Australian Journal of Crop Science AJCS

AJCS 13(07):1113-1119 (2019) ISSN:1835-2707
doi: 10.21475/ajcs.19.13.07.p1638

Genetic diversity analysis of soybean (*Glycine max (L.)* Merr.) genotypes making use of SSR markers

Keitumetse Kujane[1*], Moosa M Sedibe[1], Alina Mofokeng[2]

[1]Central University of Technology, Free State, Department of Agriculture, Private Bag x20539, Bloemfontein 9301, South Africa
[2]Agricultural Research Council-Grain Crops, Private Bag x1251, Potchefstroom, South Africa

***Corresponding author: kujane2429@gmail.com**

Abstract

In this study, we aimed to investigate the genetic diversity and polymorphism among 30 soybean genotypes maintained by the ARC using simple sequence repeat (SSR) markers. Soybean genotypes were characterized using 20 SSR primers. DNA was extracted using the standard cetyl trimethylammonium bromide method and amplified using PCR. Allele size was determined via comparison with a 100 base pair (bp) DNA ladder. Molecular data were analyzed, and a dendrogram and matrix were generated using GGT 2.0 software. A total of 216 alleles with an average of 10.8 alleles per locus were detected. The allele sizes ranged between 2 and 33 bp with an average of 18.7 bp. The polymorphic information content among genotypes varied from 0.85 (Satt001) to 0.75 (Satt43) with an average of 0.716, and heterozygosity ranged from 0.87 to 0.78 with an average of 0.7485. The most diverse genotypes were B 66 S 31, 69S 7, and R5-4-2 M, which indicated the efficiency of the SSR markers for the detection of genetic diversity. The results of the current study revealed the diversity among the soybean genotypes tested, which might aid breeders in the future in the selection of parents for breeding.

Keywords: Allele; base pairs; dendrogram; DNA; PCR.
Abbreviations: ARC_Agricultural Research Council, bp_base pairs, CTAB_cetyl trimethylammonium bromide, UPGMA_unweighted pair group method with arithmetic mean.

Introduction

Soybean (*Glycine max* (L.) *Merr*.) is one of the oldest cultivated crops. It is relatively the cheapest source of protein and is in considerable demand for food and feed supply. Soybean seeds contain 40-42 % protein, 18-20 % oil, 11 % soluble carbohydrates, and dry matter (Devi et al., 2012). Although the demand for soybean is high, there is a slight decline in the genetic improvement activities of its cultivars. It has a few limitations, such as susceptibility to diseases and pests, adverse environmental conditions, low yield, and poor handling/management. However, all these limitations can be ameliorated in many different ways, including implementing a variety improvement program, cultural practices, post-harvest technology, and field selection.

Genetic diversity plays an important role in the survival and adaptability of plant species. When the practice of farming started, farmers used selective breeding to pass on desirable traits of the crops while omitting the undesirable ones. Little or no genetic diversity makes crops extremely susceptible to plant pests. Studies have recently shown that DNA markers have become an essential tool to implement a soybean improvement program because molecular markers are robust and not affected by the environment. Moreover, several molecular maps have been generated with some morphological traits using amplified restriction fragment length polymorphism (AFLP), random amplified polymorphic DNA (RAPD), and restriction fragment length polymorphism (RFLP) markers in soybean (Apuya et al., 1988; Shoemaker and Specht, 1995; Keim et al., 1997; Cregan et al., 1999). Simple sequence repeat (SSR) markers, also known as microsatellites, are useful, reliable, and easy-to-use markers for soybean characterization. SSR markers have become important for genetic mapping and genetic diversity determination in soybean because they are co-dominant and spread throughout the genome, exhibit high levels of polymorphism, undergo PCR easily, and can be analyzed using gel electrophoresis. Thousands of SSR markers have been developed over the years from both expressed sequence tags (ESTs) and genomic DNA, and those add to the available SSRs. Soybean genome sequence have recently been developed from approximately 33,000 putative SSR markers, and some of them have been mapped (Ott et al., 2011). Marker distribution is often associated with gene distribution. To support this association, a study of microsatellites in Arabidopsis, rice, maize, wheat, and soybean clearly showed that microsatellite distribution was much higher in gene-rich, single-copy regions than in repetitive sections of the genome. A comparison between RFLP and SSR markers in soybean showed that RFLPs tended to be more closely associated with the gene-rich regions,

whereas SSR loci were closely associated with the actual genetic sequences. Ott et al. (2011) showed that many major crops exhibited clustered genes in the gene-rich regions on chromosomes, which followed the patterns of clustered markers. The variability and large number of repeat sequences makes them an excellent tool for pedigree analysis, genotype differentiation, particular genotype identification, and genetic distance evaluation among organisms. According to Guan et al. (2010) and Wang et al. (2010), SSRs have been used successfully in the identification of genetic diversity and relationships among soybean genotypes in different populations. Molecular markers aid plant breeders to indirectly select individuals from different populations that carry a gene for a favorable trait if a tight linkage exists between a marker locus and the genetic locus controlling that trait.

