Acta Botanica Hungarica 64(3–4), pp. 221–236, 2022
DOI: 10.1556/034.64.2022.3-4.1

STUDY OF EFFICIENCY OF CP-DNA GENE, IRAP, AND REMAP MARKERS FOR STUDY OF GENETIC DIVERSITY OF SOME IRANIAN WALNUT POPULATIONS

H. Aghaei[1], S. M. M. Hamdi[1], M. Sheidai[2] and F. Koohdar[2]

[1]*Department of Biology, Faculty of Biological Sciences, North-Tehran Branch, Islamic Azad University, Tehran, Iran, E-mail: hengameh.aghaei@gmail.com (corresponding author)*
[2]*Department of Plant Sciences and Biotechnology, Faculty of Life Sciences & Biotechnology, Shahid Beheshti University, Tehran, Iran, E-mails: f_koohdar@yahoo.com, msheidai@yahoo.com*

(Received: 25 January 2020; Accepted 20 December 2021)

The Persian walnut, also known *Juglans regia* of the genus *Juglans* is cultivated throughout the temperate regions of the world for its high-quality wood and edible nuts. Genetic diversity, structure and differentiation of cultivated walnut are important for effective conservation, management, and utilisation of germplasm. Recent study on genetic diversity and genotype differentiation in Persian walnut of Iran, revealed that these genotypes can be differentiated by ITS and ISSR, however, these markers show a low degree of genetic variability. ITS sequences revealed a lower degree of genetic difference of the studied Persian walnut genotypes compared to that of ISSR molecular markers. Therefore, it is important to investigate these genotypes by the other molecular markers to find out which one can produce more data on genetic structure and variability in these important genotypes. For the same reason, we continued this study by using cp-DNA (psbA-trnhG) gene, IRAP, and REMAP. To provide barcode for Iranian genotypes of *Juglans regia* is an other objective of this study. For IRAP and REMAP analyses studies, we randomly selected 60 plants from 6 populations including 3 wild and 3 cultivated populations. For cp- DNA analysis we used 21 plants of *Juglans regia* randomly selected from 6 studied populations. The present study revealed a high level of genetic variability in *Juglans regia* genotypes in those sequences investigated by IRAP and REMAP molecular markers. We reported that IRAP and REMAP molecular markers cannot be efficiently used in walnut germplasm genetic screening.

Key word: barcode, cp-DNA, IRAP, *Juglans regia*, REMAP

INTRODUCTION

The Persian walnut, also known *Juglans regia* L. of the genus *Juglans* Linnaeus (1753: 997) (Juglandaceae) is cultivated throughout the temperate regions of the world for its high-quality wood and edible nuts. Its native range extends from Iran to Turkey, India, China and Bhutan (Zohary and Hopf 1993).

Knowledge of genetic diversity, structure and differentiation of cultivated walnut is important for effective conservation, management, and utilization of germplasm (Aradhya *et al.* 2017). Molecular markers used in genetic diversity analysis and molecular phylogeny of walnut include isozymes, restriction fragment-length polymorphism (RFLP), randomly amplified polymorphic

DNA (RAPD), inter-simple sequence repeat (ISSR), amplified fragment length polymorphism (AFLP) markers, and simple sequence repeats (SSRs) (See e.g., Bayazit *et al.* 2007, Fjellstrom *et al.* 1994, Maghsoodi *et al.* 2018, Mohsenipoor *et al.* 2010, Nematollahi *et al.* 2009, Nicese *et al.* 1998, Ninot and Aletà 2003, Pop *et al.* 2013, Ruiz-Garcia *et al.* 2011, Sadat-Hosseini *et al.* 2019, Solar *et al.* 1994).

Due to continuous cultivation and human based selection, walnut genotypes may face genetic erosion. Therefore, searching for new genetic variability within available germ plasm is an important task. Genetic analysis of germplasm collections often provides insights into the complex interactions of evolutionary forces such as mutation, gene flow, selection, and drift shaping the Eco geographic structure and domestication history of a crop species (Aradhya *et al.* 2010, Sheidai *et al.* 2016).

Walnut is an out crossing species and forms many local populations. The genetic integrity of walnut species is therefore maintained by complex interactions of evolutionary forces both spatially and temporally within the range of the species. For producing a detailed picture on genetic structure of walnut in a country, we should have comprehensive germplasm collections to investigate the amount and pattern of distribution of genetic variation and estimate the role of evolutionary forces that shape the overall genetic structure of a species (Arab *et al.* 2019, Aradhya *et al.* 2010, Mohsenipoor *et al.* 2010, Nematollahi *et al.* 2009, Sheidai *et al.* 2016).

Recent study on genetic diversity and genotype differentiation in Persian walnut of Iran (Maghsoodi *et al.* 2018), revealed that these genotypes can be differentiated by ITS (internal transcribed spacer DNA) and ISSR (inter simple sequence repeats), however, these markers show a low degree of genetic variability. ITS sequences revealed a lower degree of genetic difference of the studied Persian walnut genotypes compared to that of ISSR molecular markers. In the other study, Pop *et al.* (2013) revealed that SSR and RAPD molecular marker systems are useful for genotype characterization, but SSR markers are more advisable to investigate genetic relationships, as well as Karimi *et al.* (2014) show that morphological characteristics and SSR marker were used to analyse the genetic diversity and relationships among natural populations of Persian walnut in northern and western regions of Iran. The results showed that there was a high level of genetic diversity among the walnuts, both in terms of their SSRs loci as well as morphological traits. Therefore, it is important to investigate these genotypes by the other molecular markers to find out which one can produce more data on genetic structure and variability in these important genotypes. For the same reason, we continued this study by using cp-DNA (psbA-trnhG) gene, IRAP (inter retrotransposons amplified polymorphism), and REMAP (retrotransposon-microsatellite amplified polymorphism).

