Genetic structure and marker-trait associations in parental lines of sunflower (*Helianthus annuus* L.)

Hossein Zeinalzadeh-Tabrizi[1,2]*, Arash Hosseinpour[1,2], Mehdi Ghaffari[3], Kamil Haliloglu[2]

1. Crop and Horticultural Science Research Department, Ardabil Agricultural and Natural Resources Research and Education Center, AREEO, Ardabil, Iran
2. Department of Field Crops, Faculty of Agriculture, Ataturk University, Erzurum, Turkey
3. Department of Oilseed Crops, Seed and Plant Improvement Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran

Abstract

Association mapping is a reliable tool to detect the association between phenotypic and genotypic data through linkage disequilibrium. In the present study, 21 SSR and 19 TRAP markers were applied to investigate the genetic structure and association analysis in 53 Iranian sunflower parental lines, including 23 restorer and 30 cytoplasmic male sterility lines. In the phenotypic analysis, 18 morpho-physiological traits were measured. The population structure analysis identified five and two actual subpopulations (optimum *K*) across SSR and TRAP markers, respectively. Using SSR data, population differentiation measurements (F_{ST}) between the subpopulations ranged from 0.24 to 0.43 (average 0.32); using TRAP data, F_{ST} varied between 0.23 and 0.24 (average 0.24). Association mapping analysis indicated that SSR and TRAP markers were associated with 11 and 17 traits, respectively. SSR loci *Ha 494-ar*, *Ha 806-ar*, *Ha 991-ar*, *Ha 1167-ar*, *Ha 1287-ar*, *ORS-53,* and *ORS-54* were associated with seed yield per plant, oil yield per plant, seed yield, and oil yield respectively. On the other hand, several TRAP markers, including *K11F05/TRAP03*, *K11F05/TRAP03,* and *F15O11F1/TRAP03* were associated with flowering duration, maturity, and 1000-seed weight, respectively. In conclusion, the genetic structure and marker-trait associations reported here can be exploited for marker-assisted selection (MAS) in sunflower breeding programs.

Keywords: *Helianthus annuus* L., association analysis, Bayesian clustering, SSR, sunflower, TRAP

Zeinalzadeh-Tabrizi, H., A. Hosseinpour, M. Ghaffari, K. Haliloglu . 2022. 'Genetic structure and marker-trait associations in parental lines of sunflower (*Helianthus annuus* L.)*'. Iranian Journal of Plant Physiology 12* (2), 3955-3971.

Introduction

Sunflower (*Helianthus annuus* L.) is a diploid plant (2n=2x=34) with a genome size of approximately 2,800 to 3,600 Mbp (Renaut, 2017). Sunflower oil is the fourth most important vegetable oil in world trade, with an annual production of approximately 9 million tons and a cultivated acreage of over 22 million hectares (Fernández-Martínez et al., 2009). In sunflower breeding programs, genetic variations of important traits are evaluated by quantitative methods, including offspring assessments or field trials. In quantitative

* *Corresponding Author*
E-mail Address: h.zeinalzadeh@areeo.ac.ir
Received: June, 2021
Accepted: October, 2021

genetics, complex traits that are controlled by a large number of independent genetic loci are termed polygenic traits (Bartholomé et al., 2016).

Analysis of the relationship between phenotypic variation and DNA polymorphism to elucidate the genetic architecture of variations in quantitative traits is required in terms of the number, position, influence, and identity of the loci involved (Ingvarsson et al., 2011). Association mapping (AM) or linkage disequilibrium (LD) mapping and linkage analysis (LA) are the two main approaches to detect quantitative trait loci (QTLs). In previous studies, numerous linkage maps have been identified in sunflower using several molecular markers, such as amplified fragment length polymorphism (AFLP) (Berrios et al., 2000; Mokrani et al., 2002; Ebrahimi et al., 2008; Davar et al., 2011; Haddadi et al., 2011; Arias et al., 2016), simple sequence repeats (SSRs) (*Mokrani et al.,* 2002; Al-Chaarani et al., 2004; Ebrahimi et al., 2008; Talia et al., 2010; Davar et al., 2011; Haddadi et al., 2011; Eyvaznejad et al., 2014; Arias et al., 2016; Sahari et al., 2016; Mangin et al., 2017; Zubrzycki et al., 2017), target region amplification polymorphism (TRAP) (Yue et al., 2008), and single nucleotide polymorphism (SNP) (Eyvaznejad et al., 2014; Bartholomé et al., 2016). Due to the limitations on the number of recombination events in population mapping and the cost of studying a large number of individuals, LA in plants usually locates QTLs between 10- and 20-cM intervals (Holland, 2007). In addition, restrictions on the number of QTLs identified, cloned, or labeled at the gene level is a disadvantage of QTL mapping (Price et al., 2006). AM *vs.* QTL mapping method is used for complex trait analysis, which provides data on relationships between gene-level polymorphisms and phenotypic variations in existing germplasm collections without the construction of population mapping (Fusari et al., 2012).

In recent years, AM has been developed as a tool for detecting the relationship between complex trait changes at the gene-sequence level through the use of evolutionary and historical recombination phenomena at the population level (Nordborg et al., 2002). The reduced cost and time of research, greater allele numbers, and enhanced mapping resolution are three advantages of the AM approach compared with QTL mapping (Yu et al., 2006). As the AM method was first introduced in plants (Thornsberry et al., 2001), advances in statistical methods, interest in the discovery of superior and novel alleles, and developments in the field of efficient genomic technologies have increased the popularity of the AM approach in genetic research.

AM in sunflower has been implemented by various types of molecular markers, SSRs (Darvishzadeh, 2012; Moreno et al., 2012; Filippi et al., 2015; Soleimani-Gezeljeh et al., 2018) and SNPs (Moreno et al., 2012; Mandel et al., 2013; Filippi et al., 2015; Nambeesan et al., 2015; Bartholomé et al., 2016). Hu et al. (2003) introduced a PCR-based TRAP molecular marker technique in sunflowers using gene-based sequence information (EST) to generate polymorphism. To date, only seven AM reports on sunflower have been published worldwide (Kolkman et al., 2007; Fusari et al., 2008; Fusari et al., 2012; Cadic et al., 2013; Mandel et al., 2013; Talukder et al., 2014; Nambeesan et al., 2015).

Knowledge on the genetic structure and variability levels of Iranian sunflower breeding materials using AM is still limited, indicating an incomplete diversity map of the cultivated sunflower. Moreover, no information is available on the association analysis in sunflower using TRAP markers. Hence, the objectives of this study were to assess the population structure, LD, and QTLs detected by LD-mapping approaches in cytoplasmic male sterility (CMS) and restorer parental lines in the Iranian sunflower using association analysis with SSR and TRAP markers. The QTL results for these markers were also compared to assess consistency and complementarity.

Materials and Methods

Plant materials and phenotypic assay

In this study, 53 sunflower genotypes (*Helianthus annuus* L.), including 30 CMS and 23 restorer inbred lines, were used as plant materials. To measure phenotypic data, a randomized complete block design with three replications was used at the Research Station in the Seed and Plant

Improvement Institute (SPII), Karaj, Iran (35˚79` N, 50˚93` E, 1382 m a.sl.) over three consecutive years (2014-2017). The experimental station has been characterized as having a cold and temperate climate with an average annual temperature of 10.5 ℃. The average annual precipitation is 388 mm and most rainfall occurs in the winter rather than in the summer.

Each plot consisted of one 5-m-long row with row spacing of 75 cm. Irrigations were performed based on the class A evaporation pan as described by Akbari et al. (2008). Mechanical weed control was performed two times during the experiment. The plant heads were covered to prevent bird damage during the seed-filling period.

Phenotypic data were measured using 18 morpho-physiological traits, namely flowering initiation (FI), flowering completion (FC), flowering duration (FD), maturity (MA), seed-filling period (SFP), plant height (PH), head diameter (HD), stem diameter (SD), leaf number (LN), head weight (HW), head productivity (HP), 1000-seed weight (SW), seed number per head (SN), oil content (OC), seed yield per plant (SY1), oil yield per plant (OY1), seed yield per hectare (SY), and oil yield per hectare (OY). In each plot, five plants were randomly selected to measure all traits.

Total genomic DNA isolation

Total genomic DNA was isolated from 3-week-old seedlings as described by Zeinalzadehtabrizi et al. (2015). Quantity and quality of isolated DNA were tested by NanoDrop® ND-1000 UV/Vis spectrophotometer and 0.8% agarose gel electrophoresis, respectively.

Genotyping with SSR markers

Twenty-one SSR primer pairs were designed according to (Paniego et al., 2002; Tang et al., 2002), whose sequences are shown in Table 1. Only polymorphic primers in prior sunflower research were selected and used for this study. PCR amplifications were performed in a 20-µl volume containing 1X PCR buffer (20 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$), 10 µM dNTP, 25 mM $MgCl_2$, 100 pmol of each primer, 1U Taq DNA polymerase, and 50 ng of DNA template. The amplifications were run in a Multi Gene Gradient Thermal Cycler TC9600-G-230V (Labnet International Inc.) under the following conditions: 96 ℃ for 2 min, followed by 30 cycles of 96 ℃ for 30 s, 50-60 ℃ for 45 s for annealing (based on the appropriate temperature for each pair of SSR primers), 72 ℃ for 1 min, and a final extension of 72 ℃ for 15 min. The amplified products were separated by 2% agarose gel in 1X SB buffer at 200 V for 120 min, stained with ethidium bromide (0.2 µg/ml), and visualized under a UV trans-illuminator.

