NP

GENETICS AND PLANT BREEDING

Microsatellite Based Genotyping of the *Helianthus annuus* L.

Kirandeep Kaur*, S.K. Dhillon and B.S. Gill

Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana, Punjab, India

*Corresponding author: kiran.kaur4554@yahoo.com (ORCID ID: 0000-0002-2453-4686)

ABSTRACT

Genetic diversity among 35 genotypes of sunflower was determined using microsatellites (SSR markers) in the present study. Thirty five genotypes (3 *CMS* A lines and 32 R lines) were subjected to 34 SSR markers. Thirty four primers revealed 72 alleles with average 2.20 alleles per locus. PIC value for the primer pairs varied from 0.12 (ORS 591) to 0.71 (ORS 317) with average PIC value of 0.42. The phylogenetic tree divided genotypes into two main clusters *viz.* cluster 1 and cluster 2, which further are divided into four sub-clusters. Range of genetic distances based on SSR marker analysis varied from 0.32 (6A × P188R) - 0.73 (11A × P174R). The distances measured based on SSR primers showed that the CMS 11A was most divergent genotype with a mean genetic dissimilarity coefficient of 0.59. Immense genetic diversity was observed in the present study, which can be used for further designing hybridization programme.

Highlights

⊙ The present study was aimed to elaborate the applications of SSR markers to characterize and determine genetic relationship between the parental lines by assessing their genetic distances.

Keywords: Sunflower, SSR marker, genetic distance, polymorphism, genetic diversity

Being one of the major annual, sunflower (*Helianthus annuus* L.) is grown worldwide on a large area. It's seed has high quality edible oil that is rich in polyunsaturated fatty acids and fat soluble vitamins (A, B, E and K) which are related to heart proteins (Evertt *et al.* 1987; Gossal *et al.* 1988). Due to the possibility of using its oil as raw material for manufacturing biodiesel (Backes *et al.* 2008), sunflower is lifting the interest of farmers, agriculture professionals and companies. Beg *et al.* (1984) stated that the sunflower hybrid cultivars are more uniform and high yielding than the open pollinated varieties therefore heterosis for sunflower must be used. The increased production and productivity is credited to the release of new high yielding varieties which are the result of heterosis, achieved by heterotic vigor available in the genetically diverse parental lines. The objective of heterosis breeding is to produce hybrids with higher productivity, increased oil content and

resistance to biotic and abiotic stresses. Enhanced vigour or heterosis in resulting hybrids is due to the hybridization of the genetically divergent parents will result in. So information about the genetic distances among the parental lines is prerequisite for successful hybrid breeding. The information about genetic diversity in the available germplasm and among elite breeding material which is essential in plant breeding was reported by Mosges and Friedtu (1994). The presence of genetic divergence in sunflower genotypes and the identification of heterotic groups has been carried out based on SSR primers (Darvishzadeh *et al.* 2010a). The prediction of heterosis is desirable because of the costly process of the production of hybrids. So it is convenient to screen the parental lines and then cross diverse parents for producing heterotic hybrids. Studies on the genetic distance of parental lines as the base of hybrids performance prediction have been undertaken in several crops by different research

workers and different levels of correlation were shown between the genetic distance of parental genotypes and the hybrid performance. Review of literature showed very limited works on the predication of sunflower heterosis and hybrid performance by marker based genetic distance of the parental lines (Tersac *et al.* 1994; Cheres *et al.* 2000). The present study was carried out to assess the genetic diversity among some sunflower parental lines using SSR primers.

MATERIAL AND METHODS

The study involved thirty five genotypes (Table 1) including 3 CMS A-lines (67A, 47A and 11A) and 32 restorer lines developed from the germplasm available at Punjab Agricultural University, Ludhiana. The CMS A-lines were maintained by crossing them with B-lines, while the R lines were maintained by sib mating during kharif-2014. For DNA isolation seeds of these genotypes were sown in trays and were kept in the incubators of oil seed section of the department of plant breeding and genetics, PAU, Ludhiana for proper germination and growth. Leaves of 30 days old seedling of parental lines were taken to isolate total genomic DNA using CetylTrimethyl Ammonium Bromide (CTAB) (Murray and Thomson, 1980). Thirty four polymorphic microsatellite (SSR) markers (Table 2) were then applied to all the thirty five sunflower genotypes so as to find the genetic distances and to analyze genetic diversity. Amplification was carried out in the Central Molecular Biology Laboratory, Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana. Temperature regime for particular SSR primer pair was adjusted on the basis of their neucleotide sequence. Temperature profile followed for the amplification is given in Table 2. The formula provided by Mazers *et al.* (1991) was used to calculate melting temperature (Tm), whereas annealing temperature of the primers ranged 5°C below melting temperature.

$$T_m \ (\degree C) = 4(G+C) + 2(A+T)$$

Where,

T_m = Melting temperature (°C).