Bar coding is also important for economically important crop plants. A DNA barcode is a universally accepted short DNA sequence allowing the

prompt and unambiguous identification of species. DNA barcodes have been used in the identification of medicinal plants, and differentiating true and original products from that of adulterant species/ DNA (Laiou *et al.* 2013). Based on the DNA sequencing, multiple alignment and the amount of variation displayed (sufficient to discriminate among sister species without affecting their correct assignation through intra-specific variation), three plastid loci are currently used in plants: rbcL (a universal but slowly evolving coding region), matK (a relatively fast evolving coding region) and trnH-psbA (a rapidly evolving intergenic spacer) (CBOL Plant Working Group 2009). In addition, the nuclear ribosomal internal transcribed spacer (ITS) has also been used as an efficient barcoding locus for many complex plant groups (Hollingsworth *et al.* 2011).

DNA barcoding and the original definition of a reliable DNA identification of species requires (i) recovery of a barcode sequence from the sample, (ii) representation of relevant species in the reference database, and (iii) sufficient nucleotide sequence variability to distinguish among closely related species (Fazekas *et al.* 2008).

A precise and accurate genetic conservation program needs an extensive and detailed population genetic investigation. It gives insight on the genetic structure, genetic diversity, and gene flow versus genetic fragmentation of local populations (Sheidai *et al.* 2013, 2016).

Therefore, the aims of present study are: 1). to investigate genetic diversity in Persian walnut genotypes by IRAP, REMAP, and psbA sequences of cp-DNA; and 2). comparing genetic distance as a measure of genetic difference among walnut genotypes obtained by different molecular markers; and 3). to provide barcode for Iranian genotypes of *Juglans regia*.

Both IRAP and REMAP molecular markers are highly reproducible and polymorphic molecular markers and are very efficient for genetic diversity studies and genetic finger printing in plants (Kalendar and Schulman 2006). These molecular markers require neither restriction enzyme digestion nor ligation to generate the marker bands. The IRAP products are generated from two nearby retrotransposons using outward-facing primers. In REMAP, amplification between retrotransposons proximal to simple sequence repeats (microsatellites) produces the marker bands. Efficient molecular barcodes have been produced from the plastid genome in plants (Fazekas *et al.* 2008).

MATERIAL AND METHODS

Plant materials – For IRAP and REMAP analyses studies, we randomly selected 60 plants from 6 populations including 3 wild and 3 genotypes (10 leaf per tree). For cp-DNA analysis we used 21 plants of *Juglans regia* randomly selected from 6 studied populations (Table 1).

Table 1
Juglans regia population in IRAP, REMAP and cp-DNA studies

No	Type	Population	Province	Locality	Latitude (N)	Longitude (E)	Altitude (m)
1	Wild	Nahavand	Hamadan	Nahavand	48° 23′	48° 23′	1,627
2	Cultivated	Soozani	Markazi	Tafresh	34° 42′	49° 59′	1,838
3	Cultivated	Basloghi	Markazi	Tafresh	34° 42′	49° 59′	1,837
4	Wild	Astara	Gilan	Astara	38° 20′	48° 52′	–26
5	Cultivated	Kaghazi	Markazi	Tafresh	34° 42′	49° 59′	1,838
6	Wild	Khoy	W Azarbaijan	Khoy	38° 33′	44° 57′	1,135

DNA extraction and PCR details – For molecular studies, the fresh leaves were randomly collected from 70 plants in the studied area and were dried in silica gel powder. The genomic DNA was extracted using CTAB-activated charcoal protocol (Križman *et al.* 2006). The extraction procedure was based on activated charcoal and polyvinylpyrrolidone (PVP) for binding of polyphenolics during extraction and under mild extraction and precipitation conditions. This promoted high-molecular-weight DNA isolation without interfering contaminants. Quality of extracted DNA was examined by running on 0.8% agarose gel. PCR reactions were performed in a 25 μL volume containing 10 mM Tris-HCl buffer at pH 8, 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP (Bioron, Germany), 0.2 μM of a single primer, 20 ng of genomic DNA, and 3 U of Taq DNA polymerase (Bioron).

IRAP assay – 3 IRAP primer combinations including ((3′LTR, 5′LTR2), (5′LTR1, 5′LTR2), (NIKTA, 3′LTR)) were used on 60 plants from 6 populations. Amplification reactions were performed in a Techne thermocycler (Germany) with the following program: 5 min for initial denaturation step at 94 °C, 30 s at 94 °C, 1 min at 55 °C, and 1 min at 72 °C. The reaction was completed by a final extension step of 7 min at 72 °C. The amplification products were visualized by running on 2% agarose gel, followed by ethidium bromide staining. The fragments size was estimated by using a 100-bp molecular size ladder (Fermentas, Germany).

REMAP assay – 12 REMAP primer combinations, derived from 4 single IRAP primers (3′LTR, NIKITA, 5′LTR1, 5′LTR2) with 3 ISSR primers ((CA)7GT, (GA)9T, (GT)7TG) were tested on 60 plants from 6 populations (Table 2).The amplifications, reactions were performed in Techne thermocycler (Germany) with the following program: 5 min initial denaturation step at 94 °C, followed by five cycles of 94 °C for 1 min, 35 °C for 45 s, and 72 °C for 1 min; followed by 40 cycles of 94 °C for 1 min, 56 °C for 45 s, and 72 for 1 min; followed by 7 min at 72 °C. 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).

