MECENAS PUBLISHING

Phenotypic and molecular traits diversity in soybean launched in forty years of genetic breeding

Mário do Carmo Oda[1], Tuneo Sediyama[2], Éder Matsuo[3,*]; Cosme Damião Cruz[4], Everaldo Gonçalves de Barros[5] and Marcia Flores da Silva Ferreira[6]

[1]Testing Operation Manager (TOM), Soybean Breeding, Monsanto, Sorriso, MT, Brazil. [2]Universidade Federal de Viçosa, Departamento de Fitotecnia, Viçosa, MG, Brazil. [3]Universidade Federal de Viçosa, Instituto de Ciências Exatas e Tecnológicas, Campus de Rio Paranaíba, Rio Paranaíba, MG, Brazil. [4]Universidade Federal de Viçosa, Departamento de Biologia Geral, Viçosa, MG, Brazil. [5]Universidade Federal de Viçosa, Departamento de Biologia Geral, Centro de Ciências Biológicas e da Saúde, Viçosa, MG, Brazil. [6]Universidade Federal do Espírito Santo, Departamento de Biologia, Centro de Ciências Agrárias, Alegre, ES, Brazil. Corresponding author, E-mail: edermatsuo@ufv.br

ABSTRACT

The aim of this study was to evaluate the diversity of phenotypic and molecular traits in soybean cultivars launched in forty years of breeding. The DNA was amplified with 42 microsatellite markers (SSR). Polymorphisms of 38 SSR markers were identified in polyacrylamide gel at 10%. 106 alleles were amplified with an average of 2.52 alleles per SSR locus. Polymorphism information content varied from 0 to 0.68 with an average of 0.38. Genetic dissimilarities between pairs of cultivars varied from 0.4 to 0.6, 0.8 to 1.0 and 0.0 to 0.4 for data obtained from SSR markers, coefficient of parentage and phenotypic characters, respectively. It was possible to verify the contribution of cultivars considered old, intermediate and recent as well as the genetic variability of the group of cultivars used, which remained the same over 40 years of breeding. It was also observed that, with the combination of six microsatellite primers, it was possible to distinguish the 21 cultivars used in this study; and that microsatellite markers showed less biased estimates compared to the estimates obtained by the parentage coefficient and phenotypic characters in studies on genetic diversity.

Key words: *Glycine max*, Germplasm, Genetic variability, SSR marker, Coefficient of parentage, Phenotypic characterization.

INTRODUCTION

Soybean (*Glycine max* (L.) Merr.) is one of the most important crops in Brazilian agribusiness. The expectation for the 2013/2014 Brazilian harvest is that the domestic production reaches a record level of 85,442,500,000 tons of grains (Conab 2014), becoming increasingly realistic the projection that Brazil will become the world's largest producer of soybean in the coming years (Sediyama et al. 2009). Part of this great success is due to the genetic improvement programs of various research institutions and Brazilian universities.

Studies have shown that, despite the large number of existing varieties of soybean in Brazil, there is little genetic variation among them, mainly because they originate from a few ancestors, which results in a narrow genetic base (Hiromoto and Vello 1986; Miranda et al. 2007; Wysmierski and Vello 2013). However, there are studies in the literature indicating that the genetic diversity of the soybean germplasm used in breeding programs in Brazil has remained generally constant in recent years (Priolli et al. 2004; Bonato et al. 2006). Furthermore, it is possible to detect significant genetic variability in the Brazilian soybean germplasm even among elite cultivars when using microsatellite markers selected by their informativeness (Vieira et al. 2009). Plant breeders have measured the magnitude of genetic diversity of soybean cultivars by the coefficient of parentage, morphological traits and molecular markers (microsatellite).

The coefficient of parentage has been used to study the genetic diversity and to understand models of breeding programs of soybean (Cox et al. 1985a; Hiromoto e Vello 1986; Vello, Hiromoto and Azevedo. 1988; Gizlice, Carter and Burton 1993; Gizlice, Carter and Burton 1994; Sneller, 1994; Gizlice et al., 1996; Bharadwaj et al., 2002; Priolli et al. 2002; Fu, Peterson and Morrison 2007; Wysmierski and Vello 2013), wheat (*Triticum aestivum* L.) (Cox et al. 1985b), barley (*Hordeum vulgare* L.) (Graner, Ludwing and Melchinger 1994) and coffee (Coffea arabica L.) (Setotaw et al. 2013).

