Korean J. Plant Res. 32(3):244-253(2019)
https://doi.org/10.7732/kjpr.2019.32.3.244

Print ISSN 1226-3591
Online ISSN 2287-8203

Original Research Article

# Assessment of Genetic Diversity and Population Structure on Kenyan Sunflower (*Helianthus annus* L.) Breeding Lines by SSR Markers

**Esther W. Mwangi[1], Salem Marzougui[2], Jung Suk Sung[3], Ernest C.Bwalya[4], Yu-Mi Choi[3] and Myung-Chul Lee[3]***

[1]Researcher, Food Crops Research Institute, Kenya Agricultural & Livestock Research Organization P.O. Private Bag Njoro-20107, Republic of Kenya
[2]Researcher, Institution de la Recherche et de l'Enseignement Supérieur Agricoles, National Agricultural Research Institute, Compexe Universitaire, Boulifa, ElEkf 7100, Republic of Tunisia
[3]Senior Researcher, National Agrobiodiversity Center, National Institute of Agricultural Sciences RDA, 370 Duckjin-gu, Jeonju-si, Jeollabuk-do, 54874, Korea
[4]Researcher, Zambia Agricultural Research Institute, P/Bag 7 Mt. Makulu Rd, Chilanga, Republic of Zambia

**Abstract -** In crop breeding program, information about genetic dissimilarity on breeding resources is very important to corroborate genealogical relationships and to predict the most heterozygotic hybrid combinations and inbred breeding. This study aimed to evaluate the genetic variation in Kenyan sunflower breeding lines based on simple sequence repeat (SSR). A total of 83 alleles were detected at 32 SSR loci. The allele number per locus ranged from 2 to 7 with an average of 2.7 alleles per locus detected from the 24 sunflower accessions and the average value of polymorphic information contents (PIC) were 0.384. A cluster analysis based on the genetic similarity coefficients was conducted and the 24 sunflower breeding resources were classified into three groups. The principal coordinates (PCoA) revealed 34% and 13.38% respectively, and 47.38% of total variation. It was found that the genetic diversity within the Kenyan sunflower breeding resources was narrower than that in other sunflower germplasm resources, suggesting the importance and feasibility of introducing elite genotypes from different origins for selection of breeding lines with broader genetic base in Kenyan sunflower breeding program.

**Key words** − EST-SSR, Genetic diversity, Population structure, Sunflower

## Introduction

Sunflower belongs to the genus *Helianthus* of the family *Asteraceae*, which is a diverse and large family of flowering plants. Sunflower, *Helianthus annus* L., is an important agricultural crop in most sunflower growing countries that produces edible oils, that is estimated as a12% of global oil production of vegetable oils worldwide. It takes position four after palm oil, soybean and canola in vegetable oil production (Rauf *et al*., 2017). Oil from sunflower is considered a healthy option to vegetable oils, because it has high content of polyunsaturated fatty acids with 55-60 percent of linoleic acids and 25-30 percent of oleic acids, which helps in reduction of cardiac related problems.

*Corresponding author. E-mail : mcleekor@korea.kr
Tel. +82-63-238-4900

The sunflower plant is a major oil crop and belongs in a crop rotation system in South-West Kenya due to growing well under drought condition compared with other crops in tropical semi arable condition. The sunflower crop has various varieties, but two main types are exists as tall and dwarf in Kenya. The tall varieties are open pollination as a landrace or inbred line and can grow to 1.5-2.4 m. The most grown tall varieties are Hungarian white, Kenya fedha, Kenya shaba, Record and Grey stripped. The most commonly local dwarf varieties are normally hybrids and grow to 1.2m with smaller head. The efficiency of hybrid breeding programs could be augmented, if the inbred lines could be assessed and superior crosses predicted before field evaluation (Melchinger *et al*., 1991). Open-pollinated and composite populations and inbred lines are main resources for sunflower breeding programs in Kenya. Therefore,

ⓒ 2019 by The Plant Resources Society of Korea
This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

evaluation of genetic diversity of such parental lines using molecular markers is important in sunflower breeding for successful exploration, genetic stability and enhanced heterotic effect in the available germplasm, because morphological characters are limited in number and unstable (Sahranavard *et al*., 2015).

Molecular markers play a major role in identifying variation in genomic DNA sequences. Polymorphism at molecular level in different species has played a major role in the analysis of genetic diversity, identification of phylogenetic relationships and also in transfer of target traits to elite germplasm using molecular linkage map in the aim of improving breeding activities. Different molecular markers have been used in mapping genes leading to development of sunflower linkage maps resulting in studies revealing genetic diversity in the genus *Helianthus*. Among The first molecular linkage maps was developed using Random Amplified Polymorphic DNA (RAPD) markers on wild sunflower (Rieseberg *et al*., 1993). There after other markers were generated and published using PCR based Restriction Fragment Length Polymorphic (RFLP) markers (Berry *et al*., 1995; Gentzbitel *et al.,* 1995; Jan *et al*., 1998). Most of the sunflower linkage maps reported 17 linkage groups which represent the number of haploid chromosomes present in sunflower. Simple sequence repeats (SSR) marker resources have been developed for sunflower (Paniego *et al,* 2002; Yu *et al*., 2002; Gandhi *et al.,* 2005; Pashley *et al*., 2006) and these markers were used to make new linkage maps (Tang *et al.*, 2003a,b; Dehmer and Fried, 2006) and expressed sequence tags (EST) - SSR, INDELS and single nucleotide polymorphism (SNP) markers were added on the already existing sunflower maps (Heesacker *et al*., 2008, Carla *et a*l 2015). Recently, Badouin *et al*. (2017) published draft genome sequence of sunflower.

