplasm for pistachio improvement and conservation programs. Assessing genetic diversity and relationships among cultivars of Iranian pistachio, using discriminative and robust markers, is therefore important (Mirzaei *et al.* 2005).

In the present work, 30 P. vera cultivars were characterized with 10 SCoT markers. The results confirm the efficiency of microsatellite markers for fingerprinting purposes. Our results demonstrated that the Polymorphism information content (PIC) ranged from 0.22 to 0.59 with an average value of 0.41, while the percentage of polymorphism (P%) ranged from 0.20 to 0.61 with an average value of 0.42 and also the expected heterozygosity (He) varied from 0.011 to 0.42 with an average of 0.20.These values were higher than those reported by Arabnejad *et al.* (2008), who detected an average of 3.69 alleles per primer pairs and an average PIC of 0.46 detected in 20 commercial cultivars of Iranian pistachio; and also higher than those reported by Baghizadeh *et al.* (2010) (an average of 2.75 alleles per primer pairs and an average of 0.44 for detected in 31 Iranian pistachio cultivars) and by Ahmad *et al.* (2005) (an average of 3.30 alleles per locus in 17 pistachio cultivars). Kolahi-Zonoozi (2014) assessed genetic diversity of 45 commercially Iranian cultivars using 12 nSSR markers and detected that PIC varied from 0.19–0.56 with an average of 0.33 and the mean of Ho and He were 0.49 and 0.35, respectively. Mirzaei *et al.* (2005) reported 80.00%polymorphism among 22 Iranian pistachio cultivars and wild pistachio species. In a study reported by Golan- Goldhirsh *et al.* (2004) in assessing polymorphisms among 28 Mediterranean pistachio accessions, 27 selected primers produced 259 total bands (an average of 9.59).

Some cultivars in different locations have the same name and some morphological identity, while molecular results showed differences between them. For instance, Badami-Zarand cultivar was differentiated from Badami- Kaj and Badami-Zoodras. Also, Ghazvini-Zodras showed differences with Ghazvini. These differentiations can be due to the intrinsic nature of nSSRs, since it is very unlikely that the microsatellites amplified correspond to the mutated DNA region when they have been randomly isolated from the whole genome. The results from this study showed that the studied cultivars had high genetic variation due to the species' dioeciously and cross-pollination nature (Ahmad *et al.* 2005).

CONCLUSION:

This study was aimed at evaluating the genetic diversity of Iranian pistachio in order to aid the conservation of its germplasm. The obtained information about the genetic variation between and within different populations will prepare the ground for the formulation of appropriate conservation strategies. The present analysis revealed that Iranian-cultivated pistachio germplasm is highly variable, presumably due to specific local genetic backgrounds, breeding pressure and/or limited interchange of genetic material. The unique nature of the Iranian pistachio germplasm revealed by our results, supports the case for the implementation of more intense characterization, conservation and breeding strategies. Also, the SCoT markers used were useful for determination of genetic diversity among pistachio cultivars in Iran.

ACKNOWLEDGMENT

The authors thank anonymous reviewers for valuable comments on an earlier draft.

REFERENCES

Arabnejad H, BaharM, Taj A. 2008. Genetic diversity of Pistacia khinjuk Stocks by SSRmarkers. Agri Sci Tech. 12:207–217.

Ahmad R, Ferguson L, Southwick SM 2003a. Identification of pistachio (Pistacia vera L.) nuts with microsatellite markers. J Am Soc Hortic Sci 128:898–903

Ahmad R, Struss D, Southwick SM 2003b. Development and characterization of microsatellite markers in citrus. J Am SocHorticSci 128:584–590

Ahmad R, Fergusen L. Southwick SM 2005. Molecular marker analysis of pistachio rootstocks by simple sequence sepeats and sequence-selated Amplified solymorphisms. Journal of Horticultural Science and Biotechnology, 80: 382-386.

Browiez K. 1988. Chorology of trees and shrubs in southwest Asia and adjacent regions. Polish Scientific Publication, Warsaw, Poland.

Baghizadeh A, Noroozi S, Jalali-Javaran M. 2010. Study on genetic diversity of some Iranian Pistachio (Pistacia vera L.) cultivars using random amplified polymorphic DNA (RAPD), inter sequence repeat (ISSR) and simple sequence repeat (SSR) markers: A comparative study. Afr J Biotech.9:7632–7640.

Behboodi B 2003. Ecological distribution study of wild pistachios for selection of rootstock. Options Mediterr Ser A 63:61–67.

Collard BCY, 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 27:86–93

Dollo L 1993. An isozyme study of Sicilian Pistacia species, varieties, and offspring from artificial pollination. IXth Groupe de Recherches et d'Etudes Mediterraneen pour le Pistachier et l'Amandier, Agrigento, pp 80-87.

Evanno G, S. Regnaut, J. Goudet 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14:2611-2620.