Simple sequence markers characterize genetic diversity and compare the relatedness of germplasm. They are preferred over other markers because they are highly reproducible, an important aspect in genetic analysis. SSR markers require low starter costs and small quantities of DNA for screening; in addition, they can be genotyped easily and rapidly using numerous platforms for DNA fragment analysis, and the analysis could be semi-automated (Cregan et al., 1999; Robinson et al., 2004). Simko et al. (2012) showed that SSRs exhibited high success rates in diversity studies. A study on comparative genetic diversity using SNP, DART, and SSR on 54 sugar beet cultivars showed that the success rate was the highest for SSR markers (Simko et al., 2012) owing to their highly polymorphic nature. Soybean primers were used to amplify SSRs with a success rate of 65 %, despite a lower rate of 3 to 13 % found outside the subgenus *Glycine* (Peakall et al., 1998). Wang et al. (2010) and Guan et al. (2010) showed that SSRs could be successfully used to identify genetic diversity and relationships among soybean genotypes within a population. According to Vinu et al. (2013), environmental fluctuations did not have a considerable effect on molecular profiling because of their authenticity and reliability when it comes to breeding. Although there have been studies on soybean diversity using SSR markers, there are limited studies carried out in South Africa to determine the genetic diversity of soybean using molecular markers. Hence, in this study, we aimed to investigate polymorphism among 30 soybean genotypes using SSR markers.

Results and discussion

Polymorphism and allelic diversity of SSR markers

Table 3 summarizes the number of alleles, size ranges, polymorphic information content (PIC) values, and heterozygosity. SSR markers used in this study generated 216 alleles among the 30 soybean genotypes with an average of 10.8 alleles per locus. Hipparagi et al. (2017) analyzed 75 genotypes using 21 SSR markers and reported 60 alleles with an average of 2.85 alleles per locus. In another study, 38 soybean genotypes were analyzed using 16 SSR markers resulting in 51 alleles with an average of 2.22 alleles per locus (Bisen et al., 2015). In addition, Li et al. (2008) reported 19.7 alleles per locus upon characterizing 1863 genotypes using 59 SSR markers, and a genetic diversity analysis of 205 Chinese landraces yielded 16.2 alleles per locus (Guan et al., 2010). The allele number detected in this study is higher than those reported in previous studies. Four markers (Satt156, Satt001, Satt36, and SAT1) generated 17 alleles each. The allele size ranged from 2 to 33 base pairs (bp) with an average value of 18.7 bp. The great number of alleles generated by SSR markers suggested allelic richness, a useful indicator of genetic worthiness for subsequent selection and conservation strategies (Wang et al., 2006).

SSR markers revealed marked genetic diversity among soybean genotypes. The PIC values ranged from 0.46 (GMSE0634) to 0.85 (Satt001) with an average of 0.716. High PIC values suggested a potential to detect differences among soybean genotypes. Hipparagi et al. (2017) showed an average PIC value of 0.36 with the markers, Sat554, Sat180, Sat600, and Sat478 having 4 alleles per locus each. The PIC values ranged from 0.55 to 0.66. Various studies showed PIC values ranging from 0.199 to 0.87 (Wang et al., 2006; Hisano et al., 2007; Zhang et al., 2013; Kim et al., 2014; Bisen et al., 2015). The results of this study are in line with the findings of other studies. Genotype GMSE0634 (0.5) had the lowest heterozygosity value, whereas Satt001 (0.87) had the highest value. The average heterozygosity value obtained was 0.7485 and were much higher than results reported values on soybean (Wang et al., 2006; Li et al., 2008; Liu et al., 2010; Wang et al., 2010; Zhang et al., 2013; Hipparagi et al., 2017) .

