

## Genetic Diversity in Saffron (*Crocus sativus* L.) Cultivars Grown in Iran Using SSR and SNP Markers

I. Yousefi Javan<sup>1\*</sup>, and F. Gharari<sup>1</sup>

### ABSTRACT

Saffron (*Crocus sativus* L.), one of the most expensive spices in the world, is used mainly as food coloring and flavoring in food industry and its effective components are also used in medicine. A collection of twenty-two cultivars of saffron grown in different regions of Iran was screened with 25 SSR and 5 SNP primers in order to determine genetic identities and genetic diversity in these cultivars. On an average, 50 alleles were amplified using SSR primers with scorable fragment sizes ranging from approximately 160 to 400 bp. Among these, 33 alleles were polymorphic thus revealing 72% of polymorphism. The genetic similarity estimated according to SSR data was scaled between 9.5 and 87.8%. In determination of genetic diversity, five polymorphic SNP markers were used. Since SNP markers are mainly bi-allelic, all SNPs showed two alleles only, suggesting the potential of SSR and SNP markers in discriminating among plants of distant genetic backgrounds. Un-weighted pair group method with arithmetic mean clustering grouped the cultivars into four groups. In this study, we tried to expand the genetic diversity of *C. sativus* in Iran despite their asexual reproduction. Due to the similarity of climatic conditions in Iran, a certain genetic variation was observed in saffron plants. For saffron cultivation and production of high quality crop around the world, research on genetic diversity among the large family of *C. sativus* adds value this product.

**Keywords:** Genetic identity, Pair Group method, Polymorphism, Saffron, UPGMA.

### INTRODUCTION

*Crocus sativus* L. is an autumn-flowering geophyte extensively grown in the Mediterranean basin and Near East since the Late Bronze Age (Kanakis *et al.*, 2004). *Crocus sativus* L., a member of the family Iridaceae, is a sterile autotriploid or of hybrid origin ( $2n=3x=24$ ) plant. *C. cartwrightianus* is considered to be one of the parents (Caiola *et al.*, 2010). *Crocus sativus* is traditionally grown from Spain in the west to Kashmir and China in the east. There is evidence and archaeological indications of *Crocus stigma* collection by human as early as about 2300 BC (Golmohammadi, 2014). The cultivation of saffron dates back to 1500-2500 BC in

Iran, Greece, India, China, the Mediterranean basin, and Eastern Europe (Beiki *et al.*, 2010). Iran is the largest producer, accounting for almost 80% of the total world production. Ranked first in the world, Khorasan province, Iran, is specifically the ideal place for the growth and production of cultivated saffron (Ahmad *et al.*, 2011). Iran is also the native habitat of eight *Crocus* species besides *C. sativus*, four of which are exclusively indigenous to this country. Morphological comparisons of cultivated saffron ecotypes have revealed some differences in intensity of flower color, viability, pollen size, and number of style branches and stamens (Grilli Caiola *et al.*, 2001). Molecular genetic diversity estimates are extremely useful for

<sup>1</sup> Department of Plant Production, Faculty of Agriculture, University of Torbat Heydarieh, Torbat Heydarieh, Islamic Republic of Iran.

\* Corresponding author; e-mail: i.javan@torbath.ac.ir



intellectual property protection, particularly in the determination of essential derivation. Measurements of genetic diversity can be used in breeding programs to increase the genetic variation in base germplasm populations (Lefebvre *et al.* 2001). The genetic diversity estimates based on molecular marker data may be compared to a minimum genetic distance, which indicates that two cultivars are not essentially derived. A number of molecular studies have investigated the genetic variation in *Crocus* (Harpke *et al.*, 2013; Ahrazem *et al.*, 2010; Grilli-Caiola and Canini, 2004; Frizzi *et al.*, 2007; Peterson *et al.*, 2008; Zubor *et al.*, 2004). Previous investigations in Series *Crocus* have mainly focused on interspecific variation including only few individuals of each species. However, the intraspecific variation has not been examined using molecular attempts (Bjarne *et al.* 2015). Over the past centuries, the problem of intense genetic erosion in *Crocus* due to loss of land surface allocated to this crop in many areas of the world has further reduced the genetic variation of this crop (Nemati *et al.*, 2014). During recent decades, Simple sequence repeats (SSR), also known as microsatellites, have become the most popular source of genetic markers owing to their high reproducibility, multi-allelic nature, co-dominant inheritance, abundance, and wide genome coverage. SSR markers have been successfully adopted to analyze genetic diversity in a variety of different plant species especially in saffron (Sarıkamış *et al.*, 2010). The low mutation rate of Single Nucleotide Polymorphism (SNPs) also makes them excellent markers for studying complex genetic traits and as a tool for the understanding of genome evolution (Syvanen, 2001). The SNPs are single base pair substitutions that occur within and outside genes, and they are the most common form of sequence variation (Ching *et al.*, 2001). A substantial private and public effort has been undertaken to characterize the SNPs responsible for genetic diversity. They are far more prevalent than microsatellites and, therefore, may provide a high density of

markers near a locus of interest. Limited work has been carried out to examine the occurrence of SNPs in plants, although these preliminary studies have indicated that SNPs appear to be even more abundant in plant systems than in the human genome. The sequencing results indicated that the large chromosomal regions of saffron have much mutations and different SNP haplotype. Analysis of genetic diversity in saffron will be useful, maybe in future for breeding programers and attempting to broaden the genetic base of future saffron cultivars.

Consequently, there is a rapidly growing interest in quantifiable methods for mutation screening in life science research. High-Resolution Melt (HRM) analysis is the quantitative analysis of the melt curve of a DNA fragment following amplification by PCR and can be considered the next-generation application of amplicon melting analysis. It requires a real-time PCR detection system with excellent thermal stability and sensitivity and HRM-dedicated software. The combination of improved qPCR instrumentation and saturating DNA binding dyes has permitted the identification of genetic variation in nucleic acid sequences by the controlled melting of a double-stranded PCR amplicon. New instrument calibration methods coupled with HRM-compatible software permit the rapid analysis of the resulting data sets and the discrimination of DNA sequences based on their composition, length, GC content, or strand complementarity (Garritano *et al.*, 2009). HRM experiments generate DNA melt curve profiles that are both specific and sensitive enough to distinguish nucleic acid species based on small sequence differences, enabling mutation scanning, methylation analysis, and genotyping (Garritano *et al.*, 2009). For example, HRM can be used to detect single-base sequence variations or to discover unknown genetic mutations. It can also be used to quantitatively detect a small proportion of variant DNA in a background of wild-type sequence at sensitivities approaching 5% to study somatically

acquired mutations or changes in methylation state (Malentacchi *et al.*, 2009).