Table 2
IRAP bands frequency in *Juglans regia* populations studied

Population	Pop1	Pop2	Pop3	Pop4	Pop5	Pop6
No. Bands	25	29	30	27	33	28
No. Bands Freq. (≥ 5%)	25	29	30	27	33	28
No. Private Bands	0	1	0	0	4	0
No. LComm Bands (≤ 25%)	0	0	0	0	0	0
No. LComm Bands (≤ 50%)	1	2	3	2	2	2

cp-DNA sequences analyses – The intergenic spacer of chloroplast genome psbA-trnHGUG was amplified and sequenced with universal primers following the methodology of Shaw and Small (2005) and Timmer *et al.* (2007). The psbA-trnHGUG forward primer was (trnHGUG) 5′-CGCGCATGGTGGATTCACAATCC-3′ and, the reverse primer was (psbA) 5′-GTTATGCATGAACGTAATGCTC-3′. Each 20 ml PCR mixture contained 10 ml of 2_ PCR buffer, 0.5 mM of each primer, 200 mM of each dNTP, 1 Unit of Taq DNA polymerase (Bioron, Ludwigschafen, Germany), and 1 ml of template genomic DNA at 20 ng mle1. The PCR amplification program was 5 min at 95 °C, followed by 35 cycles of 60 s at 95 °C, 60 s at 56 °C, and 60 s at 72 °C, with a final extension step of 5 min at 72 °C.

Data analyses: Molecular analysis – The IRAP and REMAP bands obtained were treated as binary characters and coded accordingly (presence = 1, absence = 0). The number of private bands versus common bands was determined. Genetic diversity parameters like: The percentage of allelic polymorphism, allele diversity (Weising *et al.* 2005), Nei' gene diversity (He), and Shannon information index (I) (Weising *et al.* 2005), were determined. We used GenAlex 6.4 for these analyses (Peakall & Smouse 2006). The Mantel test (Podani 2000) between IRAP and REMAP markers and the Nei genetic distance (Weising *et al.* 2005) was determined among the studied populations and was used for the grouping of the genotypes in GenAlex 6.4 (Peakall and Smouse 2006). Discriminating power of IRAP markers investigated by Gst analysis as implemented in POPGENE32.

Canonical discriminate analysis (CDA) was used for grouping of the population based on IRAP and REMAP markers (Podani 2000). Data analyses were performed by using PAST ver. 2.17 (Hammer *et al.* 2012).

For psbA sequences analysis, we used sequences obtained for Iranian Persian walnut samples. The sequences were aligned by MUSCLE program as implemented in MEGA 7 (Tamura *et al.* 2012). Model test was performed as implemented in MEGA, which revealed Kimura 2-parameters model as the best fit to our data. TCS network of cp-DNA sequences trees were implemented in PopART (Population Analysis with Reticulate Trees) program (http://popart.otago.ac.nz).

RESULTS

IRAP assay

IRAP primers produced 37 reproducible bands. Details of IRAP bands obtained in six studied populations are provided in Table 2. The highest number of IRAP bands (33) was observed in population 5 (Thin shelled genotypes), while population 1 (Nahavand), had the lowest number of IRAP bands (25). Only 5 private IRAP bands were obtained and the other bands were common among the studied populations (Table 2). The private bands can discriminate populations and may be related to the local populations adaptation.

Genetic diversity parameters determined in the studied populations based on IRAP data (Table 3), revealed a good level of genetic polymorphism ranging from 48% in population 6 (Khooy) to 89% in population 5 (Thin shelled genotypes). The lowest value (0.18) for Nei gene diversity occurred in population 6, while the highest value (0.30) occurred in population 3 (Basloghi genotypes).

Discriminating power of IRAP markers investigated by Gst analysis as implemented in POPGENE, produced the mean Gst value = 0.20 (Table 4), which is very low. Therefore, these markers should be used in combination with the other kinds of molecular markers for the purpose of Persian walnut genotypes differentiation. This is also supported by AMOVA which revealed non-significant difference among walnut populations based on IRAP data) PhiPT = 0.07, P = 0.12). However, as stated before, IRAP molecular markers are efficient in genetic diversity studies. All the studied IRAP loci, had Nm value >1, (Table 4) therefore, frequent gene flow has occurred in the studied populations. This may be due to out-crossing nature of *Juglans regia*.

Table 3
Genetic diversity parameters determined in Persian walnut populations studied based on IRAP marker

Pop	Na	Ne	I	He	uHe	%P
Pop1	1.351	1.532	0.414	0.289	0.315	67.57%
Pop2	1.568	1.423	0.407	0.267	0.288	78.38%
Pop3	1.541	1.580	0.444	0.309	0.344	72.97%
Pop4	1.351	1.468	0.371	0.257	0.285	62.16%
Pop5	1.784	1.469	0.447	0.292	0.311	89.19%
Pop6	1.243	1.346	0.275	0.189	0.210	48.65%

Abbreviations: Na = mean No. of alleles, Ne = mean effective allele, I = Shannon information index, He = Nei gene diversity, uHe = unbiassed Nei gene diversity, %P = polymorphism percentage

Table 4
Discriminating power of IRAP loci in *Juglans regia* populations by Gst analysis

