Physiol Mol Biol Plants (November–December 2018) 24(6):1261–1271
https://doi.org/10.1007/s12298-018-0545-0

RESEARCH ARTICLE

Molecular diversity assessment of a world collection of safflower genotypes by SRAP and SCoT molecular markers

Pooran Golkar[1] · **Niloofar Mokhtari**[2]

Received: 11 January 2018 / Revised: 6 March 2018 / Accepted: 2 May 2018 / Published online: 10 May 2018

Abstract Safflower (*Carthamus tinctorius* L.) is considered as an oil crop that is rich in medicinal and industrial properties. In this study, the genetic diversity of safflower was assessed using 12 polymorphic sequence-related amplified polymorphism (SRAPs) and 11 polymorphic start codon targeted (SCoT) markers in 100 genotypes of safflower gathered from different geographical regions of the world. The 23 primers generated a total of 227 polymorphism fragments with a mean of 68.2% within the range of 3 (SCoT 31 and SCoT 35) to 13 (SCoT 35) bands per primer. Polymorphism per primer ranged between 100% (in Me4–Em1) and 18.1% (in SCoT19), with an average of 36.76%. The polymorphism information contents of the SRAP and SCoT markers were 0.35 and 0.30, respectively, indicating that SRAP markers were more effective than SCoT markers for assessing the degree of genetic diversity of the safflower. The results of the analysis of molecular variance showed a significant difference across cultivated safflower genotypes possessing a high intra-population variation. The examined accessions were categorized into five clusters based on similarity centers: the Middle East containing Iran, Iraq, Turkey, and Tajikistan; the Far East, including India, Pakistan, and Korea; Europe; the American continent; and Africa, including Egypt, Sudan and Libya. The present study shows the effectiveness of employing the mixture of SRAP and SCoT markers in the identification of safflower genetic diversity that would be useful for conservation and population genetics of safflower improvement in further studies.

Keywords Diversity · Molecular variance · Safflower · SCAR · SCoT

Introduction

Safflower (*Carthamus tinctorius* L.) is considered an oil seed crop with diverse industrial and pharmaceutical properties (Ambreen et al. 2015; Kumar et al. 2016). It originated in the region of Fertile Crescent about 4000 years ago (Ambreen et al. 2015). There is a tremendous rise in safflower demand in various industries as a lubricant, biofuel, in soap and varnish making, cooking oil, and bird feeds (Golkar 2014, Kumar et al. 2016). It is a natural source of food flavors and cosmetic dyes (Khalid et al. 2017). Safflower seed oil contains a high content of unsaturated fatty acids (linoleic and oleic acid), which contains health-promoting properties for the treatment of trauma, heart attack, and renal thrombosis disorders (Al-Snafi 2015). Safflower leaf is a potential source of more than one hundred different bioactive substances with a wide variety of pharmaceutical and industrial applications (Khalid et al. 2017).

Iran is considered as a rich source for the *Carthamus* species (Barati and Arzani 2012; Golkar 2014). In a new type of classification, Chapman et al. (2010), determinded the presence of five genetic clusters (1: Europe; 2: Turkey–Iran–Iraq–Afghanistan; 3: Israel–Jordan–Syria; 4: Egypt –

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12298-018-0545-0) contains supplementary material, which is available to authorized users.

✉ Pooran Golkar
poorangolkar@gmail.com

[1] Research Institute for Biotechnology and Bioengineering, Isfahan University of Technology, Isfahan 84156 83111, Iran

[2] Department of Agricultural Biotechnology, College of Agriculture, Isfahan University of Technology, Isfahan 84156 83111, Iran

Ethiopia; and 5: the Far East–India–Pakistan) in safflower germplasm with EST-SSR molecular markers.

But according to first classification, safflower domesticated into seven distinguished centers of origin: the Far East, India and Pakistan, the Middle East (Iran–Afghanistan, Israel–Jordan–Iraq–Syria, and Turkey), Egypt, Sudan, Ethiopia, and Europe (Knowles 1969).

Genetic variation evaluation by various markers is decisive for the exploitation of genetic resources and its conservation in improvement programs (Poczai et al. 2013; Kiran et al. 2017). Broad genetic resources are vital for supplying valuable alleles based on a changing environment (Poczai et al. 2013; Bahmankar et al. 2017).

Genetic diversity in plant germplasms is evaluated by different markers (such as biochemical, morphological, and molecular markers), but genetic variation at the DNA level is of great importance in the study of genetics (Peng et al. 2008; Bahmankar et al. 2017). In this regard, molecular markers containing information about genealogy, genetic disorders, population structure studies and the evolutionary history of the plant are considered as an effective biotechnological tool to evaluate genetic variation (Sehgal and Raina 2005; Poczai et al. 2013). Higher polymorphism with more precision, independence from the environment and the plant growth stage is considered in molecular markers assays (Poczai et al. 2013).

The genetic diversity of safflower accessions shrinks during domestication and, therefore, flexibility against threatening surroundings, such as environmental stresses, has significantly decreased (Yang et al. 2007; Mayerhofer et al. 2011). Presently, different molecular markers have been used for genetic studies in safflower, such as genetic diversity and the quantitative trait loci mapping of important agronomic and physiological traits, including SSR (Ambreen et al. 2015; Kiran et al. 2017; Bahmankar et al. 2017), EST-SSR (Barati and Arzani 2012; Derakhshan et al. 2014), AFLP (Kumar et al. 2016), ISSR (Golkar et al. 2011; Yaman et al. 2014; Bahmankar et al. 2017), RAPD (Khan et al. 2009), and SRAP (Vilatersana et al. 2005, Peng et al. 2008, Mokhtari et al. 2013). However, the genome base studies of safflower have been restricted because of the inadequate presence of molecular markers in this plant in comparison to other major oil crops such as canola and sunflower.

According to Li and Quiros (2001), the sequence-related amplified polymorphism (SRAP) technique is a highly reproducible DNA marker that can be used for different purposes, including genetic variation (Robarts and Wolfe 2014), quantitative trait loci mapping (Zhang et al. 2006), map construction (Li and Quiros 2001), and cDNA fingerprinting and map-based cloning (Robarts and Wolfe 2014). The advantages of SRAP markers, such as simplicity, revealing of numerous co-dominant markers, and targeting of open reading frames (ORFs) (Uzun et al. 2009; Robarts and Wolfe 2014), have prioritized it over other molecular markers. Evaluation of the genetic diversity between and within the *Carthamus* species has been done by SRAP (Peng et al. 2008; Mokhtari et al. 2013).

