



**Citation:** Songpo Liu, Yuxuan Wang, Yuwei Song, Majid Khayatnezhad, Amir Abbas Minaeifar (2021) Genetic variations and interspecific relationships in *Salvia* (Lamiaceae) using SCoT molecular markers. *Caryologia* 74(3): 77-89. doi: 10.36253/caryologia-886

**Received:** March 23, 2020

**Accepted:** September 24, 2021

**Published:** December 21, 2021

**Copyright:** © 2021 Songpo Liu, Yuxuan Wang, Yuwei Song, Majid Khayatnezhad, Amir Abbas Minaeifar. This is an open access, peer-reviewed article published by Firenze University Press (<http://www.fupress.com/caryologia>) and distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**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.

**ORCID**

AAM: 0000-0002-9371-1498

## Genetic variations and interspecific relationships in *Salvia* (Lamiaceae) using SCoT molecular markers

SONGPO LIU<sup>1</sup>, YUXUAN WANG<sup>1</sup>, YUWEI SONG<sup>1,\*</sup>, MAJID KHAYATNEZHAD<sup>2</sup>, AMIR ABBAS MINAEIFAR<sup>3</sup>

<sup>1</sup>Department of Life Science and Biotechnology, Nanyang Normal University, Nanyang, 473000, China

<sup>2</sup>Department of Environmental Sciences and Engineering, Ardabil Branch, Islamic Azad University, Ardabil, Iran

<sup>3</sup>Department of Biology, Payame Noor University, P.O. Box19395-3697 Tehran, Iran

\*Corresponding author. E-mail: nanyangyws@126.com; aaminaeifar@gmail.com

**Abstract.** The genus *Salvia* includes an enormous assemblage of nearly 1000 species dispersed around the World. Iran having 19 endemic species out of 61 is regarded as one of the important regions for *Salvia* diversity in Southwest Asia. *Salvia* species are herbaceous, rarely biennial or annual, often strongly aromatic. These species are of medicinal, commercial and horticultural value. Due to the importance of these plant species, we performed a combination of morphological and molecular data for this species. For this study, we used 145 randomly collected plants from 30 species in 18 provinces. Amplification of genomic DNA using 10 primers produced 134 bands, of which 129 were polymorphic (97.78%). The obtained high average PIC and MI values revealed high capacity of SCoT primers to detect polymorphic loci among *Salvia* species. The genetic similarities of 30 collections were estimated from 0.61 to 0.93. According to the SCoT markers analysis, *S. tebesana* and *S. verticillata* had the lowest similarity and the species of *S. eremophila* and *S. santolinifolia* had the highest similarity. The aims of present study are: 1) can SCoT markers identify *Salvia* species, 2) what is the genetic structure of these taxa in Iran, and 3) to investigate the species inter-relationship? The present study revealed that SCoT markers can identify the species.

**Keywords:** Iran, species identification, structure, *Salvia*, SCoT (Start Codon Targeted).

### INTRODUCTION

Identifying the accurate boundaries of a species is critical to have a better perspective of any biological studies. Therefore, species delimitation is a subject of extensive part of studies in the framework of biology (Collard & Mackill 2009, Luo *et al.* 2011, Wu *et al.* 2013). However, defining the criterion which could address the boundaries of species is different and the place of debates (Jamzad 2012). Among different populations, genetic diversity is

non randomly distributed and is affected by various factors such as geographic variations, breeding systems, dispersal mechanisms, life span, etc. Change in environmental conditions often leads to variation in genetic diversity levels among different populations and populations with low variability are generally considered less adapted under adverse circumstances (Falk & Holsinger 1991, Olivieri *et al.* 2016). Most of the authors agree that genetic diversity is necessary to preserve the long-term evolutionary potential of a species (Falk & Holsinger 1991). In the last decade, experimental and field investigations have demonstrated that habitat fragmentation and population decline reduce the effective population size. In the same way, most geneticists consider population size as an important factor for maintaining genetic variation (Turchetto *et al.* 2016).

*Salvia* L. is known as the largest genus in Lamiaceae (Mentheae-Salviinae) with approximately 1000 species diversified in three regions of the world: Central and South America (500 spp.), Western Asia (200 spp.) and Eastern Asia (100 species) (Walker *et al.* 2004). Iran having 19 endemic species out of 61 is regarded as one of the important regions for *Salvia* diversity in Southwest Asia (Jamzad 2012). *Salvia* species are herbaceous, rarely biennial or annual, often strongly aromatic. These species are of medicinal, commercial and horticultural value (Safaei *et al.* 2016). Also, some *Salvia* species have pharmacological properties, including antiplatelet, anti-inflammatory and antithrombotic effects (Hosseinzadeh *et al.* 2003, Mayer *et al.* 2007; Fan *et al.* 2010). Some species of this genus are used in folk medicine, such as *S. miltiorrhiza* Bunge, which is used for treatment of cardiovascular diseases (Wang *et al.* 2007, 2009). *Salvia reuterana* Boiss. is an endemic species which grows in the highlands of central Iran (Jamzad 2012). Its common name in Persian is “Mariam Goli Esfahani”, and the aerial parts of the plant are traditionally used as sedative and anxiolytic herbal medicine. In addition, the antibacterial, antioxidant, free radical scavenging and anti-anxiety properties of this herb have been proved in recent studies (Erbano *et al.* 2015). The chemical composition of *Salvia* strongly indicates that the herb has potential to become an important raw material for anti-inflammatory compounds and knowledge of the diversity of wild populations will therefore be important to inform the use and conservation of this genus (Frag *et al.* 1986, Li & Quiros 2001). Genetic surveys, in particular, are key measures to efficiently access the genetic resources of species of pharmacological interest. Several markers have been previously applied to survey genetic variability within the genus *Salvia* (Song *et al.* 2010, Wang *et al.* 2011). Specifically, there are some important publications

addressing *S. miltiorrhiza*, most of them utilizing dominant markers (Wang *et al.* 2011).

Accordingly, some researchers have tried to assess this variability by ISSR and RAPD techniques in different *Salvia* species (Song *et al.* 2010, Wang *et al.* 2011, Sepehry Javan *et al.* 2012, Zhang *et al.* 2013, Peng *et al.* 2014, Urbano *et al.* 2015). Sepehry Javan *et al.* (2012) mentioned that three major factors influencing genetic variations in *Salvia* are: species, geographical distribution and selection. These factors along with cross-pollination make the taxonomy and genetic relationships of *Salvia* species unclear (Wang *et al.* 2011). Morphological characteristics are easily affected by environment that makes identification of species more complex (Chen *et al.* 2013). The conservation and suitable use of plant genetic resources require accurate monitoring of their accessions. So, genetic characterization is essential to manifest the extent of plant genetic diversity, and also to discover better genotypes; especially in the geographically differentiated genus such as *Salvia* (Song *et al.* 2010, Peng *et al.* 2014, Patel *et al.* 2014, Kharazian *et al.* 2015).