In case of genetic diversity analysis of crops, genomic-SSRs are more popularly used than EST-SSR markers due to the availability of markers, but the current trend towards the use of functional markers. ESTs derived markers have several fundamental advantages over genomic-SSRs according to embed in transcribed gene sequences. EST-SSR markers can be increase the efficiency to direct gene tagging for agronomical important traits by quantitative trait locus mapping and show a higher level of transferability to closely related species than genomic SSR markers (Gupta *et al*., 2003, Bhat *et al.,* 2005). EST-SSR markers have been used in several studies, such as evaluation of genetic diversity in little millet (Ali *et al*., 2017), foxtail millet (Ali *et al*., 2016) and perilla (Song *et al*., 2015). Furthermore, if the hybrid inbred lines *per se* could be screened by molecular marker based genetic distance, prediction of heterosis can be possible to make superior crossing before field evaluation in sunflower breeding.

In Kenya, sunflower crop is more important in semi-dried area according to climate changes, because profits from growing sunflower with three months harvesting period are higher than those from traditional crops such as maize, sorghum and tobacco. Therefore, molecular characterization and estimation of the genetic diversity of genotypes that used in breeding are required to further improve the breeding programs. The present study was designed to estimate the level of genetic diversity in newly introduced sunflower lines through SSR markers.

## Materials and Methods

### Plant material and DNA extraction

Twenty four accessions, which consists of six landraces, two modern cultiva, ten restorers, four male sterilities and two female lines, were obtained from Kenya agricultural and livestock research organization (Table 1) and analyzed for genetic diversity using 32 EST-SSR and genomic-SSR markers (Table 2) that reported by Heesacker *et al*. (2008). Ten seeds of each accession were germinated and grown in the green house. The leaves were harvested at four-leaf stage after 10 days of planting and genomic DNA was extracted according to NucleoSpin Plant II Kit protocol (Macherey-Nagel, Germany). DNA quality and quantity were determined using a UV-Vis spectrophotometer micro plate reader (Biotech instrument, USA Ltd.) and visual comparison of 1% agarose gel electrophoresis. The extracted DNA was adjusted concentration at 50 ng/µL and stored at -20℃ for further use.

### PCR amplification

Thirty two SSR markers (Table 2) that reported earlier by

Table 1. A List of 24 sunflower inbred lines, restorer and Maintainers genotypes used in this study

| Number | Genotype name | Status | Category group | Place of origin |
|---|---|---|---|---|
| 1 | PI 650627 | Landrace | Open pollination | USA |
| 2 | PI 650631 | Landrace | Open pollination | USA |
| 3 | PI 650632 | Landrace | Open pollination | USA |
| 4 | PI 650635 | Landrace | Open pollination | USA |
| 5 | PI 650637 | Landrace | Open pollination | USA |
| 6 | PI 650638 | Landrace | Open pollination | Poland |
| 7 | ARGENOSOL | Modern cultivar | Inbred line | Canada |
| 8 | KENYA FEDHA | Modern cultivar | Inbred line | Kenya |
| 9 | HVR9 | Restorer | R line | Kenya |
| 10 | R9190 | Restorer | R line | Kenya |
| 11 | R9193 | Restorer | R line | Kenya |
| 12 | R9196 | Restorer | R line | Kenya |
| 13 | R1 CM 595XRHA 859 | Restorer | R line | USA |
| 14 | R2(592XRHAX859) | Restorer | R line | USA |
| 15 | R3(RHAX383X372)CM 632 | Restorer | R line | USA |
| 16 | RHA 595XRHA 859 | Restorer | R line | USA |
| 17 | CM632 | Restorer | R line | USA |
| 18 | CM589XCM632 | Restorer | R line | USA |
| 19 | B1326 | Maintainer | Female line | Kenya |
| 20 | A6208 | Female | A lines | Kenya |
| 21 | A6213 | Female | A lines | Kenya |
| 22 | B1338 | Maintainer | B lines | Kenya |
| 23 | B1340 | Maintainer | B lines | Kenya |
| 24 | B1345 | Maintainer | B lines | Kenya |

Heesacker *et al.* (2008) were used in this study to amplified the template DNA. The SSR primers were amplified in a 20 μL total volume containing 50 ng of genomic DNA, 2 μL of each SSR primers (10 pmol), 4 μL of 5x reaction Buffer (Inclone Co, Korea), 1 U of *Taq* DNA polymerase (Inclone Co, Korea), 1.6 μL of dNTP (2.5 mM), and 11 μL nuclease-free water. DNA amplifications were performed in PTC-100 thermal controller (MJ Research Watertown, MA, USA). The PCR profile was: initial denaturation of 3 min at 94℃, followed by 35 cycles of 45s at 50-55℃ and 45s at 72℃, and a final extension of 10 min at 72℃. PCR products of clear, stable and specific bands with and expected length (100-250 bp) were considered as successful PCR amplifications. All the amplifications were separated by capillary electrophoresis using Fragment Analyzer$^{TM}$ 96-capillary Automated CE System using DNF-900 double stranded DNA Reagent Kit (Advanced analytical, USA) according to the manufacturer's instruction. SSR alleles were visualized and scored using PROSize 2.0 software version 1.2.1.1 Kit (Advanced analytical, USA).