Start codon targeted (SCoT) polymorphism, as a new dominant and/or co-dominant DNA marker, is characterized by the short conserved region flanking the start codon (ATG) (Collard and Mackill 2009). It can generate more information and better marker resolvability than random DNA markers, such as RAPD and ISSR (Mulpuri et al. 2013; Deng et al. 2015; Etminan et al. 2016), without any prior sequence information and by employing longer primers (18-mer) (Xiong et al. 2011). The SCoT markers have been used largely and successfully for the evaluation of phylogenetics (Xiong et al. 2011), genetic diversity (Satya et al. 2015), and fingerprinting (Xiong et al. 2011; Mulpuri et al. 2013).

Different molecular markers, especially random markers, have been used alone or in combination with each other to assess the genetic diversity across different plant families (Mulpuri et al. 2013; Hajibarat et al. 2015). The combination of molecular markers will be useful to fine analyze genetic diversity, phylogenetics, and fingerprinting in different plant species (Etminan et al. 2016). Application of combined molecular markers have been developed in different plant species such as SCoT and ISSR in Cicer (Amirmoradi et al. 2012) and durum wheat (Etminan et al. 2016); SSR and SCoT in chickpea (Hajibarat et al. 2015); ISSR and SRAP in *Galega officinalis* L. (Wang et al. 2012); and SSR and SRAP in orchard grass (Xie et al. 2010). Knowledge on the similarities and differences among safflower germplasms could prove to be valuable for implementing improvement strategies. A literature review showed that SCoT markers have not been used in the genetic variation of safflower genotypes thus far; this study aims to combine both SRAP and SCoT molecular markers to detect genetic variation among 100 different safflower genotypes from 30 different countries of the world in order to (1) assess the similarity centers and (2) define the patterns of genetic diversity and differentiation within safflower accessions using SRAP and SCoT markers.

Materials and methods

Plant materials

One hundred safflower genotypes including commercial cultivars, breeding genotypes and landraces originated from different countries of the world were used in this study (Table 1). The world distribution of genotypes is presented in Figure S1. The exotic genotypes were

Table 1 The collected safflower germplasm used in this study from the different geographical regions of the world

Code	Accession	Accession number	Germplasm type	Origin	Code	Accession	Accession number	Germplasm type	Origin
1	KMS36	KMS36IR	Landrace	Iran (Karaj)	52	CAR79	CAR79 J	Accession	Japan
2	A2	A2IR	Landrace	Iran (Azerbayjan)	53	CAR158	CAR158PR	Accession	Paraguay
3	K21	K21IR	Landrace	Iran (Kordestan)	54	CAR131	CAR131PR	Accession	Paraguay
4	CAR190	CAR190IR	Landrace	Iran (Karaj)	55	CAR146	CAR146EG	Accession	Egypt
5	ISF-28	ISF28IR	Landrace	Iran (Isfahan)	56	CAR70	CAR70L	Accession	Libyan
6	ISF-14	ISF14IR	Landrace	Iran (Isfahan)	57	CAR117	CAR117S	Accession	Sudan
7	Arak-2811	Arak2811IR	Cultivar	Iran (Markazi)	58	CAR201	CAR201S	Accession	Sudan
8	Il.111	Il111IR	Cultivar	Iran (Uromia)	59	CAR37	CAR37S	Accession	Sudan
9	C4110	C4110IR	Landrace	Iran (Isfahan)	60	CAR42	CAR42S	Accession	Sudan
10	Golsefid	GolsefidIR	Cultivar	Iran (Isfahan)	61	CAR24	CAR24 M	Accession	Morocco
11	CAR198	CAR198AZ	Accession	Azerbaijan	62	CAR157	CAR157 M	Accession	Morocco
12	CAR152	CAR152IQ	Accession	Iraq	63	CAR130	CAR130 M	Accession	Morocco
13	CAR83	CAR83TJ	Accession	Tajikistan	64	CAR86	CAR86T	Accession	Tunisia
14	CAR137	CAR137PK	Accession	Pakistan	65	CAR89	CAR89T	Accession	Tunisia
15	Kusamba	KusambaPK	Cultivar	Pakistan	66	CAR200	CAR200	Accession	Unknown
16	CAR124	CAR124PK	Accession	Pakistan	67	CAR188	CAR188P	Accession	Poland
17	CAR156	CAR156PK	Accession	Pakistan	68	CAR138	CAR138P	Accession	Poland
18	CAR175	CAR175IN	Accession	India	69	CAR214	CAR214P	Accession	Poland
19	CAR151	CAR151IN	Accession	India	70	CAR55	CAR55P	Accession	Poland
20	CAR118	CAR118IN	Accession	India	71	CAR19	CAR19P	Accession	Poland
21	CAR181	CAR181IN	Accession	India	72	CAR100	CAR100I	Accession	Italy
22	CAR114	CAR114IN	Accession	India	73	CAR78	CAR78H	Accession	Hungary
23	CAR116	CAR116IN	Accession	India	74	Car169	Car169H	Accession	Hungary
24	CAR215	CAR215GE	Accession	German	75	CAR87	CAR87RO	Accession	Romania
25	CAR228	CAR228GE	Accession	German	76	CAR9	CAR9CZ	Accession	Czechoslavaki
26	CAR217	CAR217GE	Accession	German	77	CAR49	CAR49SP	Accession	Spain
27	CAR68	CAR68GE	Accession	German	78	CAR94	CAR94SP	Accession	Spain
28	CAR227	CAR227GE	Accession	German	79	CAR210	CAR210SP	Accession	Spain
29	CAR67	CAR67GE	Accession	German	80	CAR106	CAR106SP	Accession	Spain
30	CAR230	CAR230GE	Accession	German	81	PI-301055	PI301055TR	Accession	Turkey
31	CAR216	CAR216GE	Accession	German	82	CAR135	CAR135PO	Accession	Portugal
32	CAR219	CAR219GE	Accession	German	83	CAR155	CAR155RU	Accession	Russia
33	CAR226	CAR226GE	Accession	German	84	CAR161	CAR161RU	Accession	Russia
34	CAR159	CAR159GE	Accession	German	85	CAR160	CAR160RU	Accession	Russia
35	CAR224	CAR224GE	Accession	German	86	CAR125	CAR125RU	Accession	Russia
36	CAR127	CAR127GE	Accession	German	87	Saffire	Saffire	Cultivar	Canada
37	CAR221	CAR221GE	Accession	German	88	Ac-suncet	Acsuncet	Cultivar	Canada
38	GE 62918	GE62918GE	Accession	German	89	Ac-stirling	CAR104Acstirling	Cultivar	Canada
39	CAR211	CAR211GE	Accession	German	90	CAR56	CAR56USA	Accession	USA
40	CAR129	CAR129GE	Accession	German	91	Hartman	HartmanUSA	Cultivar	USA
41	CAR218	CAR218GE	Accession	German	92	Jila	Jila USA	Cultivar	USA
42	CAR132	CAR132GE	Accession	German	93	MEX.17.45	MEX1745	Accession	Mexico
43	CAR126	CAR126GE	Accession	German	94	MEX. 7.	MEX7	Accession	Mexico
44	CAR126	CAR126BG	Accession	Belgium	95	MEX.22-191	MEX22191	Accession	Mexico
45	CAR74	CAR74NK	Accession	North korea	96	MEX. 6-97	MEX697	Accession	Mexico