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 gene targeted markers as well as develop many novel DNA based marker systems (Esfandani-Bozchaloyi *et al.* 2017 a, 2017b, 2017c, 2017d). 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 & Mackill 2009) translation start codon (Collard & Mackill 2009). 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 & Mackill 2009, Luo *et al.* 2011, Wu *et al.* 2013).

The present investigation has been carried out to evaluate the genetic diversity and relationships among *Salvia* species using new gene-targeted molecular markers, i.e. SCoT. This is the first study on the use of SCoT markers in *Salvia* genus; Therefore, we performed molecular study of 145 specimens of 30 *Salvia* species. We try to answer the following questions: 1) Is there infra and interspecific genetic diversity among studied species? 2) Is genetic distance among these species correlated with their geographical distance? 3) What is the

genetic structure of populations and taxa? 4) Is there any gene exchange between *Salvia* species in Iran?

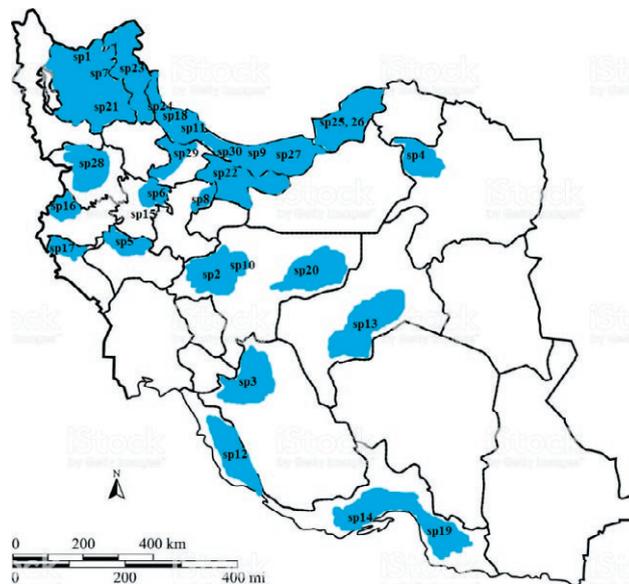
## MATERIALS AND METHODS

### Plant materials

A total of 145 individuals were sampled representing 30 geographical populations belong 30 *Salvia* species (sp1= *Salvia aristata* Aucher ex Benth; sp2= *S. eremophila* Boiss; sp3= *S. santolinifolia* Boiss; sp4= *S. tebesana* Bunge; sp5= *S. bracteata* Banks & Sol; sp 6= *S. suffruticosa* Montb. & Aucher; sp7= *S. dracocephaloides* Boiss.; sp8= *S. hydrangea* DC. ex Benth.; sp9= *S. multicaulis* Vahl.; sp10: *S. syriaca* L.; sp11: *S. viridis* L.; sp12= *S. mirzayanii* Rech. f. & Esfand.; sp13= *S. macrosiphon* Boiss.; sp14= *S. sharifii* Rech. f. & Esfand.; sp15= *S. reuterana* Boiss.; sp16= *S. palaestina* Benth.; sp17= *S. sclareopsis* Bornm. ex Hedge; sp18= *S. spinose* L.; sp19= *S. compressa* Vent.; sp20= *S. sclarea* L.; sp21= *S. aethiopsis* L.; sp22= *S. microstegia* Boiss. & Bal.; sp23= *S. xanthocheila* Boiss. ex Benth.; sp24= *S. limbata* C. A. Mey.; sp25= *S. chloroleuca* Rech. f. & Aell.; sp26= *S. virgate* Jacq.; sp27= *S. nemorosa* L.; sp28= *S. urmiensis* Bunge; sp29= *S. oligophylla* Aucher ex Benth.; sp30= *S. verticillata* L.) in East Azerbaijan, Lorestan, Kermanshah, Guilan, Mazandaran, Golestan, Yazd, Esfahan, Tehran, Arak, Hamadan, Kurdistan, Ilam, Bandar Abbas, Ghazvin, Khorasan and Ardabil Provinces of Iran during July-August 2017-2019. Out-group taxa are: *Marrubium anisodon* K. Koch and *M. cuneatum* Banks & Sol. For morphometric and SCoT analysis we used 145 plant accessions (five to twelve samples from each populations) belonging to 30 different populations with different eco-geographic characteristics were sampled and stored in -20 till further use. More information about geographical distribution of accessions are in Fig. 1.

### Morphological studies

Five to twelve samples from each species were used for Morphometry (Some endemic species were collected due to the rarity of 5 to 12 numbers). In total 22 morphological (9 qualitative, 13 quantitative) characters were studied. Data obtained were standardized (Mean= 0, variance = 1) and used to estimate Euclidean distance for clustering and ordination analyses (Podani 2000). Morphological characters studied are: corolla shape, bract shape, seed color, seed shape, bract color, corolla latex, leaf surface, calyx shape, basal leaf shape, pedicel length, calyx length, bract length, filament length,



**Figure 1.** Map of Iran shows the collection sites and provinces where *Salvia* species were obtained for this study; sp1= *Salvia aristata*; sp2= *S. eremophila*; sp3= *S. santolinifolia*; sp4= *S. tebesana*; sp5= *S. bracteata*; sp 6= *S. suffruticosa*; sp7= *S. dracocephaloides*; sp8= *S. hydrangea*; sp9= *S. multicaulis*; sp10: *S. syriaca*; sp11: *S. viridis*; sp12= *S. mirzayanii*; sp13= *S. macrosiphon*; sp14= *S. sharifii*; sp15= *S. reuterana*; sp16= *S. palaestina*; sp17= *S. sclareopsis*; sp18= *S. spinose*; sp19= *S. compressa*; sp20= *S. sclarea*; sp21= *S. aethiopsis*; sp22= *S. microstegia*; sp23= *S. xanthocheila*; sp24= *S. limbata*; sp25= *S. chloroleuca*; sp26= *S. virgate*; sp27= *S. nemorosa*; sp28= *S. urmiensis*; sp29= *S. oligophylla*; sp30= *S. verticillata*

anther length, corolla length, nut length, nut width, basal leaf length, basal leaf width, corolla color, stem leaf length and stem leaf width.