After the amplification of genomic DNA, PCR product was separated on 2.5% agarose gel prepared in 0.5X TBE buffer. The ethidium bromide at the concentration of 10ng/100ml of agarose solution in TBE buffer was used for the visualization of the fragments under UV light. To check the size of the SSR fragments, 100-bp ladder was loaded along with the amplified product (Plate 1). The total number of alleles was recorded for each microsatellite (SSR) marker in all the genotypes under the study by giving the number to amplified alleles as 1, 2, 3 and so on. The amplified bands were recorded as 1 (band present) and 0 (band absent) in a binary matrix. The computer software programme – DARwin (Perrier and Jacquemond-Collet 2006) was used for analyzing genetic diversity among the 35 parental lines. Dissimilarly matrix for SSR primers was constructed using Dice coefficient of associations to find out the genetic relationships. The data were subjected to unwieghted pair group method with arithmetic mean (UPGMA) analysis to generate dendogram. Data from 34 primers were used to estimate the dissimilarity based on the number of shared amplified bands.

RESULTS AND DISCUSSION

SSR (simple sequence repeats) are sites on DNA strand which are 2-5 base pair repeats, which are repeated 5-50 times (Turnpenny and Ellard, 2005). SSRs also known as microsatellites, occur at thousands of locations within the genome of an organism. They lead to high genetic diversity as they have a higher mutation rate than the other areas of DNA (Brinkmann *et al.* 1998). These markers are very reproducible as they are influenced by various factors, such as sequence of a primer, template quality and quantity, the type of thermocycler and polymerase concentration (Hernandez *et al.* 1999). The high discriminating power of SSRs led to uncover the possibility of multiple domestication origins in sunflower which were previously thought to be single domestication origin (Rieseberg and Seiler 1990). Tang and Knapp (2003) used SSRs to find extraordinary diversity in American landraces and wild populations of the cultivated sunflower. Sahari *et al.* (2016) characterized 59 accessions and reference lines using 30 SSR loci which generated 194 alleles in the sunflower.

In our study 34 polymorphic SSR markers (Table 2) were used to characterize thirty five genotypes of sunflower and were grouped into different clusters.