Locus	Sample size	Ht	Hs	Gst	Nm*
Locus1	36	0.4912	0.4534	0.0770	5.9912
Locus2	36	0.1390	0.1164	0.1625	2.5774
Locus3	36	0.0213	0.0201	0.0544	8.6900
Locus4	36	0.1928	0.1783	0.0752	6.1490
Locus5	36	0.4892	0.3902	0.2024	1.9704
Locus6	36	0.4892	0.3902	0.2024	1.9704
Locus7	36	0.4989	0.3787	0.2409	1.5753
Locus8	36	0.4892	0.3902	0.2024	1.9704
Locus9	36	0.3407	0.2681	0.2130	1.8477
Locus10	36	0.1658	0.1319	0.2047	1.9425
Locus11	36	0.0674	0.0552	0.1808	2.2649
Locus12	36	0.0674	0.0552	0.1808	2.2649
Locus13	36	0.0995	0.0867	0.1289	3.3793
Locus14	36	0.0929	0.0690	0.2566	1.4485
Locus15	36	0.4871	0.4395	0.0977	4.6166
Locus16	36	0.4482	0.3999	0.1077	4.1406
Locus17	36	0.4971	0.3674	0.2609	1.4168
Locus18	36	0.4971	0.3674	0.2609	1.4168
Locus19	36	0.4265	0.3875	0.0914	4.9729
Locus20	36	0.2737	0.2541	0.0714	6.5048
Locus21	36	0.0503	0.0436	0.1325	3.2748
Locus22	36	0.1899	0.1799	0.0526	9.0124
Locus23	36	0.0945	0.0895	0.0531	8.9139
Locus24	36	0.0452	0.0430	0.0477	9.9844
Locus25	36	0.1403	0.1375	0.0203	24.1628
Locus26	36	0.4999	0.3571	0.2857	1.2501
Locus27	36	0.4999	0.3571	0.2857	1.2501
Locus28	36	0.4999	0.3571	0.2857	1.2501
Locus29	36	0.4999	0.3571	0.2857	1.2501
Locus30	36	0.4999	0.3571	0.2857	1.2501
Locus31	36	0.4999	0.3571	0.2857	1.2501
Locus32	36	0.4999	0.3571	0.2857	1.2501
Locus33	36	0.4999	0.3571	0.2857	1.2501

Table 4 (continued)

Locus	Sample size	Ht	Hs	Gst	Nm*
Locus34	36	0.4999	0.3571	0.2857	1.2501
Locus35	36	0.4999	0.3571	0.2857	1.2501
Locus36	36	0.4612	0.4354	0.0559	8.4475
Locus37	36	0.1970	0.1796	0.0881	5.1723
Mean	36	0.3365	0.2670	0.2066	1.9200
St. Dev		0.0349		0.0200	

REMAP assay

REMAP primers produced 48 reproducible bands, details of which are provided in Table 4. The highest number of REMAP bands (40) was observed in population 3, while population 5, had the lowest number of bands (26). Only 5 private REMAP bands were obtained and the other bands were common among the studied populations (Table 5). Genetic diversity parameters determined based on REMAP data (Table 6) revealed that these populations have moderate to high genetic polymorphism, ranging from 35% in population 5, to 77% in population 3.

Similarly, the lowest value for Nei gene diversity, He (0.11) occurred in population 5, while population 1 and 3 had the highest value (0.25), for the same parameter. Discriminating power analysis of REMAP data also indicated that these loci have low value and cannot be used alone in walnut differentiating studies. The mean value of Gst was 0.27, while the mean Nm value was 1.28 (Table 7).

Mantel test determined between IRAP and REMAP markers produced significant association ($r = 0.50$. $P = 0.001$). This indicates that both molecular markers have similar efficiency in showing genetic variability and population relationship. Therefore, we performed canonical discriminate analysis (CDA) for grouping and relationship study of the studied *Juglans regia* populations, based on combined IRAP-REMAP data analysis. The first two CDA axes com-

Table 5
REMAP bands frequency in *Juglans regia* populations studied

Population	Pop1	Pop2	Pop3	Pop4	Pop5	Pop6
No. Bands	35	36	40	34	26	37
No. Bands Freq. (≥ 5%)	35	36	40	34	26	37
No. Private Bands	3	1	0	0	0	1
No. LComm Bands (≤ 25%)	0	0	0	0	0	0
No. LComm Bands (≤ 50%)	3	3	8	3	1	4

Table 6
Genetic diversity parameters determined in *Juglans regia* populations based on REMAP marker

Pop	Na	Ne	I	He	uHe	%P
Pop1	1.438	1.430	0.375	0.250	0.273	70.83%
Pop2	1.375	1.319	0.310	0.201	0.217	62.50%
Pop3	1.604	1.433	0.387	0.255	0.283	77.08%
Pop4	1.292	1.253	0.263	0.166	0.184	58.33%
Pop5	0.896	1.183	0.169	0.111	0.118	35.42%
Pop6	1.313	1.300	0.278	0.182	0.203	54.17%

Abbreviations: Na = mean No. of alleles, Ne = mean effective allele, I = Shannon information index, He = Nei gene diversity, uHe = unbiassed Nei gene diversity, %P = polymorphism percentage

prised about 88% of total variability. Grouping of the studied populations based on these two DA axes is presented in Figure 1. The CDA plot separated the populations 6 and 3 in distinct groups. Similarly, most of the members in populations 1 and 4 were placed close together and were separated in a distinct group. However, one plant from either population was placed

