Genetic variation among different populations of *Centaurea virgata* from Iran using Start Codon Targeted (SCoT) markers

Elaheh Ghasemi: MSc Graduate in Biochemistry, Department of Cell and Molecular Biology, Faculty of Chemistry, University of Kashan, Kashan, Iran

Zeinab Toluei✉: Assistant Prof., Department of Cell and Molecular Biology, Faculty of Chemistry, University of Kashan, Kashan, Iran (ztoluei@kashanu.ac.ir)

Abstract

The aim of the present study is the identification of the genetic variation of *Centaurea virgata* (*Asteraceae*) through Start Codon Targeted (SCoT) markers. Forty-two specimens of said species were collected from 11 different regions of Iran. Ten primers revealed 131 amplifications ranging from 200bp to 3kbp, of which 102 (77.86%) were polymorphic. Polymorphism information content (PIC) ranged from 0.37 to 0.50 with an average of 0.43, effective multiplex ration (EMR) from 1 to 4.11 and marker index (MI) from 0.006 to 0.59. Based on the analysis of molecular variance (AMOVA), the genetic variation within populations (72%) was higher than that of those among populations (28%). Overall, the highest mean for Nei's gene diversity (0.18), Shannon index (0.27) and percentage of polymorphic loci (47.83) were observed in W. Azerbaijan populations, similarly the highest mean of total heterozygosity (H_T) and subpopulation heterozygosity (H_S) were found to be 0.18 and 0.07 in W. Azerbaijan and Golestan populations, respectively. The high genetic differentiation (G_{ST} = 1) showed significant genetic variation in Razavi Khorasan, N. Khorasan, Kurdistan, and Hamedan populations. Neighbor-Joining and population structure analysis divided *C. virgata* populations into six main clusters. The current study showed that, SCoT marker was efficient in assessing the genetic variation among different populations of the studied species.

Keywords: *Asteraceae*, DNA fingerprinting, genetic distance, molecular marker, polymorphism

تنوع ژنتیکی بین جمعیت‌های مختلف گل‌گندم بوته‌ای در ایران با استفاده از نشانگرهای SCoT*

الهه قاسمی: دانش‌آموخته کارشناسی ارشد رشته بیوشیمی گروه زیست‌شناسی سلولی و مولکولی، دانشکده شیمی، دانشگاه کاشان، کاشان، ایران

زینب طلوعی✉: استادیار گروه زیست‌شناسی سلولی و مولکولی، دانشکده شیمی، دانشگاه کاشان، کاشان، ایران (ztoluei@kashanu.ac.ir)

خلاصه

هدف از مطالعه حاضر، شناسایی تنوع ژنتیکی گیاه گل‌گندم بوته‌ای (*Centaurea virgata* Lam.) متعلق به کاسنیان از طریق نشانگر SCoT است. به این منظور، ۴۲ نمونه از گونه مذکور از ۱۱ منطقه مختلف ایران جمع‌آوری شد. ده پرایمر ۱۳۱ باند از اندازه ۲۰۰ تا ۳۰۰۰ جفت باز را ایجاد کردند که ۱۰۲ باند (۷۷/۸۶٪) چندشکل بودند. محتوای اطلاعاتی چندشکلی (PIC) ۰/۳۷ تا ۰/۵۰ با میانگین ۰/۴۳، نسبت چندگانه مؤثر (EMR) از ۱ تا ۴/۱۱ و شاخص نشانگر (MI) از ۰/۰۰۶ تا ۰/۵۹ متغیر بود. براساس تجزیه و تحلیل واریانس مولکولی (AMOVA) تنوع ژنتیکی، در درون جمعیت‌ها (۷۲٪) بیشتر از بین جمعیت‌ها (۲۸٪) وجود داشت. در مجموع، بیشترین میانگین تنوع ژنی Nei (۰/۱۸)، شاخص شانون (۰/۲۷) و درصد جایگاه‌های چندشکلی (۴۷/۸۳) در جمعیت‌های آذربایجان‌غربی مشاهده شد. همچنین، بیشترین میانگین هتروزیگوتی کل (H_T) و هتروزیگوتی درون جمعیتی (H_S) به ترتیب با میزان ۰/۱۸ و ۰/۰۷ در جمعیت‌های آذربایجان‌غربی و گلستان به دست آمد. تمایز ژنتیکی بالا (G_{ST} = 1) تنوع ژنتیکی قابل‌توجهی را در جمعیت‌های خراسان‌رضوی، خراسان‌شمالی، کردستان و همدان نشان داد. آنالیز خوشه‌ای با روش NJ و تحلیل ساختار جمعیت، جمعیت‌های *C. virgata* را به شش خوشه اصلی تقسیم کرد. مطالعه حاضر نشان داد که نشانگر SCoT در ارزیابی تنوع ژنتیکی در بین جمعیت‌های مختلف گونه مورد بررسی کارآمد است.

واژه‌های کلیدی: انگشت‌نگاری دی‌اِن‌اِی، چندشکلی، فاصله ژنتیکی، کاسنیان، نشانگر مولکولی

* مستخرج از پایان‌نامه نگارنده نخست به راهنمایی خانم دکتر زینب طلوعی ارایه شده به دانشگاه کاشان

Introduction

The *Asteraceae* is one of the largest families with 1600 to 1700 genera and 24000 to 30000 species (Garcia-Jacas *et al.* 2000, Funk *et al.* 2005). *Centaurea* L. (knapweed) is one of the genera of said family that contains about 771 species worldwide (POWO 2024), which are mainly distributed around the world especially in Mediterranean region and W. Asia (Hellwig 2004). This genus consists of 88 species in the Flora Iranica (Rechinger 1987).

Genetic diversity is one of the fundamental and basic sources of biodiversity (Quinones-Perez *et al.* 2014) and its protection is vital for long-term survival of species in changing environments. Among various populations, genetic diversity is affected by different factors, such as geographic variations, breeding systems, dispersal mechanisms etc. (Huang *et al.* 2016). The changes of environmental conditions are the main causes of alteration in genetic diversity levels among various populations (Lovejoy & Hannah 2005). Information for evaluation of the genetic diversity is obtained from different factors, such as morphological, biochemical and molecular markers. Therefore, new approaches in molecular biology have been offered which extended a platform for analyzing the genetic diversity at the genome level and might be employed for assessment of inter-species or intra-species ecological, taxonomical, morphological, evolutionary, and phylogenetic relationships (Agarwal *et al.* 2008). Molecular markers play a significant role in protection of biodiversity, quantitative mapping (QTL), identification of promising cultivars and etc. (Khanam *et al.* 2012). Currently, several PCR based dominant markers, such as restriction fragment-length polymorphisms (RFLPs), amplified fragment-length polymorphisms (AFLPs), random amplified polymorphic DNA (RAPDs), single-nucleotide polymorphisms (SNPs), and simple sequence repeats (SSRs) have been employed for determination of genetic variation in plants (Kazemeini *et al.* 2020, Nurmansyah Alghamdi *et al.* 2020, Adhikari *et al.* 2022, Wu *et al.* 2022, Abd-Dada *et al.* 2023, Kader *et al.* 2023, Burridge *et al.* 2024, Wang *et al.* 2024). A novel molecular marker system known as Start Codon Targeted (SCoT) was developed as a gene-targeted DNA marker. SCoT primers are extended based on the conserved region surrounding the translation initiation codon ATG in plant genes. SCoT markers employing the 18-mer single primer as the forward and reverse primers in PCR and an annealing of 50 °C (Collard & Mackill 2009, Mukhopadhyay 2016). SCoT is a dominant marker and could be employed for analysis of genetic variation. It is associated with functional genes and corresponding traits and does not require sequence information. The SCoT marker compared to other types of DNA molecular markers, such as RAPD, ISSR, and SSR, is more stable, finds high polymorphisms and produce more reproducible and reliable bands. Recently, SCoT markers have been popularly employed in plant genetic diversity assessment and phylogenetic studies (Rai 2023).