Table 1: List of genotypes evaluated in study

R-lines	P172R, P173R, P174R, P175R, P176R, P177R, P178R, P179R, P180R, P184R, P185R, P187R, P188R, P189R, P190R, P192R, P194R, P195R, P196R, P197R, P198R, P200R, P201R, P202R, P203R, P205R, P206R, P207R, P209R, P211R, P212R, P213R
A-lines	67A, 47A, 11A
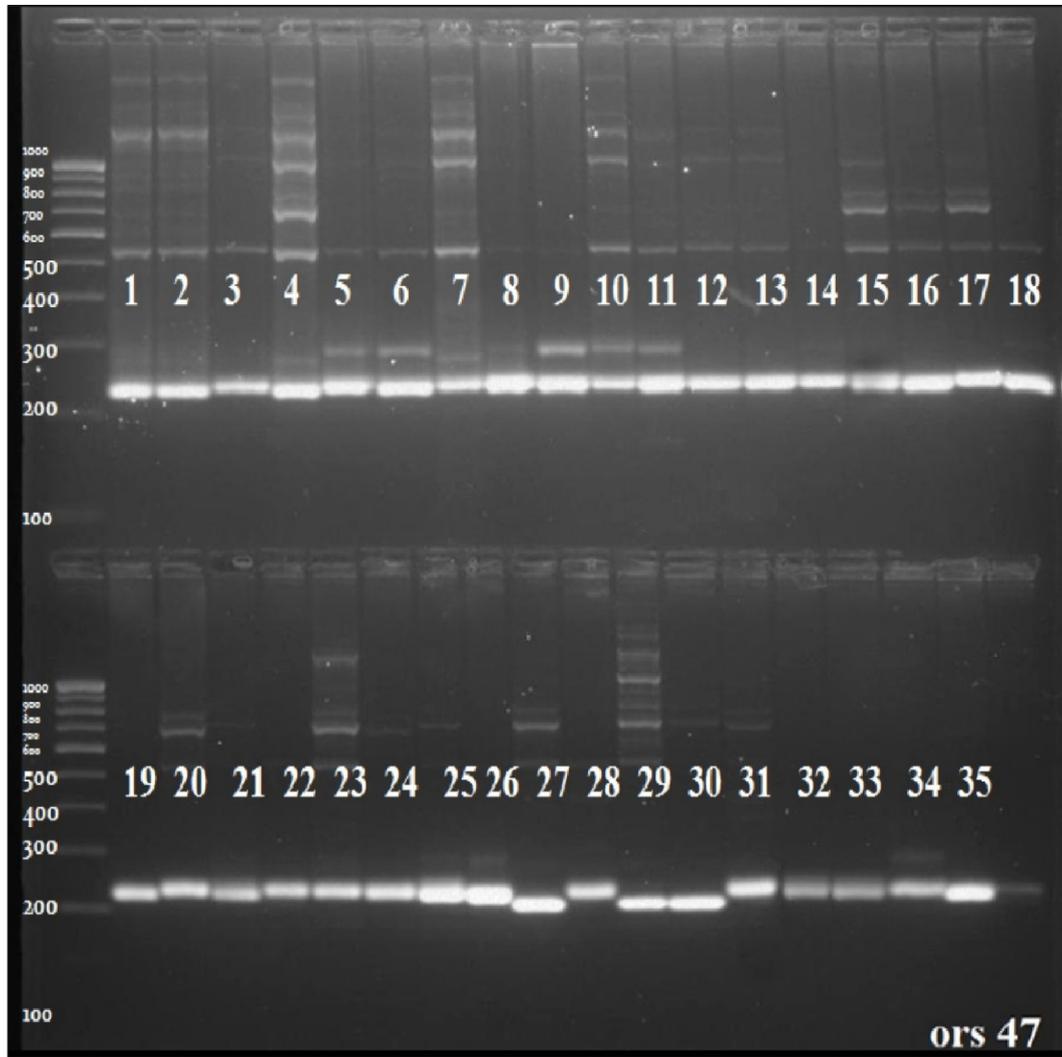
Plate 1: DNA amplification pattern of 35 genotypes of sunflower with ORS47 SSR primer

The number and size of the amplified product was strictly based on the primer pair sequence. Amplification for some primer pairs was repeated two to three times so as to check the consistency of the product, and only unique bright bands were counted for further analysis. The electrophoretic spectrum (Plate 1) of each SSR primer was analyzed. The individual SSR spectra differing by the number of amplicon and electrophoretic mobility were estimated for each line.

Number of alleles varied from 2 to 4 for each locus with total 75 alleles generated from 34 SSR primers.

The average alleles per locus observed was 2.20. These results were closer to those of Darvishzadeh *et al.* (2010), who also observed similar results in their study in which they characterized 28 genotypes on the basis of molecular markers and observed 2.32 number of allele per locus in the sunflower. Average allele per locus was lesser than the observations of Zia *et al.* (2014) who used microsatellite for estimating the genetic diversity among 40 sunflower lines and noted 6.13 alleles per locus, suggesting SSRS as powerful technology for genetic diversity analysis.

Table 2: Primers used for SSR analysis of sunflower *CMS* lines and restorer lines