Studies of genetic diversity in soybean have been conducted using morphological characteristics (Nelson et al. 1987; Nelson et al. 1988; Almeida, Peluzio and Afférri 2011; Cunha, Hamawaki and Sousa 2013).

Although morphological and agronomic traits are useful in assessing genetic diversity, they are highly influenced by the environment, and the collection of this information requires a great amount of time. DNA markers are attractive alternatives because they are virtually unlimited in number. They also have a high degree of polymorphism, independence between the environmental effects and the physiological state of the plant and can be arranged within the linkage map.

The microsatellite markers, also called SSR (Simple Sequence Reapeat), enable widespread use in breeding programs, since they are co-dominant, multi-allelic and capable of providing a high level of genetic information per locus (Lanza, Schuster and Guimarães 2000). In addition, they have proven to be an excellent tool for assessing genetic distance between individuals, for cultivar identification and for pedigree analysis (Priolli et al. 2002). There are several studies using molecular markers to estimate the genetic diversity of different characteristics among accessions (Rongwen et al. 1995; Doldi, Vollmann and Lelley 1997; Priolli et al. 2004; Fu, Peterson and Morrison 2007; Vieira et al. 2009; Guan et al. 2010; Velusamy, Toan and Park 2013; Bizari et al. 2014; Dong et al. 2014).

Studies on genetic diversity not only serve as a basis for understanding the genetic basis of soybean of different gene pools, but they also help to identify new sources of genes to increase the productivity and quality of the soybean (Fu, Peterson and Morrison 2007). This is because the success of any breeding program

depends on the complete knowledge and understanding of the diversity of the available germplasm (Setotaw et al. 2013).

Thus, the objective of this study was to evaluate the genetic contribution of old cultivars to recent cultivars that show some ancestry, to select a set of SSR primers able to distinguish 21 cultivars, to obtain information on the diversity character to be used in improvement programs, to conduct associations between estimated diversity based on phenotypic traits, genealogy and molecular markers and to estimate their genetic distance.

MATERIAL AND METHODS

Genetic material

The soybean cultivars used in this study are listed in Table 1. They are cultivars from different breeding programs, launched in the market at different times and adapted to different regions of Brazil and the world. During the selection process of these cultivars, their genealogy and launch period were used as criteria. Akin cultivars with different launch periods were selected, which had been through 40 years of soybean breeding, i.e., the difference between the oldest and the newest/most recent launch period.

DNA extraction

DNA extraction from 21 cultivars was obtained from a bulk of ten seeds taken from each cultivar through a maceration process. One sample from each of the ten seeds described by McDonald at al. (1994) was used, with some modifications.

Extraction buffer (700 µl) containing Tris-HCl 0.2 M (pH 7.5), NaCl 0.28 M, EDTA 0.25 mM and SDS 10% was added to Eppendorf tubes of 1.5 ml containing approximately 50 mg of ground seeds of each analysis unit. Samples were macerated and then centrifuged for 10 min at 14,000 rpm. The supernatants were transferred to new tubes with the addition of 10 µl of proteinase K (10 mg/ml) and 10 µl of CaCl2 1 mM and placed in water bath at 55 °C for 1.5 hours. Subsequently, 900 µl of isopropanol were added to the samples and left to stand for 2 min. Then, they were centrifuged for 10 min at 14,000 rpm. The supernatants were discarded and the pellets were washed once with ethanol 70% and a then again with alcohol 90%. After the washings, the precipitates were placed to dry for 15 min at room temperature. Subsequently, they were re-suspended in TE (Tris HCl 10 mM. EDTA1 mM. pH 8.0) containing 60 µg/ml of RNAse A and placed in a water bath for one hour. The samples were again precipitated by the addition of 900 µl of isopropanol and left for precipitation for two minutes. Immediately after, they were again centrifuged for 10 min at 14,000 rpm and the supernatants were discarded. The precipitates that were formed were resuspended in the end, in TE (Tris HCl 10 mM. EDTA 1 mM. pH 8.0). DNA quality was estimated by spectrophotometry, considering the A260/A280 ratio. The concentration was estimated from the absorbance at 260 nm, according to Sambrook, Fritsch and Maniatis (1989).

A part of each one of the samples was stored, undiluted, at -20 °C, for later use and the work solution was diluted to the concentration of 10 ng/µL and stored at 5 °C.