Genotyping with TRAP markers

Similar to SSR markers, only 19 polymorphic TRAP primer pairs based on Yue et al. (2009) were used. The sequences of these primers are presented in Table 2. PCR reactions were prepared as described for genotyping with SSR markers. PCR cycles for TRAP markers were a touchdown procedure with an initial denaturation at 94 ℃ for 2 min, followed by 5 cycles of 94 ℃ for 45 s, 35 ℃ for 45 s, and 72 ℃ for 1 min, followed by cycles of 94 ℃ for 45 s, 50 ℃ for 45 s, and 72 ℃ for 1 min, and a final extension of 72 ℃ for 15 min. The annealing temperature for all pairs of TRAP primers was fixed. The amplified products were then separated by polyacrylamide Mega-Gel dual vertical electrophoresis (Model C-DASG-400-50) at 200 V for 105 min. The gel solution consisted of 6% (w/v) acrylamide/bis-acrylamide (19:1) in 0.5X TBE buffer, 0.07% (w/v) ammonium persulfate, and 0.08% (w/v) TEMED. Finally, gels were visualized under UV light and photographed by a digital camera (Nikon Coolpix 500).

Statistical genotyping analysis

In both SSR and TRAP markers, PCR products were scored as absent (0) or present (1). Data were compiled as a binary matrix using the Analysis System (NTSYSpc version 2.0) according to the Dice similarity matrix (Dice, 1945). Using the same software, an unweighted pair-group mean average (UPGMA) tree was constructed. A Bayesian model-based clustering method was performed to evaluate the population structure and define clusters of individuals based on multi-locus genotypes using STRUCTURE ver. 2.3.4 software (Pritchard et al., 2000). The *K* values (the

Table 1.
SSR primers used in the study

No.	Primer		Sequence (5'-3')	Locus	Motif	References
1	Ha 432	F	CTTTATCCCCCACCCCCTCC	Ha432-ar	(GT)10	Paniego *et al.* (2002)
		R	GGGTTTAGTGGCCAGTAGTTGTC			
2	Ha 514	F	GGTCAACGGATTTAGAGTC	Ha514-ar	(GA)13	Paniego *et al.* (2002)
		R	GTATTGATTCCAACATCCAG			
3	Ha 1327	F	CCGTTAGGTAGTTTACTTGCGAC	Ha1327-ar	(ATT)30	Paniego *et al.* (2002)
		R	GGTGGGGGGAATATTCTGAGGTG			
4	Ha 1442	F	GCTTATGTGCTTACGTGTTCCTG	Ha1442-ar	(ATT)31	Paniego *et al.* (2002)
		R	CTAAACAGTTCGGCGAGTGTAGG			
5	Ha 1608	F	GATCTTAGGTCCGCCAC	Ha1608-ar	(ATT)25	Paniego *et al.* (2002)
		R	GATGGCATTTGGCTAGAC			
6	ORS-6	F	GTGGAGAGAGGTGTAGAGAGC	ORS-6	AGG	Tang *et al.* (2002)
		R	CACCCCTCACCCTGACAC			
7	ORS-5	F	ATCTGGAGCAGCAAATTCAG	ORS-5	AAC	Tang *et al.* (2002)
		R	CTGCTGCCCACCATACTG			
8	HNCA-2	F	TGAGACAAGCATAAGCAC	HNCA-2	GT	Tang *et al.* (2002)
		R	TAGACAAGACAAGGGACT			
9	Ha 357-ar	F	GTGGGTGTGGAAGGAAGAATC	Ha357-ar	(GA)15	Paniego *et al.* (2002)
		R	CAGACACATGCTAGTCGTCGTG			
10	Ha 360-ar	F	CAACAAGGAACCGATAACTGCT	Ha360-ar	(GA)15	Paniego *et al.* (2002)
		R	CACCCTTCATCTCCTTC			
11	Ha 494-ar	F	GCGTTGGTTAAGGCCTGAGGTC	Ha494-ar	(GA)17 A (GA)2 N12 T15	Paniego *et al.* (2002)
		R	GAGCAGCAAACAGAGGGTACACC			
12	Ha 806-ar	F	GATGTTCCTTCCTGCAC	Ha806-ar	(GT)8 N27 (GA)6	Paniego *et al.* (2002)
		R	GGTTGGATAATGGGGCAGC			
13	Ha 991-ar	F	GCCCCCTTGATGCCCTTTTC	Ha991-ar	(GA)4 T (GA)12	Paniego *et al.* (2002)
		R	GAATCGCCATTTGAATCGCCAG			
14	Ha 1167-ar	F	CGATGTCGGCGATTCGGACTGGAG	Ha1167-ar	(GT)9 N2 (GT)4	Paniego *et al.* (2002)
		R	CCCCATCTACACTTCAATACTG			
15	Ha 1287-ar	F	GATATGAGCCCATCACTCATC	Ha1287-ar	(GA)26	Paniego *et al.* (2002)
		R	GAAGATATGTCAGGTCACACCC			
16	ORS-31	F	AATTCATGCCCCAAGAGATG	ORS-31	$(AAG)_{10}$	Tang *et al.* (2002)
		R	CACAATTCATGCATTTCTCTGG			
17	ORS-53	F	GCTGGCAATTTCTGATACACGAT	ORS-53	$(T)_{30}$	Tang *et al.* (2002)
		R	CATCTAGACAACGACAGAAGATG			
18	ORS-78	F	GTTCGTCGAGTACATGTTCTGC	ORS-78	$(AAG)_{10}$	Tang *et al.* (2002)
		R	TTTCCCTCTGGAAAGTTGTCA			
19	ORS-176	F	CCCTAACTGGTTTTCTGACCC	9D7	$(TG)_{16}$	Tang *et al.* (2002)
		R	AACTTTTGTTTGTTTGTCCAGG			
20	ORS-204	F	CGTCTGGCATTATGAAATCGTC	14D3	$(GT)_{17}$	Tang *et al.* (2002)
		R	CCGCATAACAGCAATGGTCAAC			
21	ORS-254	F	AAATCCCACTTCATACAAACGT	10G9	$(TACA)_{25}$	Tang *et al.* (2002)
		R	CCTTCAGTGCTCATGCAGTG			

number of genetically distinct groups) were adjusted to change from 1 to 10. To obtain the optimal *K* value, 10 independent simulations were run using a burn-in length of 1,000,000 and a run length of 1,000,000 iterations. Following the program's settings, the no-admixture model was used, and correlated allele frequencies among populations were assumed. To infer the appropriate number of *K*, the structure output was analyzed using online STRUCTURE HARVESTER

Table 2
The TRAP primers used in the study

Fixed primers	EST ID or GenBank accession no.	Sequences 5′–3′	
MSP03	Putative male-sterile gene	GTTGCCATGGACATCAACAC	(Yue *et al.*, 2009)
A11D14F	QHA11D14	GGGGTTCTAAACAAGGTG	(Yue *et al.*, 2009)
B14G14B	QHB14G14	AATCTCAAGGACAAAAGG	(Yue *et al.*, 2009)
F15O11F1	QHF15O11	CTGGAGCCAAGACATCTG	(Yue *et al.*, 2009)
K11F05	QHK11F05	GAACCAAACTGGGGTATGTA	(Yue *et al.*, 2009)
MAX3BR	AT2G44990	ACGTTATGAGCCCCATGAAGA	(Yue *et al.*, 2009)
Arbitrary primers			
TRAP03	IR-700	CGTAGCGCGTCAATTATG	(Yue *et al.*, 2009)
Sa12	IR-700	TTCTAGGTAATCCAACAACA	(Yue *et al.*, 2009)
Sa4	IR-700	TTCTTCTTCCCTGGACACAAA	(Yue *et al.*, 2009)
TRAP13	IR-800	GCGCGATGATAAATTATC	(Yue *et al.*, 2009)
Ga3	IR-800	TCATCTCAAACCATCTACAC	(Yue *et al.*, 2009)
Ga5	IR-800	GGAACCAAACACATGAAGA	(Yue *et al.*, 2009)

(Earl, 2012). The mean of F_{ST} (F-statistics) parameter for the structure inferred groups were also estimated. An inference of the estimates of ancestry of the individuals (Q-matrix) in the selected subpopulations was followed (Pritchard et al., 2000). To identify marker-trait associations, a mixed linear model (MLM) approach was used. The MLM method incorporates both Q-matrix and kinship matrix (K-matrix) as covariates in the analysis. SSR and TRAP data were used to estimate K-matrix and the matrix of pair-wise relationship of genotypes using TASSEL version 5 software (Bradbury et al., 2007). The Q-matrix was obtained from STRUCTURE 2.3.4 adapted at K=5 and K=2 for SSR and TRAP markers, respectively. P-values for each estimated r^2 were calculated using 1000 permutations in TASSEL version 5 (Bradbury et al., 2007). Furthermore, determination of kinship coefficients and LD, as well as cluster analysis, were conducted using TASSEL version 5.