	Primer	Ta	No. of alleles	PIC	Forward Primer (5'-3')	Reverse Primer (5'-3')
1	HA 2505	41.97	2	0.23	GTGTCATGACTCGGT	GGACAATGTGATTGC
2	HA 2682	49.13	2	0.28	CACAATCGTTTCTTTCCAAAA	ACCCATATGCCCACTCATAA
3	HA 2879	42.72	2	0.47	CATACCGTTCTTGTTC	CAACCTCCTAGGTCA
4	HA 3513	47.68	2	0.47	TGACCCATTCAACTTCTTAA	TCATGGTTCCTGATGAGAAT
5	HA 3555	48.30	2	0.47	GATATCTCTCATAAGTGCCG	GGTCTTGTGATGACGAA
6	HA 3638	55.28	2	0.49	GACATAATCACTAGTTGTTGGTGC	CTCCTCCCACCTCAACAATTTC
7	HA 3651	51.02	2	0.46	GGAATTATCCATTGTAGGTTTGG	GGATGATTGATTAATTGAGGG
8	HA 3691	56.69	2	0.47	GAATGAAGCATGTGGAAGGCGG	GTGGAGGTGATGATGGTATGAG
9	HA 3878	49.86	2	0.49	TTTGTTTAGCATCATCATCATC	GAGACCCTAACCATAACATGA
10	HA 4057	50.75	2	0.36	AAACCCTTCCGACTTATCTC	TAAAGAGAGAGCCCAACAAG
11	HA 4136	52.84	4	0.37	CCTATTCCTGATAATTCACTAAGC	GGTAGCATGCTTACATTAAGATG
12	HA 4142	53.63	2	0.48	GAGTCGACATTTTCGGAAATCG	CTTCATCTTCTGACACCCAAC
13	HA 4149	54.35	2	0.30	CAAAAACCTCTCTCCGTTGGC	GACTCCAAAGTCCACCAAATC
14	ORS 10	53.56	2	0.50	CTCCACCACACAATCTCCG	TGTGGAGTTTGCCAGAACTG
15	ORS 1088	54.17	2	0.41	ACTATCGAACCTCCCTCCAAAC	GGATTTCTTTCATCTTTGTGGTG
16	ORS 123	51.48	3	0.50	GAAAACCCATGCAGGCATAC	ACATCCATCACAGTCCATTTTG
17	ORS 149	57.62	2	0.50	GCTCTCTATCTCCCTTGACTCG	TGCTCTAAGATCTCAGGCGTGC
18	ORS 154	51.48	2	0.27	GCACCTTTGGTGAGGAGATA	TGCATCAGTAGCTATTGTCTAT
19	ORS 317	51.92	4	0.70	TTTGGCAGTTTGGTGGCTTA	GGTCGTATGCTTAATTCTTTCTCT
20	ORS 331	51.92	2	0.26	TGAAGAAGGGTTGTTGATTACAAG	GCATTGGGTTCACCATTTCT
21	ORS 35	51.72	3	0.49	TTTTCTCAGCTTCAGTTTCAGC	TTTGGATCAAGAACTCATGGC
22	ORS 382	52.80	2	0.47	ACCGATGGACGATGATCCTA	TGCACCTGGAATGGCTAGAT
23	ORS 443	52.13	2	0.42	GGAAGTAGGAAGAACAAAGAA GCA	TTGGAGTGTGAATTTGGATTTG
24	ORS 47	55.06	2	0.20	TGAGATTCCTCGTACTTCATCTG	CCATAGGATGCATAGGAAGAGG
25	ORS 476	51.13	2	0.50	GCGTCGAGTCATTGGGATAA	TGCAAACGCATCTACGATATT
26	ORS 591	52.69	2	0.11	GTTGCAGGAAAGGTCGAAAG	TTTTATTCTACGCAGACAAGACG
27	ORS 607	52.41	2	0.50	GCCTGCACATGAGAATGACA	CTAATACAAAGCAACCCGACAA
28	ORS 613	52.56	2	0.46	GTAAACCCTAGGTCAATTTGCAG	ATCTCCGGAAAACATTCTCG
29	ORS 70	52.51	2	0.31	GACCCTGGTCACCGAAGTTA	ATCTGAAATCGGACAAGATTCA
30	ORS 708	52.80	2	0.36	GCACTCAACATCACGAATCC	CAACACCTTGGGAACACCTT
31	ORS 721	50.75	3	0.52	ATTCCATTCGCATTCGTGTT	TCCTCAACGCTTCACACAAG
32	ORS 809	53.04	2	0.33	GCCGGTTCTCTCATTCTTCTC	TCAGAAACAAACCACCACCA
33	ORS 822	54.37	2	0.47	CAATGCCATCTGTCATCAGCTAG	AAACAAACCTTTGGACGAAACTC
34	ORS 928	52.18	2	0.49	CATGGTTATTTTGGTTTGGGTTT	GCTATTATCATGTCCTTGTCCTTTT

Range of the amplified product varied from 50 to 350 bp. The polymorphism information content index (PIC) was estimated for polymorphic primers to analyze the amplification of SSR primers, and this varied from 0.12 (ORS 591 primer) to 0.71 (ORS 317 primer). The average PIC value for all the 34 polymorphic markers was 0.42 (Table 2). Fifty nine per cent markers had PIC value of 0.50 or greater which suggests that majority of informative SSRs were present between these accession as their PIC values were 0.5. Observation of Gedil (1999) and Yu *et al.* (2002) were similar with average PIC value (0.49 and 0.55 respectively) for polymorphic SSR markers in sunflower. Zia *et al.* (2014) observed similar range of the PIC that varied from 0.17 to 0.89.