Esfandani-Bozchaloyi S, Sheidai M, Keshavarzi M, Noormohammadi Z. 2018a. Species Relationship and Population Structure Analysis In *Geranium* Subg. *Robertium* (Picard) Rouy With The Use of ISSR Molecular Markers. Act Bot Hung, 60(1–2), pp. 47–65.

Esfandani-Bozchaloyi S, Sheidai M, Keshavarzi M, Noormohammadi Z. 2018b. Species Identification and Population Structure Analysis In *Geranium* Subg. *Geranium* (Geraniaceae) . Hacquetia, 17/2 , 235–246 DOI: 10.1515/hacq-2018-0007

Esfandani -Bozchaloyi S, Sheidai M, Keshavarzi M, Noormohammadi Z. 2018c. Morphometric and ISSR-analysis of local populations of *Geranium molle* L. from the southern coast of the Caspian Sea. Cytology and genetics, 52, No. 4, pp. 309–321.

Esfandani -Bozchaloyi S, Sheidai M. 2018d. Molecular diversity and genetic relationships among *Geranium pusillum* and *G. pyrenaicum* with inter simple sequence repeat (ISSR) regions, Caryologia, vol 71, No. 4, pp. 1-14.https://doi.org/10.1080/00087114.2018.1503500

Esfandani-Bozchaloyi S, Sheidai M, 2019. Comparison of Dna Extraction Methods from *Geranium* (Geraniaceae), *Acta Botanica Hungarica* 61(3–4), pp. 251–266.

Elshibli S, Korpelainen H 2008. Microsatellite markers reveal high genetic diversity in date palm (Phoenix dactylifera L.) germplasm fromSudan. Genetica 134:251–260

Esmail-pour A 2001. Distribution, use and conservation of pistachio in Iran. In: Padulosi S, Hadj-Hassan A (eds) In towards a comprehensive documentation and use of Pistacia genetic diversity in central and West Asia, North Africa and Europe. Report of the IPGRI workshop, 14–17 December 1998, Ibrid, Jordan. IPGRI, Rome, Italy

Faostat. 2004. FAOSTAT database. FAO statistics database on The World Wide Web. http://apps.fao.org (accessed December 2004).

Falush D, M. Stephens JK. Pritchard 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol. Ecol. Notes. 7:574–578.

Fikirie, K., Bezu, A., Eshetu, M., Bekele, D., & Rabo, M. 2020. Evaluate Technical Standards of Implemented Soil Bund in Central Rift Valley of Ethiopia: The Case of Adama, Lume and Dodota Districts. Agriculture and Food Sciences Research, 7(1), 51–57.

Freeland JR, H. Kirk, S.D. Peterson 2011. Molecular Ecology, 2nd Ed. Wiley-Blackwell, Chichester, 464 pp.

Golan-Goldhirsh A, Barazani1 O, Wangl ZS, 2004. Genetic relationships among Mediterranean Pistacia species evaluated by RAPD and AFLP markers. Plant Syst Evol. 246:9–18.

Gondal, A.H., Farooq, Q., Sohail, S., Kumar, S.S., Toor, M.D., Zafar, A., Rehman, B. 2021. Adaptability of Soil pH through Innovative Microbial Approach. Current Research in Agricultural Sciences, 8(2), 71–79.

Hammer Ø, Harper Dat. Ryan PD. 2012. PAST: Paleontological Statistics software package for education and data analysis. Palaeontologia Electronica. 4: 1–9.

Hormaza JI, Dollo L, Polito VS. 1994. Identification of RAPD marker linked tosex determination in Pistacia vera using bulked segregant analysis. Theor Appl Genet. 89:9–13.

Huson DH, D. Bryant 2006. Application of Phylogenetic Networks in Evolutionary Studies. Mol. Biol. Evol. 23: 254–267.

Hopla, G.A., Sun, Y., Sun, C., Onautshu, O. 2021. Impact of the Aerobic Mesophilic Microorganisms on Black Sigatoka of Bananas According to the Cropping Systems in the Region of Kisangani (Case of the old secondary forest). Agriculture and Food Sciences Research, 8(1), 1–9.

Ibrahim Basha A, Padulosi S, Chabane K, Hadj-Hasan A, Dulloo E, Pagnotta AM, Porceddu E 2007. Genetic diversity of Syrian pistachio (*Pistacia vera* L.) varieties evaluated by AFLP markers. Genet. Res. Crop Evol. 54: 1807-1816.

Jost L. 2008. GST and its relatives do not measure differentiation. Mol. Ecol. 17: 4015–4026.

Kolahi-Zonoozi SH, Mardi M, Zeinalabedini M, 2014. Development of 12 new SSR markers for genetic diversity and structure analysis in pistachio (Pistacia vera L.) J Hort Sci Biotech. 89:707–711.

Kafkas S, Kafkas E, Perl-Treves R 2002. Morphological diversity and germplasm survey of three wild Pistacia species in Turkey. Genet Resour Crop Evol 49:261–270.