Centaurea virgata Lam. is a perennial plant species with multiple stems, rose purple flowers and a woody base. The extract of said plant possesses pharmacological properties with chemical components including sesquiterpenes, the flavonol isokaempferide and flavones (Etminan *et al*. 2018). Intraspecific variations of *C. virgata* collected from different regions of Iran have been evaluated based on morphological data (Ghasemi & Toluei 2020). There is no study on the genetic diversity of this species. However, study on genetic diversity, population genetics or relationship of other *Centaurea* species using molecular markers has been reported. RAPD markers have been utilized for analysis of genetic variation in *C. nivea* (Bornm.) Wagenitz (Sözen & Özaydın 2009), *C. wiedemanniana* Fisch. & C.A.Mey. (Sozen & Ozaydin 2010), *C. aspera* L. and *C. seridis* L. (Ferriol *et al.* 2012), *C. ultreiae* Silva-Pando (Mallón *et al.* 2010), and *C. lycaonica* Boiss. & Heldr. (Uysal *et al.* 2012). AFLP marker has been used for evaluating of genetic variation of *C. jacea* L. (Bassin *et al.* 2004) and *C. borjae* Valdés Berm. & Rivas Goday (Lopez & Barreiro 2013). Nuclear microsatellites were used to analyze genetic variation in *C. corymbosa* Pourr. (Freville *et al.* 2001), *C. horrida* Bad. (Mameli *et al.* 2008),

Centaurea subsect. *Phalolepis* (López-Vinyallonga *et al.* 2015), and *C. alba* L. (Jèssica *et al.* 2020). ISSR primers were used to measure the level of genetic differences in *C. lycaonica* (Uysal *et al.* 2012) and *C. amaena* Boiss. & Balansa (Atasagun 2022). CAAT box-derived polymorphism (CBDP) and SCoT polymorphism, were used to analyze the genetic variation between eight wild *Centaurea* species in Egypt (Atia *et al.* 2021). In addition, genetic relationships beetween ten *Centaurea* species from Iraq have been evaluated with SCoT marker (Ismail *et al.* 2024), Nuclear DNA sequences ETS, and low-copy genes AGT1 and PgiC were analyzed for delimitation of species in *C. tentudaica* Rivas Goday & Rivas Mart. (Moreyra *et al.* 2022) and *C. calocephala* Willd. complex (Novakovic´ *et al.* 2022). There is no study about intraspecific genetic variation of *C. virgata*. The present investigation is the first attempt to determine the potential of SCoT marker method to evaluate the degree of genetic diversity among various populations of *C. virgata* from different regions of Iran.

Materials and Methods

- Plant materials

Forty-two specimens of *C. virgata* from 11 different provinces of Iran and three other species of *Centaurea* (*C. pulchella* Ledeb., *C. solstitialis* L., and *C. persica* Boiss.) as out-groups were collected from different geographic regions of Iran (Table 1).

Table 1. List of the studied *Centaurea virgata* populations along with related data

Population	Locality, altitude, and voucher specimen (UKH)	Individual No.	Abbreviation
C. virgata	Lorestan Prov.: Dorood, before Saravand village; 2125 m; Toluei 1021 (UKH)	2	dor1021
	Lorestan Prov.: Nurabad; 1855 m; Toluei 1031 (UKH)	2	nur1031
C. virgata	Isfahan Prov.: Kashan, Ghohrud village; 2235 m; Toluei 1050 (UKH)	3	gho1050
	Isfahan Prov.: Kashan, Alavi village; 1730 m; Toluei 1052 (UKH)	1	alv1052
	Isfahan Prov.: Kashan, University of Kashan; 975 m; Toluei 1053 (UKH)	1	uni1053
	Isfahan Prov.: Kashan, Eznaveh; 2690 m; Toluei 1054 (UKH)	3	ezn1054
	Isfahan Prov.: Kashan, Barzok, Vishang; 2896 m; Toluei 1058 (UKH)	3	brz1058
C. virgata	Qazvin Prov.: Avaj to Mahnian; 2205 m; Toluei 1048 (UKH)	2	qav1048
C. virgata	Golestan Prov.: Maravehtapeh to Bojnourd, 35 km after Maravehtapeh; 1184 m; Toluei & Ranjbar 1025 (UKH)	2	gol1025
C. virgata	Khorasan Razavi Prov.: Kalat to Mashhad; 1734 m; Toluei & Ranjbar 1028 (UKH)	1	kal1028
	Khorasan Razavi Prov.: Neyshabur to Kashmar, 10 km to Rivash, 15 km to Kashmar; 2032 m; Toluei & Ranjbar 1029 (UKH)	1	ney1029
C. virgata	N. Khorasan Prov.: Shirvan, Kouseh bifurcate; 1697 m; Toluei & Ranjbar 1027 (UKH)	1	shi1027
	N. Khorasan Prov.: Esfarayen to Bojnurd, Asadli neck; 1718 m; Toluei & Ranjbar 1030 (UKH)	1	esf1030

C. virgata	Kurdistan Prov.: Marivan to Saqqez, 65 km after Marivan, between Aqjeh and Qamjian; 1799 m; Toluei & Ranjbar 1032 (UKH)	1	mrv1032
	Kurdistan Prov.: Bijar, Khosroabad; 1710 m; Toluei 1033 (UKH)	1	bik1033
	Kurdistan Prov.: Bijar to Zanjan, Shirinbolagh bifurcate; 1626 m; Toluei 1034 (UKH)	1	biz 1034
	Kurdistan Prov.: Sanandaj to Divandareh, 30 km to Divandareh, before Aq Bolagh village; 1987 m; Toluei 1037 (UKH)	1	san1037
C. virgata	Zanjan Prov.: Zanjan, Mahneshan, bifurcate of Hasanabad and Hussainabad villages; 2083; Toluei 1035 (UKH)	1	mah1035
	Zanjan Prov.:Zanjan, Mahneshan to Halab; 1772 m; Toluei 1036 (UKH)	2	hal1036
C. virgata	W. Azerbaijan Prov.: Bukan to Mahabad, 25 km to Mahabad; 1805 m; Toluei 1038 (UKH)	1	buk1038
	W. Azerbaijan Prov.: Oshnavieh to Orumieh, 5 km after Aq Bolagh village; 2181 m; Toluei 1039 (UKH)	1	osh1039
	W. Azerbaijan Prov.: Oshnavieh to Orumieh, 2–3 km to Jarabad village; 1828 m; Toluei 1040 (UKH)	1	osh1040
	W. Azerbaijan Prov.: Orumieh, after Silvaneh, after Toly village; 1665 m; Toluei & Ranjbar 1041 (UKH)	2	oru1041
	W. Azerbaijan Prov.: Orumieh, Movana to Neychalan; 1715 m; Toluei & Ranjbar 1044 (UKH)	1	oru1044
	W. Azerbaijan Prov.: Chaldoran, Alimardan village; 2005 m; Toluei & Ranjbar 1045 (UKH)	2	chl1045
	W. Azerbaijan Prov.: Maku, Baduli village; 1929 m; Toluei & Ranjbar 1047 (UKH)	1	Mak1047
C. virgata	Tehran Prov.: 5 km after Polur-Firoozkooh bifurcate; 2271 m; Toluei 1024 (UKH)	1	Teh1024
C. virgata	Hamedan Prov.: Malayer; 1813 m; Toluei 1061 (UKH)	1	ham1061
	Hamedan Prov.: Malayer; 1794 m; Toluei 1062 (UKH)	1	ham1062
C. solstitialis	W. Azerbaijan Prov.: Oshnavieh to Orumieh, 2–3 km to Jarabad village; 1828 m; Toluei 1062 (UKH)	1	sol1062
C. pulchella	W. Azerbaijan Prov.: Maku, Baduli village; 1929 m; Toluei & Ranjbar 1063 (UKH)	1	pul1063
*C. persic*a	W. Azerbaijan Prov.: 30 km to Orumieh, after Silvaneh; 1680 m; Toluei & Ranjbar 1064 (UKH)	1	per1064