**Data analysis of genetic diversity and population structure**

Analysis for different parameters of variability such as number of observed alleles ($N_A$), expected homozygosity ($H_O$), expected heterozygosity ($H_E$), and polymorphism information content (PIC) were calculated by Cervus 3.0 software (Araneda *et al.,* 2004; Kalinowski *et al.,* 2007). An unweighed pair group method with arithmetic mean (UPGMA) hierarchal clustering were carried using DARWin 6.0 (Perrier and Jaccquemond-Collet, 2006) to analyze the cluster.

Among clustering method algorithms, neighbors-joining method (Saitou and Nei, 1987) was used to clustering.

Table 2. Information of 32 SSR primers used for analysis of 24 genotypes in this study

| Marker name | Marker type | Linkage group | Forward primer | Reverse Primer | Na[z] | Ho[z] | He[z] | PIC[z] |
|---|---|---|---|---|---|---|---|---|
| HA77 | gSSR | 1 | TGTAATCTGTATCACTTCCACC | GTTGTTCTGTTAGGTCGTTCCG | 2 | 0.000 | 0.156 | 0.141 |
| HA3239 | gSSR | 2 | CTTAATTCCGACATAAGCAGG | CTTTCATTATTCACAACCCCC | 3 | 0.000 | 0.511 | 0.449 |
| HT0734 | EST-SSR | 3 | GTTTCGCAGCTCTAGCATTGA | CCCCCTACAAGGCAACAAAATA | 2 | 0.000 | 0.337 | 0.275 |
| HA2920 | gSSR | 3 | AACGTGTTCATTCGGATGC | GAAGATTGGTATGATGGGGC | 3 | 0.042 | 0.451 | 0.361 |
| HT1045 | EST-SSR | 4 | GGATGAGAGACCGGAGAAATA | AAAGGAAGTATTGATTTATGTATGGA | 3 | 0.000 | 0.606 | 0.523 |
| HA991 | gSSR | 4 | GCCCCCTTGATGCCCTTTTC | GAATCGCCATTTGAATCGCCAG | 2 | 0.000 | 0.284 | 0.239 |
| HT0440 | EST-SSR | 5 | GAAGTTGGGAGGGTTGTTCAAG | CCTCCTGTTGGAACACCAAAT | 2 | 0.000 | 0.507 | 0.373 |
| HT0591 | EST-SSR | 5 | AGAGGAATGAGATCGGGTTGAT | GTGGGACAACTCAGCAACGTC | 3 | 0.000 | 0.528 | 0.420 |
| HA4103 | gSSR | 6 | TCACTTATCACCAATCTCATCCA | CTGTCTCAAATCGGGTGGTT | 2 | 0.083 | 0.422 | 0.328 |
| HA3312 | gSSR | 6 | TAACTGCAACAGACATGGAA | ATAACGATTGCACAACACAA | 2 | 0.083 | 0.479 | 0.359 |
| HA1848 | gSSR | 7 | CATCCCCTTCTGAATAGAAA | TAGGCTCGTTAAACTTACGG | 7 | 0.000 | 0.830 | 0.788 |
| ORS1041 | gSSR | 7 | AAACAAACCTTAATGGGGTCGTA | ATATTGGCTGGTTGATGCTGAT | 2 | 0.000 | 0.454 | 0.346 |
| HT0519 | EST-SSR | 8 | TTTCGTGACCGAGAAAGGTAT | TGTGGCATACATAGAAATGATCTAA | 3 | 0.000 | 0.585 | 0.510 |
| HA3581 | gSSR | 8 | GTAGAGGAGTATCGCCAATAGC | GTTCCAATGGCCAACGATTGTG | 2 | 0.000 | 0.383 | 0.305 |
| HA2063 | gSSR | 9 | CACCGCAGATGGCCAATTCAAC | GGTTCTGCTAACCTGCAATGCG | 2 | 0.000 | 0.507 | 0.373 |
| ORS510 | gSSR | 9 | CATCGCGTCCCTCTCTCTAA | CCAACCATCACAGCAATCAG | 3 | 0.000 | 0.606 | 0.523 |
| HA928 | gSSR | 10 | CCTTTGTAGTCCCTTACTGG | GGTCGATCATGTATGCGTGTTG | 3 | 0.000 | 0.613 | 0.533 |
| HA3204 | gSSR | 10 | GCCCTTCAATCCTACCATTAAC | GGATATTGAGTTGGTTGTTGGG | 2 | 0.000 | 0.223 | 0.195 |
| HT0426 | EST-SSR | 11 | AACCGTAAATGAAATCGGTGTG | GAGGGCAAAGTTGGGATACTCT | 2 | 0.000 | 0.156 | 0.141 |
| HT0382 | EST-SSR | 13 | GAAGTCTGCTTTGGTGTTGGTC | ATGGCTCTCTCATTTCCACTTG | 3 | 0.000 | 0.507 | 0.408 |
| HT0271 | EST-SSR | 13 | CCCATTGTTGCCTAATTCAAGAT | TTGGAGTGCATAACACGTATCAG | 3 | 0.000 | 0.606 | 0.523 |
| HA293 | gSSR | 14 | GGGGACATCTCCCGTCCACC | CCTCATCCATCTCCATCCAATC | 3 | 0.000 | 0.656 | 0.570 |
| HA2077 | gSSR | 14 | GATAAGGTTTTCTCTCTCCC | GAGAAAAATGAGCTGATACCG | 2 | 0.000 | 0.284 | 0.239 |
| HT1024 | EST-SSR | 15 | AAGCAGGTTGCATGAAGAGAAG | GTCGAAACGGGTCAGGTTGTAT | 2 | 0.000 | 0.223 | 0.195 |
| HT0586 | EST-SSR | 15 | TGAACTCTTGTGTTGGCATCT | AAATGTGGATTTATGTATCTCAGTAA | 2 | 0.000 | 0.422 | 0.328 |
| HA4239 | gSSR | 15 | GAATGATAGTGAATTGAGACAGG | CTGGCATCTATATCCATGGATAG | 2 | 0.000 | 0.383 | 0.305 |
| C2070 | EST-SSR | 16 | GGGTAATGCAAAGTACTAAGATGTG | GCATCATCCAACAAACTAGAAGG | 3 | 0.000 | 0.606 | 0.523 |
| HT0723 | EST-SSR | 16 | TCAAGCAATCAGACACCACATC | TGAACACAACCAAGAAATCCAA | 2 | 0.000 | 0.383 | 0.305 |
| HT0998 | EST-SSR | 17 | CATGTCCCGATCAAAGAGTTGT | CCATATCCTGGTTGTTGTGGAG | 2 | 0.000 | 0.479 | 0.359 |
| HA1402 | gSSR | 17 | GTGTGAATGTGAGTGTGAATCC | GTCACAGTTGCATATCCATCC | 3 | 0.000 | 0.638 | 0.545 |
| HA1752 | gSSR | NA | CCTTCTTCCTCAAATCCCG | CGTCATTGTCATTGAGAAGATCG | 2 | 0.000 | 0.337 | 0.275 |
| HA2147 | gSSR | NA | CCCCATAAACATTTCACACCC | CCTGAGAAATCGGTAGATAGTG | 4 | 0.000 | 0.592 | 0.523 |
| Total | | | | | 83 | 0.208 | 14.754 | 12.28 |
| Mean | | | | | 2.6 | 0.007 | 0.461 | 0.384 |