Kafkas S, Kaska A, Wassimi AN, Padulosi S 2006. Molecular characterisation of Afghan pistachio accessions by amplified fragment length polymorphisms (AFLPs). J Hortic Sci Biotechnol 81(5):864–868

Karimi HR, Zamani Z, Ebadi A, Fatahi MR 2009b. Morphological diversity of Pistacia species in Iran. Genet Resour Crop Evol 56:561–571

Katsiotis A, Hagidimitriou M, Drossou A, Pontikis C, Loukas M 2003. Genetic relationships among species and cultivars of Pistacia using RAPDs and AFLPs. Euphytica 132:279–286

Luo CXH, He H, Chen SJ, Ou MP, Gao JS, Brown CT, Tondo R, Schnell J 2011. Genetic diversity of mango cultivars estimated using SCoT and ISSR markers. Biochem Syst Ecol 39:676–684.

Lucena, A L. de M., Albuquerque, M.B. de, Alves, M.M., Araujo, R.S.R. de, Costa, C.R.G. da (2021). Cultivation and Nutritional Quality of Moringa Oleifera Lam. Produced Under Different Substrates in Semi-Arid Region in Northeast Brazil. Current Research in Agricultural Sciences, 8(1), 1–10.

Mehlenbacher S.A. 2003. Progress and prospects in nut breeding. Acta Horticulture, 622, 57–79.

Mirzaei S, Bahar M, Sharifnabi B 2005. A phylogenetic study of Iranian wild pistachio species and some cultivars using RAPD marker. Acta Hortic 726:39–43

Meirmans PG, Van Tienderen PH. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Mol. Ecol. Notes. 4:792–794.

Meirmans PG. 2012. AMOVA-based clustering of population genetic data. J. Heredity. 103: 744–750.

Minn Y, Gailing O, Finkeldey R. 2015. Genetic diversity and structure of teak (Tectona grandis L. f.) anddahat (Tectona hamiltoniana Wall.) based on chloroplast microsatellites and amplified fragment length polymorphism markers. *Genetic Resources and Crop Evolution*, 63: 961–974.

Nwankiti, A.O., Ogbonna, I.O., Eche, C.O. (2021). Bacteriological Quality and Cyanide Contents of Different Cassava Products Processed in Benue State for Use as Food for Man or Feedstock for Animals. Current Research in Agricultural Sciences, 8(1), 47–55.

Peakall R, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes. 6:288–295.

Podani J. 2000. Introduction to the Exploration of Multivariate Data. Backhuyes, Leiden, 407 pp.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus enotype Data. Genetics. 155:945–959.

Paul, S., Ara, R., Ahmad, M.R., Hajong, P., Paul, G., Kobir, M.S., Rahman, M.H. (2021). Effect of Blanching Time and Drying Method on Quality of Black Pepper (Piper nigrum). Journal of Food Technology Research, 8(1), 18–25.

Rhouma S, Mohamed Salem AO, Zehdi-Azouzi S, Chatti K.Rhouma A, Marrakchi M, Trifi M 2009. Comparative analysis of genetic diversity in Tunisian date-palm (Phoenix dactylifera L.) as revealed by RAPDs and AFLPs. In: Socias i Company R, Espiau MT, Alonso JM (eds) XII EUCARPIA symposium on fruit breeding and genetics, ISHS, Acta Horticulturae, pp 814.

Sawadogo, Y.A., Cisse, H., Oumarou, Z., Nikiema, F., Traore, Y., Savadogo, A. (2021). Reduction of Aflatoxins and Microorganisms in the Koura-Koura Produced in Burkina Faso with Spices and Aromatic Leaves. Journal of Food Technology Research, 8(1), 9–17.

TayefehAliakbarkhany S, Talaie AR, Fatahi Moghadam MR 2013. Investigation of genetic diversity among Pistacia Vera In the Khorasan by using molecular. Mod Genet 8:169–176.

Whitehouse WE. 1957. The pistachio nut. A new crop for the Western United States. Economical. Botany, 11,281-321.

Williams JGK., Kubelik AR., Levak KJ., Rafalski JA. and. Tingey, S.V.1990. DNA polymorphism amplification by arbitary primers is useful as genetics markers. Nucleic Acids Research, 18, 6531-6535.

Weising K, H., Nybom K., Wolff G., Kahl 2005. DNA Fingerprinting in Plants.

Principles, Methods, and Applications. (2nd ed.), Boca Raton, FL., USA: CRC Press, pp. 472.

Yeh Francis, C., R.C. Yang, B.J. Boyle Timothy, Z.H. Ye, X. Mao Judy, 1999. POPGENE

Version 1.32, the User-Friendly Shareware for Population Genetic Analysis, Molecular Biology and Biotechnology Centre, University of Alberta, Canada,

Wu JM, Li YR, Yang LT, Fang FX, Song HZ, Tang HQ, Wang M, Weng ML 2013. cDNA-SCoT: a novel rapid method for analysis of gene differential expression in sugarcane and other plants. AJCS 7:659–664.

Zohary M 1952. A monographic study of the genus Pistacia. Palest J Bot Jerus Ser 5:187–228.