* UKH: University of Kashan Herbarium (Kashan, Iran)

- Genomic DNA extraction

Total genomic DNA was isolated from dried leaves of each sample plant according to modified CTAB (Cetyl Trimethyl Ammonium Bromide) method (Doyle & Doyle 1987). The quantity and quality of genomic DNA were determined using UV spectrophotometer and 1% agarose gel electrophoresis, respectively.

- SCoT PCR amplification

All PCR reactions were performed in a total volume of 15 µL using a thermal cycler instrument (Biorad, USA). Each reaction contains 1 µL genomic DNA template, 1.5 µL 10x Reaction buffer (Sinagene, Iran), 1 µL 10 pM of each primer (Table 2) (Urofine, UK), 1 µL $Mgcl_2$ 50 Mm (Sinagene, Iran), 0.2 µL 5 U/µL Taq DNA Polymerase (Sinagene, Iran), 8 µL sterile ddH2O, 1 µL BSA 20 Mm (Merck, Germany), and 0.5 µL dNTP 10 Mm mixture (Sinagene, Iran). PCR

amplification were performed with a preliminary cycle and initial denaturation at 94 °C for 3 min, 36 cycles of 94 °C for 1 min, annealing temperature of primers for 1 min and 50 °C, extension at 72 °C for 2 min, and final extension at 72 °C for 5 min. For SCoT marker profiling, the amplification products were resolved in 1% agarose gels electrophoresis in 1X TAE buffer solution. The gel was photographed by a Gel Doc (TM) XR System (Molecular Imager GelDOC XR$^+$, USA).

- Data scoring and statistical analysis

Out of the eighteen SCoT primers tested, ten SCoT primers with clear and polymorphic bands were selected for final analysis (Table 2). The amplified bands produced by ten SCoT markers were scored as presence (1) or absence (0) of bands to create a binary matrix. The following genetic diversity parameters were calculated: Initially, by observing the banding patterns created by SCoT primers, total number of bands (TNB), number of polymorphic bands (NPB) and percentage of polymorphism band (PPB) were obtained. Further, potential of these molecular markers for estimation of genetic variation was assessed by measuring polymorphism information content (PIC), effective multiplex ratio (EMR) and marker index (MI). PIC values were calculated using the formula PIC = $1 - \sum_{i=1}^{n} f_i^2$, where f_i^2 is the frequency of the i^{th} allele (Smith *et al*. 1997). Marker index (MI) is the primer capability for detection of the polymorphic loci among various genotypes and was measured as EMR×PIC, where, EMR is the output of number of polymorphic loci and fraction of polymorphic loci. Genetic diversity was calculated by different parameters, such as observed number of alleles (N_a), effective number of alleles (Ne), Nei's gene diversity (H), Shannon's information index (I), total heterozygosity (H_T), subpopulation heterozygosity (H_s), genetic differentiation (G_{ST}) and gene flow (N_m), which were calculated using POPGENE (Ver. 1.32) software (Yeh 1997). The analysis of molecular variance (AMOVA) was conducted within and among populations, employing GenAlEx 6.41 software (Peakall & Smouse 2012). This software was used for measurement of Nei's genetic distance. The cluster was obtained using the PAUP software (Ver. 4) (Swofford 2003) with neighbor-joining (NJ) method [NJ search setting: Ties (if encountered) will be broken systematically, Distance measure = mean character difference]. The tree was visualized using Tree View software (Page 1996). The dominant-marker model in STRUCTURE software (Ver. 2.3.4) was employed to analyze the population structure evaluated according to the Markov chain Monte Carlo (MCMC) algorithm (Falush *et al.* 2003, Pritchard *et al.* 2000). The following parameter setting was applied: length of burning period: 10,000; Markov Chain Monte Carlo replication (MCMC Reps): 10,000 (admixture model and correlated allele frequency). A total, for each value of K (from 1 to 8), 20 independent runs were performed (Evanno *et al.* 2005).

Results

- Marker informativeness

The SCoT fingerprinting was performed to analyze the genetic variation of populations of *C. virgata* in different regions of Iran. A total of 131 distinguishable bands were generated from 10 SCoT primers, out of which 102 bands were polymorphic. The number of total bands produced by each primer ranged from 9 (SCoT 1) to 17 (SCoT 30, SCoT 31, and SCoT 14). Polymorphism percentage ranged from 45.45% (SCoT 13) to 100% (SCoT 9), with an average polymorphism of 77.86%. The size of the amplified bands ranged from 200bp to 3kbp. All of the selective primers demonstrated a high degree of reproducibility, and banding patterns were the same in each PCR repeats. The PIC values of these 10 SCoT primers ranged from 0.37 to 0.50 with an average of 0.43. Primer SCoT 41 revealed the highest discrimination power with a PIC value of 0.50, whereas SCoT 9 had the lowest PIC value of 0.37. Since the PIC values showed the differentiation ability of the primer, these 10 SCoT primers were able to efficiently differentiate among the populations

of *C. virgata*. Mean PIC, EMR, and MI values generated by SCoT primers were 0.43, 1.31, and 0.46, respectively. Among different SCoT primers used, primer SCoT 9 revealed minimum values of MI (0.06) (Table 2).

Table 2. Different selected SCoT primers for measuring the polymorphism among different populations of *Centaurea virgata*

Primer	Sequence (5′-3′)	TNB	NPB	PPB%	PIC	EMR	MI
SCOT1	CAACAATGGCTACCACCA	8	7	87%	0.47	1	0.55
SCOT9	CAACAATGGCTACCAGCA	9	9	100%	0.37	4	0.06
SCOT13	ACGACATGGCGACCATCG	11	5	45%	0.42	1	0.49
SCOT14	ACGACATGGCGACCACGC	17	13	76%	0.43	1	0.52
SCOT15	ACGACATGGCGACCGCGA	14	11	78%	0.40	1	0.46
SCOT16	ACGACATGGCGACCGCGA	19	13	68%	0.39	1	0.46
SCOT30	CCATGGCTACCACCGGCG	17	12	70%	0.43	1	0.52
SCOT31	CCATGGCTACCACCGCCT	17	15	88%	0.40	1	0.47
SCOT36	GCAACAATGGCTACCACC	10	9	90%	0.48	1	0.54
SCOT41	CAATGGCTACCACTGACA	10	9	90%	0.50	1	0.59
Total	-	131	102	-	-	-	-
Mean	-	-	-	77.86%	0.43	1.31	0.46

TNB: Total number of bands, NPB: Number of polymorphic bands, PPB: Percentage of polymorphic bands, PIC: Polymorphism information content, EMR: Effective multiplex ratio, MI: Marker index.