### DNA Extraction and SCoT Assay

Fresh leaves were used randomly from three to twelve 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 & Mackill (2009), 10 primers with clear, enlarged, and rich polymorphism bands were chosen (Table 1). PCR reactions were carried in a 25µl volume containing 10 mM Tris-HCl buffer at pH 8; 50 mM KCl; 1.5 mM MgCl<sub>2</sub>; 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 amplifications reactions were performed in Techne thermocycler (Germany) with the following program: 5 min initial denaturation step 94°C, followed by 40 cycles of 1 min at

**Table 1.** SCoT primers used for this study and the extent of polymorphism.

Primer name	Primer sequence (5'-3')	TNB	NPB	PPB	PIC	PI	EMR	MI
SCoT-1	CAACAATGGCTACCACCA	10	10	100.00%	0.36	4.86	9.55	3.45
SCoT-3	CAACAATGGCTACCACCG	9	8	84.99%	0.43	4.91	7.43	4.85
SCoT-6	CAACAATGGCTACCACGC	13	13	100.00%	0.44	4.34	11.55	3.44
SCoT-11	AAGCAATGGCTACCACCA	16	16	100.00%	0.37	3.88	8.56	1.65
SCoT-14	ACGACATGGCGACCACGC	20	20	100.00%	0.55	6.23	8.23	2.47
SCoT-15	ACGACATGGCGACCACCGA	15	14	93.74%	0.47	5.66	7.56	3.67
SCoT-16	CCATGGCTACCACCGGCC	13	12	92.31%	0.34	3.21	5.60	5.55
SCoT-17	CATGGCTACCACCGGCC	12	12	100.00%	0.47	4.32	9.55	3.45
SCoT-18	ACCATGGCTACCACCGCG	11	9	82.89%	0.43	5.56	6.34	2.11
SCoT-19	GCAACAATGGCTACCACC	15	15	100.00%	0.39	3.25	10.11	1.87
Mean		13.4	12.9	97.78%	0.46	4.9	8.4	3.6
Total		134	129					

Note: TNB - the number of total bands, NPB: the number of polymorphic bands, PPB (%): the percentage of polymorphic bands, PI: polymorphism index, EMR, effective multiplex ratio; MI, marker index; PIC, polymorphism information content for each of CDBP primers

94°C; 1 min at 52-57°C and 2 min at 72°C. The reaction was completed by final extension step of 7-10 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).

#### Data analyses

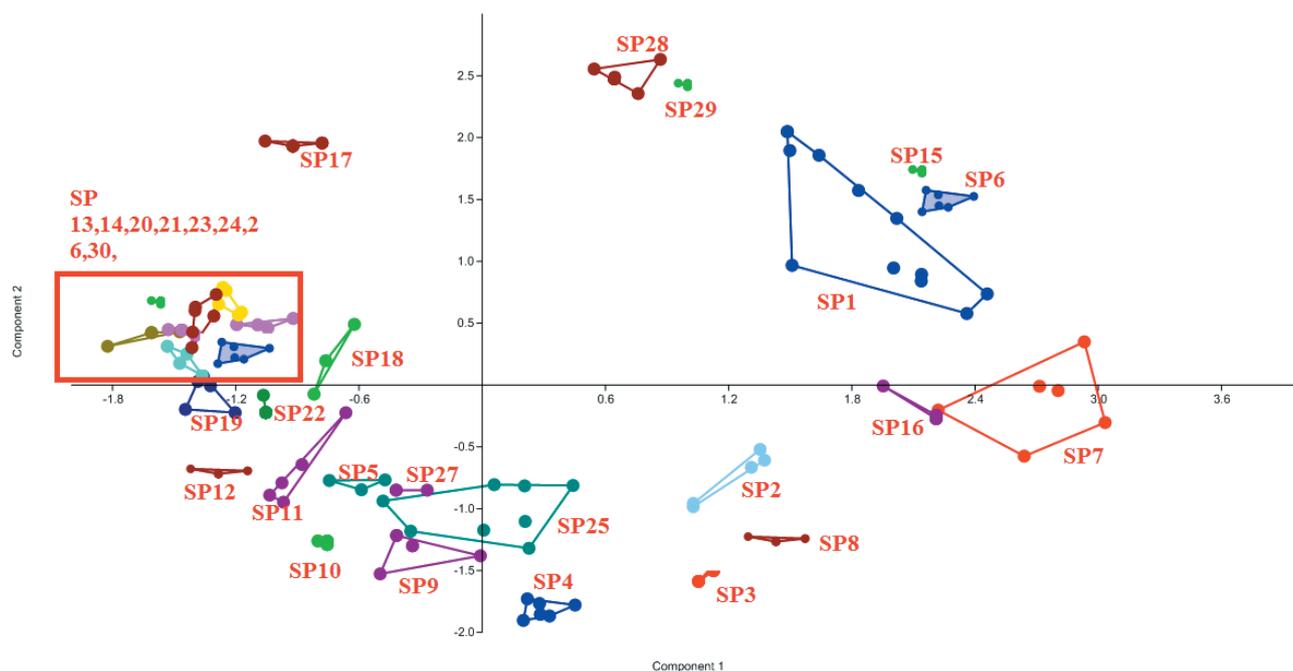
##### Morphological studies

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) ordination methods were used (Podani 2000). ANOVA (Analysis of variance) were performed to show morphological difference among the populations while, PCA (Principal components analysis) biplot was used to identify the most variable morphological characters among the studied populations (Podani 2000). PAST version 2.17 (Hammer *et al.* 2012) was used for multivariate statistical analyses of morphological data.

##### Molecular analyses

SCoT bands obtained were coded as binary characters (presence = 1, absence = 0) and used for genetic diversity analysis. Discriminatory ability of the used primers was evaluated by means of two parameters, polymorphism information content (PIC) and marker

index (MI) to characterize the capacity of each primer to detect polymorphic loci among the genotypes (Powell *et al.* 1996). MI is calculated for each primer as  $MI = PIC \times EMR$ , where EMR is the product of the number of polymorphic loci per primer ( $n$ ) and the fraction of polymorphic fragments ( $\beta$ ) (Heikrujam *et al.* 2015). The number of polymorphic bands (NPB) and the effective multiplex ratio (EMR) were calculated for each primer. Parameter like Nei's gene diversity ( $H$ ), Shannon information index ( $I$ ), number of effective alleles, and percentage of polymorphism ( $P\% = \text{number of polymorphic loci} / \text{number of total loci}$ ) were determined (Weising *et al.* 2005, Freeland *et al.* 2011). Shannon's index was calculated by the formula:  $H' = -\sum p_i \ln p_i$ .  $R_p$  is defined per primer as:  $R_p = \sum I_b$ , where " $I_b$ " 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 & Smouse 2006). Nei's genetic distance among populations was used for Neighbor Joining (NJ) clustering and Neighbor-Net networking (Huson & Bryant 2006, Freeland *et al.* 2011). Mantel test checked the correlation between geographical and genetic distances of the studied populations (Podani 2000). These analyses were done by PAST ver. 2.17 (Hammer *et al.* 2012), DARwin ver. 5 (2012) software. AMOVA (Analysis of molecular variance) test (with 1000 permutations) as implemented in GenAlEx 6.4 (Peakall & Smouse 2006) were used to show genetic difference of the populations. Gene flow was determined by (i) Calculating  $N_m$  an estimate of gene flow from  $G_{st}$  by PopGene ver. 1.32 (1997) as:  $N_m = 0.5(1 - G_{st})/G_{st}$ .



