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)

**Paper No.: 639** **Received: 14-08-2017** **Accepted: 09-12-2017**

**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. Its 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 nucleotide 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.

\[ T_m (°C) = 4(G+C) + 2(A+T) \]

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 unweighted 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.
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.
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.

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

<table>
<thead>
<tr>
<th>Primer</th>
<th>Ta No. of alleles</th>
<th>PIC</th>
<th>Forward Primer (5'-3')</th>
<th>Reverse Primer (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA 2505</td>
<td>41.97</td>
<td>2</td>
<td>0.23</td>
<td>GTGTCATGACCTCGGT</td>
</tr>
<tr>
<td>HA 2682</td>
<td>49.13</td>
<td>2</td>
<td>0.28</td>
<td>CACAATCGTTCTTTTAAAAAA</td>
</tr>
<tr>
<td>ORS 591</td>
<td>52.69</td>
<td>2</td>
<td>0.11</td>
<td>GTTTGACAGGAAAGCAGGAAAA</td>
</tr>
<tr>
<td>ORS 317</td>
<td>51.92</td>
<td>4</td>
<td>0.70</td>
<td>GTACGTTTTTACACTCGTTGTTGTC</td>
</tr>
<tr>
<td>ORS 331</td>
<td>51.92</td>
<td>2</td>
<td>0.26</td>
<td>TGAAGAAGGATGGTGGATTCAAGA</td>
</tr>
</tbody>
</table>
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
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.

ACKNOWLEDGMENTS
Authors are highly thankful to the Head, Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana-141004 for providing the necessary research facility and the staff of molecular lab and sunflower field for the required help during the entire research work.

REFERENCES


---

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

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Subcluster</th>
<th>No. of genotypes</th>
<th>Name of genotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>2</td>
<td>P198R and P200R</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>2</td>
<td>P173R and P175R</td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td>4</td>
<td>11A, 47A, P212R and P213R</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>67A</td>
<td>0.37</td>
<td>0.53</td>
<td>0.47</td>
<td>0.55</td>
<td>0.43</td>
<td>0.48</td>
<td>0.47</td>
<td>0.48</td>
<td>0.46</td>
<td>0.48</td>
<td>0.46</td>
<td>0.46</td>
<td>0.43</td>
<td>0.48</td>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td>47A</td>
<td>0.53</td>
<td>0.62</td>
<td>0.49</td>
<td>0.61</td>
<td>0.53</td>
<td>0.50</td>
<td>0.61</td>
<td>0.48</td>
<td>0.41</td>
<td>0.52</td>
<td>0.39</td>
<td>0.48</td>
<td>0.56</td>
<td>0.53</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>11A</td>
<td>0.66</td>
<td>0.69</td>
<td>0.73</td>
<td>0.51</td>
<td>0.66</td>
<td>0.63</td>
<td>0.69</td>
<td>0.65</td>
<td>0.68</td>
<td>0.46</td>
<td>0.46</td>
<td>0.56</td>
<td>0.48</td>
<td>0.53</td>
<td>0.66</td>
<td>0.48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>67A</td>
<td>0.48</td>
<td>0.51</td>
<td>0.45</td>
<td>0.57</td>
<td>0.50</td>
<td>0.45</td>
<td>0.40</td>
<td>0.34</td>
<td>0.47</td>
<td>0.67</td>
<td>0.53</td>
<td>0.56</td>
<td>0.51</td>
<td>0.57</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>47A</td>
<td>0.47</td>
<td>0.50</td>
<td>0.51</td>
<td>0.53</td>
<td>0.49</td>
<td>0.56</td>
<td>0.47</td>
<td>0.53</td>
<td>0.42</td>
<td>0.60</td>
<td>0.49</td>
<td>0.51</td>
<td>0.50</td>
<td>0.42</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>11A</td>
<td>0.50</td>
<td>0.47</td>
<td>0.54</td>
<td>0.69</td>
<td>0.61</td>
<td>0.67</td>
<td>0.51</td>
<td>0.63</td>
<td>0.66</td>
<td>0.60</td>
<td>0.63</td>
<td>0.66</td>
<td>0.54</td>
<td>0.42</td>
<td>0.61</td>
<td></td>
</tr>
</tbody>
</table>
and repetitive sequences as hybridization probes, and random primers for PCR. Plant Breed, 113: 114-124.