The purpose of the present study was to investigate the spectrum of genetic diversity within Iranian *Crocus* spp., to reveal the phylogenetic relationships of 22 *Crocus* landraces collected from different sites of Iran, using 25 microsatellite and 5 SNP markers, to propose a strategy for broadening the genetic base for future breeding of this valuable crop.

## MATERIALS AND METHODS

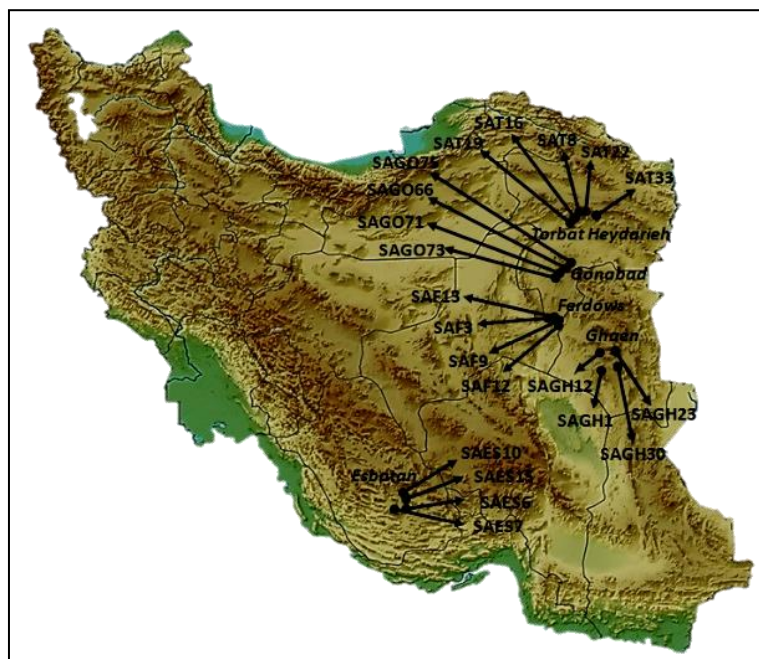
### Plant Material

The plant material for the study consisted of twenty-two various cultivars of saffron (*S. Crocus*) that had different origin and grown under Iranian environment (Table 1). The corms were kindly provided by Saffron Research Institute, Agriculture Research Center, Torbat Heydarieh, Iran. Fresh leaves were frozen using liquid nitrogen and kept at  $-80^{\circ}\text{C}$ . High molecular weight genomic DNA was extracted from fresh leaves of

**Table 1.** List of *Crocus Sativus* used in this study and collected from different regions of Iran.

No	Code	Geographical region
1	SAES6	Estahban
2	SAES7	Estahban
3	SAES10	Estahban
4	SAES15	Estahban
5	SAF3	Ferdos
6	SAF9	Ferdos
7	SAF12	Ferdos
8	SAF13	Ferdos
9	SAGH1	Ghaen
10	SAGH12	Ghaen
11	SAGH23	Ghaen
12	SAGH29	Ghaen
13	SAGH30	Ghaen
14	SAGO66	Gonabad
15	SAGO71	Gonabad
16	SAGO73	Gonabad
17	SAGO75	Gonabad
18	SAT8	Torbat Heydarieh
19	SAT16	Torbat Heydarieh
20	SAT19	Torbat Heydarieh
21	SAT22	Torbat Heydarieh
22	SAT33	Torbat Heydarieh

*Crocus* using DNeasy Plant Mini kit (Qiagen, Germany).



**Figure 1.** The different geographical places for Saffron cultivation in Iran, in this study.

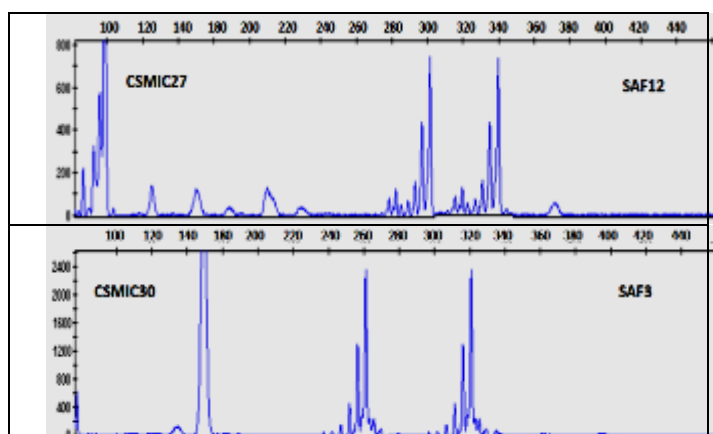


Microsatellite primers, landraces, Peak scoring, and cluster analysis of 25 microsatellite primer pairs were used for the landraces assays (Nemati *et al.*, 2014). Primer names, sequences, and corresponding annealing temperatures and the amplified fragments are listed in Table 2. The quality of the extracted DNA was verified on a 0.8% agarose gel and the amount of total genomic DNA obtained was quantified using a NanoDrop (ND-1000) Spectrophotometer. SSR markers were screened to test their polymorphisms between the landraces. Genomic DNA samples were diluted in 0.1 mM TE buffer to 25 ng  $\mu\text{L}^{-1}$  before amplification by polymerase chain reaction. The SSRs markers were amplified by PCR in different vials.