**Figure 2.** PCA plots of morphological characters revealing species delimitation in the *Salvia* sp1= *Salvia aristata*; sp2= *S. eremophila*; sp3= *S. santolinifolia*; sp4= *S. tebesana*; sp5= *S. bracteata*; sp6= *S. suffruticosa*; sp7= *S. dracocephaloides*; sp8= *S. hydrangea*; sp9= *S. multicaulis*; sp10= *S. syriaca*; sp11= *S. viridis*; sp12= *S. mirzayanii*; sp13= *S. macrosiphon*; sp14= *S. sharifii*; sp15= *S. reuterana*; sp16= *S. palaestina*; sp17= *S. sclareopsis*; sp18= *S. spinose*; sp19= *S. compressa*; sp20= *S. sclarea*; sp21= *S. aethiopsis*; sp22= *S. microstegia*; sp23= *S. xanthocheila*; sp24= *S. limbata*; sp25= *S. chloroleuca*; sp26= *S. virgate*; sp27= *S. nemorosa*; sp28= *S. urmiensis*; sp29= *S. oligphylla*; sp30= *S. verticillata*.

This approach considers the equal amount of gene flow among all populations.

## RESULTS

### *Species identification and inter-relationship*

#### Morphometry

ANOVA showed significant differences ( $P < 0.01$ ) in quantitative morphological characters among the species studied. In order to determine the most variable characters among the taxa studied, PCA analysis has been performed. It revealed that the first three factors comprised over 63% of the total variation. In the first PCA axis with 42% of total variation, such characters as seed shape, calyx shape, calyx length, bract length and basal leaf shape have shown the highest correlation ( $>0.7$ ), seed color, leaf surface, corolla length, filament length, nut width, basal leaf length, 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. 2). In general, plant samples of each species

were grouped together and formed separate groups. This result show that both quantitative and qualitative morphological characters separated the studied species into distinct groups. In the studied specimens we did not encounter intermediate forms.

### *Species Identification and Genetic Diversity*

Ten ISSR primers were screened to study genetic relationships among *Salvia* species; all the primers produced reproducible polymorphic bands in all 30 *Salvia* species. An image of the ISSR amplification generated by SCoT-11 primer is shown in Figure 3. A total of 129 amplified polymorphic bands were generated across 30 *Salvia* species. The size of the amplified fragments ranged from 100 to 2000 bp. The highest and lowest number of polymorphic bands were 20 for SCoT-14 and 8 for SCoT-3, on an average of 12.9 polymorphic bands per primer. The PIC of the 10 SCoT primers ranged from 0.36 (SCoT-1) to 0.55 (SCoT-14) with an average of 0.46 per primer. MI of the primers ranged from 1.65 (SCoT-11) to 5.55 (SCoT-16) with an average of 3.6 per primer. EMR of the SCoT primers ranged from 6.34 (SCoT-18) to 11.55 (SCoT-6) with an average of 8.4 per primer (Table 1). The primers



**Fig. 3.** Electrophoresis gel of studied ecotypes from DNA fragments produced by SCoT-16. 1= *Salvia aristata*; 2= *S. eremophila*; 3= *S. santolinifolia*; 4= *S. tebesana*; 5= *S. bracteata*; 6= *S. suffruticosa*; 7= *S. dracocephaloides*; 8= *S. hydrangea*; 9= *S. multicaulis*; 10= *S. syriaca*; 11= *S. viridis*; 12= *S. mirzayanii*; 13= *S. macrosiphon*; 14= *S. sharifii*; 15= *S. reuterana*; 16= *S. palaestina*; 17= *S. sclareopsis*; 18= *S. spinose*; 19= *S. compressa*; 20= *S. sclarea*; 21= *S. aethiopsis*; 22= *S. microstegia*; 23= *S. xanthocheila*; 24= *S. limbata*; 25= *S. chloroleuca*; 26= *S. virgate*; 27= *S. nemorosa*; 28= *S. urmiensis*; 29= *S. oligphylla*; 30= *S. verticillata*; L = Ladder 100 bp, Arrows are representative of polymorphic bands.

with the high EMR values were considered to be more informative in distinguishing the genotypes.

The genetic parameters were calculated for all the 30 *Salvia* species amplified with SCoT primers (Table 2). Unbiased expected heterozygosity ( $H$ ) ranged from 0.11 (*S. syriaca*) to 0.29 (*S. virgate*), with a mean of 0.19. A similar pattern was observed for Shannon's information index ( $I$ ), with the highest value of 0.45 observed in *S. virgate* and the lowest value of 0.12 observed in *S. syriaca* with a mean of 0.26. The observed number of alleles ( $N_a$ ) ranged from 0.214 in *S. eremophila* to 0.89 in *S. aristata*. The effective number of alleles ( $N_e$ ) ranged from 0.98 (*S. multicaulis*) to 1.440 (*S. virgate*).

AMOVA test showed significant genetic difference ( $P = 0.01$ ) among studied species. It revealed that 66% of total variation was among species and 34% was within species (Table 3). Moreover, genetic differentiation of these species was demonstrated by significant Nei's GST (0.21,  $P = 0.01$ ) and  $D_{est}$  values (0.177,  $P = 0.01$ ). These results revealed a higher distribution of genetic diversity among *Salvia* species compared to within species. *Marrubium anisodon* and *M. cuneatum* (out-groups) were separated from the other species. Two major clusters were formed in WARD tree (Fig. 4). The first major cluster (A) contained two sub-clusters: *S. sharifii* and *S. macrosiphon* are separated from the other studied species and join the others with a great distance and comprised the first sub-cluster. The second sub-cluster was formed by *S. xanthocheila*, *S. limbata*, *S. aethiopsis*, *S. sclarea* and *S. virgate*. The second major cluster also contained two sub-clusters: eight species of *S. multicaulis*; *S. syriaca*; *S. viridis*, *S. reuterana*; *S. palaestina*; *S. sclareopsis*; *S. spinose* and *S. oligphylla* were placed close

to each other, while close genetic affinity between other species. In general, relationships obtained from SCoT data agrees well with species relationship obtained from morphological. This is in agreement with AMOVA and genetic diversity parameters presented before. The species are genetically well differentiated from each other. These results indicate that SCoT molecular markers can be used in *Salvia* species taxonomy. The Nm analysis by Popgene software also produced mean Nm= 0.167, that is considered very low value of gene flow among the studied species.

Mantel test with 5000 permutations showed a significant correlation ( $r = 0.13$ ,  $p = 0.0002$ ) between genetic distance and geographical distance, so isolation by distance (IBD) occurred among the *Salvia* species studied.

Nei's genetic identity and the genetic distance determined among the studied species (Table 4). The results showed that the highest degree of genetic similarity (0.93) occurred between *S. eremophila* and *S. santolinifolia*. The lowest degree of genetic similarity occurred between *S. tebesana* and *S. verticillata* (0.66). The low Nm value (0.167) indicates limited gene flow or ancestrally shared alleles between the species studied and indicating high genetic differentiation among and within *Salvia* species.