Table 7
Discrimination power of REMAP loci in *Juglans regia* populations

Locus	Sample size	Ht	Hs	Gst	Nm*
Locus1	36	0.4872	0.2731	0.4395	0.6377
Locus2	36	0.4760	0.2970	0.3759	0.8300
Locus3	36	0.4837	0.4302	0.1106	4.0209
Locus4	36	0.4914	0.3331	0.3222	1.0518
Locus5	36	0.1390	0.1068	0.2314	1.6603
Locus6	36	0.1777	0.1648	0.0728	6.3728
Locus7	36	0.1402	0.1245	0.1125	3.9446
Locus8	36	0.0780	0.0615	0.2120	1.8583
Locus9	36	0.1390	0.1068	0.2314	1.6603
Locus10	36	0.1777	0.1648	0.0728	6.3728
Locus11	36	0.0346	0.0315	0.0896	5.0833
Locus12	36	0.4985	0.3411	0.3158	1.0834
Locus13	36	0.1164	0.1124	0.0351	13.7411
Locus14	36	0.3441	0.2567	0.2540	1.4682
Locus15	36	0.4187	0.3790	0.0950	4.7635
Locus16	36	0.4012	0.3726	0.0714	6.5009

Table 7 (continued)

Locus	Sample size	Ht	Hs	Gst	Nm*
Locus17	36	0.3299	0.3035	0.0799	5.7593
Locus18	36	0.3048	0.2422	0.2054	1.9338
Locus19	36	0.1396	0.1331	0.0469	10.1644
Locus20	36	0.0286	0.0265	0.0737	6.2863
Locus21	36	0.1689	0.1646	0.0256	19.0136
Locus22	36	0.0679	0.0630	0.0729	6.3541
Locus23	36	0.3325	0.1323	0.6020	0.3306
Locus24	36	0.4985	0.2664	0.4656	0.5739
Locus25	36	0.0286	0.0265	0.0737	6.2863
Locus26	36	0.4985	0.2664	0.4656	0.5739
Locus27	36	0.4985	0.2664	0.4656	0.5739
Locus28	36	0.3325	0.1323	0.6020	0.3306
Locus29	36	0.3325	0.1323	0.6020	0.3306
Locus30	36	0.3325	0.1323	0.6020	0.3306
Locus31	36	0.3325	0.1323	0.6020	0.3306
Locus32	36	0.3506	0.1089	0.6893	0.2253
Locus33	36	0.2224	0.2094	0.0581	8.0986
Locus34	36	0.2562	0.1959	0.2354	1.6243
Locus35	36	0.1797	0.1421	0.2093	1.8894
Locus36	36	0.1695	0.1559	0.0798	5.7618
Locus37	36	0.4170	0.4058	0.0270	17.9899
Locus38	36	0.3934	0.3351	0.1484	2.8693
Locus39	36	0.3090	0.2154	0.3029	1.1504
Locus40	36	0.2355	0.1912	0.1885	2.1530
Locus41	36	0.3001	0.2640	0.1204	3.6518
Locus42	36	0.2928	0.2807	0.0414	11.5854
Locus43	36	0.1349	0.1212	0.1021	4.3964
Locus44	36	0.3776	0.3316	0.1219	3.6012
Locus45	36	0.0929	0.0690	0.2566	1.4485
Locus46	36	0.0622	0.0580	0.0673	6.9319
Locus47	36	0.1863	0.1587	0.1482	2.8728
Locus48	36	0.1240	0.1005	0.1894	2.1395
Mean	36	0.2695	0.1942	0.2795	1.2890
St. Dev		0.0220		0.0115	

* Nm = estimate of gene flow from Gst or Gcs. E.g., Nm = 0.5(1–Gst)/Gst

Table 8
Classification result[a] of CDA, showing genetic admixture among *Juglans regia* populations

Pop	Predicted group membership						Total
	1	2	3	4	5	6	
1	5	1	0	0	0	0	6
2	0	6	0	0	1	0	7
3	0	0	5	0	0	0	5
4	1	0	0	4	0	0	5
5	0	1	0	0	7	0	8
6	0	0	0	0	0	5	5
1	83.3	16.7	0	0	0	0	100.0
2	0	85.7	0	0	14.3	0	100.0
3	0	0	100.0	0	0	0	100.0
4	20.0	0	0	80.0	0	0	100.0
5	0	12.5	0	0	87.5	0	100.0
6	0	0	0	0	0	100.0	100.0

[a] = 88.9% of original grouped cases correctly classified

inter-mixed with the other population. The plants in populations 2 and 5 were placed inter-mixed with no clear separation. These results are in agreement with Gst analysis of IRAP-REMAP molecular markers, which indicated that these markers cannot be used alone in *Juglans regia* genotypes differentiation.

Population assignment of the studied plants was investigated by CDA. The classification result (Table 8) revealed that all members in populations 3 and 6 are correctly classified, but this table shows that the other populations have some degree of genetic admixture, for example, in population 1, one plant is misclassified and it is more like population 4. Similarly, in population 2, one plant is more like population 1, while another plant is similar to population 5.

Cp-DNA analysis

Sequencing of psbA, followed by alignment and curation, produced 414 nucleotides for further analyses. The Kimura genetic distance among the studied *Juglans regia* genotypes varied from 0.00–0.08. The highest degree of genetic distance occurred between Kcm1 and the other genotypes. Similarly, genotypes Kcm2 showed 0.12 genetic distances with the rest of *Juglans regia* genotypes studied.