- Genetic diversity

Different diversity indices were measured, including observed number of alleles (Na), Effective number of alleles (Ne), Nei's gene diversity (H), Shannon's index (I), number of polymorphic loci (NPL), percentage of polymorphic loci (PPL), total heterozygosity (H_T) subpopulation heterozygosity (H_S), genetic differentiation (G_{ST}) and gene flow (N_m). The highest Na (1.48 and 1.47), Ne (1.31 and 1.28), Nei's gene diversity (0.18 and 0.17), Shannon information index (0.27 and 0.25) were observed among West Azerbaijan and Kashan populations, respectively. In addition, the highest total heterozygosity (0.18 and 0.17), PPL (47.83% and 47.20%), and NPL (77 and 76) were found in W. Azerbaijan and Kashan, respectively. The ratio of N_e/N_a was calculated within the populations of each province. The H_T value was from zero to 0.18. The H_s value was from zero to 0.07 (Golestan). A relatively high genetic differentiation (Gst = 1) and a low gene flow (N_m = 0) were observed among Razavi Khorasan, N. Khorasan, Kurdistan, and Hamedan provinces (Table 3).

Table 3. Genetic diversity parameters of *Centaurea virgata* populations based on analysis of SCoT marker data with POPGENE (Ver. 1.32) software

Province	N_a	N_e	H	I	N_e / N_a	NPL	PPL	H_T	H_s	G_{ST}	N_m
Lorestan	1.24	1.2	0.1	0.15	0.97	38	23.60	0.1	0.03	0.72	0.2
Kashan	1.47	1.28	0.17	0.25	0.87	76	47.20	0.17	0.04	0.78	0.14
Qazvin	1.06	1.04	0.026	0.038	0.98	10	6.21	0.026	0.026	0	-
Golestan	1.17	1.12	0.07	0.11	0.96	28	17.39	0.07	0.07	0	-
Razavi Khorasan	1.13	1.13	0.07	0.09	1	21	13.04	0.07	0	1	0
N. Khorasan	1.2	1.2	0.1	0.14	1	32	19.88	0.1	0	1	0
Kurdistan	1.34	1.25	0.14	0.20	0.93	54	33.54	0.14	0	1	0
Zanjan	1.14	1.11	0.061	0.09	0.97	23	14.29	0.063	0.01	0.84	0.098
W. Azerbaijan	1.48	1.31	0.18	0.27	0.89	77	47.83	0.18	0.02	0.90	0.05
Tehran	1	1	0	0	1	0	0	0	0	-	-
Hamedan	1.14	1.14	0.07	0.1	1	22	13.66	0.07	0	1	0

N_a = Observed number of alleles, N_e = Effective number of alleles, h = Nei's gene diversity, I = Shannon's information index, NPL: Number of polymorphic loci, PPL: Percentage of polymorphic loci, H_T = Total heterozygosity, H_S = Subpopulation heterozygosity, G_{ST} = Genetic differentiation, N_m = Gene flow.

- Analysis of molecular variance

The AMOVA analysis of the SCoT markers showed that, 28% of the total genetic variance was distributed among populations. A relatively high proportion of genetic diversity was attributable to differences within populations (72%) AMOVA represented the existence of high differentiation (Phi_{PT} = 0.276) among the populations of *C. virgata* (Table 4).

Table 4. Analysis of molecular variance (AMOVA) using SCoT molecular markers in population of *Centaurea virgata*

Source	Degree of freedom	Sum of square	Mean square	Variance of component	%age of variation	P value	Φ_{PT}
Among pops	10	315/935	31/593	5.017	28%	<0.01	-
Within pops	32	420/298	13/134	13.134	72%	<0.01	-
Total	42	736/233	-	18.152	100%	-	0.276

Φ_{PT} = AP / (WP + AP) = AP / TOT
Key: AP = Est. Var. Among Pops, WP = Est. Var. Within Pops

- Cluster analysis

A dendrogram was generated by Neighbor joining (NJ) method using PAUP software based on 131 polymorphic SCoT fragments. It showed six main groups in *C. virgata* populations (Fig 1) as follows: Group 1: Lorestan (dora1021, dorb1021), Tehran, and Golestan populations; Group 2: Lorestan (nura1031 and nurb1031), Zanjan, and Kurdistan populations; Group 3: W. Azerbaijan (Mak1047, Buk1038, Orua1041, Orub1041, Chla1045, and Chlb1045) populations; Group 4: Kashan and Hamedan populations; Group 5: Razavi Khorasan and N. Khorasan populations; and Group 6: Qazvin and W. Azerbaijan (Oru1044, Osh1039, and Osh1040) populations. The other species including *C. solstitialis*, *C. pulchella*, and *C. persic*a were placed in independent out-groups in separate clades. This was an indication of the high efficiency of the SCoT markers to distinguish among species.

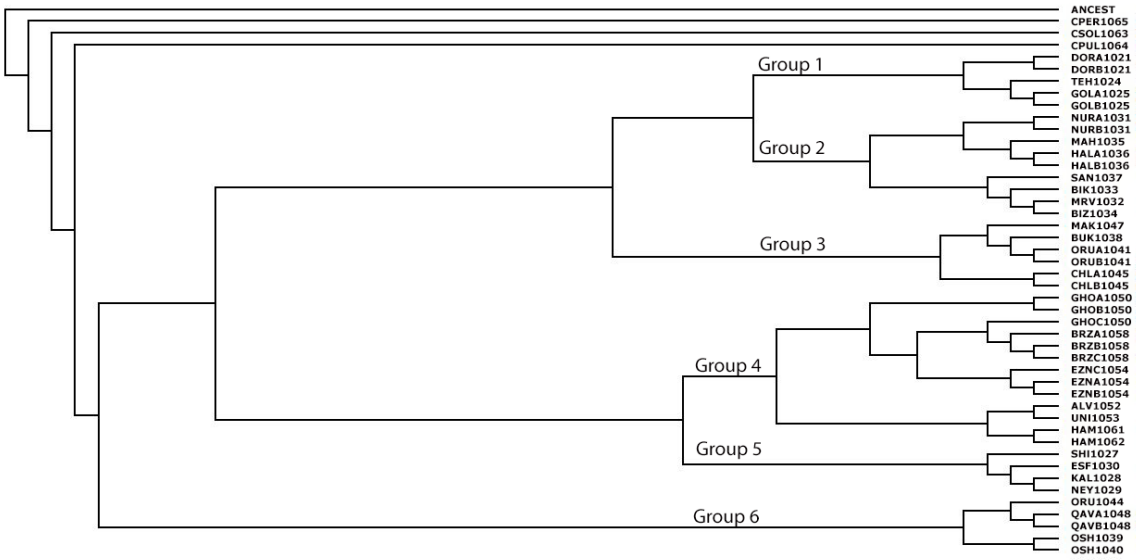

Fig. 1. Cluster analysis based on Neighbor joining (NJ) method showing genetic relationships among 42 specimens of *Centaurea virgata* collected from 11 different regions of Iran based on SCoT molecular markers.