Three  $\mu\text{L}$  genomic DNA with a concentration of 5 ng  $\mu\text{L}^{-1}$ , 5  $\mu\text{L}$  of H<sub>2</sub>O Mol Bio grade autoclaved, 1X PCR Buffer Promega (100 mM Tris-HCL – pH 8.3, 500 mM KCL, 1.5 mM MgCl<sub>2</sub>), 0.002  $\mu\text{L}$  Forward primer (→) labeled with 0.02  $\mu\text{L}$  colored fluorescent dyes or M13 tail (FAM, VIC or Hex, NED and PET) with a size, respectively, 494, 538, 535, 546 and nm (absorbance maximum), (M13 appointed a DNA fragment with the sequence complement to a tail added into 5' of the Forward primer for gene reagent or interest fragment with the two primers), 0.02  $\mu\text{L}$

unlabeled Reverse primer (←), 0.6  $\mu\text{L}$  each dNTP 100 mM (25  $\mu\text{mol}$ ), e 0.05  $\mu\text{L}$  di AmpliTaq-DNA Polymerase (Applied Biosystems, Foster City, USA). The reactions of SSR markers for screening were grouped on the basis of the annealing temperature. This was verified for each marker through the formula:  $69.3 + [0.41 \times (\%GC) - 650/L]$ , The amplification products were analyzed by means of capillary electrophoresis (ABI-3130, Figure 2.), multiplexing different fluorescent dyes. Later, we observed the peaks specificity, shown by the program Genemapper version 4.0. Indeed, This program was analyzed through GeneMapper V.4. The program Genemapper was used to identify the size of the fragments amplified through the primers used (Singer and Burke, 2003).

After amplification, the samples were loaded into a sequence for amplification reading. It is critical to the sizing method employed by the Gene Mapper, that there are at least two size standard fragments larger than the largest amplified unknown fragment. The size standards are loaded in the same capillary as the experimental samples. The sample and size standard fragments undergo the same electrophoretic forces, therefore, the relative electrophoretic mobility of any sample fragment is a good indicator of its molecular weight, as there is



**Figure 2.** Profiles of the two saffron cultivars as revealed by SSRs. A and B show the allelic segregation of the SSR markers CSMIC27 and CSMIC30, respectively, in the analyzed saffron cultivars. Lanes SAF12 and SAF3, respectively. The vertical line in each chart show DNA size marker (20-bp).

**Table 2.** Genetic diversity profile of 25 SSR primers across 22 selections of saffron.

N	Primer	Sequence	Annealing temperature (°C)	Fragment size	Fragments number
1	CSMIC8	5'TCTTGGAATGGTTAGAGCGTGT3' 5'ACCAGCAATCTTTGGAACAGTC 3'	56	300-360	2
2	CSMIC9	5'ACTGAAAGAAAAGGGGAGAAAGG 3' 5'TATATCGAATGGAGGTTCCGTC 3'	55	230-310	3
3	CSMIC10	5'AATCACACACAACATGGTTCGTT 3' 5'TGTTTAAACCAGCTAGCAGAAT 3'	55	300-380	1
4	CSMIC21	5'CTCGCTAGCCGAATCACAACT 3' 5'TGTGAGAGCACACGGTGT 3'	54	200-320	2
5	CSMIC27	5'TGGATATACGTAATCCAGACT 3' 5'ATAATTCCTGAGGTGAATT 3'	55	300-350	2
6	CSMIC28	5'TAAGGCCATGCTAGCAGAAT 3' 5'AGCAGTAATTCGTAGCGACA 3'	56	310-360	1
7	CSMIC30	5'GCAGAATCACACATGGTTACAA 3' 5'AGTGACACGACACTCACTAT 3'	54	260-320	3
8	CSMIC38	5'GTCTAAGGCCTAGCTAGCAGA 3' 5'GTTATCAAATGTTGGCCCACT 3'	56	300-360	2
9	CSMIC39	5'GCTAGCAGAATCACTACTTGA 3' 5'AATGTTGGCCCACTCACACT 3'	57	350-450	2
10	CSMIC43	5'GCAGAATCACTACTTGAAGACA 3' 5'TGAGATGGATATATTCTCTGA 3'	53	250-320	2
11	CSMIC44	5'CAGTGCTTCGGCTGAATGTGAA 3' 5'ACTGCTGGACGGTGCAACTT 3'	55	200-280	2
12	CSMIC45	5'CCGCCTAGCTAGCAGAATCACA 3' 5'GATAAGACCTGCAACTTCAACT 3'	58	160-220	2
13	CSMIC46	5'GTACAGTGTGAAGAGGAGGA 3' 5'TGGATACGCTGCACGTATCTCA 3'	54	220-300	2
14	CSMIC47	5'ACCAGGTCAGTTGATGCCTCAT 3' 5'CAGTGTAGCTACTTAGACAGT 3'	55	250-350	2
15	CSMIC49	5'ACTAGTTCACTCATCCGTTA 3' 5'TGAATCGAATGGGTAGGGAAT 3'	57	200-360	1
16	CSMIC51	5'GACGGGTAGTAGAAAGTTCTTCA 3' 5'CGAATGGGTCTCAAACCCT 3'	56	250-300	2
17	CSMIC53	5'GCAGAATCACTGCTGGACGGGT 3' 5'CAGTGCTTCGGCTGAATGTGAA 3'	55	220-280	2
18	CSMIC54	5'AGCAGCAGAGAAGTAAGACAGT 3' 5'TCAAATTCCTCAACCACTTTGA 3'	56	160-250	3
19	CSMIC55	5'AGCAACAGAGGCACACATTCA 3' 5'AGCTGTCAGTCCAATCATCAAC 3'	58	270-340	2
20	CSMIC56	5'CTTATTGGATACGCTGCA 3' 5'TAAGCCTAGCTAGCAGAA 3'	55	260-310	2
21	CSMIC57	5'GTAACCTGCTCCAGTGCTA 3' 5'TAAGCCTAGCTAGCAGAATCT 3'	54	320-400	3
22	CSMIC58	5'TCGTCATGGTCGTCGCTACTA 3' 5'AGCCTAGCTAGCAGATCATAGA 3'	55	260-340	2
23	CSMIC61	5'TTCAAGTGCTTATTGGTCCA 3' 5'CATGTTCAATGCTTCATCAAGT 3'	55	318-380	1
24	CSMIC62	5'CCAATCTGAGGACGGGCT 3' 5'AGAAGCGTGATGAAGTGA 3'	55	350-400	1
25	CSMIC7	5'GAATCACCACCTGAATTGTGAG 3' 5'AAGAGGTCGAAGAAGGGAAAAAC 3'	53	280-360	2
Total					50