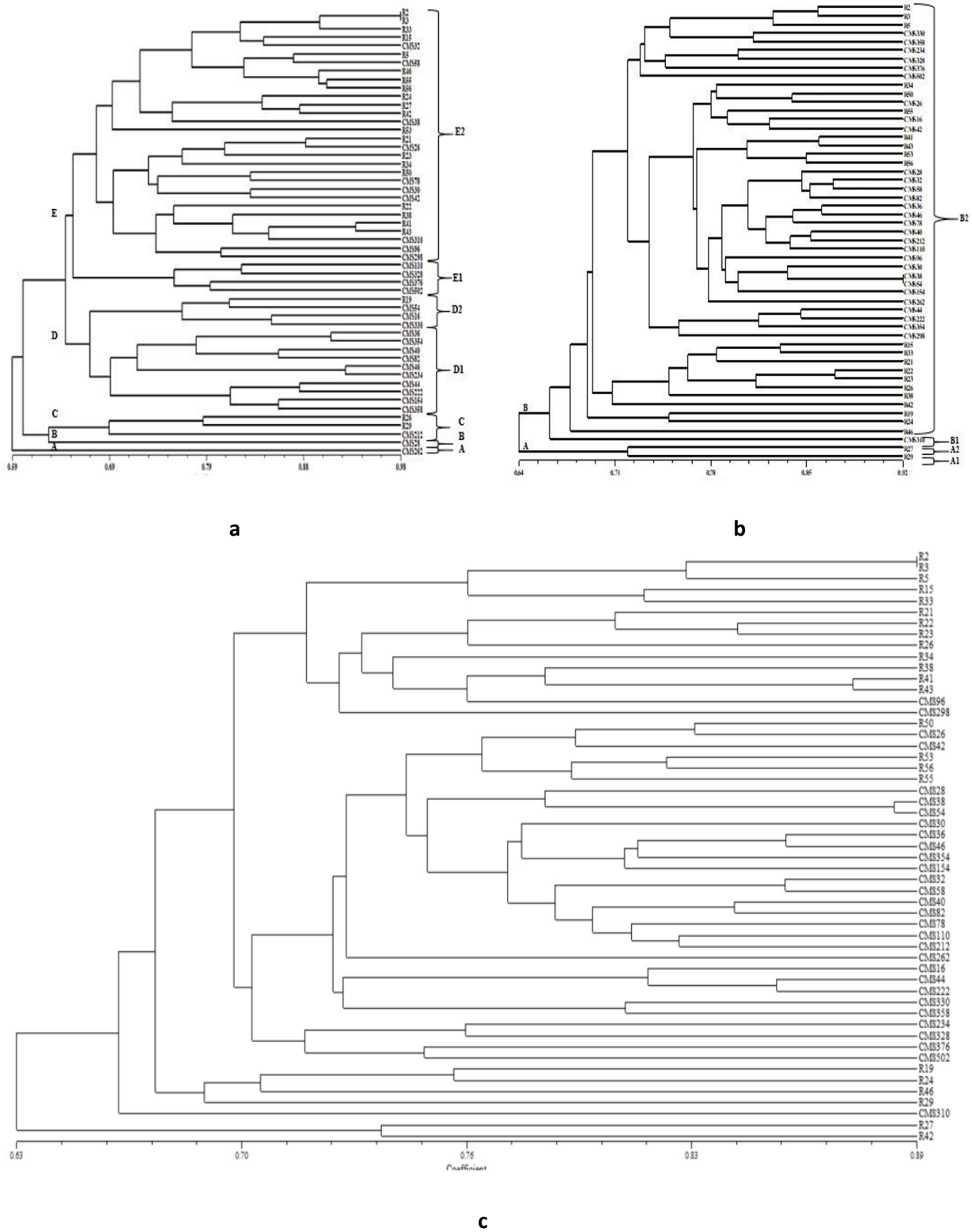
Results

Cluster analysis using SSR and TRAP markers

To group the test samples, three cluster analyses using the unweighted pair group method with arithmetic mean (UPGMA) algorithm and the Dice's genetic distance coefficient were computed. In SSR analysis, 53 sunflower lines were grouped in the following five main groups: group A consisted of CMS262; group B included CMS28; group C included two restorer lines of R29 and R26 and one CMS line of CMS 212; group D included 14 lines; group D with two subgroups consisted of 10 CMS lines (CMS358, CMS15, CMS222, CMS44, CMS234, CMS46, CMS82, CMS40, CMS354, and CMS36 as the first subgroup and three lines CMS330, CMS16, and CMS54 and also one restorer line R19 as the second subgroup); and group E included lines CMS502, CMS376, CMS 328, and CMS 110 as the first subgroup and CMS298, CMS96, CMS 310, R43, R41, R38, R22, CMS 42, CMS30, CMS78, R50, R34, R23, CMS26, R21, R53, CMS38, R42, R27, R24, R56, R55, R46, CMS58, R5, CMS32, R15, R33, R3, and R2 as the second subgroup (Fig. I. a). The results of cluster analysis using TRAP data indicated that the UPGMA dendrogram grouped all lines into two major groups. Group A included two sunflower lines (R29 and R27). Other restorer and CMS lines were located in group B (Fig. I. b), and also, SSR+TRAP data was used to perform a combined cluster analysis to compare with previous dendrograms (Fig. I. c).

Population structure using SSR and TRAP markers

The results of relationships among the investigated inbred lines and population structure showed that the optimum subpopulations (K) were 5 and 2 for SSR and TRAP markers, respectively. These results suggest that the analyzed sunflower germplasm can be divided into five and two genetically distinct groups as identified using SSR and TRAP markers, respectively (Figs. II. a and b). The results showed



a b c

Fig. I. UPGMA tree of the 53 sunflower lines using 23 SSR loci (a) and 19 TRAP loci (b) and combined SSR+TRAP (c)

a few mixed individuals using the TRAP marker. This is in contrast to the SSR marker, which showed more mixed individuals. Model-based clustering by SSR showed restorer and CMS lines were separate in all clusters. The eight restorer lines (≠ 2, ≠ 22, ≠ 3, ≠ 19, ≠ 1, ≠ 9, ≠ 13, and ≠23) and four CMS lines (≠29, ≠38, ≠50, and ≠31) have more membership probability <0.8. In addition,

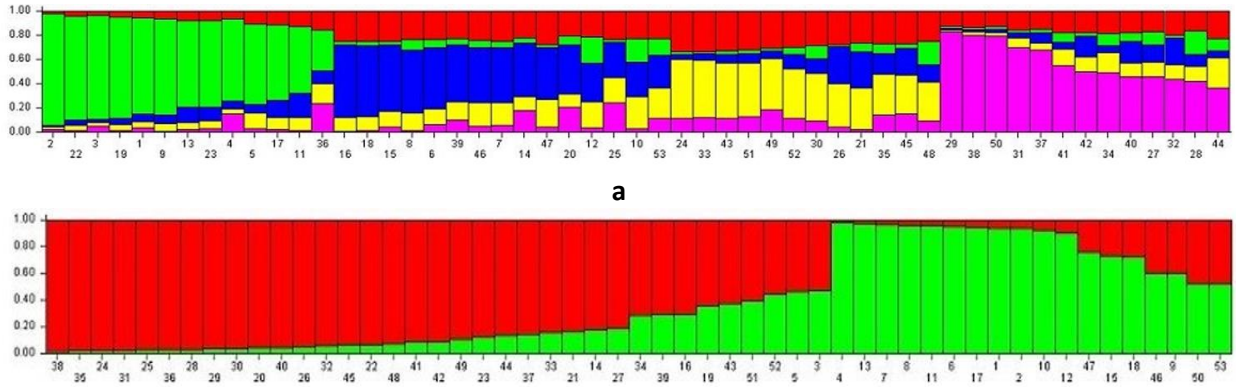


Fig. II. Genetic structure of 53 sunflower lines as inferred by structure software with 23 SSR (a) and 19 (b) TRAP marker datasets, respectively; the single vertical line represents an individual accession and different colors represent genetic stocks/gene pools.