Table 5. The pairwise population matrix of Nei genetic distance between different populations of *Centaurea virgata* based on 131 bands bvia GenAlEx 6.41

Subpopulation	Dora 1021	Dorb 1021	Nura 1031	Nurb 1031	Ghoa 1050	Ghob 1050	Ghoc 1050	Alv 1052	Uni 1053	Ezna 1054	Eznb 1054	Eznc 1054	Brza 1058	Brzb 1058	Brzc 1058	Qava 1048	Qavb 1048	Gola 1025	Golb 1025	Kal 1028	Ney 1029
Dora1021	0																				
Dorb1021	14	0																			
Nura1031	28	28	0																		
Nurb1031	29	25	9	0																	
Ghoa1050	35	41	33	38	0																
Ghob1050	36	40	26	31	11	0															
Ghoc1050	31	35	31	36	30	21	0														
Alv1052	33	43	37	40	44	39	34	0													
Uni1053	31	37	35	34	42	37	32	12	0												
Ezna1054	28	26	28	33	37	34	27	33	35	0											
Eznb1054	27	31	33	40	30	29	20	30	30	11	0										
Eznc1054	29	33	31	34	34	29	24	32	30	13	10	0									
Brza1058	34	36	24	33	37	28	21	39	41	30	31	31	0								
Brzb1058	29	37	23	28	38	33	28	38	38	29	32	30	17	0							
Brzc1058	32	38	28	33	35	30	25	35	37	32	33	33	18	7	0						
Qava1048	32	34	32	31	45	42	39	33	35	40	39	37	36	43	44	0					
Qavb1048	36	38	32	33	45	40	37	33	35	42	41	39	38	43	42	10	0				
Gola1025	27	29	35	30	44	45	48	48	42	33	38	36	43	44	47	41	45	0			
Golb1025	33	35	29	26	42	37	38	48	44	41	40	34	41	34	35	43	39	28	0		
Kal1028	31	29	29	28	34	29	28	40	38	31	32	36	39	40	37	39	37	38	36	0	
Ney1029	38	36	24	29	37	32	31	41	37	36	37	39	32	35	32	32	34	41	41	21	0
Shi1027	33	39	33	36	40	35	36	40	38	37	36	40	37	32	31	45	43	40	36	32	35
Esf1030	35	39	27	30	40	37	28	38	34	33	34	34	35	34	35	43	41	48	42	28	25
Mrv1032	38	40	26	29	45	44	41	45	39	40	39	35	38	37	36	46	42	43	35	43	38
Bik1033	35	35	29	28	46	43	38	50	42	41	44	38	39	40	39	41	41	36	36	40	35
Biz1034	34	36	22	27	43	42	37	37	33	38	35	37	38	33	34	40	36	39	35	37	32
San1037	32	34	30	29	41	42	41	47	43	40	43	41	44	39	38	44	46	39	37	35	34
Mah1035	33	31	23	20	44	39	38	36	34	41	38	36	41	36	37	35	37	40	32	30	31

Hala1036	31	31	25	20	44	39	38	42	38	41	40	32	41	34	37	33	31	40	28	32	33
Halb1036	29	29	25	20	44	41	36	40	36	39	38	36	37	34	33	29	29	36	30	30	29
Buk1038	27	31	23	24	34	31	32	38	36	33	32	34	35	32	31	37	37	38	36	28	33
Osh1039	40	44	28	31	47	42	45	41	39	44	45	37	40	39	40	28	28	47	39	39	34
Osh1040	35	39	35	34	44	41	36	38	38	43	40	38	43	48	45	33	35	40	42	40	39
Orua1041	39	35	35	32	40	39	42	42	42	37	38	36	41	42	37	41	39	38	40	42	43
Orub1041	36	36	34	31	45	42	45	45	39	44	43	39	42	39	36	38	42	33	33	43	44
Oru1044	31	39	25	30	36	37	38	36	40	41	40	42	35	36	39	25	27	40	40	38	39
Chla1045	23	35	29	32	32	33	36	34	32	37	36	38	35	30	29	41	41	32	40	34	35
Chlb1045	29	33	19	22	38	33	44	44	38	35	42	32	33	28	33	37	39	30	28	36	35
Mak1047	33	39	31	34	40	37	36	38	38	37	36	30	35	36	35	35	39	40	42	40	41
Teh1024	34	30	28	27	39	36	43	47	43	40	41	37	40	39	36	36	34	31	27	31	32
Ham1061	34	36	36	41	31	32	31	35	31	32	27	35	36	43	42	34	30	43	49	35	34

Tabla 5 (contd)

Subpopulation	**Shi 1027**	**Esf 1030**	**Mrv 1032**	**Bik 1033**	**Biz 1034**	**San 1037**	**Mah 1035**	**Hala 1036**	**Halb 1036**	**Buk 1038**	**Osh 1039**	**Osh 1040**	**Orua 1041**	**Orub 1041**	**Oru 1044**	**Chla 1045**	**Chlb 1045**	**Mak 1047**	**The 1024**	**Ham 1061**
Dora1021																				
Dorb1021																				
Nura1031																				
Nurb1031																				
Ghoa1050																				
Ghob1050																				
Ghoc1050																				
Alv1052																				
Uni1053																				
Ezna1054																				
Eznb1054																				
Eznc1054																				

Brza1058																				
Brzb1058																				
Brzc1058																				
Qava1048																				
Qavb1048																				
Gola1025																				
Golb1025																				
Kal1028																				
Ney1029																				
Shi1027	0																			
Esf1030	32	0																		
Mrv1032	43	35	0																	
Bik1033	46	40	29	0																
Biz1034	39	31	24	29	0															
San1037	37	39	34	31	34	0														
Mah1035	42	34	31	36	23	35	0													
Hala1036	40	36	33	32	27	33	20	0												
Halb1036	40	34	33	28	21	29	18	8	0											
Buk1038	38	34	37	38	33	35	28	26	24	0										
Osh1039	43	37	30	39	30	38	29	25	25	31	0									
Osh1040	42	38	35	36	35	37	34	38	32	40	33	0								
Orua1041	42	44	45	42	37	41	36	32	28	26	35	38	0							
Orub1041	41	47	38	41	32	42	35	29	25	23	32	37	23	0						
Oru1044	38	38	33	38	37	39	40	34	32	32	27	30	42	37	0					
Chla1045	32	34	41	36	35	39	38	36	30	24	43	38	38	31	34	0				
Chlb1045	32	38	31	32	37	31	34	28	30	28	27	34	32	29	30	24	0			
Mak1047	42	40	37	34	37	41	30	34	36	24	29	36	32	29	36	30	26	0		
Teh1024	33	35	34	39	32	40	29	29	25	37	32	33	35	36	35	37	31	43	0	
Ham1061	39	37	42	47	40	54	43	45	41	39	40	35	35	44	35	33	39	37	40	0

- Genetic distances between populations

The estimates of genetic distance (Nei's measure) between pairs of populations were calculated based on 131 bands via GenAlEx 6.41. Scored values ranged from 0.077 [between Alavi (Kashan) and University of Kashan populations] to 0.409 (between Hamedan and Sanandaj) (Table 5).

- Structure analysis

The number of subpopulations was determined, based on the best K value generated. The estimated membership fraction ranged from *K*1 to *K*8 and the maximum log-likelihood value obtained at $K = 6$, indicating six possible clusters in *C. virgata* through SCoT molecular markers (Fig. 2).