**Table 3.** Genetic diversity profile of 5 SNP primers across 22 selections of saffron.

N	Blast with <i>Crocus</i> <i>Sativum</i>	Sequence	SNPs	Annealing temperature (°C)	Fragment length	Fragment (T)	Fragment (P/M)
1	96%	5'TATGCCTGCCCTAACCGAAA 3' 5'CCTTCACGAGCAAGATCACG 3'	C/T	59	165	2	Poly
2	96%	5'GGTCTCCAAATCCCCTTCCA 3' 5'GGTATCCCTTACCACACCA 3'	A/C	59	239	2	Poly
3	98%	5'AGGTGCGGAGAACAACAACTC 3' 5'TCGTCGAGATTCAAGGATTCAA 3'	G/A	57	209	2	Poly
4	99%	5'CGTGTACAGCGGATTACTC 3' 5'ACATCCGCCCTTATCTAGC 3'	C/A-A/T	58	154	4	Poly
5	98%	5'AAGACCCAGATCCGGAAGTG 3' 5'TGGTCGAAGCCATACCCAAT 3'	T/C	59	239	2	Poly

no capillary-to-capillary variation (Bio-Rad Laboratories, Hercules, CA USA). The single peaks were sized and were entered into a binary matrix as discrete variables. Considering the electrophoresis results, DNA fragments were scored on a binary scale: Present (1) or absence (0). Due to the triploidy of *C. sativus*, different parameters of genetic diversity were measured using software (Clark and Jasieniuk, 2011). We used 1 for presence and 0 for absence of the character and this data matrix was subjected to further analysis. The software NTSYSpc version 2.1 (Rohlf, 2000) was used to calculate the pair wise differences matrix and plot the dendrogram among saffron cultivars. Cluster analysis was based on similarity matrix obtained with the un-weighted pair-group method using arithmetic averages (UPGMA) (Vavrek, 2016).

### SNP Primers and PCR Conditions

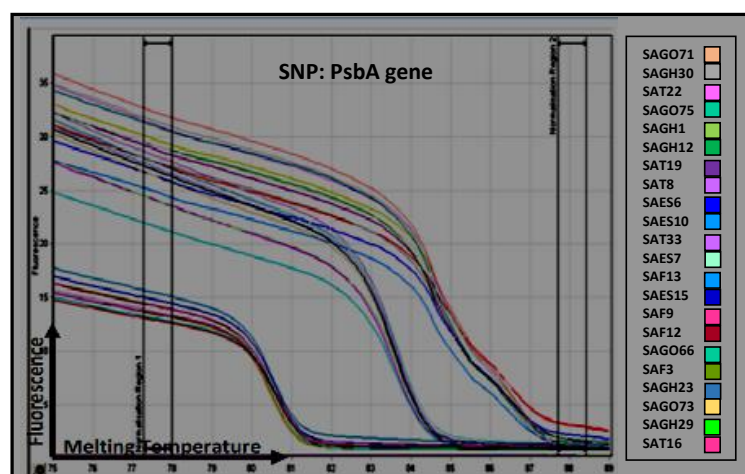
PCR amplifications were performed on a Rotor-Gene 6000, real-time PCR Thermocycler, using 1.25 units of TakaRa Taq DNA polymerase in a total volume of 25  $\mu$ L containing 10 ng of cDNA, 1X RT PCR buffer, 100  $\mu$ M of each dNTP 5 pmoles of each PCR primer,  $Mg^{2+}$  1.5 mM and 1.25  $\mu$ L of EvaGreen dye 20X. PCR reaction was performed with a cycling of 45 cycles at 95°C for 5 seconds (denaturation) and 60°C (annealing) for 10 seconds, followed by a pre-melting hold of 72°C for 2 minutes and 50°C. For 30 seconds, and a melting step with a ramp temperature of 75-95°C. For data quality

control, PCR amplifications were analyzed through the assessment of the Ct (threshold Cycle) value, end point fluorescence level, and the amplification efficiency. The melting data were normalized by adjusting the beginning and end fluorescence signals to the same level. High-resolution melting curve analysis was performed using the HRM analysis module. Different plots of the melting data were visualized by selecting a genotype for comparison and negative first-derivative melting curves were produced from the fluorescence versus temperature plots (Figure 3). An initial screening was performed probing all the primers designed on the *Crocus sativus*, in order to identify polymorphic primers. Replicates of amplification fragments resulted polymorphic, sequenced in order to identify the typology of SNP mutations occurring, using an ABI 3130xl sequencing platform (Capper *et al.*, 2015). Sequences obtained were analyzed using DNAMAN ([www/Lynnon.com](http://www.Lynnon.com)) software, validated employing RealSNP (Sequenom, Inc.). For each couple of peaks, respective bins were assigned in order to assess the polymorphism on the progeny.

## RESULTS

### Characteristics of SSR Markers

Twenty five microsatellite markers were used to test the genetic diversity of twenty-two saffron cultivars. Five markers (20%) amplified monomorphic curve patterns. The



**Figure 3.** Frequency of SNP markers and descriptive statistics of 5 SNP markers were successful in different varieties. The failure of the remaining SNPs was likely due to the sequence complexity or the presence of polymorphisms within the flanking sequences. That means several SNP variant was identified in all individuals.

remaining 20 markers (80%) generated polymorphic peak patterns (Table 2). A total of 50 alleles were detected by the SSR markers and 44 alleles were polymorphic, thus revealing 88% of polymorphism. The majority of polymorphic SSR loci (16) generated two alleles: almost an average of 2 alleles per locus, and the size of the amplified fragments ranged from 160 to 400 bp. The identification of polymorphism and monomorphism for the used markers was done by screening all the markers. We read and analyzed the peaks obtained from Genemapper software. The monomorphic markers showed a single peak with a size equal in all geographical variations, though the height of peaks could be different (y-axis), but the size of the amplified fragments was the same (x-axis), in contrast to the polymorphic markers carrying peaks with the same shape but with different sizes. All SSR markers used in the present research had two fragments. Some minor peaks were produced by some SSR markers (Figure 2). The existence of these minor peaks may have been affected, but were not considered during the allele scoring process.

Moreover, the SSR markers CSMIC7, CSMIC8, CSMIC21, CSMIC27, CSMIC38, CSMIC39, CSMIC43, CSMIC44, CSMIC45, CSMIC46, CSMIC47,

CSMIC51, CSMIC53, CSMIC55, CSMIC56 and CSMIC58 (Table 2-3 and Figure 2) were among the twenty-five SSR markers that showed two alleles. The four markers showed 3 alleles and the five e markers showed one allele (data not shown).