[z]Abbreviations; PIC: polymorphic information content, Na: the number of observed alleles, Ho: observed of heterozygosity, He: expected heterozygosity.

Principle coordinate analysis (PCoA) was performed using GenAlex version 6.5 (Peakall and Smouse, 2012) with 999 times boost-strapping to analyze the molecular variance (AMOVA). Genetic structure was analyzed by STRUCTURE 2.3.4 (Pritchard *et al*., 2000) with 10 independent run and 100,000 Markov Chain Monte Carlo (MCMC) repetitions after a burn-in period of 100,000 interactions for each group number K. The actual number of subpopulations was determined by the logarithm of likehood for each K;Ln P (D) = L (K) and the optimum value $\Delta K$ was obtained by $\Delta K$ = [L'' (K)]/ Stdev according to Evanno *et al*. (2005) as to determine the most likely number of groups.

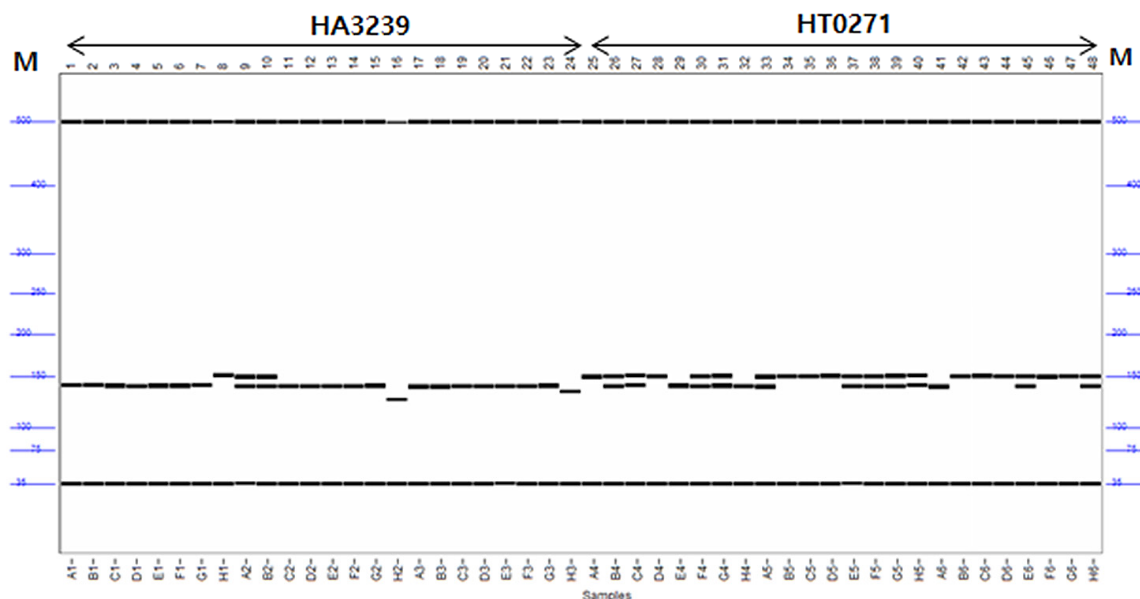


Fig. 1. Amplicons of 24 sunflower accessions using HA3239 and HT0271 primers by Fragment Analyzer$^{TM}$ 96-capillaryAutomated CE System. M: size markers of 500bp and 35bp, Lanes 1-24, 25-48: The name of accessions as listed in Table 1.