Cluster analysis

Figure 1 shows a dendrogram for the 30 soybean genotypes constructed using the unweighted pair group method with arithmetic mean (UPGMA) clustering algorithm based on 20 SSR markers. Clustering analysis aids to substantiate the results of pairwise analysis. The dendrogram divided the cultivars into two major clusters, 2.6 (a) and 0.12 (b). The closest distance (similarity) was between genotypes Santa Rosa (020) and PR 165-52 (030), whereas the greatest genetic distance (dissimilarity) was between genotypes 69 S 7 (001) and Yeluanda (015), as well as R-5-4-2 M (006) and B 66 S 31 (002). The two major clusters were further divided into two sub-clusters, 0.16 and 0.33. Sub-cluster 0.16 consisted of two sub-sub clusters, which were 2.3, having one genotype (R-5-4-2 M), and 0.10 consisting of 16 genotypes (Lee Ex RHOD (003), Oribi (016), Yeluanda (015), B 66 S 37 (024), B 66 S 41 (021), 61 S 156 (019), B 66 S 24 (026), 15/06/2012 (008), Dundee (009), Kahala (028), Egret (022), Crawford (023), Columbia M 8 A (004), B 66 S 387 (025), PR 165-52 (030), and Santa Rosa (020)). The dendrogram also showed sub-sub clusters 0.28 and 0.15, which were formed from sub-cluster 0.33. The first sub-sub cluster comprised five genotypes (IBIS (005), B 66 S 256 (027), Solar 12 (010), Egret (007), and ND 85 (017)). In addition, sub-sub cluster 0.15 had seven genotypes (69 S 7 (001), AGS 239 (018), B 66 S 8 (029), DB 1601 (014), Hawkeye USSR (011), Maksura (013), and N69-2774 (012). Recently, Bisen et al. (2015) reported two major clusters, which were further divided into two sub-groups in the analysis of 38 soybean genotypes using 16 SSR markers. In addition, Hipparagi et al. (2017) found three distinct clusters in 75 genotypes using 21 SSR markers, and Hirota et al. (2012) reported two distinct clusters. Moreover, Tantasawat et al. (2011) found four major clusters in 25 soybean genoty-

Table 1. List of accessions obtained from the Potchefstroom Agricultural Research Council Grain Crop Institute gene bank used in this study.

No.	Code	Origin/Place of collection	Accession Name	Growth Habit
1	434-001	Unknown (ARC-GCI)	69 S 7	Indeterminate
2	434-002	Unknown (ARC-GCI)	B 66 S 31	Indeterminate
3	434-003	Unknown (ARC-GCI)	Lee Ex RHOD	Indeterminate
4	434-004	China	Columbia M 8 A	Semi-determinate
5	434-005	Unknown (ARC-GCI)	IBIS	Indeterminate
6	434-006	USA	R-5-4-2 M	Indeterminate
7	434-007	Unknown (ARC-GCI)	Egret	Indeterminate
8	434-008	Unknown (ARC-GCI)	15/06/2012	Indeterminate
9	434-009	Unknown (ARC-GCI)	Dundee	Indeterminate
10	434-010	Unknown (ARC-GCI)	Solar 12	Determinate
11	434-011	USA	Hawkeye (USSR)	Indeterminate
12	434-012	USA	N69-2774	Indeterminate
13	434-013	Asia	Maksura	Determinate
14	434-014	USA	DB 1601	Determinate
15	434-015	USA	Yeluanda	Indeterminate
16	434-016	Zimbabwe	Oribi	Determinate
17	434-017	Unknown (ARC-GCI)	ND 85	Determinate
18	434-018	USA	AGS 239	Determinate
19	434-019	Unknown (ARC-GCI)	61 S 156	Determinate
20	434-020	Brazil	Santa Rosa	Indeterminate
21	434-021	Unknown (ARC-GCI)	B 66 S 41	Determinate
22	434-022	Unknown (ARC-GCI)	Egret	Determinate
23	434-023	Unknown (ARC-GCI)	Crawford	Determinate
24	434-024	Unknown (ARC-GCI)	B 66 S 37	Indeterminate
25	434-025	Unknown (ARC-GCI)	B 66 S 387	Indeterminate
26	434-026	Unknown (ARC-GCI)	B 66 S 24	Indeterminate
27	434-027	Unknown (ARC-GCI)	B 66 S 256	Indeterminate
28	434-028	USA	Kahala	Semi-determinate
29	434-029	Unknown (ARC-GCI)	B 66 S 8	Indeterminate
30	434-030	USA	PR 165-52	Indeterminate