### Genotype Identification by SSR Markers as Unique Markers

Unique markers are defined as peaks that specifically identify varieties from the others by their presence or absence. The peaks that are present in one variety but not found in the others are termed Positive Unique Markers (PUM), opposite to the Negative Unique Markers (NUM). Unique DNA markers were obtained by SSR and were used to characterize the twenty-two saffron cultivars of different origins. In the present study, five primers out of the 25 revealed eleven unique SSR alleles (4 positive and 7 negative) as recorded in Table 4. SAT8 was characterized with the highest number of unique markers (7), two positive with primer CSMIC39 and one with CSMIC56. On the other side four negative as 2 with CSMIC38 primer and 2 with CSMIC61 primer. While the cultivar SAGH12 was characterized by

**Table 4.** Saffron cultivars characterized by positive unique markers (PUM) and/or negative unique markers (NUM), marker size and total number of markers identifying each cultivar.

cultivar	PUM		NUM		Total
	primer	Size of the marker plck (bp)	primer	Size of the marker pick (bp)	
SAT8	CSMIC39, CSMIC56	354 and 380; 270	CSMIC38, CSMIC61	318 and 346 352 and 362	7
SAGH12	-		CSMIC53	242 and 254	2

two negative unique marker with primer CSMIC53.

#### Genotype identification by SNP markers

After reviewing similar sequences in the saffron family, five pair primers were designed of SNP molecular markers in saffron. We concluded that all five markers were polymorphic forms in *Crocus sativus*. In these twenty-two Saffron genomes, SNPs were classified into transitions (C/T and A/G) and transversions (A/T and C/G) according to their substitution types, and the transitions were more common than the transversions among the sequence changes. A total of 5 candidate SNPs were validated using direct sequencing of PCR products. The SNPs were chosen based on a range of redundancy and cosegregation scores and predicted expression of multiple genes. Although every effort is made to produce and submit sequence of only the highest quality, the high-throughput nature of the sequencing programs inevitably leads to the submission of inaccuracies. The frequency of occurrence of a polymorphism at a particular locus provides a measure of confidence in the SNP representing a true polymorphism and is referred to as the SNP redundancy score. By examining SNPs that have a redundancy score of two or greater,

i.e. two or more of the aligned sequences represent the polymorphism, the vast majority of sequencing errors are removed. Although some true genetic variation is also ignored due to its presence only once within an alignment, the high degree of redundancy within the data permits the rapid identification of large numbers of SNPs from data collated from a variety of sources.

#### Genetic Diversity and Relationships of Saffron Cultivars

To examine the genetic relationships among the twenty-two saffron cultivars under study based on the SSR results, the data scored from the 30 primers were compiled and analyzed using the Dice similarity coefficient. The Genetic Dissimilarity (GD) matrices based on the Dice coefficients was calculated. Similarities among the twenty-two saffron cultivars ranged from 9.5 to 88.7%. The highest value of 88.7% was observed between SAT 33 and both of SAF3 and SAT8, while the lowest value of 9.5% was observed between SAGH23 with both SAGH30 and SAES15. The relationships between all the Iranian saffron cultivars were estimated based on the Dice coefficients. In fact, the correlation between different regions of Iran were expressed.



The local cultivar SAT19 is relatively closely related to those of SAGO75 (84.7%) and SAT33 (85.9%), but distantly related to SAES10 and SAES7 (52.7%). SAT33 showed highest similarity with SAT8 (86.1%), followed by SAF3 (82.5), but distantly related to SAES6 (34.3%) and SAGH12 (43.5%). The distance matrix based on SSR data sets was used to construct a dendrogram, which is shown in Figure 4. The genetic closeness among the selections can be explained by the high degree of commonness in these selections, due to lack of sexual reproduction and geographically closed locations.

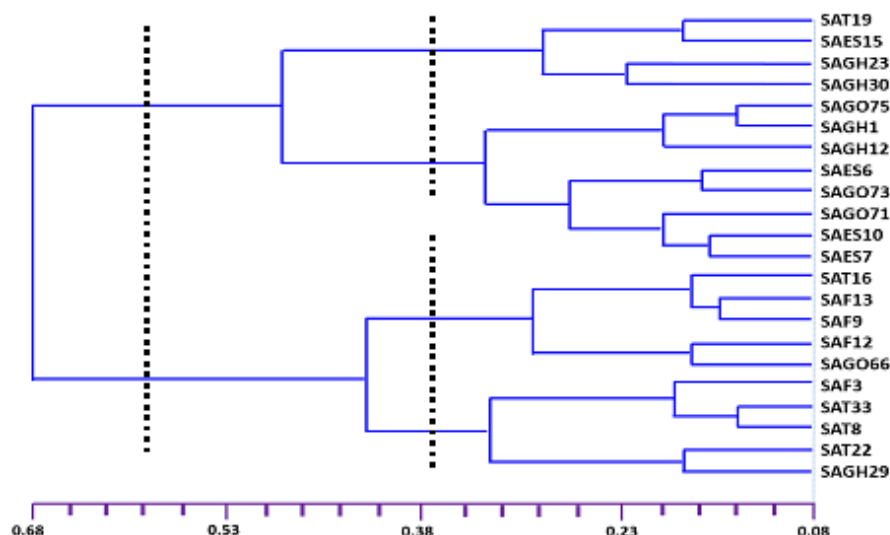
### Analysis of Saffron Varieties in Clusters

Considering the constructed dendrogram for different species from various geographical regions (Figure 4), results of the UPGMA, the existence of genetic diversity among *C. sativus* was observed. All the 22 selections formed four main clusters with some degree of sub-clustering within. Clusters I, II, III and IV represent four

selections (SAT19, SAES15, SAGH23 and SAGH30), 8 (SAGO75, SAGH1, SAGH12, SAES6, SAGO73, SAGO71, SAES10 and SAES7), 5 (SAT16, SAF13, SAF9, SAF12 and SAGO66) and 5 (SAF3, SAT33, SAT8, SAT22 and SAGH29), respectively.