## Results

### Allelic diversity of SSR marker

A total of 98 SSR primer pairs, which were earlier reported by Heesacker *et al.* (2008) to make the molecular genetic map of sunflower, were used to select the polymorphic primers about 24 Kenyan sunflower breeding lines. The 32 SSR primers were displayed clear and repeatable polymorphic bands on capillary electrophoresis (Table 2 and Fig. 1). A total number of 83 alleles were detected with an average of 2.5 alleles per locus. The number of polymorphic alleles ranged from two to seven in used sunflower accessions (Table 2). The HA1848 of genomic SSR marker recorded the highest number of alleles of the markers as seven polymorphic alleles, while two or four polymorphic alleles were observed in most of amplified product. The average value of polymorphic information content (PIC) for all the 32 polymorphic primers was 0.38 and PIC values ranged from 0.141 (HA77) to 0.788 (HA 1848). No heterozygosity deficiency (Ho) was observed quite low t the mean of 0.007. The value of expected heterozygosity (*He*) ranged from 0.156 (HA77 and HT0426) to 0.830 (HA 1848) with a mean of 0.461. The selected SSR markers were showed some high level of polymorphism in the genetic parameters.

### Establishment of genetic diversity and phylogenetic relations

The unweighted neighbor-joining dendrogram was constructed based on Nei's similarity coefficient of 24 genotypic data and showing the genetic relationship among the accession (Fig. 2). The 24 accessions were clustered into three distinct clusters (I, II and III). Cluster I was nine accessions such as two A-lines (females) and three B lines (maintainers) from Kenya, two landraces from USA, one Restorer and one variety. Cluster II contained five accessions, two R- lines from Kenya and three landraces from USA. Cluster III had ten accessions that consisted with seven accessions from Kenya and four from USA based on geographical collection region, but this cluster was mixed with their usage types such as landrace, maintainer, restore and female lines.

### Population structure

STRUCTURE software (Pritchard *et al*., 2000) was used to determine the subdivisions of the accessions. The distribution of ΔK values was determined by evaluating logarithmic likelihood [L(K)] (Evano *et al.,* 2005) and the LnP(D) values for different ΔKs ranged from 1 to 10. The structre harvester

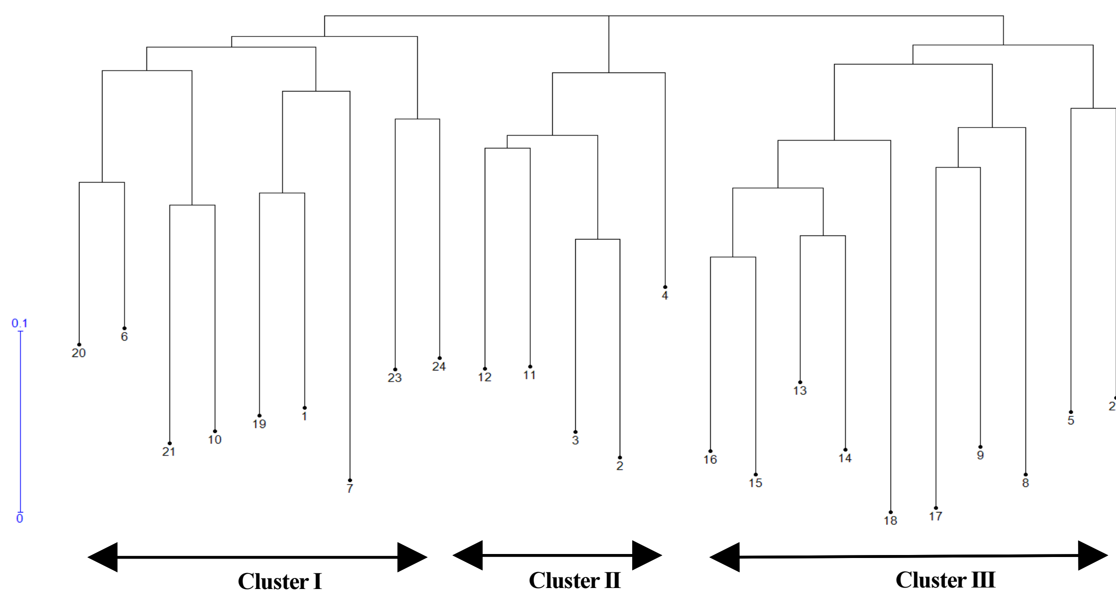


Fig. 2. Unweighted neighbor-joining dendrogram showing genetic relationship among the 24 sunflower accessions based on the genetic dissimilarity matrix data of SSR markers alleles. All the accessions were divided into three clusters.

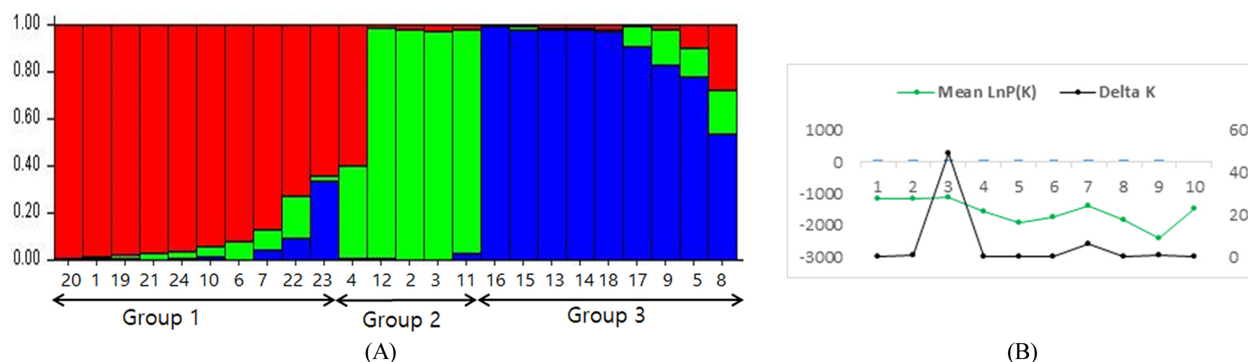


Fig. 3. Population structure of 24 accessions of sunflower based on 34 SSR-markers. (A) Structure analysis (STRUCTURE K =3) distributed the population into groups according to the clusters obtained by the UNJ analysis. Numbers in the 'y' axis show the subgroup membership and the groups are represented by different colors. (B) Average log-likelihood values (mean lnP (D) for 3 iterations) and ad-hoc statistic △k for K values ranging from 1 to 10.