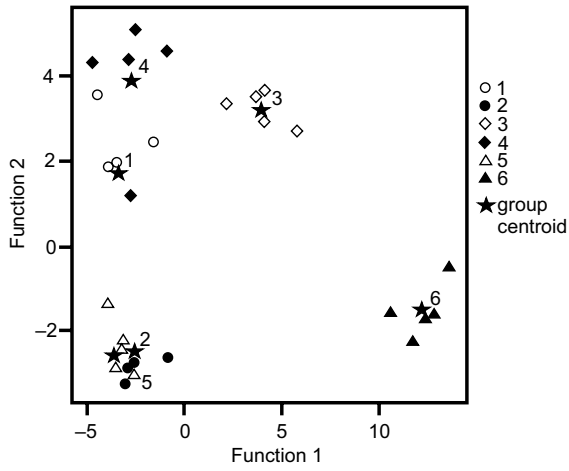

Fig. 1. CDA plot of the *Juglans regia* populations studied based on combined IRAP-REMAP data

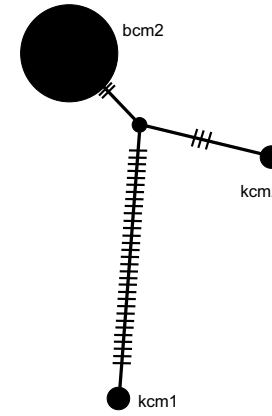

Fig. 2. TCS network of *Juglans regia* genotypes based on PsbA sequences, revealing sequence uniformity in most of the genotypes studied

TCS network of cp-DNA sequences (Fig. 2) also differentiated Kcm1 and Kcm2 from the rest of *Juglans regia* genotypes studied. The other genotypes had similar sequences and were grouped together.

The studied genotypes differed in 35 nucleotides out of 414 nucleotides (0.08) (Fig. 3), which indicates very stable sequences in psbA region. This uniformity in cp-DNA sequences can be used in barcoding of *Juglans regia* (Fig. 3).

Species/Abbrev.	*											*			*					*				*			*	*	*	*			
1. KY574026.1_J_regia_isolate	T	A	G	A	C	T	C	A	T	G	G	A	A	G	C	T	G	T	T	C	T	A	A	A	A	A	A	T	G	G	A	G	T	G
2. kcm1	T	A	A	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	T	G	G	A	C	C	T
3. kcm2	T	C	C	C	T	A	T	T	C	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
4. bcm2	T	A	A	A	T	A	T	T	C	A	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
5. scm1	T	A	A	A	T	A	T	T	C	A	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
6. wa1	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
7. bcm4	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
8. bcm5	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
9. wh2	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
10. bcm1	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
11. wg3	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
12. wh3	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
13. wh4	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
14. wh1	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
15. wh4	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
16. wa4	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
17. bcm3	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
18. wa2	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
19. AY231167.1_J_regia	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
20. wg2	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
21. wg1	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
22. scm2	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
23. wa3	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T
24. scm3	T	A	A	A	T	A	T	T	C	A	A	A	A	T	C	T	T	A	T	C	T	A	G	A	T	C	C	T	G	G	A	A	T	T

Fig. 3. Part of PsbA sequences showing variable versus uniform nucleotides in *Juglans regia* genotypes studied

DISCUSSION

Juglans regia is a member of the section *Juglans*, which evolutionary history is associated with widespread extinctions, geographic isolations, and bottlenecks during the Quaternary glaciations. Its subsequent expansion and human selection have greatly influenced the genetic structure of the *Juglans regia* (Beug 1975, Popov 1929).

The present study revealed a high level of genetic variability in *Juglans regia* genotypes in those sequences investigated by IRAP and REMAP molecular markers. In a similar study, Aradhya *et al.* (2010), reported little genetic differentiation among the 5 *Juglans regia* genotype groups and noticed that 90% of the total genetic variation resides within the studied groups.

Wang *et al.* (2008) also reported a moderate level of genetic diversity in five populations from *Juglans regia* in central and southwestern China, with the number of effect alleles per locus ranging from 1.75 to 3.35 (average 2.39) and the proportion of polymorphic loci (P) equalling 100.0%. The proportion of genetic variation due to among populations difference accounted for 18.6% of the total genetic diversity.

The present study indicated frequent occurrence of gene flow and genetic admixture among the study *Juglans regia* genotypes. Wang *et al.* (2008) also reported a high degree of gene flow (Nm) (1.10) among populations. Ehteshamnia *et al.* (2009) investigated genetic variability in five populations of Iran walnut (*Juglans regia* L.) by using SSR markers. The number of effective alleles obtained was between 1.57 and 5.32. The mean number of effective alleles for each locus was 3.77.

We reported here that combined IRAP and REMAP molecular markers can be efficiently used in walnut germplasm genetic screening. In a similar study, Ahmed *et al.* (2012) used a combination of RAPD and SSR molecular markers to investigate genetic diversity in 82 walnut genotypes adapted to the Northwestern Himalayan region of Jammu and Kashmir, India. They reported a high level of genetic diversity within populations with the number of alleles per locus ranging from one to five in case of SSR primers and two to six in case of RAPD primers. They illustrated that the combined SSR and RAPD markers are efficient in discriminating the walnut genotypes.