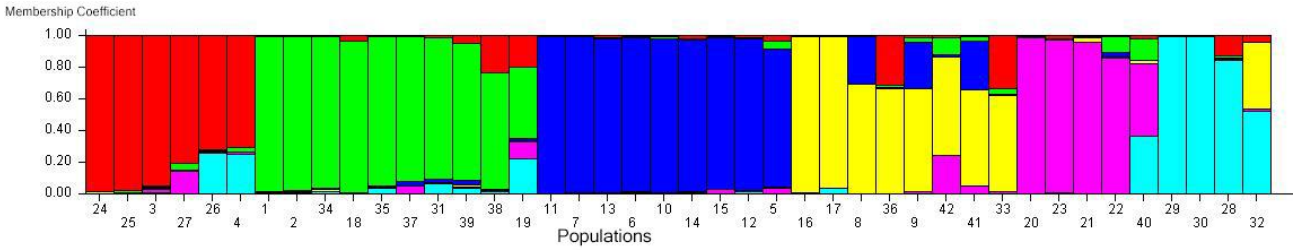

Fig. 2. Graphical summary of population structure analysis using STRUCTURE for 42 samples. Maximum number of subpopulations was inferred at K = 6 (1. dora1021, 2. dorb1021, 3. nura1031, 4. nurb1031, 5. ghoa1050, 6. ghob1050, 7. ghoc1050, 8. alv1052, 9. uni1053, 10. ezna1054, 11. eznb1054, 12. eznc1054, 13. barza1058, 14. barzb1058, 15. barzc1058, 16. qava1048, 17. qavb1048, 18. gola1025, 19. golb1025, 20. kal1028, 21. ney1029, 22. shi1027, 23. esf1030, 24. mrv1032, 25. bik1033, 26. biz1034, 27. san1037, 28. mah1035, 29. hala1036, 30. halb1036, 31. buk1038, 32. osh1039, 33. osh1040, 34. orua1041, 35. orub1041, 36. oru1044, 37. chla1045, 38. chlb1045, 39. mak1047, 40. teh1024, 41. ham1061, 42. ham1062).

Discussion

This is the first report on the identification and genetic comparison among different populations of *C. virgata* using SCoT markers. In the present study, a total of 131 sharp bands were identified through the amplification of 10 SCoT primers in different populations of *C. virgata*. The result of this investigation showed the SCoT molecular marker efficiency in evaluating the genetic variation among the populations of *C. virgata* according to different genetic diversity indices, including PIC, EMR and MI values.

A dendrogram was generated with Neighbor joining (NJ) method based on 131 polymorphic SCoT fragments, which showed six main groups. Among the groups of some populations, including Tehran, Golestan, Zanjan, Kurdistan, Kashan, Hamedan, Razavi Khorasan, and N. Khorasan, the relationship between genetic distance and geographical distance was significantly clear. However, for populations of Lorestan and W. Azerbaijan, the relationship between genetic distance and geographical distance was not significant. Therefore, the populations from each province of the Lorestan and W. Azerbaijan are separated in different groups. The high genetic variation of the populations might be probably due to ecological reasons, environmental conditions and etc.

Moreover, this clustering among 11 populations of *C. virgata* was also established by Bayseian clustering algorithm using STRUCTURE software. The delta *K* method was found to be best at K = 6, which described types of diversity in the clustering.

The PCR amplified product was used for genetic distance index's coefficient calculation that was based on Nei matching coefficient. The scored values ranged from 0.077 [between Alavi (Kashan) and University of Kashan populations] to 0.409 (between Hamedan and Sanandaj). A value <0.1 means small genetic distances, 0.10–0.15 means

moderate genetic distances, 0.15–0.2 means high genetic distances, and > 0.2 means very high genetic distances (Nei & Li 1979) (in the present studied populations, small to high genetic distances were observed).

The PIC value is useful for marker informativeness and is a substantial factor to recognize population genetic diversity. According to Botstein *et al.* (1980), for codominant markers, when the PIC value is higher than 0.5, it is indicative of high polymorphism whereas when PIC values range between 0.25 and 0.5, it demonstrates medium polymorphism. The lowest polymorphism is indicated by PIC values below 0.25. If the PIC value is zero, there is no allelic variation and PIC value of 1.0 is the maximum value (Gulsen *et al.* 2009). For dominant markers, the PIC value ranges from zero for monomorphic markers to 0.5 for markers present in 50% of individuals and absent in the remaining 50%. When the PIC value is 0 to 0.10, it is indicative of low infarmativeness, 0.10 to 0.25 medium, 0.30 to 0.40 high and 0.40 to 0.50 for very high informativeness (Serrote *et al.* 2020). In the present study, PIC ranged from 0.37 to 0.50 with an average of 0.43. Therefore, the populations showed high to very high polymorphism in the present research.

The SCoT markers have been used in genetic diversity analysis and diagnostic fingerprinting in some species of the *Asteraceae*, such as *Taraxacum* sect. *Erythrosperma* species from Poland with 19 SCoT primers. The average percentage of polymorphism was 94% (Wolanin *et al.* 2023), *Lactuca sativa* L. and *L. serriola* L. with exhibiting a notable average polymorphism of 67.55% with 5 SCoT primers (Essa *et al*. 2024).

About using SCoT markers in *Centaurea* species the genetic variation between eight wild *Centaurea* species in Egypt were evaluated with seventeen SCoT primers. These primers generated 80.2 polymorphic amplicons. The values range of the PIC was between 0.364 and 0.482, with an average of 0.387. The Rp values were between 0.123 and 0.864 (Atia *et al.* 2021). In addition, genetic relationships of ten *Centaurea* species growing naturally in the Duhok City, Kurdistan region of Iraq were studied with 10 SCoT primers. The polymorphism percentage was 100% in all primers. The PIC value was ranged from 0.24 to 0.36 with an average of 0.319. The Rp values was ranged from 3.4 to 12 with an average of 5.74. Thus, the SCoT markers were indicated as an efficient marker for genetic variation analysis for the *Centaurea* (Ismail *et al.* 2024).

Despite the SCoT molecular markers studies on other species of *Centaurea*, the present investigation is the first study about genetic variation of *C. virgata*. Moreover, here a considerable genetic variation (such as other genetic studies on other species of *Centaurea*) existed among populations of *C. virgata*. In addition, the average polymorphic value was here found to be 77.86%. The AMOVA analysis of the SCoT markers showed a high genetic diversity between populations (Φ_{PT} = 0.276). The AMOVA analysis showed 28% of the total genetic variance among populations of *C. virgata* and the highest Na (1.48), Ne (1.31), Nei's gene diversity (0.18), Shannon information index (0.27) gene diversity index (H_T) (0.18), PPL (47.83%) and NPL (77) were found in W. Azerbaijan populations.

The SCoT marker was an effective tool to estimate the genetic diversity of *C. virgata* populations. Moreover, a high number of polymorphic bands and high polymorphic fragment percentage was obtained. The investigations showed the capability of SCoT marker in diversity analyses and fingerprinting. Therefore, it can be used for detecting the genetic variation in other species of *Centaurea*.

Conclusion

This was the first report of genetic variation and population structure study on different populations of *C. virgata* using SCoT marker technique. The SCoT method was reliable and very suitable for characterization and evaluation of genetic relationships among different populations of *C. virgata*. The results of the current study showed a significant genetic diversity in the studied populations. In addition, understanding the genetic diversity among populations of *C.*

virgata as a widespread and invasive species is important to select the efficient strategy for employing in breeding programs and germplasm conservation. SCoT marker targets a highly conserved region in plant genes that flanking the start codon, thus can determine genetic variations in a specific gene that link to a specific trait. It is a simple, highly polymorphic and reproducible molecular marker for which there is no need for prior sequence information. Therefore, SCoT is an efficient marker for evaluation of genetic variation in plant species.

Acknowledgments

The authors gratefully acknowledge the financial supports granted by the University of Kashan (Kashan, Iran).

References

Abd-dada, H., Bouda, S., Khachtib, Y. Bella, Y.A. & Haddioui, A. 2023. Use of ISSR markers to assess the genetic diversity of an endemic plant of Morocco (*Euphorbia resinifera* O. Berg). Journal of Genetic Engineering and Biotechnology 21(1): 91. DOI: 10.1186/s43141-023-00543-4.