Table 3
Fst values and heterozygosity of five and two sunflower subpopulations using SSR and TRAP markers, respectively

Marker type	Subpopulation (K)	Fst value	Expected heterozygosity
SSR	1	0.2774	0.3446
	2	0.4153	0.2358
	3	0.2619	0.3151
	4	0.2366	0.3398
	5	0.4349	0.2599
	Average	0.3252	0.2990
TRAP	1	0.2395	0.2915
	2	0.2326	0.2944
	Average	0.2360	0.2929

Segments of each vertical line show extent of admixture in an individual

model-based clustering using the TRAP marker revealed that cluster A (49.05% probability) contained 26 lines (≠ 38, ≠ 35, ≠ 24, ≠ 31, ≠ 25, ≠ 36, ≠ 28, ≠ 29, ≠ 30, ≠ 20, ≠ 40, ≠ 26, ≠ 32, ≠ 45, ≠ 22, ≠ 48, ≠ 41, ≠ 42, ≠ 49, ≠ 23, ≠ 44, ≠ 37, ≠ 33, ≠ 21, ≠ 14, and ≠2 7), and cluster B (26.41% probability) included 14 lines (≠ 4, ≠ 13, ≠ 7, ≠ 8, ≠ 11, ≠ 6, ≠ 17, ≠ 1, ≠ 2, ≠ 10, ≠ 12, ≠ 47, ≠ 15, and ≠ 18). Furthermore, ≠ 34, ≠ 39, ≠ 16, ≠ 19, ≠ 43 ≠, 51, ≠ 52, ≠ 5, ≠ 3, ≠ 46, ≠ 9, ≠ 50, and ≠ 53 were observed in both cluster A and cluster B (24.52%) (membership probability <0.8) (Fig. II. b).

The F_{ST} value, which examines genetic divergence of population substructure, ranged from 0.2366 (in the fourth subpopulation) to 0.4349 (in the fifth subpopulation) with a mean of 0.3252 (Table 3). The relatively high estimation of this value in the SSR marker indicates a good separation of all the subpopulations and their diversity. However, the F_{ST} value for the TRAP marker ranged from 0.2326 (second subpopulation) to 0.2395 (first subpopulation) with a mean of 0.2360 (Table 3), which was similar to the result for the SSR marker. Moreover, the expected heterozygosity (an estimate of two randomly chosen individuals being heterozygous at a given locus) for the SSR marker ranged from 0.2358 (second subpopulation) to 0.3446 (first subpopulation) with a mean of 0.2990. The expected heterozygosity for the TRAP marker ranged from 0.2915 (first subpopulation) to 0.2944 (second subpopulation) with an average of 0.2929 (Table 3). Genetic variation based on F_{ST} values among sunflower subpopulations detected by population structure analysis employing SSR and TRAP markers is presented in Table 4. Result showed that the maximum F_{ST} among subpopulations was 0.1162 in Pop B and Pop E. The minimum value (0.0035) was also observed in Pop A and Pop D using the SSR marker. In this situation, the value using the TRAP marker between Pop A and Pop B was 0.0769 (Table 4).

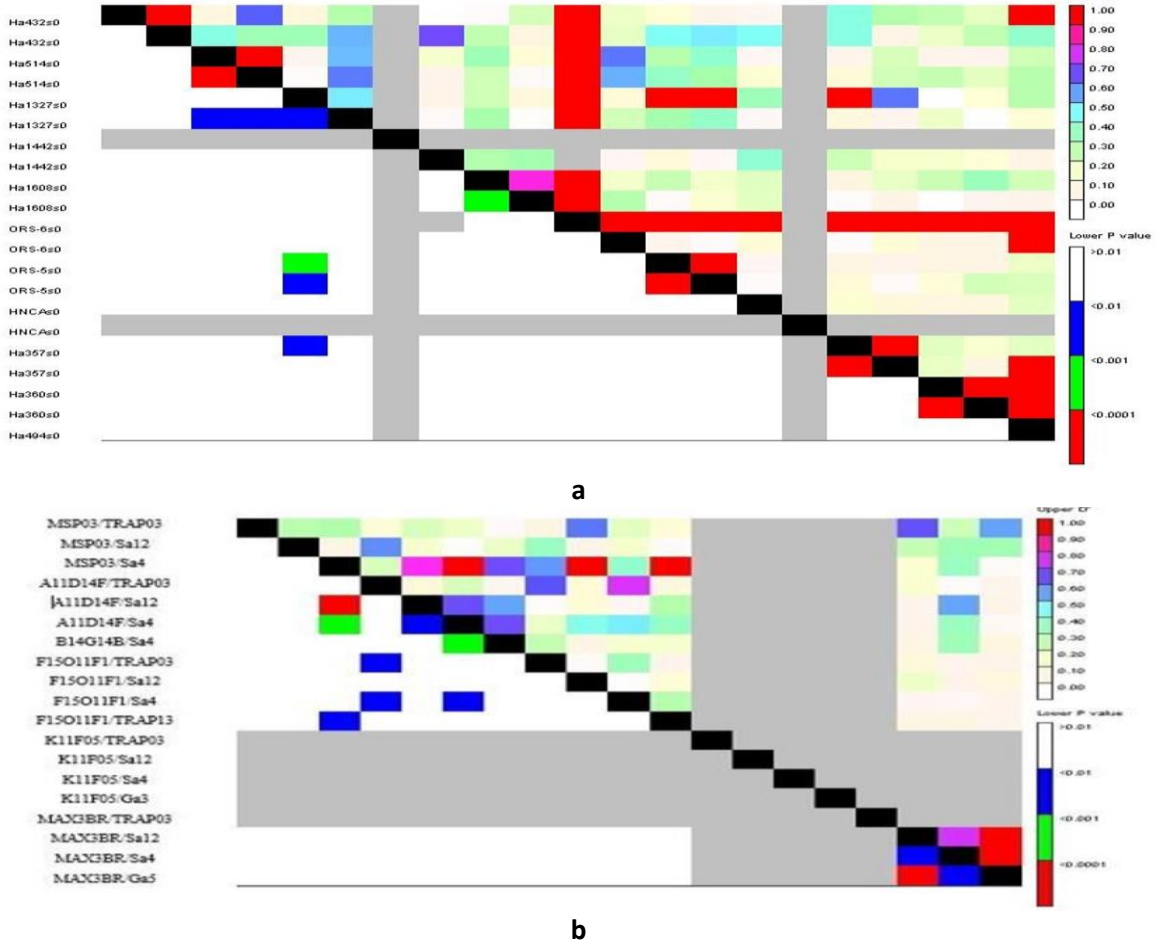


Fig. III. r^2 values (Linkage Disequilibrium [LD]) throughout the investigated sunflower genome for SSR (a) and TRAP (b) markers; marker numbers correspond to those in Tables 2 and 3. Heat map cells indicate a single marker pair. The r^2 values for each marker pair are on the top half of the heat map and represent from 0.0 (white) in increasing equal increments of 0.1 to 1.0 (red). The *P*-values of each r^2 estimate are in the bottom half of the heat map and represent from non-significant (P>0.05; white) to highly significant (P<0.0001; red).

Table 4.
Genetic divergence based on F_{ST} values among five and two sunflower subpopulations identified by population structure analysis using SSR and TRAP markers

Marker type		Pop A	Pop B	Pop C	Pop D	Pop E
SSR	Pop A	-	0.0567	0.0153	0.0035	0.0427
	Pop B	0.0567	-	0.0618	0.0605	0.1162
	Pop C	0.0153	0.0618	-	0.0297	0.0734
	Pop D	0.0035	0.0605	0.0297	-	0.0560
	Pop E	0.0427	0.1162	0.0734	0.0560	-
TRAP	Pop A	-	0.0769	-	-	-
	Pop B	0.0769	-	-	-	-

LD decay analysis using SSR and TRAP markers

r^2 was calculated to estimate the LD extent in the studied sunflower genotypes with both SSR and TRAP markers (Fig. III). In SSR analysis, r^2 ranged from 0.0 (on linkage group 9) to 0.84 (on linkage group 12) with a mean of 0.046 (Table 5). With the TRAP marker, r^2 ranged from 0.67 (on linkage group 16) to 0.0 (on linkage group 1) with a mean of 0.052 (Table 6).

Table 5.
List of SSR markers linked to various traits and their r^2 and associated *P*-value