SSR markers and amplification conditions. The 42 microsatellite markers used in this study (Table 2) were selected based on information from articles and theses on polymorphisms in soybean. The primers used in the characterization of cultivars were synthesized by Invitrogen Life Technologies, GibcoBRL and MWG-Biotech.

The microsatellite reactions were performed in micro tubes of 0.2 ml at a total reaction volume of 15 µL, containing PCR buffer (100 mM of Tris-HCl and 500 mM of KCl, pH 8.0), 2.5 mM of each deoxyribonucleotide (dATP, dTTP, dGTP e dCTP), 20 mM of MgCl2, 6 µM of each primer (forward and reverse), a unit of Taq polymerase enzyme and 40 ng of template DNA.

The amplifications were carried out in Perkin-Elmer thermal cyclers (GeneAmp PCR System 9600) using a touch down program. This program consisted of one denaturation cycle at 94 °C for 4 min, an annealing stage at 65 °C for 40 seconds, followed by 10 touchdown cycles decreasing 1o C every cycle to 55 °C, followed by 30 cycles of 55 °C for 40 seconds each. In each cycle, the respective denaturation (94 °C / 40 seconds) and polymerization (72 °C / 1 minute) temperatures were maintained. The final stage of the program consisted of one cycle of polymerization at 72 °C for 7 minutes. The separation of DNA fragments, which were amplified in PCR, was performed in a 10% polyacrylamide gel.

Parentage coefficient estimation. Parentage coefficient was estimated between the 21 cultivars used in this study, combining them in pairs, and using a total of 210 combinations of cultivars. Some assumptions have been adopted to calculate the coefficient of parentage (CP): (1) ancestral cultivars were considered unrelated - CP = 0; (2) cultivar derived from simple crossover receives half of its genes from each parent - CP = 0.5; (3) all parents were considered homozygous and homogeneous; (4) parents whose genealogy is not known were considered uncorrelated - CP = 0; (5) considered CP = 1 between a cultivar and itself; (6) considered CP = 0.75 between a cultivar and another obtained from this selection; (7) considered CP = 0.56 between two selections obtained from the same cultivar (Bowman, May and Calhoun 1997).

Using the parentage coefficient of each formed pair, the genetic contribution of each cultivar in relation to a group of cultivars formed according to the release date was calculated. Moreover, with the data of parentage coefficient decreased by one, the dissimilarity matrix was obtained.

Data analysis

The genetic diversity of each microsatellite locus was obtained from the allele frequency using the following formula:

$$Polymorphism\ information\ content\ (PIC) = 1 - j = 1 - \sum_{j=1}^{n} p_{ij}^2$$

where p is the frequency of the jth allele for the primer i (Anderson et al. 1993). The value of the genetic diversity of the locus is similar to the heterozygosity which is commonly used to describe the informativeness of a molecular marker in outcrossing plants.

The genetic distances between cultivars obtained through information generated by microsatellite markers (SSR) were evaluated from a dissimilarity matrix constructed using the complement of the similarity index (SI) for codominates/multi allelic variables. The scores 0, 1 and 2 were used for the absent allele in heterozygous and homozygous, respectively. The index was obtained by dividing the total number of microsatellite loci containing common alleles by the total number of analyzed loci.

The genetic distances obtained by phenotypic values were calculated considering the characteristics as multicategoric, giving scores for each scale of variation of the considered phenotypic information. In addition, genetic distances were obtained considering the coefficient of parentage, as described above.

The methods used to perform the cluster analysis based on the dissimilarity matrix obtained by SSR markers, coefficient of parentage and phenotypic characteristics were the UPGMA (unweighted pair-group mean average) and the Tocher optimization method. The Genes Program (Programa Genes) was used during the analyses (Cruz 2013).

RESULTS AND DISCUSSION

Of the 42 evaluated markers, 4 were monomorphic and 38 showed polymorphism among the 42 analyzed soybean cultivars. 106 alleles were amplified with an average of 2.52 alleles per SSR locus (Table 3). The primers that showed higher polymorphism

Table 1. Soybean cultivars, genealogy and some phenotypic characteristics.