Table 1 continued

Code	Accession	Accession number	Germplasm type	Origin	Code	Accession	Accession number	Germplasm type	Origin
46	CAR72	CAR72NK	Accession	North korea	97	MEX.2-138	MEX2138	Accession	Mexico
47	CAR199	CAR199 K	Accession	North korea	98	MEX.13-216	MEX13216	Accession	Mexico
48	CAR75	CAR75NK	Accession	North korea	99	MEX. 7-38	MEX738	Accession	Mexico
49	CAR77	CAR77NK	Accession	North korea	100	CAR64	CAR64SL	Accession	Sloveni
50	CAR80	CAR80NK	Accession	North korea					
51	CAR76	CAR76NK	Accession	North korea					

obtained from Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany and native breeding genotypes were deposited at Agricultural Research Center of Isfahan (IUT), Iran. During the growth period, young leaves were transferred to the laboratory and stored in − 80 °C freezer for further experiments.

Extraction of genomic DNA

The young leaves were used for DNA extraction according to modified CTAB protocol described by (Murray and Thompson 1980). DNA concentration was measured with a Nano Drop, spectrophotometer (Nano Drop Technologies Inc Wilmington, DE, USA), and DNA was diluted to a working concentration of 50 ng/µl.

Molecular marker analysis

SRAP–PCR amplification

The molecular analysis of SRAP markers was done according to the procedure described by Li and Quiros (2001) with minor modifications. All SRAP primer combinations were initially screened using a group of samples according to Uzun et al. (2009) (Table 2). Twelve primer combinations with scorable polymorphic bands were used for accessions. PCR reactions were performed in a total volume of 15.0 µl as track: 8 µl of Master Mix (Amplicon), 2 µl of DNA template, 2 µl of forward primer and reverse primer (10 µmol/l primers), and 3 µl of ddH_2O. The SRAP markers were amplified on following parameters: 5 min at 94 °C, 5 cycles of 94 °C for 1 min, 35 °C for 1 min, 72 °C for 2 min, 30 cycles of 94 °C for 1 min, 50 °C for 1 min, 72 °C for 2 min, and final extension of 5 min at 72 °C. The amplification reaction products were analyzed by electrophoresis on a 2% (w/v) Meta Phore agarose-Lonza gel using 1 × TBE running buffer and detected by staining with ethidium bromide.

SCoT–PCR amplification

Twelve polymorphic SCoT primers (Table 2) were also used to analyze and verify the genetic diversity between safflower genotypes according to Collard and Mackill (2009). PCR reactions were performed in 15 µl volumes containing 8 µl of Master Mix (Amplicon), 20 ng of genomic DNA and 10 µmol/l primers. SCoT-PCR amplification was performed by Gradient thermocycler (Bio Rad, USA) under initial denaturation for 5 min at 95 °C, followed by 35 cycles of 40 s at 95 °C, 40 s at 50 °C and 2 min at 72 °C, followed by final extension of 7 min at 72 °C. The amplification reaction products were analyzed by electrophoresis on a 2% (w/v) MetaPhore Agarose-Lonza gel using 1× TBE running buffer and detected by staining with ethidium bromide.

Data analysis

Amplified SRAP and SCoT amplicons were scored in a binary matrix as present (1) or absent (0) for each sample based on the relative standard size (100 bp DNA Ladder) (Collard and Mackill 2009). Vague bands that could not be easily detected were not scored. POPGENE ver 1.32 (Yeh et al. 1999) software was used to compute different diversity indices.

Analysis of molecular variance (AMOVA) was done to determine genetic differentiation among and within populations using the Arlequin v 3.1 software (Excoffier et al. 1992). Cluster analyses were calculated according to standardized Jaccard's similarity index (Jaccard 1908) and dendrogram were constructed by NTSYS-pc through Unweighted Pair group Method with Arithmetic mean

Table 2 The molecular marker characteristics used to assess genetic variation in different safflower accessions

SRAP	Sequence (5′ → 3′)	SCoT	Sequence (5′ → 3′)
Me1	5′-TGAGTCCAAACCGGATA-3′	SCoT-3	5′-CAACAATGGCTACCACCG-3′
Me2	5′-TGAGTCCAAACCGGAGC-3′	SCoT-12	5′-ACGACATGGCGACCAACG-3′
Me3	5′-TGAGTCCAAACCGGAAT-3′	SCoT-18	5′-ACCATGGCTACCACCGCC-3′
Me4	5′-TGAGTCCAAACCGGACC-3′	SCoT-19	5′-ACCATGGCTACCACCGGC-3′
Me5	5′-TGAGTCCAAACCGGAAG-3′	SCoT-21	5′-ACGACATGGCGACCCACA-3′
Em1	5′-GACTGCGTACGAATTAAT-3′	SCoT-22	′5′-AACCATGGCTACCACCAC-3′
Em2	5′-GACTGCGTACGAATTTGC-3′	SCoT-27	5′-ACCATGGCTACCACCGTG-3′
Em3	5′-GACTGCGTACGAATTGAC-3′	SCoT-30	5′-CCATGGCTACCACCGGCG-3′
Em4	5′-GACTGCGTACGAATTTGA-3′	SCoT-31	5′-CCATGGCTACCACCGCCT-3′
Em5	5′-GACTGCGTACGAATTAAC-3′	SCoT-32	5′-CCATGGCTACCACCGCAC-3′
Em6	5′-GACTGCGTACGAATTGCA-3′	SCoT-35	′5′-CATGGCTACCACCGGCCC-3′
		SCoT-36	5′-GCAACAATGGCTACCACC-3′

(UPGMA). Principal coordinate analysis (PCoA) was used to construct 3D eigenvectors using the DCENTER module of the NTSYS 2.0 software, version 2.0 (Rohlf 1992) to add complementary information to the cluster analysis.