Cluster analysis (Fig. 1) divided all the parental lines into two main clusters (A and B) depending upon their genetic distances. Cluster A had two
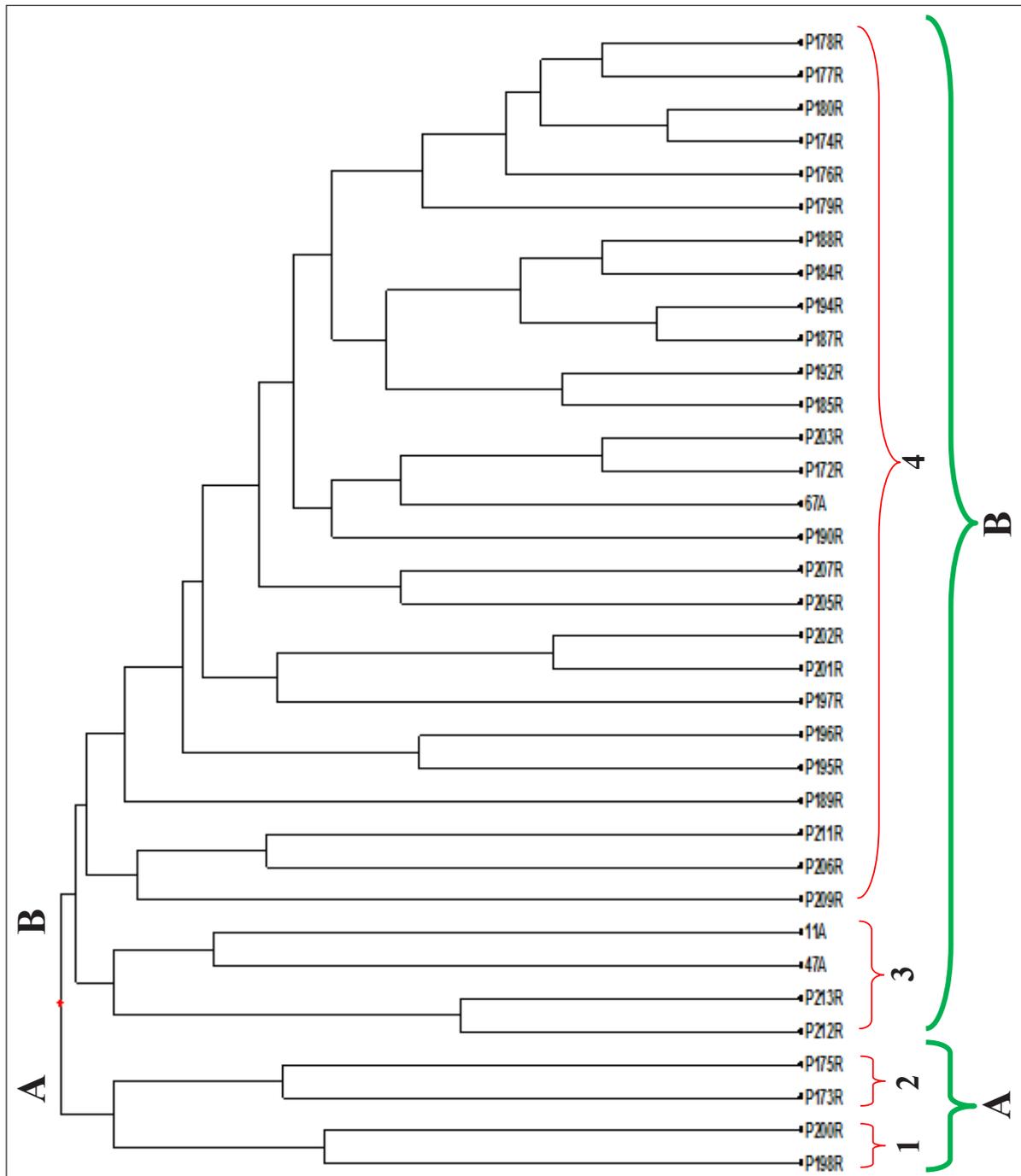
Fig. 1: Dendrogram constructed for 35 sunflower parental lines based on distances using 34 polymorphic SSR primers