Number	Origin	Cultivars	Genealogy*	Color of flower	Color of hilum	Color of pubescence	Growth habit	Cicle
1	USA	Improved Pelican	PI 548461	Purple	Brown	Brown	Undetermined	Average
2	UFV	Viçoja	D492491 x I. Pelican	Purple	Brown	Brown	Determined	Average
3	UFV	UFV-1	Selection in Viçoja	Purple	Brown	Brown	Determined	Tardy
4	USA	Davis	D 49-2573 X N 45-1497	White	Brown	Grey	Determined	Precocious
5	MONSOY	FT-12 (Nissei)	FT 9510 x Prata	White	Brown	Brown	Determined	Average
6	IAC	IAC-8	F5 strain obtained from the crossing Bragg x E 70 -51	Purple	Brown	Brown	Determined	Average
7	MONSOY	FT-Cristalina	Natural crossing in UFV-1	Purple	Brown	Grey	Determined	Tardy
8	Embrapa	Doko	F7 progeny. Obtained from the population RB 72-1, derived from six crossovers (E 70-46 x Pickett. E 70-47 x F 65-1376 e Davis x IAC 79308)	White	Black	Brown	Determined	Tardy
9	UFV	UFV-17	FT - 12 x IAC-8	Purple	Brown	Brown	Determined	Tardy
10	UFV	UFV-19	FT - 12 x IAC-8	White	Black	Brown	Determined	Average
11	UFV	UFV-18	FT-Cristalina x IAC-8	Purple	Black	Brown	Determined	Tardy
12	UFV	UFVS-2007	FT-Cristalina x Doko	Purple	Brown	Brown	Determined	Tardy
13	Embrapa	BRS - Tuiuiu	FT-Cristalina (4) x Doko	Purple	Brown	Brown	Determined	Tardy
14	MONSOY	FT - Cristalina RCH	FT-Cristalina (5) x Embrapa-20	Purple	Brown	Grey	Determined	Tardy
15	Embrapa	Conquista	Lo76-4484 x Numbaira	Purple	Black	Brown	Determined	Average
16	Embrapa	Valiosa RR	Conquista (n) x RR	Purple	Black	Brown	Determined	Average
17	Embrapa	Santa Rosa	D49-772 x La41-1219	White	Brown	Brown	Determined	Average
18	USA	Bragg	Jackson x D49-2491	White	Black	Brown	Determined	Precocious
19	Embrapa	BR-IAC-21	IAC-8(6) x FT-Cristalina	Purple	Black	Brown	Determined	Average
20	UFV	UFV-10 (Uberaba)	Santa Rosa x UFV-1	Purple	Brown	Brown	Determined	Tardy
21	UFV	UFVS-2301	[FT-Cristalina (6) x Doko] x FT-72285	White	Black	Brown	Determined	Tardy

* Information obtained from the literature and supplemented with personal information.

were: Satt263, Satt192, Satt070 and Sct_189 with 5, 4, 4 and 4 alleles per locus, respectively. Priolli et al. (2002) in a study on the characterization of 186 Brazilian soybean cultivars found an average number of alleles per locus of 5.3. Yamanaka et al. (2007) using 12 pairs of microsatellite primers in a study on the genetic relationship between cultivars of Brazilian, Chinese and Japanese soybeans had 82 alleles for a group of 272 cultivars,with an average of 6.83 alleles per locus. And, Vieira et al. (2009) obtained 124 alleles, with an average of 2.34 per locus, using 53 SSR markers on 53 cultivars widely cultivated and used in Brazilian breeding programs.

The polymorphism information content (PIC) calculated to estimate the informativeness of each microsatellite locus ranged from 0 to 0.68, with an average of 0.38 (Table 3). Studies conducted by Rongwen et al. (1995), using seven microsatellite primers to characterize 96 genotypes, showed genetic diversity from 0.71 to 0.95. The authors linked this high diversity with the use of the input material and other species (Glycine max and Glycine soja). Narvel et al. (2000), working with 74 SSR primers to estimate the diversity of some accessions and elite soybean varieties, found PIC values ranging from 0 to -0.84, with an average of 0.56 in the approach used and from 0 to 0.79 with an average of 0.50 in elite varieties. This, according to the authors, served as a good indication that genotype groups with a narrow genetic basis show lower genetic diversity. Priolli et al. (2004) in a study to estimate the genetic diversity between periods and between breeding programs in Brazil found an average genetic diversity value of 0.63 for different improvement programs. Vieira et al. (2009) obtained a PIC value between 0.16 and 0.66, with an average of 0.47. Velusamy, Toan and Park (2013) obtained from 178 accessions of soybean collected in Korea and submitted to 9 SSR markers, PIC ranging from 0.7447 (Satt423) to 0.8585 (Satt155) with an average equal to 0.8040. And, Dong et al. (2014) obtained from 100 accessions of vegetable soybean (Edamame) in

Table 2. Microsatellite primers used in the study.