Results and discussion

SRAP analysis

In this study, 12 primers from 30 primer combinations showed polymorphism (Table 3). The percentage of polymorphic bands across the primers ranged from 55.5% (Me5–Em6) to 100% (Me4–Em1), with an average of 77.28% polymorphism. The primer bands size ranged from 80 to 1200 base pairs. The ambiguous bands in the upper and the lower position of the gel were removed because of the faintness and inconsistency of amplification. The number of amplicons produced by each primer set ranged from 6 (Me3–Em2 and Me5–Em4) to 14 (Me5–Em1), with an average of 9.41 amplicons per primer set. On the other hand, the number of polymorphic amplicons ranged from 4 (Me5–Em4) to 13 (Me5-Em1) (Table 3). Out of the 113 amplicons, 88 polymorphic fragments (with an average of 7.3 polymorphic bands per primer combination < 10) were identified (Table 3). This finding was similar to the reports of *Coffea arabica* (Kumar Mishra et al. 2011), Citrus (Uzun et al. 2009), and *G. officinalis* L. (Wang et al. 2012). The percentage of polymorphic markers produced by each primer pair was within the range of 55.55% (Me5–Em6) to 100% (Me4–Em1), with a mean of 77.28% (Table 3). The average polymorphism information content (PIC), as an index for gained information from each marker values, was calculated as 0.35 (Table 3). The lowest (0.10) and the highest (0.68) values of PIC were attributed to the Me4–Em6 and the Me5–Em6 primer combination, respectively (Table 3). The SRAP markers were located in the medium level of PIC ($0.25 < PIC < 0.5$) in *C. tinctorius* (Xie et al.

Table 3 Polymorphism characteristics revealed by SRAP molecular markers in safflower genotypes

Primer combination	No. total bands	No. polymorphic bands	Polymorphism (%)	PIC value
Me1–Em2	11	7	63.63	0.13
Me3–Em2	6	5	83.33	0.25
Me3–Em6	10	9	90	0.42
Me4–Em1	12	12	100	0.40
Me4–Em3	7	5	71.42	0.26
Me4–Em5	13	8	61.53	0.44
Me4–Em6	8	7	87.5	0.10
Me5–Em1	14	13	92.85	0.48
Me5–Em2	7	6	85	0.31
Me5–Em4	6	4	66.66	0.51
Me5–Em5	10	7	70	0.24
Me5–Em6	9	5	55.55	0.68
Total	113	88	–	–
Means	9.41	7.33	77.28	0.35

2010) as well as in other species such as *Lolium multiflorum* (Huang et al. 2014) and *G. officinalis* L. (Wang et al. 2012). In addition, different PIC values were reported for SSR (0.30) (Ambreen et al. 2015) and SRAP (0.37– 0.57) markers (Peng et al. 2008; Mokhtari et al. 2013) among different *C. tinctorius* species.

SCoT analysis

A total of 36 SCoT primers were used to screen polymorphic markers using safflower accessions. Among the studied primers, a total of 11 primers exhibited distinct and reliable band patterns (Table 4). The number of amplified bands of each primer ranged between 7 (SCoT 31) and 14 (SCoT 30), with an average of 10.36 bands per primer, whereas the number of polymorphic amplicons ranged from 2 (SCoT 19) to 11 (SCoT 30), with an average of 6.18 amplicons per primer. The percentage of polymorphism varied from 37.5% (SCoT 35) to 100% (SCoT 27), with 3–11 polymorphic bands per primer (Table 4). Polymorphism information content was within the range of 0.12 (SCoT 35) to 0.49 (SCoT 3 and SCoT 12), with an average of 0.31 (Table 4). The amplified band sizes ranged from 100 to 2000 bp, most of which ranged from 500 to 1500 bp. Previous reports declared different polymorphism percentages such as 90.24% in orchard grass (Xie et al. 2010), 96.6%, in diospyros (Deng et al. 2015), and 66.67% in peanut (Xiong et al. 2011), which demonstrated high SCoT markers reproducibility in genetic diversity studies.

Analysis of molecular variance

A combination of all the SRAP and SCoT data was carried out to obtain more accurate genetic estimates. The identification of genetic variance among and within populations was identified by an analysis of molecular variance (AMOVA) ($P < 0.05$). Genetic inbreeding within subpopulations (F_{IS}) could range from − 1.0 (all individuals are heterozygous) to + 1.0 (no observed heterozygous individuals) (Wright 1984) and the fixation index (F_{ST}), which calculates the extent of genetic differentiation among subpopulations, can range from 0.0 (no differentiation) to 1.0 (complete differentiation—subpopulations fixed for different alleles) (Wright 1984). The result represents the existence of heterozygosity ($F_{IS} = -0.65$) and differentiation ($F_{ST} = 0.16$) among the *C. tinctorius* population. It was declared that the most significant variation was attributed to the within-individual contribution (Table 5). Though, safflower have high cross-pollinating percentage of about 30–35%, this significant high variation within-individual is expected.

AMOVA analysis was also performed in safflower by other markers such as EST-SSR (Barati and Arzani 2012) and ISSR (Golkar et al. 2011). A high within-population diversity indicates that these safflower populations should be conserved in their natural habitat for preserving genetic diversity (Satya et al. 2015). The result demonstrated medium genetic differentiation. This result reveals the potential of the new marker (SCoT) in basic evolutionary researches, such as map construction and population genetics (Ellis and Burke 2007).