	2015		2016		2017	
	P-value	r^2	*P*-value	r^2	*P*-value	r^2
FC						
Ha 1287-ar	0.048	0.109	-	-	-	-
ORS-53	0.052	0.081	-	-	-	-
ORS-204	-	-	0.036	0.055		
FD						
Ha 991-ar	0.046	0.061	-	-	-	-
ORS-204	-	-	0.079	0.043	-	-
SD						
Ha 1442	-	-	0.052	0.132	-	-
LN						
Ha 806-ar	-	-	0.055	0.127	-	-
Ha 991-ar	-	-	0.024	0.085	-	-
Ha 1167-ar	-	-	0.053	0.097	-	-
HP						
Ha 494-ar	0.047	0.178	-	-	-	-
Ha 991-ar	0.010	0.150	-	-	-	-
Ha 1167-ar	0.029	0.158	-	-	-	-
ORS-53	0.033	0.152	-	-	-	-
ORS-54	0.049	0.176	-	-	-	-
SW (gr)						
Ha 1442	0.045	0.077	-	-	-	-
Ha 494-ar	0.052	0.096	-	-	-	-
Ha 806-ar	0.068	0.088	-	-	-	-
Ha 991-ar	0.032	0.056	-	-	0.024	0.066
Ha 1167-ar	0.034	0.084	-	-	0.056	0.075
Ha 1287-ar	0.083	0.082	-	-	-	-
ORS-53	0.047	0.075	0.39	0.68	0.048	0.079
SN						
Ha 494-ar	-	-	0.018	0.169	0.028	0.157
Ha 806-ar	-	-	0.017	0.171	0.028	0.156
Ha 991-ar	0.010	0.090	0.002	0.164	0.003	0.151
Ha 1167-ar	0.033	0.091	0.007	0.169	0.011	0.156
Ha 1287-ar	-	-	0.013	0.184	0.019	0.174
ORS-53	0.030	0.094	0.005	0.182	0.007	0.172
ORS-54	-	-	0.014	0.181	0.022	0.166
SY1 (gr)						
Ha 494-ar	0.038	0.112	0.047	0.118	0.037	0.126
Ha 806-ar	0.053	0.101	0.045	0.120	0.036	0.128
Ha 991-ar	0.009	0.092	0.006	0.114	0.004	0.122
Ha 1167-ar	0.027	0.095	0.021	0.116	0.016	0.125
Ha 1287-ar	0.058	0.098	0.034	0.129	0.030	0.134
ORS-53	0.027	0.096	0.015	0.127	0.012	0.133
ORS-54	0.059	0.097	0.040	0.124	0.035	0.128
OY1 (gr)						
Ha 494-ar	0.038	0.112	0.061	0.119	0.045	0.125
Ha 806-ar	0.053	0.101	0.055	0.123	0.045	0.125
Ha 991-ar	0.009	0.092	0.009	0.112	0.007	0.116
Ha 1167-ar	0.027	0.095	0.030	0.114	0.022	0.119
Ha 1287-ar	0.058	0.098	0.043	0.133	0.034	0.136
ORS-53	0.027	0.096	0.020	0.128	0.016	0.131
ORS-54	0.059	0.097	0.056	0.122	0.046	0.124
SY (kg/ha)						
Ha 1442	0.019	0.088	0.047	0.118	0.037	0.126
Ha 494-ar	0.004	0.157	0.045	0.120	0.036	0.128
Ha 806-ar	0.005	0.152	0.006	0.114	0.004	0.122

Table 5. (Continued)
List of SSR markers linked to various traits and their r^2 and associated *P*-value

	2015		2016		2017	
	P-value	r^2	*P*-value	r^2	*P*-value	r^2
Ha 991-ar	0.001	0.139	0.021	0.116	0.016	0.125
Ha 1167-ar	0.002	0.151	0.034	0.129	0.030	0.134
Ha 1287-ar	0.005	0.151	0.015	0.127	0.012	0.133
ORS-53	0.002	0.148	0.040	0.124	0.035	0.128
ORS-54	0.007	0.139	0.047	0.118	0.037	0.126
OY (kg/ha)						
Ha 494-ar	0.038	0.112	0.061	0.119	0.045	0.125
Ha 806-ar	0.053	0.101	0.055	0.123	0.045	0.125
Ha 991-ar	0.009	0.092	0.009	0.112	0.007	0.116
Ha 1167-ar	0.027	0.095	0.030	0.114	0.022	0.119
Ha 1287-ar	0.058	0.098	0.043	0.133	0.034	0.136
ORS-53	0.027	0.096	0.020	0.128	0.016	0.131
ORS-54	0.059	0.097	0.056	0.122	0.046	0.124

Marker-trait association analysis using SSR and TRAP markers

According to the SSR marker results, associated markers were found for 11 out of the 18 studied traits. Only one associated marker was identified for each trait of SD. More than one marker identified other linked traits. No linked SSR markers were detected for FI, MA, SFP, PH, HD, HW, and OC. Accordingly, two, three, and five associated markers were identified for FD, FC, and HP, respectively. Only eight markers were associated with SY while seven of the same markers were associated with SW, SN, SY1, OY1, and OY in three years. The greatest number of associated markers (eight markers) was found for seed yield. Markers *Ha 1287-ar*, *ORS-53,* and *ORS-204* associated with FC explained 10.9%, 8.1%, and 5.5% of the variation in this trait in 2015, 2015, and 2016, respectively. Markers *Ha 991-ar* and *ORS-204* showed a significant association with FD in 2015 and 2016, explaining 6.1% and 4.3% of its variance, respectively. The marker *Ha 1442* was associated with SD, explaining 13.2% of its variation in 2016. Markers *Ha 806-ar*, *Ha 991-ar,* and *Ha 1167-ar* revealed significant associations with LN and explained 12.7%, 8.5%, and 9.7% of its total variation in 2016, respectively. Markers *Ha 494-ar*, *Ha 991-ar, Ha 1167-ar, ORS- 53,* and *ORS-54* were associated with HP and explained 17.8%, 15.0%, 15.8%, 15.2%, and 17.6% of the variation of this trait in 2015, respectively. Markers *Ha 1442*, *Ha 494-ar*, *Ha 806-ar*, *Ha 991-ar*, *Ha 1167-ar*, *Ha 1287-ar,* and *ORS-53* were associated with SW and explained 7.7%, 9.6%, 8.8%, 5.6%, 8.4%, 8.2%, and 7.5% of the variation of this trait in 2015, respectively. However, *Ha 991-ar* and *Ha 1167-ar* explained 6.6% and 7.5% of the variation of this trait in 2016, respectively while *ORS -53* explained 6.8% and 7.9% of the variation in 2016 and 2017, respectively. Markers *Ha 991-ar*, *Ha 1167-ar,* and *ORS-53* were associated with SN and explained 9.0%, 9.1%, and 9.4% of the variation of this trait in 2015, respectively. However, the associated markers *Ha 494-ar*, *Ha 806-ar*, *Ha 991-ar*, *Ha 1167-ar*, *Ha 1287-ar*, *ORS-53,* and *ORS-54* also explained 16.9%, 17.1%, 16.4%, 16.9%, 18.4%, 18.2%, and 18.1% of the variation in 2016, respectively, and 15.7%, 15.6%, 15.1%, 15.6%, 17.4%, 17.2%, and 16.6% of the variation in 2017, respectively. Markers *Ha 494-ar*, *Ha 806-ar*, *Ha 991-ar*, *Ha 1167-ar*, *Ha 1287-ar*, *ORS-53,* and *ORS-54* were associated with SY1, OY1, SY, and OY in 2015, 2016, and 2017. *ORS-53* was associated with most of the traits examined, namely FC, FD, LN, HP, SW, SN, SY1, OY1, SY, and OY. Consequently, the *ORS-53* markers associated with traits SW, SN, SY1, OY1, SY, and OY were the same for 3 years (Table 5).

Seventeen significant associations were found according to the TRAP markers. Only two associated markers were identified for each trait of MA and OY1. The other linked traits recognized in more than one marker were the same as SSR markers. No linked TRAP markers were detected for SFP. As a result, two (HP and OC), three (FI, FC, SW, and LN), four (FD, SY1, and SY), five (HD, SD,

Table 6
List of TRAP markers linked to various traits and their r^2 and associated *P*-value

	2015		2016		2017	
	P-value	r^2	*P*-value	r^2	*P*-value	r^2
FI						
MSP03/Sa12	0.0992	0.05534	-	-	-	-
A11D14F/Sa12	0.06559	0.06949	-	-	-	-
A11D14F/Sa4	0.00411	0.17741	0.01932	0.11458	-	-
FC						
A11D14F/Sa4	0.01478	0.12503	0.05399	0.07636	-	-
F15O11F1/Sa4	0.02451	0.10546	-	-	-	-
K11F05/TRAP03	0.03669	0.13876	-	-	-	-
FD						
F15O11F1/Sa4	-	-	-	-	0.0427	0.08481
K11F05/TRAP03	0.03509	0.14077	0.03164	0.14543	0.04204	0.1526
K11F05/Sa12	-	-	-	-	0.0477	0.12705
MA						
K11F05/TRAP03	0.02165	0.16271	0.02162	0.16275	0.02627	0.15386
PH						
F15O11F1/Sa12	0.05169	0.07164	-	-	-	-
K11F05/Sa12	0.00988	0.19923	-	-	-	-
K11F05/Sa4	0.00962	0.20054	-	-	-	-
K11F05/Ga3	0.01043	0.19667	-	-	-	-
MAX3BR/TRAP03	0.00825	0.20784	-	-	-	-
MAX3BR/Ga5	-	-	0.05632	0.07486	-	-
HD						
MSP03/Sa12	-	-	0.02862	0.0996	-	-
F15O11F1/Sa4	-	-	0.01995	0.11334	-	-
K11F05/Ga3	-	-	0.05198	0.12325	-	-
SD						
MSP03/TRAP03	-	-	0.03375	0.09345	-	-
MSP03/Sa12	-	-	0.03205	0.09537	-	-
MAX3BR/Ga5	-	-	0.01407	0.12696	-	-
LN						
F15O11F1/TRAP13	0.05363	0.07654	-	-	-	-
MAX3BR/Sa12			0.01622	0.12137	-	-
HW						
F15O11F1/Sa12	-	-	0.05353	0.07665	-	-
F15O11F1/Sa4	-	-	-	-	0.0569	0.07447
K11F05/Sa4	0.01595	0.12204	-	-	0.01992	0.16648
SW						
F15O11F1/TRAP03	0.00165	0.2171	0.00337	0.18588	0.0068	0.15633
MAX3BR/Ga5	0.03133	0.09621	-	-	0.03876	0.08833
SN						
K11F05/TRAP03	0.05851	0.11109	-	-	-	-
MAX3BR/Sa4	-	-	0.05565	0.06946	-	-
MAX3BR/Ga5	-	-	0.04184	0.08555	-	-
HP	-	-	-	-	-	-
F15O11F1/Sa4	0.05233	0.07128	-	-		-
MAX3BR/Sa4	0.04782	0.08071	-	-		
OC			-	-		
MSP03/Sa12	-	-	-	-	0.05653	0.06899
SY1						
K11F05/TRAP03	0.04246	0.13223			-	-
MAX3BR/Ga5	0.04313	0.08444	-	-	-	-
SY						
K11F05/TRAP03	0.04346	0.1312	-	-	-	-
MAX3BR/Ga5	0.04299	0.08456	-	-	-	-