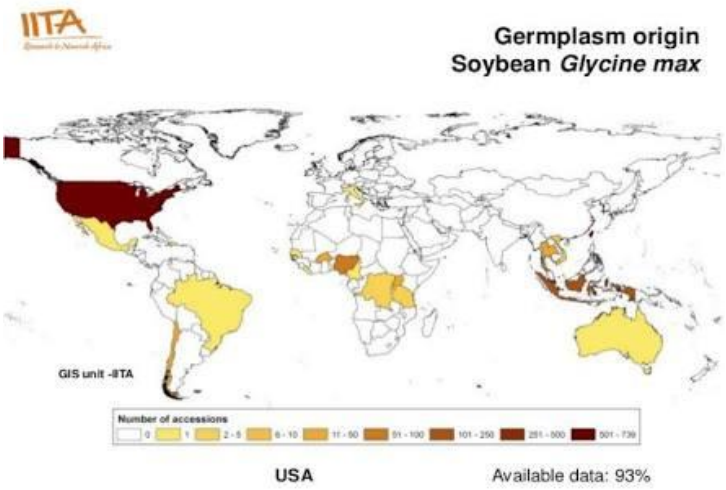

Fig 1: Germplasm collections, conservation and distribution Source: IITA (2008)

Table 2. Description of the 20 SSR markers used to analyze diversity of soybean genotypes (http://www.soybase.org).

Marker	Forward primer (5'→3)	Reverse Primer (3'→5)	Repeat Type
Satt001	AAAGTCTTTAAAAGTGTGTCTTA	TTAAAAGAAAAATGCAACAT	(ATT)2
Satt173	TGCGCCATTTATTCTTCA	AAGCGAAATCACCTCCTCT	(ATT)18
S45035	TTTGTGAACGATAGAAATTTAT	AGGGGAAAATTTTTAAAGA	
Satt373	TCCGCGAGATAAATTCGTAAAAT	GGCCAGATACCCAAGTTGTACTTGT	(ATT)21
SoyPRP1	CGTGCCAAATTACATCA	TGATGGGAACAAGTACATAA	(TAT)20
Satt534	CTCCTCCTGCGCAACAACAATA	GGGGGATCTAGGCCATGAC	(ATT)25
Satt005	TATCCTAGAGAAGAACTAAAAAA	GTCGATTAGGCTTGAAATA	(ATT)19
Satt242	GCGTTGATCAGGTCGATTTTTATTTGT	GCGAGTGCCAACTAACTACTTTTATGA	(ATT)26
Satt009	CCAACTTGAAATTACTAGAGAAA	CTTACTAGCGTATTAACCCTT	(ATT)14
Sat001	GCGGATACGACCAAAAATTGTT	GCGAACTGCGAAGATACTACCC	(AT)17
BarcSatt100	ACCTCATTTTGGCATAAA	TTGGAAAACAAGTAATAATAACA	(ATT)33
SATT9	ATTACTAGAGAAATTAGTTTA	CTTACTAGCGTATTAACCCTT	(AAT)12
Satt038	GGGAATCTTTTTTTCTTTCTATTAAGTT	GGGCATTGAAATGGTTTTAGTCA	(ATT)17
Satt002	TGTGGGTAAAATAGATAAAAAT	TCATTTTGAATCGTTGAA	(ATT)25
GMSE0634	TGGGTAGGTTTTTCAGCAATG	GCAAAGGGACCCAAAGGTAT	
Sat417	GCGAATATGGCGTTGAAAATAGTGAT	GCGACCCAGATTCTGTGCTAAGA	(AT)16
Satt156	GCGGTGTGGATCCAAAAACTCAAACTT	GCGTGCTAGTTCGATCAGCTTAGTTTC	(AT)17
Satt36	AAAGTCATAAGTGGCACTCCAAGTTT	GAACATAACAATAATAAATATAGCTC	(AT) I9
SAT1	CTGGTGGACTATTGATACGACC	AACTGCGAAGATACTACCCTCC	(AT)17
Satt43	AAATTCTGTTCATTGTCCGTC	CATTTTAATATCCCGAGTAGG	(AT)20