Genetic diversity indices and genetic similarity analysis

Different diversity indices including polymorphism information content (PIC) value, Shannon's information index (I), number of alleles (Na), effective number of alleles (Ne) and expected heterozygosity (He) are estimated by using appropriate mathematical derivations of population studies (Yeh et al. 1999). The polymorphic information content (PIC) was calculated by the formula [$PIC_i = 2f_i\ (1 - f_i)$]

Table 4 Polymorphism characteristics revealed by SCoT primer pairs in safflower genotypes

Primer	No. total bands	No. polymorphic bands	Polymorphism (%)	PIC value
SCoT 3	13	5	38.46	0.49
SCoT 12	12	8	66.66	0.49
SCoT 18	12	10	83.33	0.2
SCoT 19	11	2	18.18	0.22
SCoT 21	8	5	62.50	0.26
SCoT 22	10	5	50	0.28
SCoT 27	8	8	100	0.42
SCoT 30	14	11	78.57	0.48
SCoT 31	7	3	42.85	0.30
SCoT 32	11	8	72.72	0.21
SCoT 35	8	3	37.5	0.12
Total	114	68	–	–
Mean	10.36	6.18	59.16	0.31

Table 5 Analysis of molecular variance (AMOVA) in safflower genotypes grouped according to populations

Source of variation	Degree of freedom	Sum of squares	Variance components	Percentage of variance	*P* value	F-statistic
Among groups	4	27.42	0.16 Va	15.03	< 0.0001	FCT = 0.15
Among populations within groups	24	14.02	0.05 Vb	4.82		FSC = 0.056
Among individuals within populations	71	20.818	− 0.56 Vc	− 52.63		FIS = − 0.65
Within individuals	100	141.5	1.41 Vd	132.8		FIT = − 0.32
Total	199	203.8	1.065			

(Anderson et al. 1993); where f_i is the percentage of the ith amplified. The percentage of polymorphism was calculated by dividing the number of polymorphic bands with the total number of regenerated bands. The Shannon index (I) is an information statistic index for measure of gene diversity is calculated based on the formula ($I = -\sum p_i \ln p_i$), where p_i is the proportion of (n/N), such that n is number of polymorphic loci, N is total number of loci, ln is the natural log, Σ is the sum of the calculations, and s is the number of species (Shannon 1948). The populations with more Shannon index, have higher diversity. Expected hetrozygosity (He) (or Nei's genetic diversity) estimates as $H = -\sum p_i^2$ which p_i = frequency of the ith allele in a locus (Hedrick 1999) and ranged between 0 and 1. It is maximized when there are many alleles at equal frequencies. The number of alleles (Na) counts the number of alleles with non-zero frequency (Hedrick 1999). The number of effective alleles, which estimates the homozygosity, is calculated as $Ne = 1/\sum p_i^2$ (Nei 1983), in which the number of samples had a large effect on this parameter.

Diversity indices were calculated among five distinct regions (Table 6). The results showed that the highest Shannon index (0.42) and number of observed alleles (Na) (1.87) were observed in Europe. Based on this, it could be concluded that gene flow factors or stable mutations were predominant in the European habitats of the safflower. The highest heterozygosity index (H_e) (0.29) denoted to groups of Europe and American content. The highest value of Ne (1.52) was observed in the Far East (Table 6). The mean of Na, Ne, He, and the Shannon index were, respectively, 1.66, 1.44, 0.26, and 0.35, which showed relatively high polymorphism in these genotypes (Table 6). The population from Africa exhibited lower genetic polymorphism than other populations. The range of the similarity coefficient among accessions varied from 0.53 to 1, with a mean value of 0.76 (data not shown). Different genetic similarities have been reported previously among *C. tinctorius* populations (Zhang et al. 2006; Peng et al. 2008; Mokhtari et al. 2013). Different genetic factors, such as variation in climate, genome structure, and environmental effects, could be involved in these differences (Sehgal and Raina 2005; Barati and Arzani 2012; Hajibarat et al. 2015), but identifying a subset of accessions that represents the highest possible genetic distance from the whole collection would enable a better exploitation of genetic resources for breeding as well as removing useful genes (Reeves et al. 2012). Different ranges of genetic similarities have been reported in the safflower germplasm. In terms of genetic similarity analysis, the same conclusion can be drawn as the conclusions of Mokhtari et al. (2013), Kiran et al. (2017), and Bahmankar et al. (2017), indicating that there is high level of genetic diversity between the different genotypes of *C. tinctorius*. The highest similarity coefficients may be caused by their common origin, possibly as a result of a genetic mutation that occurred in an initial variety (Chapman et al. 2010; Mokhtari et al. 2013). It is not surprising that the similarity between cultivated accessions were close to each other by genetic similarity > 0.80, which is similar to the reports of Peng et al. (2008), which indicated lower values across different *Carthamus* species (Mokhtari et al. 2013).

Table 6 Summary of genetic variation statistics in different *C. tinctorius* genotypes with combined two different molecular markers (*SRAP and SCoT)*

Population	Sample size	*Na*	*Ne*	*He*	*I*
Middle East	14	1.68	1.48	0.27	0.38
Far East	19	1.68	1.52	0.28	0.39
Europe	41	1.87	1.51	0.29	0.42
American content	15	1.68	1.44	0.29	0.37
Africa	11	1.40	1.26	0.21	0.22
Total means	100	1.66	1.44	0.26	0.35

Principal coordinate analysis (PCoA)

Principal coordinate analysis was performed based on the genetic similarity matrix to visualize the relationships among safflower accessions. A two-dimensional plot was also performed in terms of its position relative to two coordinate axes (Fig. 1). The first three eigenvectors explained 49.22% of the total molecular variation, which accounted for 18, 15, and 14% of the observed variation, respectively. PCoA and UPGMA analyses resulted in similar groupings in this study (Figs. 1, 2). Similarities in genotypes grouped in common clusters could be due to the participation of a common parentage, convergent evolution, and subsequent selection (Barati and Arzani 2012).

Cluster analysis

The cluster analysis dendrogram of the 100 genotype-based, categorized the samples into five clusters (A, B, C, D, and E) based on 166 polymorphic SRAP and SCoT fragments (Fig. 2). The UPGMA results were relative to the centers of similarity. Group A contained accessions that were predominantly from the Middle East (Iran, Iraq, Turkey, and Tajikistan), which were grouped in the same cluster. In this project, Turkish accession is closer to Middle Eastern accessions. Chapman et al. (2010) classified Turkish accessions along with all the accessions from Iran and Afghanistan with microsatellite markers. Cluster B consisted of accessions which were nearly entirely from the Far East (India, Pakistan, and Korea). This finding confirmed the data of Chapman et al. (2010) data, but it was different from the classification of Knowles (1969). These discrepancies could be referred to different used molecular markers and genotypes in different studies.