## DISCUSSION

Genetic diversity is a basic component of biodiversity and its conservation is essential for long term survival of any species in changing environments (Mills & Schwartz 2005, Tomasello *et al.* 2015, Miao *et al.* 2019;

**Table 2.** Genetic diversity parameters in the studied *Salvia* species.

SP	N	Na	Ne	I	He	UHe	%P
<i>S. aristata</i>	6.000	0.892	1.138	0.221	0.141	0.165	38.63%
<i>S. eremophila</i>	6.000	0.244	1.032	0.26	0.23	0.18	55.53%
<i>S. santolinifolia</i>	4.000	0.314	1.044	0.16	0.18	0.23	43.38%
<i>S. tebesana</i>	8.000	0.201	1.00	0.33	0.17	0.12	42.23%
<i>S. bracteata</i>	5.000	0.341	1.058	0.24	0.27	0.20	53.75%
<i>S. suffruticosa</i>	3.000	0.567	1.062	0.24	0.224	0.113	44.73%
<i>S. dracocephaloides</i>	5.000	0.336	1.034	0.23	0.25	0.19	51.83%
<i>S. hydrangea</i>	4.000	0.344	1.042	0.20	0.23	0.20	57.53%
<i>S. multicaulis</i>	5.000	0.369	1.011	0.15	0.18	0.12	42.15%
<i>S. syriaca</i>	8.000	0.566	1.014	0.45	0.10	0.11	32.58%
<i>S. viridis</i>	9.000	0.432	1.049	0.18	0.22	0.25	55.05%
<i>S. mirzayanii</i>	8.000	0.313	1.026	0.144	0.13	0.26	49.23%
<i>S. macrosiphon</i>	12.000	1.244	1.322	0.28	0.284	0.192	50.91%
<i>S. sharifii</i>	5.000	0.358	1.117	0.28	0.15	0.12	44.30%
<i>S. reuterana</i>	6.000	0.458	1.039	0.28	0.18	0.23	49.38%
<i>S. palaestina</i>	5.000	0.455	1.077	0.377	0.24	0.22	55.05%
<i>S. sclareopsis</i>	8.000	0.499	1.067	0.14	0.101	0.14	49.26%
<i>S. spinose</i>	9.000	0.261	1.014	0.142	0.33	0.23	43.15%
<i>S. compressa</i>	6.000	0.555	1.021	0.39	0.25	0.28	43.53%
<i>S. sclarea</i>	10.000	0.431	1.088	0.33	0.22	0.13	57.53%
<i>S. aethiopsis</i>	3.000	0.255	1.021	0.15	0.18	0.12	42.15%
<i>S. microstegia</i>	3.000	0.288	1.024	0.23	0.15	0.17	64.30%
<i>S. xanthocheila</i>	9.000	0.352	1.083	0.23	0.22	0.14	45.05%
<i>S. limbata</i>	8.000	0.333	1.016	0.122	0.12	0.22	48.23%
<i>S. chloroleuca</i>	12.000	1.247	1.199	0.271	0.184	0.192	55.91%
<i>S. virgata</i>	5.000	0.358	1.440	0.114	0.30	0.29	66.50%
<i>S. nemorosa</i>	6.000	0.299	1.029	0.231	0.18	0.23	44.38%
<i>S. urmiensis</i>	5.000	0.462	1.095	0.288	0.25	0.22	62.05%
<i>S. oligophylla</i>	8.000	0.399	1.167	0.259	0.234	0.133	32.88%
<i>S. verticillata</i>	8.000	0.477	1.167	0.356	0.233	0.148	31.26%

Abbreviations: N = number of samples, Na= number of different alleles; Ne = number of effective alleles, I= Shannon's information index, He = gene diversity, UHe = unbiased gene diversity, P%= percentage of polymorphism, populations.

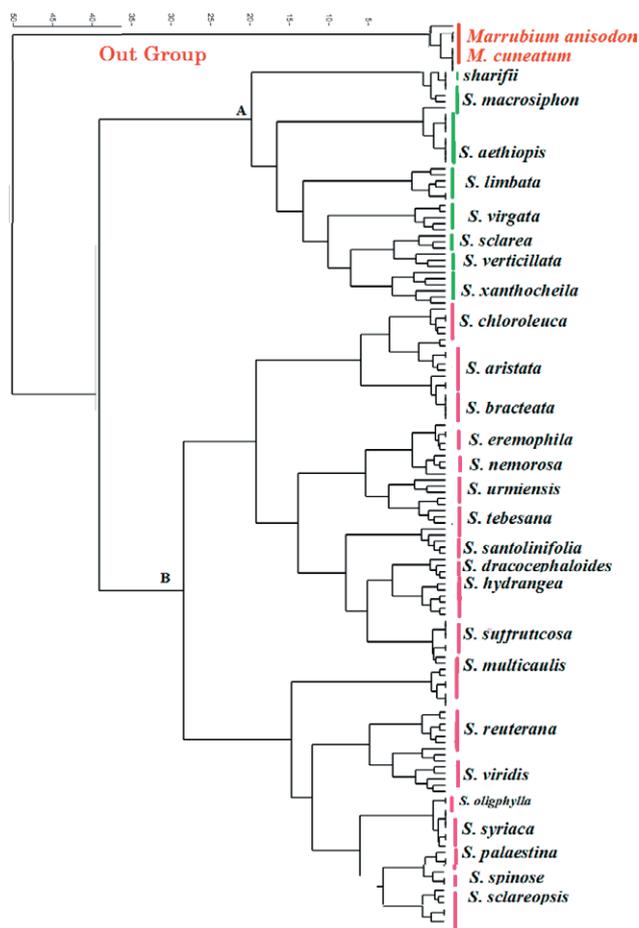
**Table 3.** Analysis of molecular variance (AMOVA) of the studied species.

Source	df	SS	MS	Est. Var.	%	$\Phi_{PT}$
Among Pops	28	1801.364	75.789	12.154	66%	
Within Pops	129	334.443	3.905	2.888	34%	66%
Total	144	1955.807		14.060	100%	

**df:** degree of freedom; **SS:** sum of squared observations; **MS:** mean of squared observations; **EV:** estimated variance;  **$\Phi_{PT}$ :** proportion of the total genetic variance among individuals within an accession, ( $P < 0.001$ ).

Xu *et al.* 2021, Zou *et al.* 2019, Wang *et al.* 2020). This is very important in fragmented populations because are more vulnerable due to the loss of allelic richness and increased population differentiation by genetic drift (decreases heterozygosity and eventual fixation of alleles)

and inbreeding depression (increases homozygosity within populations; Frankham 2005). Therefore, knowledge of the genetic variability and diversity within and among different populations is crucial for their conservation and management (e.g. Esfandani-Bozchaloyi *et al.*



**Figure 4.** WARD tree of SCoT data revealing species delimitation in the *Salvia*.

2018a, 2018b, 2018c, 2018d; Salari *et al.* 2013; 2020; Jahani *et al.* 2019).

In the present study we used morphological and molecular (SCoT) data to evaluate species relationship in *Salvia*. Morphological analyses of the studied *Salvia* species showed that they are well differentiated from each other both in quantitative measures (the ANOVA test result) and qualitative characters (The PCA plot result). In addition, PCA analysis suggests that characters like bract length, stipule length, bract shape, calyx shape, petal shape, length and width of stem-leaf, length and width of petal could be used in species groups delimitation. This morphological difference was due to quantitative and qualitative characters.