The first cluster belongs to the central part of Iran and is divided into two sub-clusters. First sub-cluster contains the two cultivars SAT19 and SAES15, each in one branch. The second sub-cluster contains the other two cultivars, i.e. SAGH23 and SAGH30. The second cluster, was further divided into two sub-clusters, each of which was divided again into two clusters. One of these sub-clusters contained only the cultivar SAGH12, while the second was divided into several branches, the first contained SAES6 and SAGO73 and the second contained SAGO71, SAES10 and SAES7. The third cluster was divided into two branches on the other side SAF13 and SAF9 together with the second branch contained SAF12 and SAGO66. Noticeably, the cultivars SAT22 and SAGH29 together were in the second sub-cluster of the fourth main cluster.

Cluster I shows average similarity of 59.5,



**Figure 4.** Dendrogram for the 22 Saffron cultivars constructed from SSRs data analysis using un-weighted pair group with arithmetic average similarity matrices computed according to Dice coefficients.



38.6, and 44.3% with clusters II, III and IV, respectively. However, within the cluster I, the average similarity was 78.4% and within clusters II, III, and IV it was 76.8, 80.7, and 78%, respectively.

## DISCUSSION

The present study revealed the genetic diversity within a group of saffron genotypes of different origins cultivated in Iran using molecular markers (SSR and SNP) approaches (Figure 1). The possibility and application of the SSR technique in varietal identification of saffron have been well explored (Pritesh *et al.*, 2010). Molecular markers can serve as useful tools for detecting the level of polymorphism in *C. sativus* populations studied. Presently, there are no reports of the study of genetic diversity in Saffron through SNP markers. It would be interesting to extend the study of genetic diversity using a larger number of SSR loci and sample size (Eckert *et al.*, 2008). The transmissibility of SSR loci isolated from different varieties of *C. sativus* can implicate the conservation of genic regions through the *Crocus* genus. This result can significantly help to reduce costs of genetic analysis in future breeding researches (Moretzsohn *et al.*, 2004).

Many of the selected primers produced amplicons, even when using modified amplification conditions. But some markers, five SSR markers that produced reproducible were monomorphic, and in this study, all markers had polymorphic SSR loci and generated two fragments. There was high percentage of polymorphism in the present research. It is reported that, in iridaceae plants, there is a low frequency of polymorphism among cultivars and intraspecific (Mir *et al.*, 2015), probably because saffron is a triploid and sterile plant. In particular, this plant has asexual reproduction.

Since the present markers did not reveal high significant polymorphism at genetic level between different saffron selections,

there is the need to develop more markers, particularly SSR and SNP markers, to be used as new markers for high density profiling. The detection of minor and nonspecific products that could be shadow, heteroduplex, or faint peak may affect the allele scoring process and increases the difficulty of legitimate allele identification. We did not consider these minor peaks during allele scoring, however, Wang *et al.* (2005) and Rodriguez *et al.* (2001) reported that the minor peaks can be useful during gel scoring for genotype verification, because they are generally consistent.

In this research, it was proved that, although the regional and climatic conditions of the investigated saffron clones were close to each other, but in terms of genetic make up, they were different. In our study, *Crocus sativus* accessions from different geographical regions were included in different groups. Unlike previous research of Alavi-Kia *et al.* (2008) and Rubio-Moraga *et al.* (2009) who have reported that there is no genetic variation within *C. sativus*, microsatellites indicated the genetic diversity among *C. sativus* from different regions of Iran. Based on genetic similarity, the cultivated genotypes may have originated from common parents. Spontaneous mutation and artificial selection may have led to existing cultivars. And also, contrary to reports of Namayandeh *et al.* (2013), although some of *Crocus sativus* accessions tested in this study had a similar geographic origin, but in terms of SNP and SSR markers, they were grouped in different clusters. That issue again establishes the genetic diversity between samples of this plant. Selections chosen for present study showed variation at molecular level, which can further be deciphered by using more number of molecular markers for high throughput scanning and high density profiling of saffron genome. Further study needs to be done to identify divergent selections with respect to different traits for improvement of this crop. The close relationships, and sometimes high genetic diversity, among

saffron accessions revealed in this study can be due to vegetative propagation, human selection of superior genotypes, and existence of narrow genetic base of saffron. Our results confirmed that the SSR markers, and specially SNP markers, are powerful tools and effective marker system for evaluation of genetic diversity among saffron accessions.

To slow down or stop geographical decline in saffron cultivation and, even, to achieve an appreciable expansion of saffron-producing areas, a substantial increase in productivity by selecting superior clones is the need of the hour. The creation of genetic map and new cultivars based on the identification of clones with higher second metabolites potential will facilitate this process. There is an urgent necessity to study the genetic and selection potential of various populations of saffron in Iran, as well as the populations in each region separately. Further research of the genetic diversity and population structure of Iranian saffron using SSR markers is needed to identify the origin and wild relatives of this valuable species. Also, it would be useful to breeders to determine how genetic diversity, as analyzed with molecular markers, relates to phenotypic variability and, more importantly, how it reflects the variability of the important traits.

### CONCLUSIONS

The present study revealed the genetic diversity within a group of saffron genotypes of different origins and cultivated in Iran, using molecular markers (SSR and SNP) approaches (Figures 1, 2, and 4). This genetic difference created the saffron corms tested from different regions of the country, even though most of these areas were close. There are two main results: first, although saffron is a sterilization plant, saffron population in one area can have high genetic diversity. Second, plants cultivated in a region can also, benefit from high genetic diversity. The possibility and application of

the SSR and SNP technique in varietal identification of saffron have been well explored (Namayandeh *et al.* 2013). Another aim of this study, which we achieved, was to use microsatellite markers isolated to evaluate the efficiency of these SSR markers in detecting genetic polymorphisms in Iranian saffron.

### ACKNOWLEDGEMENTS

This work was funded by the University of Torbat Heydariyeh in Iran (UTH) (Project number: 80024-225) and Saffron Institute of Torbat Heydariyeh. The authors thank Prof. Karbasi and Prof. Alipanah for their assistance in the acceptance and performance of the project, respected reviewers for their valuable comments and discussions on the manuscript, and the agricultural assistant by Dr. Feizi for his help.