The majority accessions from Group C were from Europe, except one accession from the USA, two from Paraguay, one from Portugal, one from Morocco, and all the Russian genotypes. It seems that Russian genotypes are closer to Europe genotypes than the Eastern accessions. Group D included accessions from the American content, and finally, Group E included Egyptian, Sudanese, and Libyan accessions. Custer analysis revealed the high discrimination power of the mixture of the molecular markers from distinct geographical regions. Moreover, an unknown accession of safflower (Car200) was found among the Far East genotypes. The majority of American accessions were grouped into one cluster and it seems that these genotypes (from Canada, Mexico, and USA) have the same ancestor. In this project, only one of the accessions from USA and all the genotypes from Paraguay belonged to Cluster C. The results of cluster analysis showed a considerable connection between geographical origin and genetic variation among safflower accessions. However, there were certain inconsistencies between the origin countries of the genotypes placed in different subgroups, which could be due to the plant germplasm substitution across the regions (Vilatersana et al. 2005), high percentage of cross-pollination (Khan et al. 2009), or stable genetic mutations (Poczai

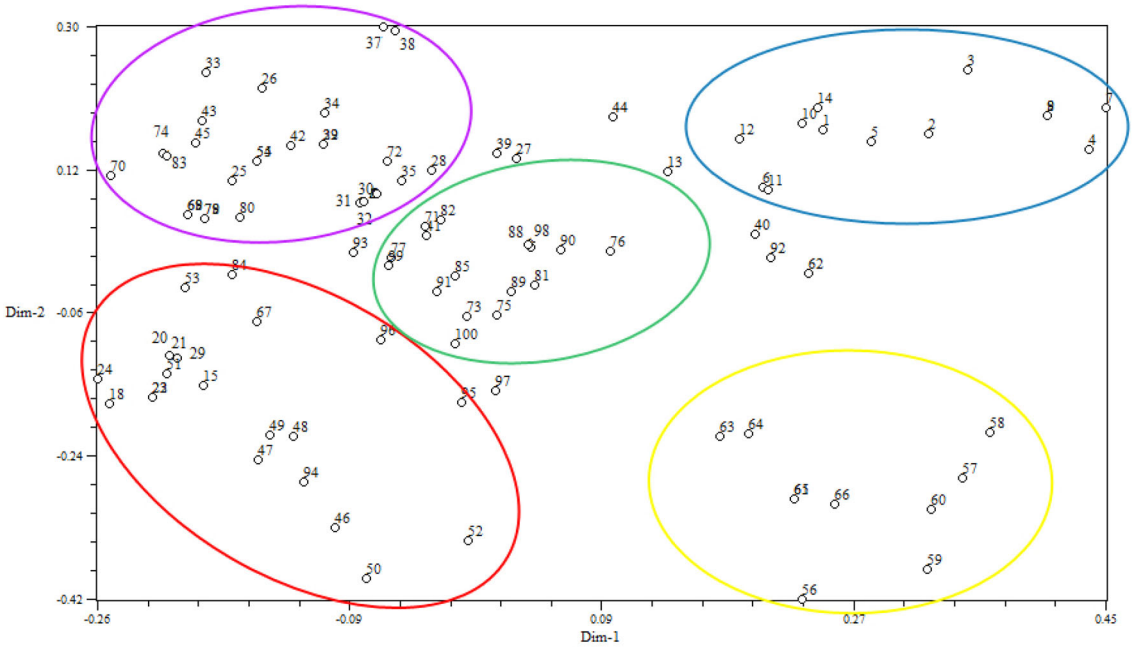

Fig. 1 Principal coordinate analysis (PCoA) of the 100 safflower genotypes based on combined SRAP and SCoT markers (the genotypes number are based on the coding in manuscript)