subclusters *i.e.* 1 and 2, two genotypes fell in subcluster 1 (P198R and P200R), two in subcluster 2 (P173R and P175R). Whereas cluster B contained rest of the 31 genotypes which were further grouped into two subclusters *i.e.* 3 and 4, subcluster 3 contained four genotypes *viz.* 47A, 11A, P212R and P213R, and subcluster 4 contained the remaining 27 genotypes (Table 3).

Dissimilarity matrix results revealed high genetic diversity among all the parental lines (Table 4). The range of the genetic distances estimated for all the combinations varied from 0.32 (67A × P188R) to 0.73 (11A × P174R). The distances measured based on SSR primers showed that the CMS 11A was the most divergent genotype with a mean genetic dissimilarity coefficient of 0.59. The genotypes

Table 3: Grouping of parental lines into clusters on the basis of SSR primer analysis

Cluster	Subcluster	No. of genotypes	Name of genotypes
A	1	2	P198R and P200R
	2	2	P173R and P175R
	3	4	11A, 47A, P212R and P213R
B	4	27	P178R, P177R, P180R, P174R, P176R, P179R, P188R, P184R, P194R, P187R, P192R, P185R, P203R, P172R, 67A, P190R, P207R, P205R, P202R, P201R, P197R, P196R, P195R, P189R, P211R, P206R and P209R

Table 4: Genetic distances based on simple sequence repeat markers data for all pair-wise combinations of thirty five parental genotypes

Genotypes	P172R	P173R	P174R	P175R	P176R	P177R	P178R	P179R	P180R	P184R	P185R	P187R	P188R	P189R	P190R	P192R
67A	0.37	0.53	0.47	0.55	0.43	0.48	0.47	0.48	0.46	0.48	0.50	0.43	**0.32**	0.67	0.43	0.66
47A	0.53	0.62	0.49	0.61	0.53	0.50	0.56	0.61	0.48	0.41	0.52	0.39	0.48	0.56	0.53	0.38
11A	0.66	0.69	**0.73**	0.51	0.66	0.63	0.69	0.65	0.68	0.48	0.46	0.56	0.48	0.53	0.66	0.48

Genotypes	P194R	P195R	P196R	P197R	P198R	P200R	P201R	P202R	P203R	P205R	P206R	P207R	P209R	P211R	P212R	P213R
67A	0.48	0.51	0.45	0.57	0.50	0.50	0.45	0.40	0.34	0.47	0.67	0.53	0.56	0.51	0.57	0.54
47A	0.47	0.50	0.51	0.53	0.49	0.56	0.47	0.42	0.53	0.42	0.60	0.49	0.51	0.50	0.42	0.65
11A	0.50	0.47	0.54	0.69	0.61	0.67	0.51	0.53	0.63	0.66	0.60	0.63	0.66	0.54	0.42	0.61

having highest genetic distance can be used for hybridization as they are more diverse from each other. Pankovic *et al.* (2004a) and Pankovic-Saftic (2007) observed the genetic distance (GD=7-75%) between sunflower inbred lines, obtained with RAPD and SSR markers, that indicate large variability.

REFERENCES

Backes, R.L., Souza, A.M., Balbinot, A.A., Gallotti, G.J.M. and Bavaresco, A. 2008. Desempenho de cultivares de girassolemduasepocas de Plantio de safrinha no planaltonortecatarinense. *Sci. Agric.,* **9**: 41-48.

Beg, A., Rana, M.A. and Aslam, M. 1984: Sunflower production practices. *Prog. Fmg.,* **4**: 14-19.

Brinkmann, B., Klintschar, M., Neuhuber, F., Hühne, J. and Rolf, B. 1998. Mutation Rate in Human Microsatellites: Influence of the Structure and Length of the Tandem Repeat. *American J Human Genetics*, **62(6):** 1408–1415.