#### Genetic structure and gene flow

PIC and MI characteristics of a primer help in determining its effectiveness in genetic diversity analysis.

Sivaprakash *et al.* (2004) suggested that the ability of a marker technique to resolve genetic diversity may be more directly related to the degree of polymorphism. Generally, PIC value between zero to 0.25 imply a very low genetic diversity among genotypes, between 0.25 to 0.50 shows a mid-level of genetic diversity and value  $\geq 0.50$  suggests a high level of genetic diversity (Tams *et al.* 2005). In this research, the SCoT primers' PIC values ranged from 0.36 to 0.55, with a mean value of 0.46, which indicated a mid-ability of SCoT primers in determining genetic diversity among the *Salvia* species. Comparable but low PIC values have been reported with other markers like RAPD and AFLP in African plantain (Ude *et al.* 2003), ISSR and RAPD in *Salvia* species (Yousefiazar-Khanian *et al.* 2016), AFLP in wheat (Bohn *et al.* 1999) and SCoT markers (Etminan *et al.* 2018, Pour-Aboughadareh *et al.* 2017, 2018). In Heikrujam *et al.* (2015), CBDP markers were found to be more effective than SCoT markers with regard to the average PIC which was higher. In our study, the SCoT markers were found to be effective in the estimation of different *Salvia* species genetic diversity with regard to average percentage polymorphism (97.78%), average PIC value of SCoT markers (0.46), average MI (3.6) and average EMR of SCoT markers (8.4), which were higher than other reported markers on *Salvia* (Wang *et al.* 2009, Song *et al.* 2010, Yousefiazar-Khanian *et al.* 2016, Etminan *et al.* 2018, Gholamin and Khayatnezhad 2020 a, b, c, d). However, various marker techniques were found to have different resolution of the genome regions and the number of loci that cover the whole genome for estimating of genetic diversity (Souframanien & Gopalakrishna 2004). A diverse level of polymorphism in *Salvia* species using ISSR, CoRAP, SRAP, SCoT and RAPD markers had been reported earlier by Wang & Zhang (2009), Song *et al.* (2010), Yousefiazar-Khanian *et al.* (2016) and Etminan *et al.* (2018). Gene flow is inversely correlated with the gene differentiation but is very important for population evolution, and takes place by pollen and seeds between populations (Song *et al.* 2010). In the current study, detected gene flow ( $N_m$ ) among *Salvia* species was 0.167, showed low genetic differentiation among *Salvia* species.

As a general rule, insects are the pollinators of *Salvia* in Old World (Claßen-Bockhoff *et al.* 2004, Khayatnezhad and Gholamin, 2012a, b). At the lower elevations, bees and at the higher altitudes insects like flies are the dominate pollinators among bilabiate flowers such as *Salvia* (Pellissier *et al.* 2010).

According to Moein *et al.* (2019) genetic structure of SRAP marker showed that despite the presence of a limited gene flow, two distinct ecotypes were formed which may be the consequences of reproductive isolation



caused by altitude gradient and different niches through parapatric speciation. The heterozygosity ( $H$ ) and Shannon index ( $I$ ) reflect diversity and differentiation among and within the germplasm collections, respectively (Que *et al.* 2014), and the higher the indices, the greater the genetic diversity. The magnitude of variability among  $N_a$ ,  $N_e$ ,  $H$  and  $I$  indices using studied SCoT markers demonstrated a high level of genetic diversity among and within *Salvia* species.

The similar results reported in *Salvia miltiorrhiza* based on ISSRs (Zhang *et al.*, 2013) and other *Salvia* species using AFLP markers (Sajadi *et al.*, 2010) as 95% and 99% polymorphism, respectively. Also, polymorphism index (PI) in RAPD primers was higher; whereas, other indices like PIC, EMR and MI were somewhat high in ISSRs. On the other hand, RP index was approximately equal in both techniques. In general, small differences in terms of calculated indices showed that both techniques had similar efficiency to differentiate the closely related ecotypes of *Salvia*. Chen *et al.* (2013) reported PIC values about 0.20 in *ocimum* species by ISSR and RAPD markers and also showed the RP values as 1.39 and 5.13, respectively. PIC analysis can be used to select the most appropriate markers for genetic mapping. Also, the high MI reflects the marker efficiency to simultaneously analyze a large number of bands (Powell *et al.*, 1996; Patel *et al.*, 2014). The high average Simpson's coefficients (about 0.80) indicate high genetic variability among studied accessions of *Salvia*, too. This finding was similar to the study by Manica-Cattani *et al.* (2009) on accessions of *Lippia alba* by ISSR and RAPD. In their study on *Salvia lachnostachys* ecotypes by ISSR primers, Erbaro *et al.* (2015) showed a range of 0.66-0.86 for Simpson's index. Comparison of Nei's similarity coefficients between ISSRs and RAPDs showed that both markers had high diagnostic capability. This is consistent with the results of ISSR markers in Mint accessions by Kang *et al.* (2013) and *Salvia miltiorrhiza* germplasms studied by Zhang *et al.* (2013); while the genetic similarity derived from SRAPs and ISSRs represented high

proximity among *Salvia miltiorrhiza* populations (Song *et al.*, 2010). Cluster analysis could group all 21 ecotypes and the results showed reasonable congruency in RAPD and ISSR in terms of species topology. Zhang *et al.* (2013) showed five major clusters for *S. miltiorrhiza* germplasms based on Nei's similarity coefficient for ISSRs; which did not indicate any clear pattern according to their locations. Patel *et al.* (2014) reported that in dendrograms of ISSR and RAPD, the genotypes of each *Ocimum* species were grouped, separately. Similar studies in populations of *S. japonica* and some other *Salvia* species (Sudarmono and Okada 2008) did not show cor-

relation between morphological variations and allozyme and DNA sequences. It was concluded that *S. japonica* is still at the early stage of speciation process. Sympatry or co-occurrence of closely related species can either result from a sympatric speciation process or from secondary contact due to range expansion after speciation. Under the allopatric scenario, genetic variation tends to be uniform across the genome due to a large proportion of the genome changing through a combination of divergent selection, differential response to similar selective pressures and genetic drift (see for example Strasburg *et al.* 2012). In contrast, in the extreme case of sympatric speciation, gene flow between the incipient species can homogenize most of the genome, except for loci that experience strong divergent selection pressures or regions that are tightly linked with these loci (see for example, Strasburg *et al.* 2012, Via 2012).

In conclusion, the results of this study showed that to evaluate the genetic diversity of the *Salvia* genus, the primers derived from SCoT were more effective than the other molecular markers. Also, *Salvia* ecotypes/species were clearly separated from each other in the dendrogram and MDS, indicating the higher efficiency of SCoT technique in *Salvia* species identification.

#### ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (U1404303) and Postgraduate Education Reform and Quality Improvement Project of Henan Province (YJS2021JD17).