Adhikari, S., Biswas, A., Saha, S., Bandyopadhyay, T.K. & Ghosh, P. 2022. AFLP-based assessment of genetic variation in certain Indian elite cultivars of *Cymbopogon* species. Journal of Applied Research on Medicinal and Aromatic Plants 29: 100372. DOI: 10.1016/j.jarmap.2022.100372.

Agarwal, M., Shrivastava, N. & Padh, H. 2008 Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Reports 27: 617–631. DOI: 10.1007/s00299-008-0507-z.

Atasagun. B. 2022. Assessment of the genetic diversity of a critically endangered species *Centaurea amaena* (Asteraceae). Archives of Biological Sciences 74(4): 325–332. DOI: 10.2298/ABS220826031A.

Atia, M.A.M., El-Moneim, D.A., Abdelmoneim, T.K., Reda, E.H., Shakour, Z.T.A., El-Halawany, A.M., El-Kashoury, E.S.A., Shams, K.H., Abdel-Azim, N.S. & Hegazy, M.E.F. 2021. Evaluation of genetic variability and relatedness among eight *Centaurea* species through CAAT-box derived polymorphism (CBDP) and start codon targeted polymorphism (SCoT) markers. Biotechnology & Biotechnological Equipment 35(1): 1230–1237. DOI: 10.1080/13102818.2021.1960891.

Bassin, S., Kölliker, R., Cretton, C., Bertossa, M., Widmer, F., Bungener, P. & Fuhrer, J. 2004. Intra-specific variability of ozone sensitivity in *Centaurea jacea* L., a potential bioindicator for elevated ozone concentrations. Environmental Pollution 131(1): 1–12. DOI: 10.1016/j.envpol.2004.02.014.

Burridge, A.J., Winfield, M., Przewieslik-Allen, A., Edwards, K.J., Siddique, I., Barral-Arca, R., Griffiths, S., Cheng, S., Huang, Z., Feng, C., Dreisigacker, S., Bentley, A.R., Brown-Guedira, G. & Barker, G.L. 2024. Development of a next generation SNP genotyping array for wheat. Plant Biotechnology Journal 22(8): 2235–2247. DOI: 10.1111/pbi.14341.

Collard, B.C.Y. & Mackill, D.J. 2009. Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. Plant Molecular Biology Reports 27: 86–93. DOI: 10.1007/s11105-008-0060-5.

Essa, N.M., Ibrahim, A.A. & Soliman, M.I. 2024. Association study between some cultivated species and their wild relatives from Apiaceae, Asteraceae and Brassicaceae families based on molecular and DNA barcoding in Egypt. Genetic Resources and Crop Evolution 71: 1125–1143. DOI: 10.1007/s10722-023-01681-x.

Etminan, A., Pour-Aboughadareh, A., Noori, A., Ahmadi-Rad, A., Shooshtari, L., Mahdavian, Z. & Yousefiazar-Khanian, M. 2018. Genetic relationships and diversity among wild *Salvia* accessions revealed by ISSR and SCoT markers. Biotechnology & Biotechnological Equipment 32(3): 610e617. DOI: 10.1080/13102818.2018.1447397.

Evanno, G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology 14: 2611e2620. DOI: 10.1111/j.1365-294X.2005.02553.x.

Falush, D., Stephens, M. & Pritchard, J.K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164(4): 1567–1587. DOI: 10.1093/genetics/164.4.1567.

Ferriol, M., Garmendia, A., Ruiz, J.J., Merle, H. & Boira, H. 2012. Morphological and molecular analysis of natural hybrids between the diploid *Centaurea aspera* L. and the tetraploid *C. seridis* L. (Compositae). Plant Biosystems 146(supll.): 86–100. DOI: 10.1080/11263504.2012.727878.

Freville, H., Justy, F. & Olivieri, I. 2001. Comparative allozyme and microsatellite population structure in a narrow endemic plant species, *Centaurea corymbosa* Pourret (Asteraceae). Molecular Ecology 10(4): 879–89. DOI: 10.1046/j.1365-294x.2001.01249.x.

Funk, V.A., Bayer, R.J., Keeley, S., Chan, R., Watson, L., Gemeinholzer, B., Schilling, E., Panero, J.L., Baldwin, B.G., Garcia- Jacas, N., Susanna, A. & Jansen, R.K. 2005. Everywhere but Antarctica: using a super tree to understand the diversity and distribution of the Compositae. Biologiske Skrifter 55: 343–374. https://repository.si.edu/handle/10088/11397.

Garcia-Jacas, N., Susanna, A., Mozaffarian, V. & Ilarslan, R. 2000. The natural delimitation of *Centaurea* (Asteraceae: Cardueae): ITS sequence analysis of the *Centaurea jacea* group. Plant Systematics and Evolution 223(3): 185–199. DOI: 10.1007/BF00985278.

Ghasemi, E. & Toluei, T. 2020. Intraspecific variations and biological relationships of different populations of *Centaurea virgata* Lamark (Asteraceae) in Iran. Iranian Journal of Genetics and Plant Breeding 9(1): 61–69. DOI: 20.1001.1.22519610.2020.9.1.7.9.

Gulsen, O., Sever-Mutlu, S., Mutlu, N., Tuna, M., Karaguzel, O., Shearman, R.C., Riordan, T.P. & Heng-Moss, T.M. 2009. Polyploidy creates higher diversity among *Cynodon* accessions as assessed by molecular markers. Theoretical and Applied Genetics 118: 1309–1319. DOI: 10.1007/s00122-009-0982-9.

Hellwig, F.H. 2004. Centaureinae (Asteraceae) in the Mediterranean-history of ecogeographical radiation. Plant Systematics and Evolution 246(3–4): 137–162. DOI: 10.1007/s00606-004-0150-2.

Huang, W., Zhao, X., Li, Y. & Lian, J. 2016. Effects of environmental factors on genetic diversity of *Caragana microphylla* in Horqin sandy land, northeast China. Ecology and Evolution 6(22): 8256–8266. DOI: 10.1002/ece3.2549.

Ismail, R.Y., Muhammed, J.J. & Sinjare, D.Y. 2024. Phylogenetic study of ten species from *Centaurea* (Asteraceae) in Duhok city, Kurdistan region-Iraq. Science Journal of University of Zakho 12(3): 277–284. DOI: 10.25271/sjuoz.2024.12.3.1279.

Jedrzejczyk, I. 2020. Genome size and SCoT markers as tools for identification and genetic diversity assessment in *Echinacea* genus. Industrial Crops and Products 144: 112055. DOI: 10.1016/j.indcrop.2019.112055.

Jèssica, R., López-Pujol, J., Carnicero, P., Susanna, A. & Garcia-Jacas, N. 2020. The *Centaurea alba* complex in theIberian Peninsula: gene flow, introgression, and blurred genetic boundaries. Plant Systematics and Evolution 306(2): 1–21. DOI: 10.1007/s00606-020-01669-5.

Kader, A., Sinha, S.N. & Ghosh, P. 2023. ISSR markers endorsed genetically stable plants regeneration of neem (*Azadirachta indica* A. Juss.) through indirect organogenesis using different explants. Molecular Biology Reports 50(9): 7305–7317. DOI: 10.1007/s11033-023-08655-3.

Kazemeini, F., Asri, Y., Mostafavi, G., Kalvandi, R. & Mehregan, I. 2020. RAPD-based evaluation of genetic diversity among populations of the Iranian endemic species *Rhabdosciadium aucheri* Boiss. (Apiaceae). Molecular Biology Reports 47(12): 9345–9352. DOI: 10.1007/s11033-020-06028-8.