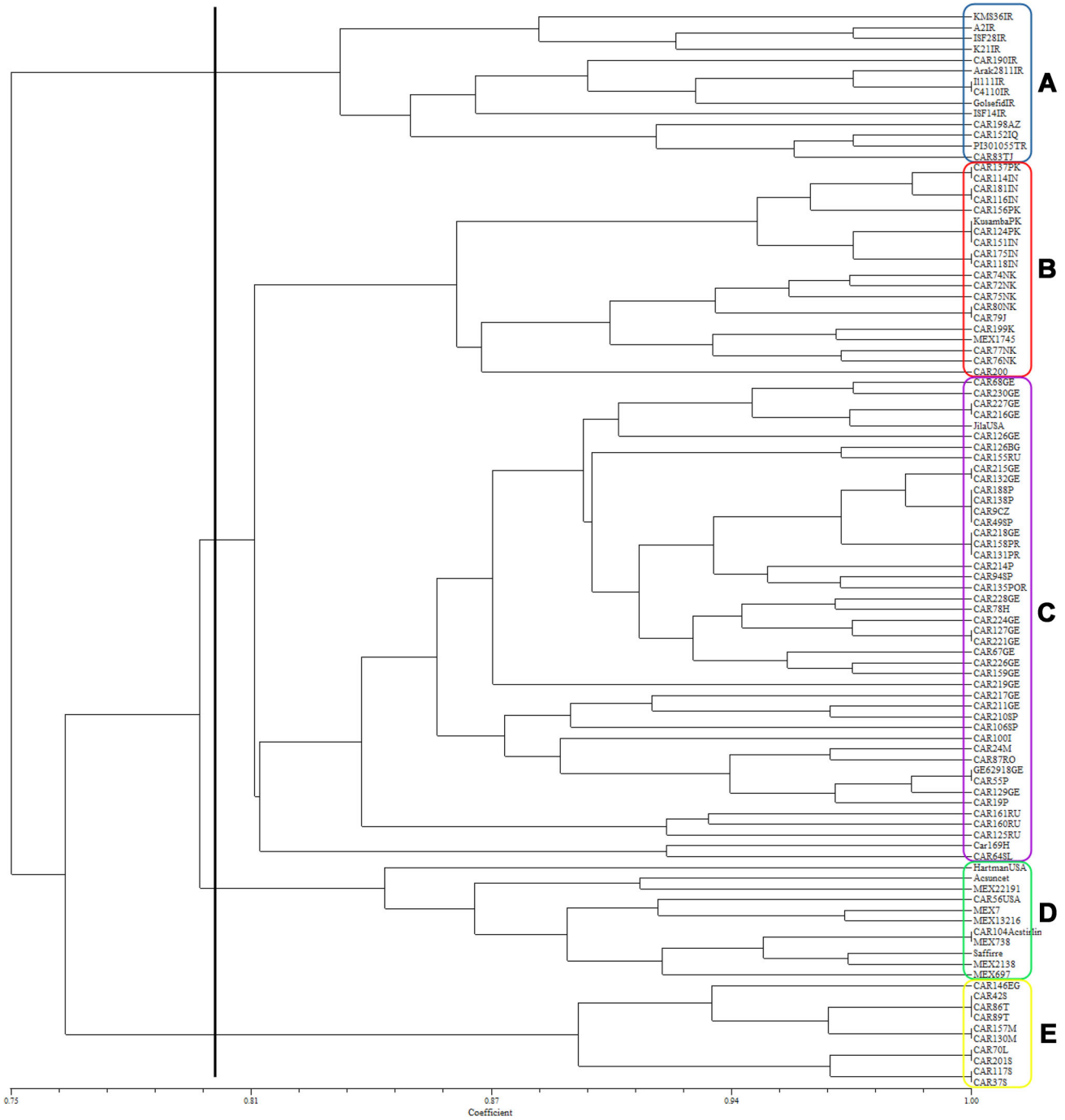

Fig. 2 UPGMA dendrogram of 100 Iranian safflower genotypes based on SRAP genetic similarities; the geographic origin of each genotype is indicated by symbols

et al. 2013). This result is similar to the result of Barati and Arzani (2012) and Golkar et al. (2011), all of whom obtained the same conclusion using different molecular markers. Considering the problems of the current classifications in the *Carthamus* genus (Vilatersana et al. 2005), the novel SCoT and SRAP markers will also be beneficial in accelerating the taxonomic classification of this genus. However, to confirm the available pattern in safflower genotypes, it is essential to use more number of accessions from each geographical location.

Conclusion

Broadening the genetic base of breeding materials in safflower requires exploiting diverse genetic resources available in the world. Developing newer molecular markers is beneficial for the identification of genetic diversity, interspecies crossbreeding, and producing new cultivars in safflower. The genetic diversity of different safflower species was evaluated using SCoT molecular markers at first. Both SRAP and SCoT markers showed high efficiency for the identification of genetic diversity in safflower accessions, but SRAP markers showed a little more efficiency in diversity identification. This study illustrated a high level of genetic diversity in safflower genotypes. It was found that genetic variance mainly exists within populations. The UPGMA cluster showed that molecular classification was correlated with the geographical distribution of genotypes. High polymorphism primers related to SCoT and SRAP markers were more useful for DNA fingerprinting, population structure analysis, reconstruction of the breeding history of the domesticated genotypes, and the effective management of genetic resources in safflower. These new achievements will help the future selection program to produce new cultivars and interspecies crossbreeding lines. Whether polymorphic bands from SCoT create from the functional region or not; this study requires profound analysis through methods such as sequencing. Conversion of these diagnostic markers into SCARs is suggested in future studies.

Acknowledgements The authors would like to thank Research Institute for Biotechnology and Bioengineering, Isfahan University of Technology, Isfahan, Iran. We appreciated Radan English Edit Company for editing the paper.

References

Al-Snafi AE (2015) The chemical constituents and pharmacological importance of *Carthamus tinctorius*—an overview. J Pharm Biol 5(3):143–166

Ambreen H, Kumar S, Variath MT, Joshi G, Bali S, Agarwal M, Kumar A, Jagannath A, Goel S (2015) Development of genomic microsatellite markers in *Carthamus tinctorius* L. (Safflower) using next generation sequencing and assessment of their cross-species transferability and utility for diversity analysis. PLoS ONE 10(8):e0135443

Amirmoradi B, Talebi R, Karami E (2012) Comparison of genetic variation and differentiation among annual Cicer species using start codon targeted (SCoT) polymorphism, DAMD-PCR, and ISSR markers. Plant Syst Evol 298:1679–1688

Anderson JA, Churchill GA, Autrique JE, Sorells ME, Tanksley SD (1993) Optimizing parental selection for genetic-linkage maps. Genome 36:181–186

Bahmankar M, Nabati DA, Dehdari M (2017) Genetic relationships among Iranian and exotic safflower using microsatellite markers. J Crop Sci Biotechnol 20(3):159–165

Barati M, Arzani A (2012) Genetic diversity revealed by EST-SSR markers in cultivated and wild safflower. Biochem Syst Ecol 44:117–123

Chapman MA, Hvala J, Strever J, Burke JM (2010) Population genetic analysis of safflower (*Carthamus tinctorius*; Asteraceae) reveals a Near Eastern origin and five centers of diversity. Am J Bot 97:831–840

Collard BC, Mackill DJ (2009) Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. Plant Mol Biol Rep 27:86

Deng L, Liang Q, He X, Luo C, Chen H, Qin Z (2015) Investigation and analysis of genetic diversity of diospyros germplasms using SCoT molecular markers in Guangxi. PLoS ONE 10:e0136510

Derakhshan E, Majidi MM, Sharafi Y, Mirlohi A (2014) Discrimination and genetic diversity of cultivated and wild safflowers (*Carthamus* spp.) using EST-microsatellites markers. Biochem Syst Ecol 54:130–136

Ellis J, Burke J (2007) EST-SSRs as a resource for population genetic analyses. Heredity 99:125–132

Etminan A, Pour-Aboughadareh A, Mohammadi R, Ahmadi-Rad A, Noori A, Mahdavian Z, Moradi Z (2016) Applicability of start codon targeted (SCoT) and inter-simple sequence repeat (ISSR) markers for genetic diversity analysis in durum wheat genotypes. Biotechnol Biotechnol Equip 30(6):1075–1081

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Golkar P (2014) Breeding improvements in safflower (*Carthamus tinctorius* L.): a review. Aust J Crop Sci 8:1079–1085

Golkar P, Arzani A, Rezaei AM (2011) Genetic variation in safflower (*Carthamus tinctorious* L.) for seed quality-related traits and inter-simple sequence repeat (ISSR) markers. Int J Mol Sci 12:2664–2677

Hajibarat Z, Saidi A, Hajibarat Z, Talebi R (2015) Characterization of genetic diversity in chickpea using SSR markers, start codon targeted polymorphism (SCoT) and conserved DNA-derived polymorphism (CDDP). Physiol Mol Biol Plant 21:365–373

Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. Evolution 53:313–318

Huang LK, Jiang XY, Huang QT, Xiao YF, Chen ZH, Zhang XQ, Miao JM, Yan HD (2014) Genetic diversity and relationships in cultivars of *Lolium multiflorum* Lam. using sequence-related amplified polymorphism markers. Genet Mol Res 13(4):10142–10149