#### REFERENCES

- Al-Quran S. 2008. Taxonomical and pharmacological survey of therapeutic plants in Jordan. *Journal of Natural Products*, 1(1):10-26. doi: 10.1556/034.59.2017.3-4.3
- Bohn M., Utz H. F. Melchinger AE. 1999. Genetic similarities among wheat cultivars determined on the basis of RFLPs, AFLPs and SSRs and their use for predicting progeny variance. *Crop Sci.* 39, 228-237.
- Chen SY, Dai TX, Chang YT, Wang SS, Ou SL, Chuang WL, Cheng CY, Lin YH, Lin LY, Ku HM 2013. Genetic diversity among *Ocimum* species based on ISSR, RAPD and SRAP markers. *Australian Journal of Crop Science* 7(10):1463-1471.
- 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.

- Claßen-Bockhoff R, Speck T, Tweraser E, Wester P, Thimm S, Reith M. 2004. The staminal lever mechanism in *Salvia* L. (Lamiaceae): a key innovation for adaptive radiation? *Org. Divers. Evol.* 4(3):189-205. <https://doi.org/10.1016/j.ode.2004.01.004>
- Erbano M, Schnell e Schühli G, Pereira dos Santos É. 2015. Genetic variability and population structure of *Salvia lachnostachys*: implications for breeding and conservation programs. *Int. J. Mol. Sci.* 16(4):7839-7850. <https://doi.org/10.3390/ijms16047839>
- Etminan A, Pour-Aboughadareh A, Nooric A. 2018. Genetic relationships and diversity among wild *Salvia* accessions revealed by ISSR and SCoT markers. *Biotechnol. Biotechnol. Equip.* 32:610-617
- Esfandani-Bozchaloyi S, Sheidai M, Keshavarzi M, Noor-mohammadi Z, 2017a. Genetic Diversity and Morphological Variability In *Geranium Purpureum* Vill. (Geraniaceae) Of Iran. *Genetika* 49:543-557. <https://doi.org/10.2298/GENSR1702543B>
- Esfandani-Bozchaloyi S, Sheidai M, Keshavarzi M, Noor-mohammadi Z, 2017b. Species Delimitation In *Geranium* Sect. *Batrachioidea*: Morphological And Molecular. *Acta Botanica Hungarica* 59(3-4):319-334. doi: 10.1556/034.59.2017.3-4.3
- Esfandani-Bozchaloyi S, Sheidai M, Keshavarzi M, Noor-mohammadi Z, 2017c. Genetic and morphological diversity in *Geranium dissectum* (Sec. Dissecta, Geraniaceae) populations. *Biologia* 72(10):1121-1130. DOI: 10.1515/biolog-2017-0124
- Esfandani-Bozchaloyi S, Sheidai M, Keshavarzi M, Noor-mohammadi Z, 2017d. Analysis of genetic diversity in *Geranium robertianum* by ISSR markers. *Phytologia Balcanica* 23(2):157-166.
- Esfandani-Bozchaloyi S, Sheidai M, Keshavarzi M, Noor-mohammadi 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):47-65.
- Esfandani-Bozchaloyi S, Sheidai M, Keshavarzi M, Noor-mohammadi 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, Noor-mohammadi 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(4):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, *Caryologi* 71(4):1-14. <https://doi.org/10.1080/00087114.2018.1503500>
- Farag RS, Salem H, Badei AZMA and Hassanein DE, 1986. Biochemical studies on the essential oil of some medicinal plants. *Fette Seifen Anstrichmittel* 88:69-72.
- Freeland JR, Kirk H, Peterson S.D. 2011. *Molecular Ecology* (2nded). Wiley-Blackwell, UK, 449 pp.
- Falk D.A, Holsinger K.E. (Eds.). 1991. *Genetics and conservation of rare plants*. Oxford Univ. Press, New York.
- Frankham R. 2005. Stress and adaptation in conservation genetics. *J. Evol. Biol.* 18:750-755.
- Gholamin R, Khayatnezhad M. 2020a. The Effect of Dry Season Stretch on Chlorophyll Content and RWC of Wheat Genotypes (*Triticum Durum* L.). *Biosc. Biotech. Res. Comm.* 13(4):55-66.
- Gholamin R, Khayatnezhad M. 2020b. The Study of Path Analysis for Durum wheat (*Triticum durum* Desf.) Yield Components. *Biosc. Biotech. Res. Comm.* 13(4):112-118.
- Gholamin, R. and M. Khayatnezhad. 2020c. "Assessment of the Correlation between Chlorophyll Content and Drought Resistance in Corn Cultivars (*Zea Mays*)." *Helix* 10(5):93-97.
- Gholamin, R. and M. Khayatnezhad 2020d. "Study of Bread Wheat Genotype Physiological and Biochemical Responses to Drought Stress." *Helix* 10(5):87-92.
- Huson D.H., Bryant D. 2006. Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution* 23:254-267.
- Hammer O., Harper D.A., Ryan P.D. 2012. PAST: Paleontological Statistics software package for education and data analysis. *Palaeo Electro* 4:9.
- Heikrujam M., Kumar J, Agrawal V. 2015 Genetic diversity analysis among male and female *Jobba* genotypes employing gene targeted molecular markers, start codon targeted (SCoT) polymorphism and CAAT box-derived polymorphism (CBDP) markers. *Meta Gene* 5:90-97.
- Jahani M., M. Azadbakht, H. Rasouli, R. Yarani, D. Rezazadeh, N. Salari, and K. Mansouri 2019. L-arginine/5-fluorouracil combination treatment approaches cells selectively: Rescuing endothelial cells while killing MDA-MB-468 breast cancer cells, *Food and Chemical Toxicology*, 123:399-411.
- Jamzad Z. 2012. *Lamiaceae*. – In: Assadi, M., Maassoumi, A. and Mozaffarian, V. (eds): *Flora of Iran*. Vol. 76. Research Institute of Forests and Rangelands, Tehran, 810 pp.
- Kharazian N, Rahimi S, Shiran B. 2015. Genetic diversity and morphological variability of fifteen *Stachys* (Lamiaceae) species from Iran using morphological and ISSR molecular markers. *Biologia* 70(4):438-452.