### REFERENCES

1. Ahmad, M., Zaffar, G., Mir, S. D., Razvi, S. M., Rather, M. A. and Mir, M. R. 2011. Saffron (*Crocus sativus* L.) Strategies for Enhancing Productivity. *Res. J. Med. Plant*, **5**: 630–649.
2. Ahrazem, O., Rubio-Moraga, Á. and Gomez Gomez, L. 2010. The Expression of Chromoplast Specific Beta Lycopene Cyclase Gene is Involved in the High Production of Saffron Precursors. *J. Exp. Bot.*, **61**: 105–119.
3. Alavi-Kia, S. S., Mohammadi, S. A., Aharizad, S. and Moghadam, M. 2008. Analysis of Genetic Diversity and Phylogenetic Relationships in *Crocus* Genus of Iran Using Inter-Retrotransposon Amplified Polymorphism. *Biotechnol.*, **22**: 795–800.
4. Batley, J., Barker, G., O'Sullivan, H., Edwards, K. J. and Edwards, D. 2003. Mining for Single Nucleotide Polymorphisms and Insertions/Deletions in Maize Expressed Sequence Tag Data. *Plant Physiol.*, **132**(1): 84–91.
5. Beiki, A.H., Keifi, F. and Mozafari, J. 2010. Genetic Differentiation of *Crocus*



- Species by Random Amplified Polymorphic DNA. *Genet. Eng. Biotechnol. J.*, **18**: 1–10.
6. Bjarne, L., Jihad, O., Carsten, P., Marian, Ø. 2015. Large Intraspecific Genetic Variation Within the Saffron-Crocus Group (*Crocus* L., Series *Crocus*; Iridaceae), *Plant Syst. Evol.*, **301**:425-437.
  7. Capper, R. L., Jin Y. K., Lundgren P. B., Peplow L.M., Matz M. V. and Van Oppen, M. J. H. 2015. Quantitative High Resolution Melting: Two Methods to Determine SNP Allele Frequencies from Pooled Samples. *BMC Genet.*, **16(62)**: 1–13.
  8. Caiola, M.G., Leonardi, D. and Canini, A. 2010. Seed Structure in *Crocus sativus* L. × *C. cartwrightianus* Herb., *C. thomasi* Ten., and *C. hadriaticus* Herb. at SEM. *Plant Syst. Evol.*, **285(1)**: 111–120.
  9. Ching, A., Jung, M., Dolan, M., Smith, O. S., Tingey, S., Morgante, M. and Rafalski, A. J. 2001. SNP Frequency, Haplotype Structure and Linkage Disequilibrium in Elite Maize Inbred Lines. *BMC Genet.*, **3(19)**: 3–19.
  10. Clark, L.V., Jasieniuk, M. 2011. Polysat: an R Package for Polyploid Microsatellite Analysis. *Mol. Ecol. Res.*, **11**: 562–566.
  11. D'Agostino, N., Pizzichini, D., Chiusano, M. L. and Giuliano, G. 2007. An EST Database from Saffron Stigmas. *BMC Plant Biol.*, **7**: 53–53.
  12. Eckert, C. G., Samis, K. E. and Lougheed, C. 2008. Genetic Variation Across Species Geographical Ranges: The Central Marginal Hypothesis and Beyond. *Mol. Ecol.*, **17**: 1170–1188.
  13. Frizzi, G., Miranda, M., Pantani, C. and Tammaro, F. 2007. Allozyme Differentiation in Four Species of the *Crocus cartwrightianus* Group and in Cultivated Saffron (*Crocus sativus* L.). *Biochem. Syst. Ecol.*, **35**: 859–868.
  14. Garritano, S., Gemignani, F., Voegele, C., Nguyen-Dumont, T., Le Calvez-Kelm, F., De Silva, D., Lesueur, F., Landi, S., Tavtigian, S.V. 2009. Determining the Effectiveness of High Resolution Melting Analysis for SNP Genotyping and Mutation Scanning at the TP53 Locus. *BMC Genet.* **10(5)**: 5. doi.org/10.1186/1471-2156
  15. Golmohammadi, F. 2014. Saffron and its Farming, Economic Importance, Export, Medicinal characteristics and Various Uses in South Khorasan Province- East of Iran. *Int. J. Farm. Allied Sci.*, **3-5**: 566–596
  16. Grilli Caiola, M., Di Somma, D. and Lauretti, P. 2001. Comparative Study on Pollen and Pistil of *Crocus sativus* L. (Iridaceae) and Its Allied Species. *Ann. Bot. Roma.*, **1**: 93–103.
  17. Grilli-Caiola, M. and Canini, A. 2004. Ultrastructure of Chromoplasts and Other Plastids in *Crocus sativus* L. (Iridaceae). *Plant Biosyst.*, **138**: 43–52.
  18. Harpke, D., Meng, S., Kerndorff, H., Rutten, T. and Blattner, F. R. 2013. Phylogeny of *Crocus* (Iridaceae) Based on one Chloroplast and Two Nuclear Loci: Ancient Hybridization and Chromosome Number Evolution. *Mol. Phylogenet Evol.*, **66**: 617–627.
  19. Kanakis, C. D., Daferera, D. J., Tarantilis, P. A. and Polissiou, M. G. 2004. Qualitative Determination of Volatile Compounds and Quantitative Evaluation of Safranal and 4-Hydroxy-2, 6, 6-Trimethyl-1 Cyclohexene-1-Carboxaldehyde (HTCC) in Greek Saffron. *J. Agr. Food Chem.*, **52**: 4515–4521.
  20. Larsen, B., Orabi, J., Pedersen, C. and Orgaard, M. 2015. Large Intraspecific Genetic Variation within the Saffron-Crocus Group (*Crocus* L., Series *Crocus*; Iridaceae). *Plant Syst. Evol.*, **301(1)**: 425–437.
  21. Lefebvre, V., Goffinet, B., Chauvet, J. C., Caromel, B., Signoret, P., Brand, R. and Palloix, A. 2001. Evaluation of Genetic Distances between Pepper Inbred Lines for Cultivar Protection Purposes: Comparison of AFLP, RAPD and Phenotypic Data. *Theor. Appl. Genet.*, **102**: 741-750.
  22. Malentacchi, F., Forni, G., Vinci, S., and Orlando, C. 2009. Quantitative Evaluation of DNA Methylation by Optimization of a Differential-high Resolution Melt Analysis Protocol. *Nucleic Acids Research*, **37(12)**: e86.
  23. Mir, J. I., Ahmed, N., Mudasir Khan, H., Mokhdomi, T.A., Wani, S. H., Bukhari, S., Amin, A. and Qadri R. A. 2015. Molecular Characterization of Saffron-Potential Candidates for Crop Improvement. *Not. Sci Biol.*, **7(1)**: 81–89.