Aradhya *et al.* (2017) suggested close association between SW Asian, Caucasus, and Eastern European walnut groups, while Central Asian and East Asian walnut are somewhat separate groups. They concluded that, walnut probably expanded from SW Asia into other regions following glaciations, which latter on were subjected to either local genetic differentiation after recolonization or separate expansion events from different refugia. Furthermore, human mediated dispersal and local domestication events since Greek and Roman times perhaps significantly influenced the current distribution of

genetic diversity in walnut. These authors believe that Central Asian mountain ranges such as the Pamir, Kopet Dagh, and Tien Shan are important centres of biological diversity and believed to be a centre of origin and diversity of walnut. Similarly, the Kopet Dagh riparian forests along the southern and southwestern shores of the Caspian Sea, is the place where walnut has been subjected to over harvesting and intense grazing and formed sparse isolated populations (Popov 1994). Walnut was introduced to China from Iran, Tibet, and Kashmir region of India. It is suggested that historical migration of walnut from this region occurred through early trade along the ancient silk route connecting these two regions (Aradhya *et al.* 2017). Barcoding of economically important crops is of immediate importance as it can help for genotype discrimination and any sort of adulteration attempts. We here reported a high level of genotype uniformity in PsbA sequences. Previously Maghsoodi *et al.* (2018) reported high degree of uniformity in ITS sequences among *Juglans regia* genotypes. Therefore, combined ITS-PsbA sequences can be ideally used as a barcode for *Juglans regia* genotypes.

REFERENCES

Ahmed, N., Mir, J. I., Mir, R. R., Rather, N. A., Rashid, R., Wani, S. H., Shafi, W., Mir, H. and Sheikh, M. A. (2012): SSR and RAPD analysis of genetic diversity in walnut (Juglans regia L.) genotypes from Jammu and Kashmir, India. – *Physiol. Mol. Biol. Plants* **18**(2): 149–160 (2012). https://doi.org/10.1007/s12298-012-0104-z

Arab, M. M., Marrano, A., Abdollahi-Arpanahi, R., Leslie, C. A., Askari, H., Neale, D. B. and Vahdati, K. (2019): Genome-wide patterns of population structure and association mapping of nut-related traits in Persian walnut populations from Iran using the Axiom J. regia 700K SNP array. – *Sci. Rep.* **9**(1): 6376. https://doi.org/10.1038/s41598-019-42940-1

Aradhya, M., Woeste, K. and Velasco, D. (2010): Genetic diversity, structure and differentiation in cultivated walnut (Juglans regia L). – *Acta. Hort.* **861**: 127–132. https://doi.org/10.17660/ActaHortic.2010.861.16

Aradhya, M., Velasco, D., Ibrahimov, Z., Toktoraliev, B., Maghradze, D., Musayev, M., Bobokashvili, Z. and Preece, J. E. (2017): Genetic and ecological insights into glacial refugia of walnut (Juglans regia L.). – *PLoS ONE* **12**(10): 0185974. https://doi.org/10.1371/journal.pone.0185974

Bayazit, S., Kazan, K., Gülbitti, S., Cevik, V., Ayanoglu, H. and Ergül, A. (2007): AFLP analysis of genetic diversity in low chill requiring walnut (Juglans regia L.) genotypes from Hatay, Turkey. – *Sci: Hortic.* **111**: 394–398. https://doi.org/10.1016/j.scienta.2006.11.006

Beug, H. J. (1975): *Man as a factor in the vegetational history of the Balkan Peninsula*. – In: Jordanov, D., Bondev, I., Kozuharov, S., Kuzmanov, B. and Palamarev, E. (eds): Problems of Balkan flora and vegetation. Publishing House of the Bulgarian Academy of Sciences, Sofia, Bulgaria, pp. 72–78.

CBOL Plant Working Group (2009): *A DNA barcode for land plants*. – Communicated by Daniel H. Janzen, University of Pennsylvania, Philadelphia, PA.

Ehteshamnia, A. A., Sharifani, M., Vahdati, K. and Erfani Moghadam, V. (2009): Investigation of genetic diversity among some native populations of walnut (Juglans regia L.) in Golestan Province by SSR markers. – *J. Plant Prod.* **4**: 39–58.

Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W. and Newmaster, S. G. (2008): Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. – *PLoS ONE* **3**: e2802. https://doi.org/10.1371/journal.pone.0002802

Fjellstrom, R. G., Parfitt, D. E. and Mcgranahan, G. H. (1994): Genetic relationship and characterization of Persian walnut (Juglans regia L.) genotypes using restriction fragment length polymorphisms (RFLPs). – *J. Amer. Soc. Hortic. Sci.* **119**: 833–839. https://doi.org/10.21273/jashs.119.4.833

Hammer, Ø., Harper, D. A. T. and Ryan, P. D. (2012): PAST: Paleontological statistics software package for education and data analysis. – *Pal. Electronica* **4**: 9.

Hollingsworth, P. M., Graham, S. W. and Little, D. P. (2011): Choosing and using a plant DNA barcode. – *PLoS ONE* **6**(5): e19254. https://doi.org/10.1371/journal.pone.0019254

Kalendar, R. and Schulman, A. H. (2006): IRAP and REMAP for retrotransposon-based genotyping and fingerprinting. – *Nat Protoc.* **1**(5): 2478–2484. https://doi.org/10.1038/nprot.2006.377

Karimi, R., Ershadi, A., Vahdati, K., Ehtesham Nia, A., Sharifani, M., Rasouli, M. and Ebrahimi, A. (2014): Morphological and molecular evaluation of Persian walnut populations in northern and western regions of Iran. – *J. Nuts* **5**(2): 21–31. https://doi.org/10.22034/jon.2014.515686

Križman, M., Jakše, J., Baričevič, D., Javornik, B. and Prošek, M. (2006): Robust CTAB-activated charcoal protocol for plant DNA extraction. – *Acta Agric. Sloven.* **87**: 427–433.