computed best value of K at 3. The genetic structure of 24 genetic accessions showed that the most probable number of subpopulation was three (Fig. 3) based on the ΔK values and the subgroups was in agreement the same accessions were placed under discrimination subgroups by unweighted neighbor-joining dendrogram except B1338 (number 22). B1338 was belonged to Cluster III by unweighted neighbor-joining, but it was located group 1 by STRUCTURE. The results of STRUCTURE analysis indicated that some of accession in Group 1 and 3 showed the admixture type, but group 2 was showed as pure lines. The admixtures observed by STRUCTURE analysis may be a result of cross-pollination in breeding process and have the same parental background.

## Principal coordinate analysis

Principal coordinate analysis (PCoA) was conducted to further assess the population subdivisions identified using SRUCTURE. The principal coordinates explained 34% and 13.38% respectively, and 47.38% of total variation. The PCoA was largely consistent with results of STRUCTURE and the unweighted neighbor-joining dendrogram, the first principal coordinate (PCo1) clearly separated 24 sunflower

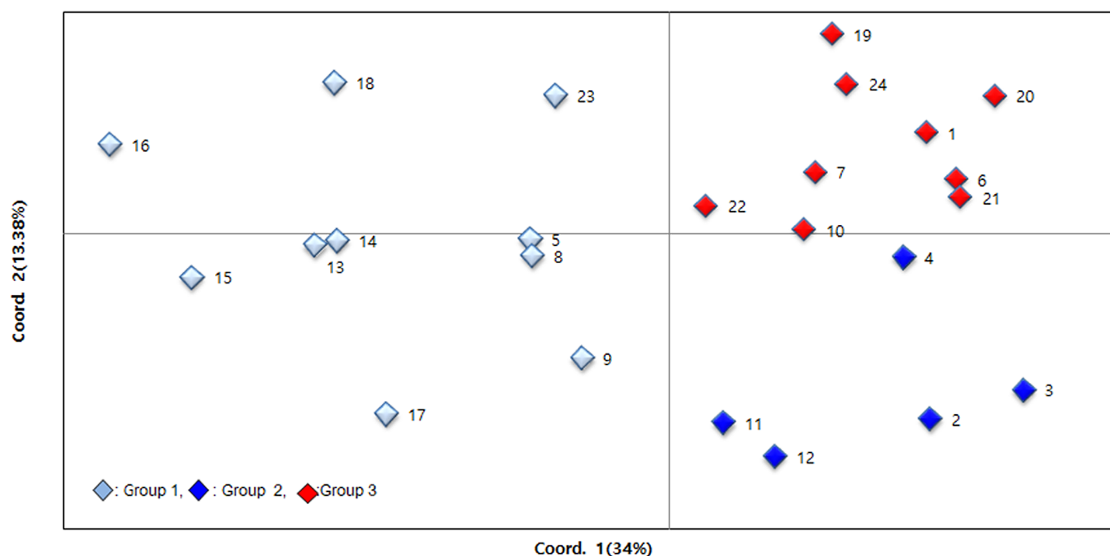


Fig. 4. Principal coordinates analysis (PCoA) of 24 sunflower accessions. Coord. 1 (34%) and Coord. 2 (13.38%) refer to the first and second principal component, respectively.

genotypes into two population and one of group could divide into two subpopulations by the PCo2. All subpopulations were not consisted with collection sites of geographical regions, but the accessions belong to cluster I of UPGMA and group 1 of STRUCTURE showed the more diversity than the other subpopulations (Fig. 4).

## Discussion

Estimate the genetic diversity of plant breeding resources is one of the important pre-breeding activities in crop breeding. The analysis of genetic diversity could be used for three reasons; one is to identify genotypes that underlie important phenotypic and genetic shifts during domestication and breeding using the approach of elimination method (Vigouroux *et al*., 2008). Second is to introduce aspects of the history and timing of domestication and the third is to identify distinct genetic group for retention of germplasm (Agrama and Eizenga, 2009). Identification of sunflower cultivars, lines and hybrids of Kenyan genetic resources have been based on morphological traits and morphological data cannot provide reliable information of calculation of genetic distance in term validation of pedigree (Stanton *et al.,* 1994). In this study, 24 Kenyan sunflower accessions that were mainly used in inbreeding and hybrid breeding system in Kenya is to evaluate the genetic to constitutes backbone to improve the breeding program through the introduction of foreign genes.

In most cases population structure and genetic diversity are commonly estimated using SSR derived information. Genomic SSR are essential markers for population diversity because of their high level of polymorphism. Recently, the use of EST-SSRs with gSSR markers has been an increase to capture high level of reliable diversity. In this study EST-SSR and gSSR markers that were reported position of chromosome map by Heesacker *et al.* (2008) were used to analysis of diversity of sunflower gemplasm accessions. The rage of PIC value was from 0.141 to 0.788 and a mean *He* of 0.41. This was very important, because high PIC value indicates high polymorphism. Markers of a polymorphism higher than 0.5 are useful in genetic studies for they help in distinguishing polymorphism rates of a marker of a specific loci (DeWoody *et al.,* 1994). Observed heterozygosity ranged from 0.00 to 0.083. This was very low as it is expected for inbred lines. *Ho* values were only recorded for HA292, HT410 and HA331 Expected heterozygosity values for the EST-SSR markers ranged from 0.156 to 0.830 (Table 2). There were differences in the number of alleles and the number of private alleles was detected among the subpopulations, but there were no detectable differences in terms of expected heterozygosity.