- Khayatnezhad, M. and Gholamin, R., 2012a. The effect of drought stress on leaf chlorophyll content and stress resistance in maize cultivars (*Zea mays*). *African Journal of Microbiology Research*, 6(12):2844-2848.
- Khayatnezhad, M. and Gholamin, R., 2012b. Effect of nitrogen fertilizer levels on different planting remobilization of dry matter of durum wheat varieties Seimareh. *African Journal of Microbiology Research*, 6(7):1534-1539.
- Li G. Quiros CF, 2001. Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: Its application to mapping and gene tagging in *Brassica*. *Theoretical and Applied Genetics* 103:455-461.
- 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.
- Miao R, Ma J, Liu Y, Liu Y, Yang Z.,... Guo, M. 2019. Variability of Aboveground Litter Inputs Alters Soil Carbon and Nitrogen in a Coniferous–Broadleaf Mixed Forest of Central China. *Forests*, 10(2): 188-192.
- Mills M. Schwartz M. 2005. Rare plants at the extremes of distribution: broadly and narrowly distributed rare species. *Biodivers. Conserv.* 14:1401-1420
- Olivieri I, Tonnabel J, Ronce O, Mignot A. 2016. Why evolution matters for species conservation: perspectives from three case studies of plant metapopulations. *Evol. Appl.* 9:196-211.
- Peakall R, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288-295.
- Pellissier L, Pottier J, Vittoz P, Dubuis A. Guisan A. 2010. Spatial pattern of floral morphology: possible insight into the effects of pollinators on plant distributions. *Oikos* 119(11):1805-1813. <https://doi.org/10.1111/j.1600-0706.2010.18560.x>
- Podani J. 2000. Introduction to the Exploration of Multivariate Data English translation. Backhuyes publisher, Leide, 407 pp.
- Peng L, Ru M, Wang B, Wang Y, Li B, Yu J, Liang Z. 2014. Genetic diversity assessment of a germplasm collection of *Salvia miltiorrhiza* Bunge. based on morphology, ISSR and SRAP markers. *Biochemical Systematics and Ecology* 55:84-92.
- Powell W, Morgante M, Doyle JJ, McNicol JW, Tingey SV, Rafalski A J. 1996. Gene pool variation in genus *Glycine* subgenus *Soja* revealed by polymorphic nuclear and chloroplast microsatellites. *Genetics* 144:793-803.
- Pour-Aboughadareh A, Ahmadi J, Mehrabi A, Etminan A, Moghaddam M. 2017. Assessment of genetic diversity among Iranian *Triticum* germplasm using agro-morphological traits and start codon targeted (SCoT) markers. *Cereal Res. Commun.* 45:574-586.
- Pour-Aboughadareh A, Ahmadi J, Mehrabi A, Etminan A, Moghaddam M. 2018. Insight into the genetic variability analysis and relationships among some *Aegilops* and *Triticum* species, as genome progenitors of bread wheat, using SCoT markers. *Plant Biosys. For.* 152:694-703.
- Que Y, Pan Y, Lu Y, Yang C, Yang Y, Huang N. 2014. Genetic analysis of diversity within a Chinese local sugarcane germplasm based on start codon targeted polymorphism. *Biomed Res. Int.* 2014:1-10.
- Salari N., Mohammadi M., Vaisi-Raygani A. et al. 2020. The prevalence of severe depression in Iranian older adult: a meta-analysis and meta-regression. *BMC Geriatr* 20, 39. <https://bmgeriatr.biomedcentral.com/articles/10.1186/s12877-020-1444-0>
- Salari N., Shohaimi S., Najafi F. et al. 2013. Application of pattern recognition tools for classifying acute coronary syndrome: an integrated medical modeling. *Theor Biol Med Model* 10,57. <https://tbiomed.biomedcentral.com/articles/10.1186/1742-4682-10-57>
- Song Z, Li X, Wang H, Wang J. 2010. Genetic diversity and population structure of *Salvia miltiorrhiza* Bge in China revealed by ISSR and SRAP. *Genetica* 138(2):241-249.
- Sepehry Javan Z, Rahmani F, Heidari R. 2012. Assessment of genetic variation of genus *Salvia* by RAPD and ISSR markers. *Australian Journal of Crop Science* 6(6):1068-1073.
- Sivaprakash KR, Prasanth SR, Mohanty BP, Parida A. 2004. Genetic diversity of black gram landraces as evaluated by AFLP markers. *Curr. Sci.* 86:1411-1415.
- Souframanien J, Gopalakrishna T. 2004. A comparative analysis of genetic diversity in blackgram genotypes using RAPD and ISSR markers. *Theor. Appl. Genet.* 109:1687-1693.
- Tams SH, Melchinger AE, Bauer E. 2005. Genetic similarity among European winter triticale elite germplasms assessed with AFLP and comparisons with SSR and pedigree data. *Plant Breed.* 124:154-160.
- Tomasello S, Álvarez I, Vargas P, Oberprieler C. 2015. Is the extremely rare Iberian endemic plant species *Castrilanthemum debeauxii* (Compositae, Anthemideae) a 'living fossil'? Evidence from a multi-locus species tree reconstruction. *Mol. Phylogenet. Evol.* 82:118-130.
- Turchetto C, Segatto ALA, Mäder G, Rodrigues DM, Bonatto S, Freitas LB. 2016. High levels of genetic diversity and population structure in an endemic and rare species: implications for conservation. *AoB Plants* 8:plw002.

- Ude G, Pillay M, Ogundiwin E. 2003. Genetic diversity in an African plantain core collection using AFLP and RAPD markers. *Theor. Appl. Genet.* 107:248-255.
- 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.
- Wang X, Gao P, Liu Y, Li H, Lu F. 2020. Predicting Thermophilic Proteins by Machine Learning. *Current Bioinformatics*, 15(10):493-502.
- Wang M, Li J, Zhang L, Yang RW, Ding CB, Zhou YH, Yin ZQ 2011. Genetic diversity among *Salvia miltiorrhiza* Bunge and related species using morphological traits and RAPD markers. *Journal of Medicinal Plants Research* 5(13):2687-2694.
- Wang B, Zhang Y, Chen CB, Li XL, Chen RY, Chen L. 2007 Analysis on genetic diversity of different *Salvia miltiorrhiza* geographical populations in China. *Chin. Med. J.* 32:1988-1991.
- Wang O, Zhang B, Lu L. 2009. Conserved region amplification polymorphism (CoRAP), a novel marker technique for plant genotyping in *Salvia miltiorrhiza*. *Plant Mol. Biol. Rep.* 27:139-143.
- Weising K, Nybom H, Wolff K, Kahl G. 2005. DNA Fingerprinting in Plants. Principles, Methods, and Applications. 2nd ed. CRC Press, Boca Rayton, 472 pp.
- Walker JB, Sytsma KJ, Treutlein J, Wink M. 2004. *Salvia* (Lamiaceae) is not monophyletic: implications for the systematics, radiation, and ecological specializations of *Salvia* and tribe Mentheae. *Amer. J. Bot.* 91(7):1115-1125. <https://doi.org/10.3732/ajb.91.7.1115>
- Xu L, Jiang S, Wu J, Zou Q. 2021. An in silico approach to identification, categorization and prediction of nucleic acid binding proteins. *Briefings in bioinformatics*, 22(3).
- Yousefiazar-Khanian M, Asghari A, Ahmadi J. 2016. Genetic diversity of *Salvia* species assessed by ISSR and RAPD markers. *Not. Bot. Horti. Agrobi.* 44:431-436.
- Zou Q, Xing P, Wei L, Liu B. 2019. Gene2vec: gene sub-sequence embedding for prediction of mammalian N6-methyladenosine sites from mRNA. *RNA (Cambridge)*, 25(2): 205-218.