Jaccard P (1908) Nouvelles recherches sur la distribution florale. Bull Soc Vaud Sci Nat 44:223–270

Khalid N, Khan RS, Hussain MI, Farooq M, Ahmad A, Ahmed I (2017) A comprehensive characterization of safflower oil for its potential applications as a bioactive food ingredient—a review. Trends Food Sci Technol 66:176–186

Khan MA, von Witzke-Ehbrecht S, Maass BL, Becker HC (2009) Relationships among different geographical groups, agro-morphology, fatty acid composition and RAPD marker diversity in safflower (*Carthamus tinctorius*). Genet Resour Crop Evol 56:19–30

Kiran BU, Mukta N, Kadirvel P, Alivelu K, Senthilvel S, Kishore P, Varaprasad KS (2017) Genetic diversity of safflower (*Carthamus tinctorius* L.) germplasm as revealed by SSR markers. Plant Genet Resour 15(1):1–11

Knowles P (1969) Centers of plant diversity and conservation of crop germplasm: safflower. Econ Bot 23:324–329

Kumar Mishra M, Suresh N, Bhat AM, Suryaprakash N, Satheesh Kumar S, Kumar A (2011) Genetic molecular analysis of *Coffea*

arabica (Rubiaceae) hybrids using SRAP markers. Rev Biol Trop 59(2):607–617

Kumar S, Ambreen H, Variath MT, Rao AR, Agarwal M, Kumar A, Goel S, Jagannath A (2016) Utilization of molecular, phenotypic and geographical diversity to develop compact composite core collection in the oilseed crop, Safflower (*Carthamus tinctorius* L.) through maximization strategy. Front. Plant Sci 7:1554

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. Theor Appl Genet 103:455–461

Mayerhofer M, Mayerhofer R, Topinka D, Christianson J, Good AG (2011) Introgression potential between safflower (*Carthamus tinctorius*) and wild relatives of the genus *Carthamus*. BMC Plant Biol 11:47

Mokhtari N, Rahimmalek M, Talebi M, Khorrami M (2013) Assessment of genetic diversity among and within *Carthamus* species using sequence-related amplified polymorphism (SRAP) markers. Plant Syst Evol 299:1285–1294

Mulpuri S, Muddanuru T, Francis G (2013) Start codon targeted (SCoT) polymorphism in toxic and non-toxic accessions of *Jatropha curcas* L. and development of a codominant SCAR marker. Plant Sci 207:117–127

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4326

Nei M (1983) Genetic polymorphism and the role of mutation in evolution. Evol Genes Proteins 71:165–190

Peng S, Feng N, Guo M, Chen Y, Guo Q (2008) Genetic variation of *Carthamus tinctorius* L. and related species revealed by SRAP analysis. Biochem Syst Ecol 36:531–538

Poczai P, Varga I, Laos M, Cseh A, Bell N, Valkonen JPT, Hyvönen J (2013) Advances in plant gene-targeted and functional markers: a review. Plant Methods 9:6

Reeves PA, Panella LW, Richards CM (2012) Retention of agronomically important variation in germplasm core collections: implications for allele mining. Theor Appl Genet 124:1155–1171

Robarts DW, Wolfe AD (2014) Sequence-related amplified polymorphism (SRAP) markers: a potential resource for studies in plant molecular biology. Appl Plant Sci 2:1400017

Rohlf FJ (1992) NTSYS-PC: numerical taxonomy and multivariate analysis system version 2.0. State University of New York, Stony Brook

Satya P, Karan M, Jana S, Mitra S, Sharma A, Karmakar P, Ray D (2015) Start codon targeted (SCoT) polymorphism reveals genetic diversity in wild and domesticated populations of ramie (*Boehmeria nivea* L. Gaudich.), a premium textile fiber producing species. Meta Gene 3:62–70

Sehgal D, Raina SN (2005) Genotyping safflower (*Carthamus tinctorius*) cultivars by DNA fingerprints. Euphytica 146:67–76

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423, 623–656

Uzun A, Yesiloglu T, Aka-Kacar Y, Tuzcu O, Gulsen O (2009) Genetic diversity and relationships within Citrus and related genera based on sequence related amplified polymorphism markers (SRAPs). Sci Hortic 121:306–312

Vilatersana R, Garnatje T, Susanna A, Garcia-Jacas N (2005) Taxonomic problems in *Carthamus* (Asteraceae): RAPD markers and sectional classification. Bot J Linn Soc 147:375–383

Wang Z, Wang JE, Wang XM, Gao HW, Dzyubenko NI, Chapurin VF (2012) Assessment of genetic diversity in *Galega officinalis* L. using ISSR and SRAP markers. Genet Resour Crop Evol 59(5):865–873

Wright S (1984) Evolution and the genetics of populations: experimental results and evolutionary deductions, vol 3. University of Chicago Press, Chicago

Xie W, Zhang X, Cai H, Huang L, Peng Y, Ma X (2010) Genetic maps of SSR and SRAP markers in diploid orchardgrass (*Dactylis glomerata* L.) using the pseudo-testcross strategy. Genome 54:212–221

Xiong F, Zhong R, Han Z, Jiang J, He L, Zhuang W, Tang R (2011) Start codon targeted polymorphism for evaluation of functional genetic variation and relationships in cultivated peanut (*Arachishy pogaea* L.) genotypes. Mol Biol Rep 38:3487–3494

Yaman H, Tarıkahya-Hacıoglu B, Arslan Y, Subasi I (2014) Molecular characterization of the wild relatives of safflower (*Carthamus tinctorius* L.) in Turkey as revealed by ISSRs. Genet Resour Crop Evol 61:595–602

Yang YX, Wu W, Zheng YL, Chen L, Liu RJ, Huang CY (2007) Genetic diversity and relationships among safflower (*Carthamus tinctorius* L.) analyzed by inter-simple sequence repeats (ISSRs). Genet Resour Crop Evol 54(5):1043–1051

Yeh F, Yang R, Boyle T (1999) POPGENE. Microsoft Window based freeware for population genetic analysis. Version 1.32. Centre for International Forestry Research, University of Alberta, Edmonton

Zhang L, Huang B-B, Kai G-Y, Guo M-L (2006) Analysis of intraspecific variation of Chinese *Carthamus tinctorius* L. using AFLP markers. Acta Pharm Sin 41(1):91–96