Number	Loci	Repeated Structure	Group Connection	Number	Loci	Repeated Structure	Group Connection
1	Satt521	(ATT)12	N	22	Satt077	(ATT)12	D1a
2	Satt526	(ATT)9	A1	23	Satt335	(ATT)12	F
3	Satt531	(ATT)14	D1a	24	Satt215	(ATT)11	J
4	Satt180	(ATT)16	C1	25	satt263	(ATT)19	E
5	Satt181	(ATT)18	H	26	Satt492	(ATT)15	O
6	Satt302	(ATT)12	H	27	Satt211	(ATT)10	A1
7	Satt102	(ATT)11	K	28	Satt242	(ATT)26	K
8	Satt571	(ATT)14	I	29	Satt285	(ATT)19	J
9	Satt417	(ATT)18	K	30	Satt215	(ATT)11	J
10	Satt108			31	Satt309	(ATT)13	G
11	Satt237	(ATT)17	N	32	Satt487	(ATT)22	O
12	Satt192	(ATT)32	H	33	Satt182	(ATT)17	L
13	Satt070	(ATT)24	B2	34	Satt177	(ATT)16	A2
14	Satt200	(ATT)17	A1	35	Sct_189	(CT)17	I
15	Satt336	(ATT)14	M	36	Satt406	(ATT)31	J
16	Satt464	(ATT)16	D2	37	Satt173	(ATT)18	O
17	Satt191	(ATT)18	G	38	Satt256	(ATT)10	D2
18	Satt079	(ATT)13	C2	39	Satt001	(ATT)25	K
19	Satt100	(ATT)33	C2	40	Satt113		
20	Satt304	(ATT)29	B2	41	Satt170	(ATT)10	C2
21	Satt142	(ATT)21	H	42	Satt146	(ATT)17	F

Table 3. Number of alleles per microsatellite locus and polymorphic information content (PIC) for 21 soybean genotypes released over 40 years of improvement.

Number	Loci	Number of alleles	PIC	Number	Loci	Number of alleles	PIC
1	Satt521	2	0.3744	22	Satt077	3	0.4065
2	Satt526	2	0.3658	23	Satt335	3	0.5578
3	Satt531	2	0.0866	24	Satt215	3	0.5313
4	Satt180	2	0.3119	25	satt263	5	0.6808
5	Satt181	2	0.3457	26	Satt492	2	0.3698
6	Satt302	2	0.3648	27	Satt211	2	0.2149
7	Satt102	2	0.1575	28	Satt242	3	0.4898
8	Satt571	3	0.2051	29	Satt285	2	0.3290
9	Satt417	3	0.3267	30	Satt215	3	0.5313
10	Satt108	3	0.4783	31	Satt309	2	0.3457
11	Satt237	3	0.5400	32	Satt487	3	0.3586
12	Satt192	4	0.5915	33	Satt182	2	0.3515
13	Satt070	4	0.5854	34	Satt177	3	0.5594
14	Satt200	3	0.5594	35	Sct_189	4	0.5471
15	Satt336	3	0.5668	36	Satt406	3	0.5564
16	Satt464	2	0.3739	37	Satt173	3	0.5439
17	Satt191	3	0.3092	38	Satt256	2	0.3687
18	Satt079	3	0.5547	39	Satt001	1	0
19	Satt100	3	0.4236	40	Satt113	1	0
20	Satt304	2	0.3374	41	Satt170	1	0
21	Satt142	2	0.3515	42	Satt146	1	0

China by means of analysis using 53 SSR markers, PIC values in the magnitude of 0.071 to 0.831, with an average of 0.573, and 38 markers showed PIC values greater than 0.6.

With six primers (Satt192, Satt263, Satt070, Satt100, Satt108 and Satt215 of the 42) it was possible to differentiate 21 soybean cultivars (Table 4). Cultivars UFV-19 and BR-IAC-21 showed 2 alleles in the Satt215 and Satt192 locus, respectively, which could be an indication that the varieties are not an inbred line or there was an impurity in the samples.