LN, and HW) and seven (PH) associated markers were identified. Only three markers were associated with FD (*K11F05/TRAP03*), MA (*K11F05/TRAP03*), and SW (*F15O11F1/TRAP03*) in three years, while seven markers were associated with FI (*A11D14F/Sa12*), FC (*A11D14F/Sa4*), PH (*MSP03/Sa12*), OC (*MAX3BR/Sa12*), SY1 (*MAX3BR/Ga5*), OY1 (*MAX3BR/Ga5*), and OY (*MAX3BR/Ga5*) in 2 years. Markers *MSP03/Sa12*, *A11D14F/Sa12,* and *A11D14F/Sa4* associated with FI elucidated 5.5%, 6.9%, and 17.7% of the variation of this trait in 2015, respectively. In addition, *A11D14F/Sa4* associated with FI explained 11.45% of the variation in 2016. Markers *A11D14F/Sa4*, *F15O11F1/Sa4,* and *K11F05/TRAP03* displayed significant associations with FC in 2015 and 2016, explaining 12.5%, 10.5%, 13.8%, and 7.6% of its variance, respectively. Marker *F15O11F1/Sa4* was associated with FD and explained 8.4% of the variation in 2017. Marker *K11F05/TRAP03* revealed a significant association with FD and explained 14.0%, 14.5%, and 15.2% of its total variation in 2015, 2016, and 2017, respectively. Marker *K11F05/Sa12* was associated with FD and explained 12.7% of the variation of this trait in 2017. Marker *K11F05/TRAP03* was associated with MA and explained 16.2%, 16.2%, and 15.3% of the variation of this trait in 2015, 2016, and 2017, respectively. Markers *F15O11F1/Sa12*, *K11F05/Sa12*, *K11F05/Sa4*, *K11F05/Ga3,* and *MAX3BR/TRAP03* were associated with PH and explained 7.1%, 19.9%, 20.0%, 19.6%, and 20.7% of its total variation in 2015, respectively. *MAX3BR/Ga5* was also associated with PH, explaining 7.4% of the variation of this trait in 2016. Markers *MSP03/Sa12*, *F15O11F1/Sa4,* and *K11F05/Ga3* were associated with HD and explained 9.9%, 11.3%, and 12.3% of its total variation in 2016, respectively. Markers *MSP03/TRAP03*, *MSP03/Sa12,* and *MAX3BR/Ga5* were associated with SD and explained 9.3%, 9.5%, and 12.6% of its total variation in 2016, respectively. The marker *F15O11F1/TRAP13* was associated with LN and explained 7.6% of its total variation in 2015. Moreover, *F15O11F1/TRAP13* was also associated with LN in 2016, explaining 12.1% of its total variation. *F15O11F1/Sa12* had a significant association with HW, explaining 7.6% of its total variation in 2016. Marker *F15O11F1/Sa4* was associated with HW, explaining 7.4% of the variation of this trait in 2017. Marker *K11F05/Sa4* was associated with HW and explained 12.2% and 16.6% of the variation of this trait in 2015 and 2017, respectively. The *F15O11F1/TRAP03* marker was associated with SW for all 3 years, explaining 21.7%, 18.5%, and 15.6% of the variation of this trait, respectively. Marker *K11F05/TRAP03* showed a significant association with SN and explained 11.1% of its total variation in 2015; also associated with SN were *MAX3BR/Sa4* and *MAX3BR/Ga5,* explaining 6.9% and 8.4% of the variation of this trait in 2016 and 2017, respectively. Markers *F15O11F1/Sa4* and *MAX3BR/Sa4* were associated with HP, accounting for 7.1% and 8.0% of the variation of this trait in 2015, respectively. Marker *MSP03/Sa12* was associated with OC and explained 6.8% of the variation of this trait in 2017. The *K11F05/TRAP03* and *MAX3BR/Ga5* markers were associated with SY1 and explained 13.2% and 8.4% of the variation of this trait, respectively. The same result was also obtained in SY traits (13.1% and 8.4% in 2015) (Table 6).

Discussion

The existence of genetic diversity in the breeding materials allows breeders to develop new genotypes with favored traits, from the perspective of both breeders and farmers. The invention of modern biotechnological tools and techniques has accelerated the genetic modification process of breeding materials and reduced the duration of breeding cycles. Thus far, various molecular markers have been used in sunflower; SSR and SNP are the most popular while TRAP is less as popular. To date, more than 2000 SSR markers have been developed from EST (EST-SSR) and genomic (gSSR) libraries (Dehmer et al., 1998; Yu et al., 2002). The use of the TRAP marker is more recent (Yue et al., 2009). In conventional breeding, the accuracy of cultivar development has been increased with marker-assisted selection (MAS) for relatively large-effect genes. In this study, grouping obtained with UPGMA cluster analysis by SSR and TRAP markers was not concordant with CMS and restorer lines and did not yield a reasonable category. This might be due to the gene flow in CMS and restorer lines. Similar results were previously observed by Yue et

al. (2009), who reported that one maintainer inbred line clustered into the restorer group and seven restorer inbred lines clustered into the maintainer inbred line group. This finding confirms that some inbred lines from different groups have similar genetic backgrounds. Genetic structure analysis divided the studied germplasm into five and two genetically distinct groups in SSR and TRAP markers, respectively. Moreno et al. (2012) reported that the number of genetic clusters to be expected was K=9. Consistent with a nested pattern of population sub-structuring, three local maxima of K=2, K=6, and K=9 were identified in Evanno's analysis (Evanno et al., 2005). Filippi et al. (2015) and Mandel et al. (2013) stated that the genetic structure analysis of sunflower associations as assessed by SSR yielded K=3. In the ΔK method, the number of subdivisions in their dataset showed a greater probability to be K=2. In population genetics, the F_{ST} value is an estimation of population substructure, which depends on the allele frequencies at a locus. F_{ST} is regularly measured from genetic polymorphism data such as SNPs, SSRs (Jakobsson et al., 2013), or TRAP marker systems. Mandel et al. (2013) showed an F_{ST} value of 0.047 with the SSR marker. In most studies, SSR-derived information was used for population structure and genetic diversity analyses. Being adequate, reliable, and highly polymorphic, genomic SSR is an interesting marker for population diversity studies. However, regarding the TRAP marker system, few studies have explored differences in information content and other evaluations of genetic variation. As concerns with sunflower, comparisons between SSR and TRAP markers have confirmed that TRAP markers can show higher levels of diversity than SSR (Hu et al., 2003; Zeinalzadeh-Tabrizi et al., 2015). In another investigation, one particularly intriguing example was the region surrounding markers *ORS331* and *ORS143* on LG7, which harbors five selection candidates and contains QTL for flowering time and the number of main stem leaves produced. Additionally, four selection candidates mapped to the interval between markers ORS878 and ORS613 on LG10, a region that contains QTL for seed size in three different mapping populations as well as numerous other traits (Wills et al., 2007). According to the MLM, two, five, and 11 loci were significantly (P<0.01) associated with seed-related traits under natural, mild, and severe drought stress conditions, respectively (Jannatdoust et al., 2017). Until now, no reports have been published on sunflower genotype structure analysis by the TRAP marker. In this study, SSR and TRAP markers provided an opportunity to identify marker-trait associations.