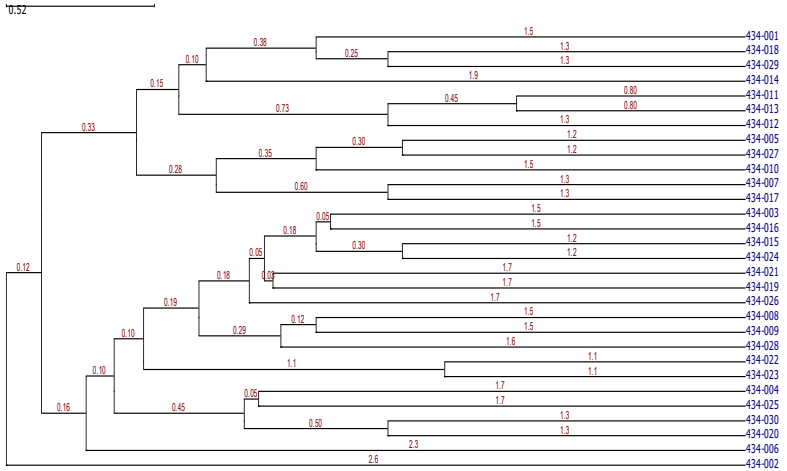

Fig 2: Dendrogram showing genetic relationships among 30 soybean genotypes evaluated using 20 SSR markers

Table 3. Genetic information of 30 soybean genotypes obtained using 20 SSR markers.

Marker	Allele number	Allele size (*bp)	*PIC	Heterozygosity
Satt001	10	2	0.85	0.87
Satt173	16	18	0.79	0.81
S45035	5		0.51	0.59
Satt373	16	21	0.80	0.82
SoyPRP1	4	20	0.65	0.70
Satt534	9	25	0.72	0.75
Satt005	11	19	0.81	0.83
Satt242	7	26	0.73	0.76
Satt009	6	14	0.65	0.70
Sat001	19	17	0.83	0.85
BarcSatt100	6	33	0.80	0.82
Satt9	5	12	0.51	0.54
Satt038	6	17	0.61	0.66
Satt002	7	25	0.71	0.75
GMSE0634	4		0.46	0.50
Sat417	13	16	0.72	0.76
Satt156	22	17	0.80	0.82
Satt36	17	19	0.82	0.84
SAT1	17	17	0.80	0.82
Satt43	16	20	0.75	0.78
Average	10.8	18.7	0.716	0.748

*PIC – Polymorphic information content *bp – base pairs

pes using SSR markers, whereas Wang et al. (2006) and Ghosh et al. (2014) observed two major clusters. A previous study on revolutionary relationship between *Glycine soja* and *Glycine max* revealed two clusters (Wen et al., 2009). The results of the present study are consistent with the findings of Wang et al. (2006), Wen et al. (2008), Hirota et al. (2012), Ghosh et al. (2014), and Bisen et al. (2015). The genotypes used in this study showed genetic variation because of the different source and/or area from which they were collected; hence, they belonged to different clusters. Farmers usually concentrate on mass production rather than maintaining gene purity, resulting in the production of new lines.

Various studies have assessed genetic diversity in soybean using SSR markers. At the Embrapa Research Institute in Brazil, Mulato et al. (2010) used SSR markers to assess genetic relationships between soybean cultivars. Twenty SSR markers and 10 EST SSR markers were used to analyze 79 soybean accessions, and results revealed high levels of genetic diversity among them. A total of 259 alleles were obtained, ranging from 2 to 21 alleles per locus, with an average of 8.63. The genotypes were assigned to five major clusters and numerous subgroups. In addition, Wang et al. (2006) assessed the genetic diversity of 129 accessions from the Chinese core collection by using 60 SSR markers, which suggested that the material was quite divergent. A total of 732 alleles were observed, and the PIC values varied from 0.05 to 0.91 with a mean of 0.23. These accessions were divided into five major clusters according to geographical origin (two clusters from the Northern ecotypes, one cluster from the Yellow River Valley ecotypes, one cluster from a mixture of the Yellow River Valley and Northern ecotypes, and one cluster from the Southern ecotypes). The accessions obtained from the Yellow River Valley exhibited the highest allelic richness and were simultaneously highly dispersed in their clustering pattern (Wang et al., 2006). These results were in line with the findings of Li et al. (2008), who reported 1,160 alleles among 1,863 landraces using 59 SSR markers and found seven clusters collected from the Yellow River Valley with a high genetic variation. These findings provided evidence that the Yellow River Valley might be the origin of the cultivated soybean. In Japan, a study was carried out to compare between 1,305 wild soybean and 53 cultivated soybean using 20 SSR markers, and 28 and 5 alleles per locus were obtained for wild and cultivated soybean, respectively, indicating that cultivated soybean had less polymorphism than its progenitor.