To get information on genetic diversity, dissimilarity matrices were obtained from information of SSR markers, genealogy and phenotypic characters. The data were compiled and are presented in Figure 1. Dissimilarity values indicate how a pair of cultivars is different in genetic content. When analyzing the dissimilarity matrix obtained from molecular data, it was observed that the calculated genetic dissimilarity ranged from 0.0769 to 0.7631. The lowest dissimilarity value was obtained by the cultivars Viçoja and UFV-, which is consistent with the genealogies (UFV-1 is a selection of Viçoja). The pair of cultivars with the highest genetic distance was UFV-1 and Conquista (Conquista presents no family relationship with the cultivar UFV-1, which was attested by the coefficient of dissimilarity of the pair of cultivars).

Figure 1A shows that the dissimilarities between the pair of cultivars obtained by microsatellite markers are about 0.4 to 0.6. When using the coefficient of parentage (Figure 1B), dissimilarity values are around 0.8 to 1.0 and when using phenotypic traits information (Figure 1C) they are around 0.0 to 0.4. Therefore, for the group of cultivars used in the study, the greater divergence between them was detected by using the coefficient of parentage. Probably the genealogy of each cultivar is overestimating the values of genetic dissimilarity, where many papers state that the genetic basis of soybean is narrow.

Results from other studies have also detected greater genetic diversity by using the coefficient of parentage. Bertini (2004), for instance, found greater genetic diversity using coefficients of parentage when working with cotton. By evaluating and characterizing 100 accessions of an active soybean germplasm bank using microsatellite markers, morphological evaluation and pedigree information, Alcântara Neto (2005) concluded that the coefficient of parentage provided larger groups, however, it can be biased when compared with other methods of evaluation.

With the purpose of facilitating the interpretation of dissimilarity results using microsatellite markers, a table with dissimilarity mean values was created (Table 5). First, the 21 cultivars were grouped into three groups (older, intermediate and recent cultivars) using the time when each cultivar was developed as a criterion. Next, the average dissimilarity for each cultivar in each group was obtained.

It can be inferred that by comparing cultivars from the old group with itself, the intermediate and recent groups, there was a slight decrease in the dissimilarity mean value over the years, which may indicate a narrowing of the genetic base. However, by observing the dissimilarity mean values of recent cultivars group between themselves and the other two groups, it can be seen that the Conquista cultivar, considered recent, shows relatively high dissimilarity mean value, indicating that the cultivar should be included in the blocks of crosses to generate an important dissimilarity in a soybean improvement program. Genetic variability

Table 4. Differentiation of 21 soybean cultivars based on six microsatellite markers.

Numbers	Genotypes	Locus[1]					
		Satt263	Satt192	Satt070	Satt100	Satt108	Satt215
1	Improved Pelican	A	A	A	A	A	A
2	Viçoja	B	A	*	*	B	B
3	UFV - 1	B	A	A	B	B	B
4	Davis	C	B	B	B	C	B
5	FT - 12 Nissei	B	B	C	B	B	C
6	IAC - 8	D	C	D	C	A	B
7	FT - Cristalina	B	A	A	B	C	B
8	Doko	E	C	D	B	A	C
9	UFV - 17	D	B	C	C	A	C
10	UFV - 19	B	B	D	B	B	B/C
11	UFV - 18	D	C	A	C	B	B
12	UFVS - 2007	B	A	A	B	C	C
13	BRS - Tuiuiu	B	C	A	B	B	B
14	FT -Cristalina RCH	B	A	D	*	B	C
15	Conquista	A	B	C	A	A	A
16	Valiosa RR	C	B	*	B	B	B
17	Santa Rosa	D	C	D	C	B	B
18	Bragg	E	C	D	B	B	C
19	BR-IAC- 21	D	B/C	D	B	B	C
20	UFV-10	A	A	D	B	B	A
21	UFVS-2301	B	B	D	B	B	*

* Lost data.

[1] Those cultivars followed by the same letter in the column have alleles in common. The letters are assigned in descending order according to the size of the alleles, namely, the letter A refers to the larger allele and E to the smaller allele for the locus Satt263. The same goes for the other loci.