The K+Q model, which was corrected for both kinship and population structure, appeared to be the most conservative model across traits and showed the lowest tendency for false positives. This model was thus selected for all subsequent association analyses. AM needs a complete understanding of the population structure and LD of the genetic material evaluated (Fusari et al., 2008). This analysis cannot be done in the absence of LD, which shows the nonrandom relation of alleles at different genetic loci on a single chromosome (Mackay et al., 2007). Moreover, the results of an association analysis are determined by LD structure across a genome (Remington et al., 2001). Our results showed 18.57% and 8.18% of the possible SSR and TRAP locus pairs had a significant LD level (P<0.05)**.** Similarly, Soleimani-Gezeljeh et al. (2018) reported that in the collection under investigation, a significant LD level (P<0.05) was observed at 8.97% of the possible SSR locus pairs. In our study, the r^2 yielded an average of 0.046 and 0.052 with SSR and TRAP markers, respectively. The result revealed a greater LD extent for TRAP than for SSR. In addition, the result showed a low level of LD in the entire collection in this study, although it is not reliable because of the low number of SSR and TRAP markers used to estimate r^2. Fusari et al. (2008) and Talukder et al. (2014) found that due to the low level of LD in germplasm resources of the sunflower, successful AM requires many markers available for analysis. Factors such as type of traits investigated, type and size of population, number of environments, and type and genome coverage of molecular markers affect the results of AM (Ruggieri et al., 2014). To identify associated markers with a low level of interactions with the environment, phenotyping was performed in 3 years. However, more phenotyping data acquired over several years and environments are needed for the identification of reliable associated markers for breeding programs. Since previous

investigations have confirmed the efficiency of the K+Q model in correcting spurious associations in sunflower populations (Fusari et al., 2008), we employed the MLM method to detect highly significant associations. According to the SSR marker results, associated markers were found for 11 out of the 18 studied traits. Markers *Ha 494-ar*, *Ha 806-ar*, *Ha 991-ar*, *Ha 1167-ar*, *Ha 1287-ar*, *ORS-53,* and *ORS-54* were associated with SY1, OY1, SY, and OY in 2015, 2016, and 2017. The marker *ORS-53* was associated with most of the traits examined, such as SW, SN, SY1, OY1, SY, and OY, over the three years. According to the TRAP marker results, associated markers were found for 17 out of the 18 studied traits. Only three markers were associated with FD (*K11F05/TRAP03*), MA (*K11F05/TRAP03*), and SW (*F15O11F1/TRAP03*) in the 3 years. Marker *K11F05/TRAP03* revealed a significant association with FD, explaining 14.0%, 14.5%, and 15.2% of its total variation in 2015, 2016, and 2017, respectively. Marker *K11F05/TRAP03* was associated with MA and explained 16.2%, 16.2%, and 15.3% of the variation of this trait in 2015, 2016, and 2017, respectively. The *F15O11F1/TRAP03* marker was associated with SW for the three years and explained 21.7%, 18.5%, and 15.6% of the variation of this trait, respectively. Genetic linkage or the pleiotropic effects of the same genes already revealed by QTLs may explain such localized associations (Lecomte et al., 2004). The markers highly associated with the targeted traits may be promising in marker-assisted selection in sunflower breeding programs and warrant further studies. Validation of the results in mapping populations or different germplasms should also be considered.

Conclusion

DNA molecular markers have been extensively and successfully used for mapping QTLs. Population structure analysis clearly differentiated between SSR and TRAP markers. The AM analysis identified 21 SSR and 19 TRAP markers associated with 11 and 17 traits, respectively. The use of markers highly associated with a given trait for three years could be a valuable starting point for marker-assisted selection. This strategy relies on detecting LD between genetic markers and genes controlling the targeted phenotype by exploiting the historical recombination events and thus offers increased mapping resolution and should facilitate marker-assisted selection in plant breeding. In recent years, the AM approach has been extensively used for complex quantitative trait analysis and to determine candidate genes affecting such traits (Hall et al., 2010). The findings suggest that use of SSR and TRAP markers and K+Q, a highly reliable statistical model, are appropriate for the identification of associations between the traits investigated. Thus, it may well be argued that SSR and TRAP markers might be utilized for targeting the specific desirable phenotype and marker-assisted selection in sunflower breeding programs. Finally, the adaptation of genotypes to diverse environments is controlled via varied QTLs that can be usefully pyramided through MAS. Furthermore, overlapping QTLs can be assessed to genetically determine trait associations, which are helpful for indirectly selecting yields through yield-related traits. Nonetheless, before MAS application as a practical strategy, the identified QTLs must be corroborated in the same way as any other quantitative traits.

Acknowledgements

The authors thank the Department of Oilseed Crops, Seed and Plant Improvement Institute, Agricultural Research, Education and Extension Organization (AREEO), Iran for providing the facilities and the field experiments to conduct this research.

References

Akbari G., H. Jabbari, J. Daneshian, I. Alahdadi And N. Shahbazian, 2008. 'The effect of limited irrigation on seed physical characteristics in sunflower hybrids'. *Journal of Crop Production and Processing*. 12(45):513-523 (in Persian).

Al-Chaarani G.R., L. Gentzbittel, X. Huang And A. Sarrafi. 2004. 'Genotypic variation and

identification of qtls for agronomic traits, using AFLP and SSR markers in RILs of sunflower (*Helianthus annuus* L.) '. *Theoretical and Applied Genetics*. 109(7):1353-1360.

Arias M., M. Hernandez, N. Remondegui, K. Huvenaars, P. Van Dijk and E. Ritter. 2016. 'First genetic linkage map of *Taraxacum koksaghyz rodin* based on AFLP, SSR, COS and EST-SSR markers'. *Scientific Reports*. 6:31031.

Bartholomé J., M.C. Bink, J. Van Heerwaarden, E. Chancerel, C. Boury, I. Lesur, F. Isik, L. Bouffier and C. Plomion. 2016. 'Linkage and association mapping for two major traits used in the maritime pine breeding program: Height growth and stem straightness'. *PloS one*. 11(11): e0165323.

Berrios E.F., L. Gentzbittel, H. Kayyal, G. Alibert and A. Sarrafi. 2000. 'AFLP mapping of QTLs for in vitro organogenesis traits using recombinant inbred lines in sunflower (*Helianthus annuus* L.) '. *Theoretical and Applied Genetics*. 101(8):1299-1306.

Bradbury P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss and E.S. Buckler. 2007. 'TASSEL: Software for association mapping of complex traits in diverse samples'. *Bioinformatics*. 23(19):2633-2635.

Cadic E., M. Coque, F. Vear, B. Grezes-Besset, J. Pauquet, J. Piquemal, Y. Lippi, P. Blanchard, M. Romestant and N. Pouilly. 2013. 'Combined linkage and association mapping of flowering time in sunflower (*Helianthus annuus* L.) '. *Theoretical and Applied Genetics. 126*(5):1337-1356.

Darvishzadeh R. 2012. 'Association of SSR markers with partial resistance to *Sclerotinia sclerotiorum* isolates in sunflower (*Helianthus annuus* L.) '. *Australian Journal of Crop Science*. *6*(2):276.

Davar R., A. Majd, R. Darvishzadeh and A. Sarrafi. 2011. 'Mapping quantitative trait loci for seedling vigour and development in sunflower (*Helianthus annuus* L.) using recombinant inbred line population'. *Plant Omics*. *4*:418-427.

Dehmer K. and W. Friedt. 1998. 'Evaluation of different microsatellite motifs for analysing genetic relationships in cultivated sunflower (*Helianthus annuus* L.) '. *Plant Breeding*. *117*(1):45-48.

Dice L.R. 1945. 'Measures of the amount of ecologic association between species'. *Ecology*. *26*(3):297-302.

Ebrahimi A., P. Maury, M. Berger, S. Poormohammad Kiani, A. Nabipour, F. Shariati, P. Grieu and A. Sarrafi. 2008. 'QTL mapping of seed-quality traits in sunflower recombinant inbred lines under different water regimes'. *Genome*. *51*(8):599-615.

Evanno G., S. Regnaut and J. Goudet. 2005. 'Detecting the number of clusters of individuals using the software structure: A simulation study'. *Molecular Ecology*. *14*(8):2611-2620.

Eyvaznejad N. And R. Darvishzadeh. 2014. 'Identification of QTLs for grain yield and some agro-morphological traits in sunflower (*Helianthus annuus* L.) using SSR and SNP markers'. *Journal of Plant Molecular Breeding*. *2*(2):68-87.

Fernández-Martínez J., B. Pérez-Vich and L. Velasco. 2009. Sunflower. In 'oil crops'. (eds j vollmann, i rajcan) pp. 155–232. *Springer*: New York.

Filippi C.V., N. Aguirre, J.G. Rivas, J. Zubrzycki, A. Puebla, D. Cordes, M.V. Moreno, C.M. Fusari, D. Alvarez and R.A. Heinz. 2015. 'Population structure and genetic diversity characterization of a sunflower association mapping population using SSR and SNP markers'. *BMC Plant Biology*. *15*(1):52.

Fusari C.M., J.A. Di Rienzo, C. Troglia, V. Nishinakamasu, M.V. Moreno, C. Maringolo, F. Quiroz, D. Álvarez, A. Escande and E. Hopp. 2012. 'Association mapping in sunflower for sclerotinia head rot resistance'. *BMC Plant Biology*. *12*(1):93.

Fusari C.M., V.V. Lia, H.E. Hopp, R.A. Heinz and N.B. Paniego. 2008. 'Identification of single nucleotide polymorphisms and analysis of linkage disequilibrium in sunflower elite inbred lines using the candidate gene approach'. *BMC Plant Biology*. *8*(1):7.

Haddadi P., B. Yazdi-Samadi, M. Naghavi, A. Kalantari, P. Maury and A. Sarrafi. 2011. 'QTL analysis of agronomic traits in recombinant inbred lines of sunflower under partial irrigation'. *Plant Biotechnology Reports*. *5*(2):135-146.

Hall D., C. Tegström and P.K. Ingvarsson. 2010. 'Using association mapping to dissect the genetic basis of complex traits in plants'. *Briefings in Functional Genomics*. *9*(2):157-165.