Cheres, M.T., Miller, J.F., Crane, J.M. and Knapp, S.J. 2000. Genetic distance as a predictor of heterosis and hybrid performance within and between heterotic groups in sunflower. *Theor. Appl. Genet.,* **100:** 889-894.

Darvishzadeh, R., Azizi, M., Hatami-Maleki, H., Bernousi, I., Abdollahi Mandoulakani, B., Jafari, M. and Sarrafi, A. 2010a. Molecular characterization and similarity relationships among sunflower (*Helianthus annuus* L.) inbred lines using some mapped simple sequence repeats. *African J. Biotech.,* **9:** 7280-7288.

Evertt, N.P., Robinson, K.E. and Mascarenhas, D.1987. Genetic engineering of sunflower (*Helianthus annuus* L.). *Nat. Biotech.,* **5:** 1201-1204.

Gedil, M.A. 1999. Marker development, genome mapping, and cloning of candidate disease resistance genes in sunflower (*Helianthus annuus* L.) Ph.D thesis, Oregon State University, Corvallis.

Gossal, S.S., Vasiljevic, L. and Brar, D.S. 1988. Plant biotechnology and sunflower improvement. Proc 12th International Sunflower Conference, pp. 599, Novisad.

Hernández, P., Martín, A. and Dorado, G. 1999. Development of SCARs by direct sequencing of RAPD products: a practical tool for the introgression and marker-assisted selection of wheat. *Mol. Breed.,* **5:** 245-253.

Mazers, G.R., Moyret, C., Jeanteur, P. and Theillet, C.G. 1991. Direct sequencing by thermal asymmetric PCR. *Nucleic Acids Res.,* **19:** 47-83.

Mosges, G. and Friedtu, W. 1994. Genetic 'Fingerprinting' of sunflower lines and F1 hybrids using. Isozymes, simple

and repetitive sequences as hybridization probes, and random primers for PCR. *Plant Breed,* **113:** 114-124.

Murray, M.G. and Thompson, W.F. 1980. Rapid isolation of high molecular weight DNA. *Nucleic Acids Res.,* **8:** 4321-4325.

Pankovic, D., Jocic, S., Lacok, N., Sakac, Z. and Scoric, D. 2004a. The use of PCR-based markers in the evaluation of resistance to downy mildew in NS-breeding material. *Helia,* **27(40):** 149- 158.

Pankovic-Saftic, D. 2007. The application of molecular markers in sunflower breeding. *Genetika,* **39(**1): 1 - 11.

Perrier, X. and Jacquemond, C.J. 2006. DARwin software. http://darwin.cirad.fr/darwin/

Rieseberg, L.H. and Seiler, G.J. 1990. Molecular evidence and the origin and development of the domesticated sunflower (*Helianthus annuus*, Asteraceae). *Econ. Bot.,* **44**: 79-91.

Sahari, K., Nicolas, P., Stephane, M., Aurelie, B., Faycal, B.J., Patrick, V. and Dominique, B. 2016. Genetic diversity and core collection cotitution for subsequent creation of new sunflower varieties in Tunisia. *Helia***,** **39**(64): 123-137.

Tang, S. and Knapp, S.J. 2003. Microsatellites uncover extraordinary diversity in native American land races and wild populations of cultivated sunflower. *Theor. Appl. Genet.,* **106(6):** 990-1003.

Tersac, M., Blanchard, P., Brunel, D. and Vincourt, P. 1994. Relationships between heterosis and enzymatic polymorphism in populations of cultivated sunflowers (*Helianthus annuus* L.). *Theor. Appl. Genet.,* **88**: 49-55.

Turnpenny, P. and Ellard, S. 2005. *Emery's Elements of Medical Genetics* (12th ed.). London: Elsevier.

Yu, J.K., Mangor, J., Thompson, L., Edwards, K.J., Slabaugh, M.B. and Knapp, S.J. 2002. Allelic diversity of simple sequence repeats among elite inbred line of cultivated sunflower. *Genome,* **45**: 1652-1660.

Zia, Z.U., Sadaqat, H.A.,Tahir, M.H.N., Sadia, B., Bushman, B.S., Hole, D., Michael, L. and Malik, W. 2014. Estimation of Genetic Diversity Using SSR Markers in Sunflower. *Russian J. Genet.,* **50:** 498-507.