Khanam, S., Sham, A., Bennetgen, J.L. & Mohammed, A.M.A. 2012. Analysis of molecular marker-based characterization and genetic variation in date palm (*Phoenix dactylifera* L.). Australian Journal of Crop Science 6(8): 1236–1244. https://www.cropj.com/khanam_6_8_2012_1236_1244.pdf.

Lopez, L. & Barreiro, R. 2013. Genetic guidelines for the conservation of the endangered polyploid *Centaurea borjae* (Asteraceae). Journal of Plant Research 126(1): 81–93. DOI: 10.1007/s10265-012-0497-3.

López-Vinyallonga, S., López-Pujol, J., Constantinidis, T., Susanna, A. & Garcia-Jacas, N. 2015. Mountains and refuges: Genetic structure and evolutionary history in closely related, endemic *Centaurea* in continental Greece. Molecular Phylogenetics and Evolution 92: 243–54. DOI: 10.1016/j.ympev.2015.06.018.

Lovejoy, T.E & Hannah, L.J. 2005. Climate Change and Biodiversity, Yale University Press, New Haven, CT.

Mallón, R., Rodríguez-Oubiña, J. & González, M.L. 2010. In vitro propagation of the endangered plant *Centaurea ultreiae*: assessment of genetic stability by cytological studies, flow cytometry and RAPD analysis. Plant Cell, Tissue and Organ Culture 101(1): 31–39. DOI: 10.1007/s11240-009-9659-y.

Mameli, G., Filigheddu, R., Binelli, G. & Meloni, M. 2008. The genetic structure of the remnant populations of *Centaurea horrida* in Sardinia and associated islands. Annals of Botany 101(5): 633–640. DOI: 10.1093/aob/mcn012.

Moreyra, L.D., Márquez, F., Susanna, A., Garcia-Jacas, N., Vázquez, F.M. & López-Pujol, J. 2021. Genesis, evolution, and genetic diversity of the hexaploid, narrow endemic *Centaurea tentudaica*. Diversity 13(2): 72. DOI: 10.3390/d13020072.

Mukhopadhyay, M., Mondal, T.K. & Chand, P.K. 2016. Biotechnological advances in tea (*Camellia sinensis* [L.] O. Kuntze): a review. Plant Cell Reports 35(2): 255–287. DOI: 10.3390/d13020072.

Nei, M. & Li, W.H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. The Proceedings of the National Academy of Sciences 76(10): 5269–5273. DOI: 10.1073/pnas.76.10.5269.

Novakovi´c, J., Jana´ckovi´c, P., Susanna, A., Lazarevi´c, M., Borši´c, I., Milanovici, S., Lakuši´c, D., Zlatkovi´c, B., Marin, P.D. & Garcia-Jacas, N. 2022. Molecular insights into the *Centaurea Calocephala* complex (Compositae) from the Balkans - Does phylogeny match systematics? Diversity 14(5): 394. DOI: 10.3390/d14050394.

Nurmansyah Alghamdi, S.S., Migdadi, H.M., Khan, M.A. & Afzal, M. 2020. AFLP-based analysis of variation and population structure in mutagenesis induced *Faba* Bean. Diversity 12(8): 303. DOI: 10.3390/d12080303.

Page, R.D. 1996. Tree View: an application to display phylogenetic trees on personal computers. Bioinformatics 12(4): 357–358. DOI: 10.1093/bioinformatics/12.4.357.

Peakall, R. & Smouse, P.E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28(19): 2537–2539. DOI: 10.1093/bioinformatics/bts460.

Plants of the World Online (POWO). Facilitated by the Royal Botanic Gardens, Kew [Published on the Internet, cited 5 Aug. 2024]. https://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:330045-2.

Pritchard, J.K., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multi locus genotype data. Genetics 155(2): 945–959. DOI: 10.1093/genetics/155.2.945.

Quinones-Perez, C.Z., Saenz Romero, C. & Wehenkel, C. 2014. Genetic diversity and conservation of *Picea chihuahuana* Martinez: a review. African Journal of Biotechnology 13(28): 2786–2795. DOI: 10.5897/AJB2014.13645.

Rai, MK. 2023. Start codon targeted (SCoT) polymorphism marker in plant genome analysis: current status and prospects. Planta 257(2): 34. DOI: 10.1007/s00425-023-04067-6.

Rechinger, K.H. 1987. Flora iranica: Akademische Druck-u. Verlagsanstalt: Graz.

Doyle, J.J. & Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19: 11–15. https://worldveg.tind.io/record/33886.

Samarina, L.S., Malyarovskaya, V.I., Reim, S., Yakushina, L.G., Koninskaya, N.G., Klemeshova, K.V., Shkhalakhova, R.M., Matskiv, A.O., Shurkina, E.S., Gabueva, T.Y., Slepchenko, N.A. & Ryndin, A.V. 2021. Transferability of ISSR, SCoT and SSR markers for *Chrysanthemum* × *Morifolium* Ramat and genetic relationships among commercial Russian cultivars. Plants 10(7): 1302. DOI: 10.3390/plants10071302.

Serrote, C.M.L., Reiniger, L.R.S., Silva, K.B., Rabaiolli, S.M.D.S. & Stefanel, C.M. 2020. Determining the polymorphism information content of a molecular marker. Gene 726: 144175. DOI: https://doi.org/10.1016/j.gene.2019.144175.

Smith, J.S.C., Chin, E.C., Shu, H., Smith, O.S., Wall, S.J., Senior, M.L., Mitchell, S.E., Kresovich, S. & Ziegle, J. 1996. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.) comparisons with data from RFLPs and pedigree. Theoretical and Applied Genetics 95: 163–173. DOI: 10.1007/s001220050544.

Sözen, E. & Özaydın, B. 2009. A preliminary study on the genetic diversity of the critically endangered *Centaurea nivea* (Asteraceae). Annales Botanici Fennici 46(6): 541–548. DOI: 10.5735/085.046.0606.

Sozen, E. & Ozaydin, B. 2010. A study of genetic variation in endemic plant *Centaurea wiedemanniana* by using RAPD markers. Ecology 19(77): 1–8. DOI: 10.5053/ekoloji.2010.771.

Swofford, D.L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Uysal, T., Özel, E., Bozkurt, M. & Ertuğrul, K. 2012. Genetic diversity in threatened populations of the endemic species *Centaurea lycaonica* Boiss. & Heldr. (Asteraceae). Research Journal of Biological Sciences 2(3): 110–116. https://api.semanticscholar.org/CorpusID:73718811.

Wang, H., Bai, Y. & Biligetu, B. 2024. Effects of SNP marker density and training population size on prediction accuracy in alfalfa (*Medicago sativa* L.) genomic selection. The Plant Genome 17: e20431. DOI: 10.1002/tpg2.20431.

Wolanin, M., Klichowska, E., Jedrzejczyk, I., Rewers, M. & Nobis, M. 2023. Taxonomy and distribution of *Taraxacum* sect. *Erythrosperma* (Asteraceae) in Poland. PhytoKeys 224: 1–88. DOI: 10.3897/phytokeys.224.99463.

Wu, K., Liu, Y., Yang, B., Kung, Y., Chang, K. & Lee, M. 2022. Rapid discrimination of the native medicinal plant *Adenostemma lavenia* from its adulterants using PCR-RFLP. Peer Journal 10: e13924. DOI: 10.7717/peerj.13924.

Yeh, F.C. & Boyle, T.J.B. 1997. Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belgian Journal of Botany 129: 157–163. https://www.sid.ir/paper/571184/en.