Holland J.B. 2007. 'Genetic architecture of complex traits in plants'. *Current Opinion in Plant Biology*. *10*(2):156-161.

Hu J. And B.A. Vick. 2003. 'Target region amplification polymorphism: A novel marker technique for plant genotyping'. *Plant Molecular Biology Reporter*. *21*(3):289-294.

Ingvarsson P.K. and N.R. Street. 2011. 'Association genetics of complex traits in plants'. *New Phytologist*. *189*(4):909-922.

Jakobsson M., M.D. Edge and N.A. Rosenberg. 2013. 'The relationship between fst and the frequency of the most frequent allele'. *Genetics*. *193*(2):515-528.

Jannatdoust M., R. Darvishzadeh, H. Azizi, M. Ebrahimi, R. Ziaefard and E. Gholinezhad. 2017. 'Identification of retrotransposon markers associated with agromorphological traits in confectionary sunflower (*Helianthus annuus* L.) under normal and drought stress conditions'. *Journal of Crop Breeding*. *8*:183-197. (in Persian)

Kolkman J.M., S.T. Berry, A.J. Leon, M.B. Slabaugh, S. Tang, W. Gao, D.K. Shintani, J.M. Burke and S.J. Knapp. 2007. 'Single nucleotide polymorphisms and linkage disequilibrium in sunflower'. *Genetics*. *177*(1):457-468.

Lecomte L., V. Saliba-Colombani, A. Gautier, M. Gomez-Jimenez, P. Duffe, M. Buret and M. Causse. 2004. 'Fine mapping of QTLs of chromosome 2 affecting the fruit architecture and composition of tomato'. *Molecular Breeding*. *13*(1):1-14.

Mackay I. And W. Powell. 2007. 'Methods for linkage disequilibrium mapping in crops'. *Trends in Plant Science*. *12*(2):57-63.

Mandel J.R., S. Nambeesan, J.E. Bowers, L.F. Marek, D. Ebert, L.H. Rieseberg, S.J. Knapp and J.M. Burke. 2013. 'Association mapping and the genomic consequences of selection in sunflower'. *PLoS Genetics*. *9*(3): e1003378.

Mangin B., N. Pouilly, M.-C. Boniface, N.B. Langlade, P. Vincourt, F. Vear and S. Muños. 2017. 'Molecular diversity of sunflower populations maintained as genetic resources is affected by multiplication processes and breeding for major traits'. *Theoretical and Applied Genetics*. *130*(6):1099-1112.

Mokrani L., L. Gentzbittel, F. Azanza, L. Fitamant, G. Al-Chaarani and A. Sarrafi. 2002. 'Mapping and analysis of quantitative trait loci for grain oil content and agronomic traits using AFLP and SSR in sunflower (*Helianthus annuus* L.) '. *Theoretical and Applied Genetics*. *106*(1):149-156.

Moreno, Mv., Cm. Fusari, V. Nishinakamasu, D. Alvarez, Dij. Rienzo, Eh. Heinz, R. Paniego, Vv Lia. 2012. 'Association mapping of drought tolerance in sunflower: phenotypic and genetic characterization of INTA inbred lines'. In 18th International Sunflower Conference. February 2012 (Vol. 27).

Nambeesan S.U., J.R. Mandel, J.E. Bowers, L.F. Marek, D. Ebert, J. Corbi, L.H. Rieseberg, S.J. Knapp and J.M. Burke. 2015. 'Association mapping in sunflower (*Helianthus annuus* L.) reveals independent control of apical vs. basal branching'. *BMC Plant Biology*. *15*(1):84.

Nordborg M. and S. Tavaré, 2002. 'Linkage disequilibrium: What history has to tell us'. *Trends in Genetics*. *18*(2):83-90.

Paniego N., M. Echaide, M. Muñoz, L. Fernández, S. Torales, P. Faccio, I. Fuxan, M. Carrera, R. Zandomeni and E.Y. Suárez. 2002. 'Microsatellite isolation and characterization in sunflower (*Helianthus annuus* L.) '. *Genome*. *45*(1):34-43.

Price A.L., N.J. Patterson, R.M. Plenge, M.E. Weinblatt, N.A. Shadick and D. Reich. 2006. 'Principal components analysis corrects for stratification in genome-wide association studies'. *Nature Genetics*. *38*(8):904-909.

Pritchard J.K., M. Stephens and P. Donnelly. 2000. 'Inference of population structure using multilocus genotype data'. *Genetics*. *155*(2):945-959.

Remington D.L., J.M. Thornsberry, Y. Matsuoka, L.M. Wilson, S.R. Whitt, J. Doebley, S. Kresovich, M.M. Goodman and E.S. Buckler. 2001. 'Structure of linkage disequilibrium and phenotypic associations in the maize genome'. Proceedings of the National Academy of Sciences. *98*(20):11479-11484.

Renaut S. 2017. 'Genome sequencing: Illuminating the sunflower genome'. *Nature Plants*. *3*(7):17099.

Ruggieri V., G. Francese, A. Sacco, A. D'alessandro, M.M. Rigano, M. Parisi, M. Milone, T. Cardi, G. Mennella and A. Barone. 2014. 'An association mapping approach to identify favourable alleles for tomato fruit quality breeding'. *BMC Plant Biology*. *14*(1):337.

Sahari K., P. Nicolas, M. Stéphane, B. Aurélie, B.J. Fayçal, V. Patrick and B. Dominique. 2016. 'Genetic diversity and core collection constitution for subsequent creation of new sunflower varieties in tunisia'. *Helia*. *39*(64):123-137.

Soleimani-Gezeljeh A., R. Darvishzadeh, A. Ebrahimi and M.R. Bihamta. 2018. 'Identification of SSR and retrotransposon-based molecular markers linked to morphological characters in oily sunflower (*Helianthus annuus* L.) under natural and water-limited states'. *Journal of Genetics*. *97*(1): 189-203.

Talia P., V. Nishinakamasu, H.E. Hopp, R.A. Heinz and N. Paniego. 2010. 'Genetic mapping of EST-SSR, SSR and INDEL to improve saturation of genomic regions in a previously developed sunflower map'. *Electronic Journal of Biotechnology*. *13*(6):7-8.

Talukder Z.I., B.S. Hulke, L. Qi, B.E. Scheffler, V. Pegadaraju, K. Mcphee and T.J. Gulya. 2014. 'Candidate gene association mapping of sclerotinia stalk rot resistance in sunflower (*Helianthus annuus* L.) uncovers the importance of coi1 homologs'. *Theoretical and Applied Genetics*. *127*(1):193-209.

Tang S., J.-K. Yu, M. Slabaugh, D. Shintani and S. Knapp. 2002. 'Simple sequence repeat map of the sunflower genome'. *Theoretical and Applied Genetics*. *105*(8):1124-1136.

Thornsberry J.M., M.M. Goodman, J. Doebley, S. Kresovich, D. Nielsen and E.S. Buckler IV. 2001. 'Dwarf8 polymorphisms associate with variation in flowering time'. *Nature Genetics*. *28*(3):286-289.

Wills D.M. and J.M. Burke. 2007. 'Quantitative trait locus analysis of the early domestication of sunflower'. *Genetics*. *176*(4):2589-2599.

Yu J.-K., J. Mangor, L. Thompson, K.J. Edwards, M.B. Slabaugh and S.J. Knapp. 2002. 'Allelic diversity of simple sequence repeats among elite inbred lines of cultivated sunflower. *Genome'*. *45*(4):652-660.

Yu J. And E.S. Buckler. 2006. 'Genetic association mapping and genome organization of maize'. *Current Opinion in Biotechnology*. *17*(2):155-160.

Yue B., X. Cai, B.A. Vick and J. Hu. 2009. 'Genetic diversity and relationships among 177 public sunflower inbred lines assessed by trap markers'. *Crop Science*. *49*(4):1242-1249.

Yue B., S. Radi, B. Vick, X. Cai, S. Tang, S. Knapp, T. Gulya, J. Miller and J. Hu. 2008. 'Identifying quantitative trait loci for resistance to sclerotinia head rot in two USDA sunflower germplasms'. *Phytopathology*. *98*(8):926-931.

Zeinalzadeh-Tabrizi H., K. Haliloglu and A. Razban-Haghighi. 2015. 'Genetic diversity of sunflower genotypes (*Helianthus annuus* L.) using TRAP markers'. *Crop Biotechnology*. *12*:39-53 (in Persian).

Zeinalzadeh-Tabrizi H., A. Hosseinpour, M. Aydin and K. Haliloglu. 2015. 'A modified genomic DNA extraction method from leaves of sunflower for PCR based analyzes'. *J Biol Env Sci*. *7*:222-225.

Zubrzycki J.E., C.A. Maringolo, C.V. Filippi, F.J. Quiroz, V. Nishinakamasu, A.F. Puebla, J.A. Di Rienzo, A. Escande, V.V. Lia and R.A. Heinz. 2017. 'Main and epistatic QTL analyses for sclerotinia head rot resistance in sunflower'. *PloS one*. *12*(12): e0189859.