Laiou, A., Aconiti Mandolini, L., Piredda, R., Bellarosa, R. and Simeone, M. C. (2013): DNA barcoding as a complementary tool for conservation and valorisation of forest resources. – *ZooKeys* **365**: 197–213. https://doi.org/10.3897/zookeys.365.5670

Maghsoodi, M., Sheidai, M. and Koohdar, F. (2018): Population genetic study in Juglans regia L. (Persian walnut) and its taxonomic status within the genus Juglans L. – *Phytotaxa* **376**: 154–166. https://doi.org/10.11646/phytotaxa.376.4.1

Mohsenipoor, S., Vahdati, K., Amiri, R. and Mozaffari, M. (2010): Study of the genetic structure and gene flow in Persian walnut (Juglans regia L.) using SSR markers. – *Acta Hortic*. **861**: 133–142. https://doi.org/10.17660/actahortic.2010.861.17

Nematollahi, A., Vahdati, K., Golein, B., Nabipour, A. and Danesh, M. (2009): Assessment of genetic variation in citrus germplasm of south of Iran using SSR markers. – *Acta Hortic.* **839**: 549–554. https://doi.org/10.17660/actahortic.2009.839.75

Nicese, F. P., Hormaza, J. I. and Mcgranahan, G. H. (1998): Molecular characterization and genetic relatedness among walnut (Juglans regia L.) genotypes based on RAPD markers. – *Euphytica* **101**: 199–206.

Ninot, A. and Aletà, N. (2003): Identification and genetic relationship of Persian walnut genotypes using isozyme markers. – *J. Amer. Pomol. Soc.* **57**: 106–114.

Peakall, R. and Smouse, P. E. (2006): GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. – *Mol. Ecol. Notes* **6**: 288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Podani, J. (2000): *Introduction to the exploration of multivariate data*. – Backhuys, Leiden, 407 pp.

Pop, I. F., Vicol, A. C., Botu, M., Raica, P. A., Vahdati, K. and Pamfil, D. (2013): Relationships of walnut genotypes in a germplasm collection: Comparative analysis of phenotypic and molecular data. – *Sci. Hortic.* **153**: 124–135. https://doi.org/10.1016/j.scienta.2013.02.013

Popov, K. P. (1994): *Trees, shrubs, and semishrubs in the mountains of Turkmenistan*. – In: Fet, V. and Atamuradov, K. I. (eds): Biogeography and ecology of Turkmenistan. Springer, Dordrecht, pp. 173–186.

Popov, M. G. (1929): Wild growing fruit trees and shrubs of Asia Minor. – *Bull. Appl. Bot. Pl. Breed.* **22**: 241–483. [in Russian]

Ruiz-Garcia, L., Lopez-Ortega, G., Denia, A. F. and Frutos Tomas, D. F. (2011): Identification of a walnut (Juglans regia L.) germplasm collection and evaluation of their genetic variability by microsatellite markers. – *Spanish J. Agric. Res.* **9**: 179–192. https://doi.org/10.5424/sjar/20110901-227-10

Sadat-Hosseini, M., Vahdati, K. and Leslie, C. A. (2019): Germination of Persian walnut somatic embryos and evaluation of their genetic stability by ISSR fingerprinting and flow cytometry. – *HortScience* **54**(9): 1576–1580. https://doi.org/10.21273/hortsci13787-18

Shaw, J. and Small, R. L. (2005): Chloroplast DNA phylogeny and phylogeography of the North American plums (Prunus subgenus Prunus section Prunocerasus, Rosaceae). – *Amer. J. Bot.* **92**: 2011–2030. https://doi.org/10.3732/ajb.92.12.2011

Sheidai, M., Taban, F., Talebi, S. M. and Noormohammadi, Z. (2016): Genetic and morphological diversity in Stachys lavandulifolia (Lamiaceae) populations. – *Biologija* **62**(1): 9–24. https://doi.org/10.6001/biologija.v62i1.3286

Sheidai, M., Zanganeh, S., Haji-Ramezanali, R., Nouroozi, M., Noormohammadi, Z. and Ghasemzadeh-Baraki, S. (2013): Genetic diversity and population structure in four Cirsium (Asteraceae) species. – *Biologia* **68**: 384–397. https://doi.org/10.2478/s11756-013-0162-x

Solar, A., Smole, J., Štampar, F. and Viršček-Marn, M. (1994): Characterization of isozyme variation in walnut (Juglans regia L.). – *Euphytica* **77**: 105–112. https://doi.org/10.1007/bf02551471

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2012): MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. – *Mol. Biol. Evol.* **28**(10): 2731–2739. https://doi.org/10.1093/molbev/msr121

Timmer, J. C., Enoksson, M., Wildfang, E., Zhu, W., Igarashi, Y., Denault, J. B., Ma, Y., Dummitt, B., Chang, Y. H., Mast, A. E., Eroshkin, A., Smith, J. W., Tao, W. A. and Salvesen, G. S. (2007): Profiling constitutive proteolytic events in vivo. – *Biochem. J.* **407**(1): 41–48. https://doi.org/10.1042/bj20070775

Wang, H., Pei, D., Gu, R.-S. and Wang, B.-Q. (2008): Genetic diversity and structure of walnut populations in Central and Southwestern China revealed by microsatellite markers. – *J. Amer. Soc. Hort. Sci.* **133**(2): 197–203. https://doi.org/10.21273/jashs.133.2.197

Weising, K., Nybom, H., Wolff, K. and Kahl, G. (2005): *DNA Fingerprinting in plants*. – In: Principles, methods, and applications. (2nd ed.). CRC Press, Boca Rayton, 472 pp.

Zohary, D. and Hopf, M. (1993): *Domestication of plants in the old world*. – Clarendon Press, Oxford, 249 pp.