24. Moretzsohn, M. C., Hopkins, M. S., Mitchell, S. E., Kresovich, S., Valls, J. F. M. and Ferreira, M. E. 2004. Genetic Diversity of Peanut (*Arachis hypogaea* L.) and Its Wild Relatives Based on the Analysis of Hyper Variable Regions of the Genome. *BMC Plant Biol.*, **4**: 11–17.
25. Namayandeh, A., Nemati, Z., Kamelmanesh, M. M., Mokhtari, M. and Mardi, M. 2013. Genetic Relationships among Species of Iranian Crocus (*Crocus* spp.). *Crop Breed. J.*, **3(1)**: 61–67.
26. Nemati, Z., Mardi, M., Majidian, P., Zeinalabedini, M., Pirseyedi, S. M. and Bahadori, M. 2014. Saffron (*Crocus sativus* L.), a Monomorphic or Polymorphic Species?. *Spanish J. Agric. Res.*, **12(3)**: 753–762.
27. Petersen, G., Seberg, O., Thorsøe, S., Jørgensen, T., Mathew, B. 2008. A Phylogeny of The genus *Crocus* (Iridaceae) Based on Sequence Data from Five Plastid Regions. *Taxon*, **57**: 487–499.
28. Pritesh, P., Vishal, P., Oza, V., Chauhan, A. D. and Patel, K. B. 2010. Kathiria Subramanian R.B. Genetic Diversity and DNA Fingerprint Study of Tomato Discerned by SSR Markers. *Int. J. Biotechnol. Biochem.*, **6(5)**: 657–666.
29. Rodriguez, S., Visedo, G. and Zapata, C. 2001. Detection of Errors in Dinucleotide Repeats Typing by Non-Denaturing Electrophoresis. *Electrophoresis*, **22**: 2656–2664.
30. Rohlf, F. J. 2000. *NTSYS-PC Numerical Taxonomy and Multivariate Analysis System, Version 2.1*. Exeter Software, New York.
31. Rubio-Moraga, Á., Castillo-López, R., Gómez-Gómez, L. and Ahrazem, O. 2009. Saffron is a Monomorphic Species as Revealed by RAPD, ISSR and Microsatellite Analyses. *BMC Res. Notes*, **2**: 189.
32. Sarikamiş, G., Yanmaz, R., Ermis, S. and Bakir, M. 2010. Genetic Characterization of Pea (*Pisum sativum*) Germplasm from Turkey Using Morphological and SSR Markers. *Genet. Mol. Res.*, **9**: 591–600.
33. Singer, T. and Burke, E. 2003. High-throughput TAIL-PCR as a Tool to Identify DNA Flanking Insertions. *Methods Mol. Biol.*, **236**: 241–272.
34. Syvanen, A. C. 2001. Genotyping Single Nucleotide Polymorphisms. *Nat. Rev. Genet.*, **2**: 930–942.
35. Vavrek M. J. 2016. A Comparison of Clustering Methods for Biogeography with Fossil Datasets. *Peer J.*, 8–19. doi: 10.7717/peerj.1720
36. Wang, J. R., Wei, Y. M., Yan, Z. H. and Zheng, Y. L. 2005. Detection of Single Nucleotide Polymorphisms in 24 kDa Dimeric Amylase Inhibitors from Cultivated Wheat and Its Diploid Putative Progenitors. *Biochimica et Biophysica Acta*, **1723(1-3)**: 309–320.
37. Zubor, Á. A., Surányi, G., Gyóri, Z., Borbély, G., Prokisch, J. 2004. Molecular Biological Approach of the Systematics of *Crocus sativus* L. and Its Allies. *Acta Hort.*, **650**: 85–93.

## بررسی تنوع ژنتیکی ارقام کشت شده زعفران (*Crocus sativus* L.) در ایران با استفاده از نشانگرهای SSR و SNP

۱. یوسفی جوان، و ف. قراری

### چکیده

زعفران (*Crocus sativus* L.) یکی از گران ترین ادویه های موجود در جهان است که عمدتاً به عنوان رنگ آمیزی و طعم دهنده در صنایع غذایی مورد استفاده قرار می گیرد. از متابولیت های ثانویه



این گیاه در پزشکی استفاده می شود. مجموعه ای از بیست و دو رقم زعفران که در مناطق مختلف ایران کشت شدند، با ۲۵ نشانگر SSR و ۵ نشانگر SNP به منظور تعیین هویت و تنوع ژنتیکی در این ارقام مورد بررسی قرار گرفتند. طی استفاده از نشانگرهای SSR به طور متوسط ۵۰ عدد آلل، و قطعاتی با اندازه تقریبی بین ۱۶۰ تا ۴۰۰ نوکلئوتیدی تمایز یافت. ۳۳ آلل پلی مورفیک بودند که ۷۲٪ پلی مورفیسم را نشان دادند. شباهت ژنتیکی بر اساس داده های SSR، ۹.۵ تا ۸۷.۸ درصد تخمین زده شد. در طی این تنوع ژنتیکی، از پنج نشانگر SNP نیز استفاده گردید، که هر پنج نشانگر SNP نیز عمدتاً به دلیل حضور تنها دو آلل پلی مورفیک بود. از پتانسیل نشانگرهای SSR و SNP جهت تشخیص فواصل ژنتیکی در بین ارقام گیاهی استفاده می گردد. پس از تحلیل داده ها، و آنالیز خوشه ای، ارقام کشت شده زعفران به دو گروه اصلی، و هر یک از آن دو گروه، به دو گروه فرعی تقسیم شدند. در این مقاله سعی شده است تا تنوع ژنتیکی زعفران کشت شده، در ایران را با وجود تکثیر غیرجنسی آن نشان دهیم. با توجه به شباهت اقلیمی برخی از مناطق ایران، تغییرات ژنتیکی خاصی در بین گونه های زعفران آن مناطق مشاهده گردید. برای کشت زعفران و تولید با کیفیت بالا در سراسر جهان، تحقیقات در مورد تنوع ژنتیکی خانواده بزرگ *C. sativus* بر ارزش این محصول، شدیداً می افزاید.