According to Jeffreys et al. (1994), SSR markers could produce very high allelic variations and are highly polymorphic. The use of SSRs and other molecular markers in germplasm diversity studies was investigated by Park et al. (2009), who concluded that SSRs exhibited the highest heterozygosity and genetic variation. Among 61 genotypes studied, 1 to 37 alleles were also observed. The genetic diversity of 303 accessions of *Glycine max* was analyzed using 99 SSR markers, which showed high gene diversity of 0.77 (Li et al., 2008). High levels of heterozygosity suggest a high proportion of genetic diversity, which increases selection response in breeding programs. The high level of heterozygosity observed among the genotypes used in this study suggested that they were collected from various geographical areas with different levels of selection pressure. Farmers maintain a large number of landraces on a single plot to cope with the diverse environmental conditions, resulting in a continuous exchange of genes through pollen flow (Manzelli et al., 2007; Barnaud et al., 2008). In addition, farmers exchange seeds as gifts and *via* markets to renew old seed stocks or to acquire new varieties. Consequently, there may be a continuous exchange of genes among genotypes (Mofokeng et al., 2014).

The history of soybean begins off in Southeast Asia were it was domesticated by Chinese farmers during the 1100 BC and was later grown in Japan as well as many other countries. Research studies revealed that soybean seed first reached America in 1765 when it was planted as a garden crop in Georgia around 1770. Later on the seeds were distributed to farmers in Illinois and gained popularity leading to farmers growing them as forage for their livestock. The plants performed very well under hot and humid temperatures of North Carolina causing the United States Department of Agriculture to perform tests on them and also encourage farmers to cultivate them mainly for animal feed. The valuable source of protein and oil from soybean was further discovered and also led to the realization of benefits it had on preserving the quality of soil. Soybean grew so much that the US started off with 20 proven varieties only to realize later on that there is more than 10000 varieties that encouraged agricultural scientists to study this plant and this meant new and improved varieties for better production for farmers all over the world. Soybean farming reached a peak in the United States around the 1940's while production in China came to a halt due to World War II. Studies show that there thirty one states in the U.S that are involved in soybean production and they include Indiana, Illinoi, Minnesota and Iowa being the top producers, supplying to Africa, Asia and other parts of the world. One-tenth of the volume of soybean produced in Iowa compares to that of North Carolina but as a net importer of soybean meals and soybean itself, North Carolina ranks as high as many other countries (NC Soybean Producers Association, 2018).

Materials and methods

Plant materials and study site

The experiment was conducted at the Agricultural Research Council – Grain Crops Institute, South Africa located at 26°44'43.16"S–27°04'47.71"E with an altitude of 1340 m above sea level. Thirty soybean genotypes that are sourced and maintained by the Agricultural Research Council were grown under controlled conditions in a growth chamber until the 4th leaf stage. Day and night temperatures of the growth chamber were kept constant at 29 °C, and plants were irrigated to field capacity every fourth day depending on soil moisture depletion. Two seeds of each accession were planted in a 5-L pot containing locally obtained loamy soil. This procedure was replicated twice for each genotype. The genotypes were coded as illustrated in Table 1.

DNA extraction and PCR analysis

After reaching the fourth leaf stage, leaves were harvested and kept in 1.5-mL polypropylene sample tubes prior to freeze-drying for three days. Freeze-dried leaf samples of each genotype were grounded into powder using a Qiagen tissue lyser grinder (Rogstad, 2003). DNA was extracted using the standard cetyl trimethylammonium bromide (CTAB) method (Saghai-Maroof et al., 1984). DNA quality was determined using 1 % agarose, and DNA concentration was quantified using a NanoDrop machine. DNA samples were diluted to a final concentration of 50 ng/µL using Tris-EDTA (TE) buffer. DNA was amplified by PCR using a Gene Amp PCR system 9700 thermal cycler (Applied Biosystems, Foster City, CA, USA). Twenty primers were used for diversity analysis of the soybean genotypes. Amplified products were electrophoresed on 3.5 % agarose gel and detected using ethidium bromide staining. The primers and their sequences are shown in Table 2. The alleles were quantitatively scored as present or absent. Allele sizes were estimated in comparison with a 100 bp DNA ladder.

Data analysis

Data were captured using the GGT 2.0 software (Van Berloo, 2007). The resulting fragments were analyzed, and the alleles were scored using GeneMapper® version 4.1 software (Applied Biosystems). A dissimilarity matrix was generated using DARwin 5.0 software (Perrier and Jacquemoud-Collet, 2006). The data matrices of the genetic distances were used to create the dendrogram using the UPGMA algorithm. The assay efficiency index, referred to as the PIC value, was calculated using the formula PIC = 1 - ∑fi, where fi is the frequency of the allele (Smith et al., 1997).