An UPGMA branch indicated the existence of 3 genetic

groups in these set of germplasm, however few of the genotypes had no clear corresponding relationship between clustering groups as it was also observed by Ali *et al*. (2016) in their study of diversity in foxtail millet. Used sunflower accessions were not consisted with geographical collected region, but used accessions were divided into three subgroups according to their usage of F1 hybrid and inbreeding system STRUCTURE and PCoA. Several studies have been done to evaluate the levels and distribution of genetic diversity in different sunflower accessions, and reported that landraces of sunflower are divide into their collection area, but inbred lines and F1 hybrid materials were not classified by geographically (Mandel *et al*., 2011; Moreno *et al*., 2013). This reason is that cytoplasmic male sterility (CMS) has been used to hybrid breeding in sunflower and resources of CMS are very restricted.

B1326 has been used maintainer to A6213 of female line and the two accessions were located the same cluster, and these two lines are expected the best recombination to make the F1 hybrid. Because the maintainer lines have the restorer fertile and sterile on some female lines that have not fertility, which have involved the maintaining of some gene pools to maximize heterosis in hybrid crosses (Dimitrijevic and Horn, 2017; Fehr, 1991). Only one B line (B1340) was belong to cluster I that consisted with most of the R lines were clustered. It could be due to the breeding history and timing of the development of the two types. We conclude that the estimation of genetic diversity and population structure of 24 main breeding materials of Kenyan sunflower using SSR primer pairs may provide more accurate information to sunflower breeders than the classical pedigree method. The 34 SSR primer pairs used in this study may also be of potential value for further research on genetic mapping, segregation analysis and phylogenetic status of newly introduced germplasm.

## Acknowledgment

This study was carried out with the support of “Research Program for Agricultural Science & Technology Development (Project No. PJ0130832019)” and KAFACI program, National Institute of Agricultural Sciences, Rural Development Administration, Republic of Korea.

## References

Agrama H. and G. Eizenga. 2008. Molecular diversity and genome-wide linkage disequilibrium patterns in a world-wide collection of and its wild relatives. Euphytica 160(3): 339-355. doi: 10.1007/s10681-007-9535-y.

Araneda, C., C. Correa, N. Lam, E. Uribe, M. Camiruaga and P. Iturra. 2004. Asignación de paternidad-maternidad y parentesco en un plantel de avestruces (*Struthio camelus*), utilizando seis loci microsatélites polimórficos. Av. Produc. Anim. 29:3-14.

Ali, A., Y.M. Choi, D.Y. Hyun, S. Lee, J.H. Kim, S.S. Oh and M.C. Lee. 2017. Development of EST-SSRs and assessment of genetic diversity in little millet (*Panicum sumatrense*) germplasm. Korean J. Plant Res. 30(3):287-297.

Ali, A., Y.M. Choi, D.Y. Hyun, S. Lee, S. Oh, H.J. Park, Y.H. Cho and M.C. Lee. 2016. EST-SSR based genetic diversity and population structure among Korean landraces of foxtail millet (*Setaria italica*). Korean J. Plant Res. 29(3):322-330.

Berry, S.T, A.J. Leon, C.C. Hanfrey, P. Challis and S.R. Barnes *et al*.. 1995. Molecular marker analysis of *Helianthus annus* L 2. Construction of an RLFP linkage map for cultivated sunflower. Theor. Appl. Genet. 911:195-199.

Badouin, H., J. Gouzy, C.J. Grassa, F. Murat, D.E. Staton and L. Cottret *et al*.. 2017. The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. Nature 546:148-152. doi: 10.1038/nature22380.

Bhat, P.R., K.V. Krishn, P.S. Hendre, P.K. Rajendra, R.K. Varshney and L.R. Aggarwa. 2005. Identification and characterization of expressed sequence tags-derived simple sequence repeats, markers from robusta coffee variety ‘C×R’ (an interspecific hybrid of *Coffea canephora* × *Coffea congensis*). Mol. Ecol. Notes 5:80-83.

Carla, V.F., A. Natalia, G.R. Juan, Z. Jeremias and P. Andrea *et al*.. 2015. Population structure and genetic diversity characterization of a sunflower association mapping population using SSR and SNP markers. BMC Plant Biol. 15:52 DOI 10.1186/s12870-014-0360-x.

Dehmer, K.J. and W. Fried. 2006. Evaluation of different microsatellite motifs for analysing genetic relationships in cultivated sunflower (*Helianthus annuus* L.). Plant Breeding 117(1):45-48.

DeWoody, J.A., L. Honeycutt and L.C. Skow. 1994. Microsatellite markers in white tailed deer. The J. of Heredity 86(4):317-319.

Dimitrijevic, A. and R. Horn. 2017. Sunflower hybrid breeding: from makers to genomic selection. Front. Plant Sci. 8:2238. Doi: 10.3389/fpls.2017.002238.

Evanno, G, S. Regnaut and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611-2620.

Fehr, W.R. 1991. Principles of Cultivar Development, Theory and Technique. Macmillan, New York Felsenstein, USA.