Conclusion

Analysis of 30 soybean genotypes using 20 SSR markers revealed that there was genetic diversity among these genotypes. All 30 genotypes were obtained from the ARC-GCI gene bank and particularly used for breeding new lines; hence, there was a need to study its genetic diversity. Three major clusters and two sub-clusters were detected. The first sub-cluster comprised 17 genotypes that were closely related, whereas the second sub-cluster included 12 genotypes that also exhibited close genetic relation. The study was carried out to aid breeders and farmers in selecting for desirable traits of the crops while omitting the undesirable ones in a breeding program. Assessment of genetic diversity is important for efficient management and protection of available genetic variability, as well as for crop improvement. The preferred method for breeding is molecular profiling because this method is authentic, reliable, and less affected by environmental changes. Diversity studies play a major role in categorizing the population into diverse groups, which results in the development of gene pool.

Acknowledgements

This project was supported by the Central University of Technology, Free State Research Grant Scheme, National Research Foundation fund, and Agricultural Research Council- Grain Crops for the use of their facilities.

References

Apuya R, Frazier NL, Keim B, Jill P, Roth E, Lark K (1988) Restriction fragment length polymorphisms as genetic markers in soybean, *Glycine max* (*L.*) Merr. Theor Appl Genet. 75:889-901.

Barnaud A, Trigueros G, McKey D, Joly HI (2008) High outcrossing rates in fields with mixed sorghum landraces: how are landraces maintained? Heredity. 101:445-452.

Bisen A, Khare D, Nair P, Tripathi N (2015) SSR analysis of 38 genotypes of soybean (*Glycine Max* (*L.*) *Merr.*) genetic diversity in India. Physiol Mol Biol Plants. 21:109-115.

Cregan PB, Jarvik T, Bush AL (1999) An integrated genetic linkage map of the soybean genome. Crop Sci. 39:1464-1490.

Devi KN, Singh LNK, Singh MS, Singh SB, Singh KK (2012) Influence of sulphur and boron fertilization on yield, quality, nutrient uptake and economics of soybean (*Glycine max*) under upland Cconditions. J Agric Sci. 4:1-10.

Ghosh J, Ghosh PD, Choudhury PR (2014) An assessment of genetic relatedness between soybean [*Glycine max* (*L*.) *Merrill*] cultivars using SSR markers. Am J Plant Sci. 5:3089-3096.

Guan R, Chang R, Li Y, Wang L, Liu Z, Qiu L (2010) Genetic diversity comparison between Chinese and Japanese soybeans (*Glycine max* (L.) Merr.) revealed by nuclear SSRs. Genet Resour Crop Evol. 57:229-242.

Hipparagi Y, Singh R, Choudhary DR, Gupta V (2017) Genetic diversity and population structure analysis of Kala bhat (*Glycine max* (*L*.) *Merrill*) genotypes using SSR markers. Hereditas. 154:9.

Hirota T, Sayama T, Yamasaki M, Sasama H, Sugimoto T, Ishimoto M, Yoshida S (2012) Diversity and population structure of black soybean landraces originating from Tanba and neighboring regions. Breed Sci. 61:593-601.

Hisano H, Sato S, Isobe S (2007) Characterisation of soybean genome using EST-derived microsatellite markers. DNA Res. 14:271-281.

Jeffreys AJ, Tamaki K, MacLeod A (1994) Complex gene conversion events in germ line mutations at human mini satellites. Nat Genet. 6:136-145.

Keim P, Schupp JM, Travis SE, Clayton K, Zhu T, Shi L, Ferreira A, Webb DM (1997) A high-density soybean genetic map based on AFLP markers. Crop Sci. 37:537-543.

Kim KS, Chirumamilla A, Hill CB, Hartman GL, Diers BW (2014) Identification and molecular mapping of two soybean aphid resistance genes in soybean PI 587732. Theor Appl Genet. 127:1251-1259.

Kuroda Y, Tomooka N, Kaga A, Wanigadeva SM, and Vaughan DA (2009) Genetic diversity of wild soybean (*Glycine soja* Sieb. *et Zucc*.) and Japanese cultivated soybeans [*Glycine max* (*L*.) *Merr*.] based on microsatellite (SSR) analysis and the selection of a core collection. Genet Resour Crop Evol. 56:1045-1055.

Li Y, Guan R, Liu Z, Ma Y, Wang L (2008) Genetic structure and diversity of cultivated soybean [*Glycine max* (L.) Merrill] landraces in China. Theor Appl Genet. 117:857-871.

Liu YL, Li YH, Zhou GA, Uzokwe N, Chang RZ, Chen SY, Qui LJ (2010) Development of soybean EST-SSR markers and their use to assess genetic diversity in the *Subgenus soja*. Agr Sci China. 9:1423-1429.

Manzelli M, Pileri L, Lacerenza N, Bendettelli S, Vecchio V (2007) Genetic diversity assessment in Somalia sorghum (*Sorghum bicolor* L. Moench) accessions using microsatellite markers. Biodiversity Conserv. 16:1715-1730.

Mofokeng MA, Shimelis H, Toongona P, Laing M (2014) A genetic diversity analysis of South African sorghum genotypes using SSR markers. S Afr J Plant Soil. 31:145-152.

Mulato BM, Möller M, Zucchi MI, Quecini V, Pinheiro JB (2010) Genetic diversity in soybean germplasm identified by SSR and EST-SSR markers. Pesqui Agropecu Bras. 45:276-283.

North Carolina Soybean Producers Association (2018) History of soybeans (online). Available at http://ncsoy.org/media-resources/history-of-soybeans/ (Accessed 18 August 2018).

Ott A, Trautschold B, Sandhu D (2011) Using microsatellites to understand the physical distribution of recombination on soybean chromosomes. PLoS One. 6:e22306.

Park YJ, Lee JK, Kim NS (2009) Simple sequence repeat polymorphisms (SSRPs) for evaluation of molecular diversity and germplasm classification of minor crops. Molecules. 14:4546-4569.

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: Implications for transferability of SSRs in plants. Mol Biol Evol. 15:1275-1287.

Perrier X, Jacquemoud-Collet JP (2006) DARwin software. Available at http://www.darwin.cirad.fr/darwin.html [accessed 10 January 2018].

Robinson AJ, Love CG, Batley J, Barker G, Edwards D (2004) Simple sequence repeat marker loci discovery using SSR primer. Bioinformatics. 20:1475-1476

Rogstad SH (2003) Plant DNA extraction using silica. Plant Mol Biol Report. 21:463.

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNAsepacer-length polymorphism in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci U S A. 81:8014-8019.

Shoemaker RC, Specht JE (1995) Integration of the soybean molecular and classical genetic linkage groups. Crop Sci. 35:436-446.

Simko I, Eujayl I, van Hintum TJ (2012) Empirical evaluation of DArT, SNP and SSR marker- systems for genotyping, clustering and assigning sugar beet hybrid varieties into populations. Plant Sci. 184:54-62.

Smith JSC, Chin ECL, Shu H, Smith OH, Wall SJ, Senior ML, Mitchell SE, Kresovich S, Ziegle J (1997) An evaluation of the utility of SSR loci as molecular makers in maize (*Zea mays* L.): comparison with data from RFLPs and pedigree. Theor Appl Genet. 95:163-173.

Tantasawat P, Trongchuen J, Prajongjai T, Jenweerawat S, Chaowiset W (2011) SSR analysis of soybean (*Glycine max* (L.) *Merr*.) genetic relationship and variety identification in Thailand. Aust J Crop Sci. 5:283-290

Van Berloo R (2007) GGT graphical genotypes. Laboratory of plant breeding. Wageningen University (online). Available at http://www.dpw.wau.nl/pv/pub/ggt/ (Accessed 23 January 2018). Wageningen Agricultural University, the Netherlands.

Vinu V, Singh N, Vasudev S, Yadava DK, Kumar S, Naresh S, Bhat SR, Prabhu KV (2013) Assessment of genetic diversity in *Brassica juncea* (*Brassicaceae*) genotypes using phenotypic differences and SSR markers. Rev Biol Trop. 61:1919-1934.

Wang L, Guan R, Zhangxiong L, Chang R, Qiu L (2006) Genetic diversity of Chinese cultivated soybean revealed by SSR markers. Crop Sci. 46:1032-1038.

Wang M, Li RZ, Yang WM, Du WJ (2010) Assessing the genetic diversity of cultivars and wild soybeans using SSR markers. Afr J Biotechnol. 9:4857-4866.

Wen ZX, Zhao TJ, Zheng YZ, Liu SH, Wang CE, Wang F, Gai JY (2008) Association analysis of agronomic and quality traits with SSR markers in *Glycine max* and *Glycine soja* in China: I. Population structure and associated markers. Acta Agron Sin. 34:1169-1178.

Zhang GW, Xu SC, Mao WH, Hu QZ and GYM (2013) Determination of the genetic diversity of vegetable soybean [*Glycine max* (*L*.) *Merr*.] using EST-SSR markers. J Zhejiang Univ Sci. 4:279-288.