Gandhi, S.D., A.F. Heesacker, C.A. Freeman, K. Argyris, K. Bradford and S.J. Knapp. 2005. The self-incompatibility locus (S) and quantitative trait loci for self-pollination and seed dormancy in sunflower. Theor. Appl. Genet. 111:619-629.

Gentzbitel, L., F. Vear, Y.X. Zhang, A. Bevilla and P. Nicolas. 1995. Development of concentrated linkage RFLP map of cultivated sunflower *Helianthus annus* L Theor. Appl. Genet. 90:1079-1086.

Gupta, P.K., S. Rustg, S. Sharma, R. Singh, N. Kumar and H.S. Balyan. 2003. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. Mol. Genet. Genomics 270:315-23.

Heesacker, A., V.K. Kishore, W. Gao, S. Tang, J.M Kolkman, A. Gingle, M. Matvienko, A. Kozik, R.M. Michelmore, Z. Lai, L.H. Rieseberg and S.J. Kanpp. 2008. SSRs and INDELs mined from the sunflower EST database: abundance, polymorphisms and cross pyaxa utility. Theor. Appl. Genet. 117:1021-1029.

Jan, C.C., B.A. Vick, J.K. Miller, A.L. Kahler and ET.I. Butler. 1998. Construction of an RFLP linkage map for cultivated sunflower. Theor. Appl. Genet. 96(1):15-22.

Kalinowski, S.T., M.L. Taper & T.C. Marshall. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol. Ecol., 16: 1099-1106.

Mandel, J.R., J.M. Dechaine, L.F. Marek and J.M. Burke. 2011. Genetic diversity and population structure in cultivated sunflower and a comparison to its wild progenitor, *Helianthus annuus* L. Theor. Appl. Genet. 123:693-704.

Melchinger, A.E., M.M. Messmer, M. Lee, W.L. Woodman and K.R. Lamkey. 1991, Diversity and relationships among U.S. maize inbreds revealed by restriction by restriction fragments length polymorphisms. Crop Sci. 31:6669-678.

Moreno, M.V., V. Nishinakamasu, M.A. Loray, D. Alvarez, J. Gieco, A. Vicario, H.E. Hopp, R.A. Heinz, N. Paniego and V.V. Lia. 2013. Genetic characterization of sunflower breeding resources from Argentina: assessing diversity in key open-pollinated and composite populations. Plant Genetic Resources 11(3):238-249.

Paniego, M., M. Echaid, M. Munoz, L. Fernandez, S. Torales, P. Faccio, I. Fuxan, M. Carrera, R. Zandomeni, E.Y. Suarez and H. Hopp. 2002. Microsatellite isolation and characterization in sunflower (*Helianthus annus* L.). Genome 45:34-43.

Pashley, C.H, J.R. Ellis, D.E. McCauley and J.M. Burke. 2006. EST databases a source for molecular markers: Lessons from *Helianthus*. J. Hered. 97:381-388.

Perrier, X. and J.P Jacquamoud-collect. 2006. DARWin software htt://Darwin.cirad.fr/.

Peakall, R and P.E. Smouse. 2012. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-An update. Bioinformatics 28:2537-2539.

Pritchard, J.K., M. Stephens and P. Donnelly. 2000. Inference of population structure using multi. locus genotype data. Genetics 155:945-959.

Rauf, S.J, N. Jamil, S.T. Ali, M. Khan and M. Kausal. 2017. Progress in modification of sunflower oil to expand its industrial value. J. Sci. Food Agri. 97:1997-2006.

Rieseberg, L.H., H. Choi, R. Chan and C. Spore. 1993. Genomic map of diploid hybrid species. Heredit. 70:285-285.

Sahranavard, A.F., R. Darvishzadel, M. Ghadimzadel, H. Azizi and Z. Aboulghasemi. 2015. Identification of SSR loci related to some important agro morphological traits in different oil sunflower (*Helianthus annuus* L.) lines using association mapping. J. Plant Gene. Res. 2:15-32.

Saitou, N and M. Nei. 1987. The neighbor-joining method: a new method for construction phylogenetic tree. Mol. Biol. Evol. 4:406-425.

Song, J.Y., J.R. Lee, S. Oh, C.Y. Kim, C.H. Bae, G.A. Lee, K.H. Ma, Y.M. Choi, H.J. Park and M.C. Lee. 2012. Assessment of genetic diversity and fatty acid composition of perilla (*Perilla frutescens* var. frutescens) germplasm. Korean J. Plant Res. 25(6):762-772.

Stanton, M.A., J.M. Stewart, A.E. Percival and J.F. Wandel. 1994. Morphological diversity and relationships in A-genome cottons, *Gossypium arboreum* and *Gossypium herbaceum*.

Crop Sci. 34:519-527.

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

Tang, S., V.K. Kishore and S.J. Knapp. 2003b. PCR-multiplexes for a genome-wode framework of simple sequence repeat marker loci in cultivated sunflower. Theor. Appl. Genet. 107:6-19.

Vigouroux, Y., J. Glaubitz, Y. Matsuoka, M.M. Goodman, G.J. Sanchez and J. Doebley. 2008. Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. Am. J. Bot. 95:1240-1253.

Yu, J.K., J. Mangor, L. Thompson, K.J. Edward, M.B. Slabaugh and S.J. Knapp. 2002. Allelic diversity of simple sequence repeat markers among elite inbred lines in cultivated sunflower. Genome 45:652-660.

(Received 7 June 2019 ; Revised 21 June 2019 ; Accepted 21 June 2019)