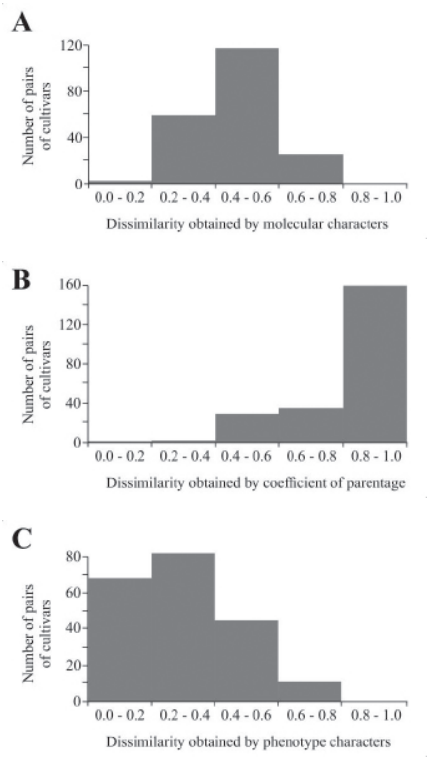

Figure 1. Graphical distribution of genetic distances for 210 pairs of cultivars assessed by microsatellite markers (A), parentage coefficient (B), and phenotypic characteristics (C).

Table 5. Average dissimilarity between soybean groups (old, intermediate and newer/most recent) based on microsatellite markers.

	Cultivars	Old	Intermediate	Newer
Old	Improved Pelican	0.6792	0.6566	0.6495
	Bragg	0.6151	0.5792	0.6198
	Viçoja	0.5242	0.6170	0.5571
	UFV-1	**0.5019**	**0.5830**	**0.5259**
	Davis	0.6830	0.6000	0.5764
	Santa Rosa	0.6358	0.5453	0.5717
	Average	**0.6065**	**0.5969**	**0.5834**
Intermediate	UFV-10	0.6116	0.6439	0.6052
	FT-12 (nissei)	0.6509	0.6391	0.5604
	IAC-8	0.5786	0.6108	0.5660
	FT-Cristalina	0.5377	0.6557	0.5137
	Doko	0.6006	0.6157	0.6066
	Average	**0.5959**	**0.6330**	**0.5704**
Newer	UFV-17	0.6085	0.5500	0.5949
	UFV-18	0.5204	0.5547	0.5241
	UFV-19	0.5629	0.4849	0.5220
	BR-IAC-21	0.6061	0.5632	0.5681
	UFVS-2007	0.5079	0.5283	0.5278
	BRSMS-Tuiuiu	0.5000	0.5302	0.5441
	FT-Cristalina RCH	0.5645	0.5415	0.5026
	Valiosa RR	0.5723	0.6170	0.6195
	Conquista	0.7170	0.6698	0.7248
	UFVS-2301	0.6745	0.6641	0.6247
	Average	**0.5834**	**0.5704**	**0.5753**

has remained the same for over 40 years of breeding in the group of cultivars used in this study.

By analyzing 44 publications, through meta-analysis, regarding the 20th Century diversity trend for eight different crops (including soybeans), Van der Wouw et al. (2010) reported that, in general, trends were not identified by pointing out to loss of genetic diversity in cultivars released by breeders in the last century. However, they do not exclude the possibility of diversity loss in cultures and/or specific regions. Moreover, as the possible techniques in plant breeding have advanced rapidly, it is unclear what will happen in the future with the level of diversity in cultures (Van der Worw et al. 2010).

Through the analysis of a group of 184 soybean cultivars developed by public and private companies in Brazil, by means of 12 SSR markers, Priolli et al. (2004) concluded that the Brazilian soybean germplasm kept constant genetic variability over the last 30 years of culture expansion and breeding. Furthermore, studies by Vieira et al. (2009) showed that it is possible to detect significant variability in the evaluated Brazilian soybean germplasm, even among elite cultivars, when using microsatellite markers

Table 6. Groups obtained for the 21 soybean cultivars by the method of average distances (UPGMA) based on the dissimilarity measures calculated using information from microsatellites, parentage coefficient and phenotypic characters.

Microsatellite Markers	
Group	Cultivars
1	FT-Cristalina, UFVS-2007, BRS-Tuiuiú, UFV-18, FT-Cristalina RCH, Viçoja and UFV-1
2	Davis
3	ValiosaRR
4	BR-IAC-21
5	UFV-10
6	IAC-8, Santa Rosa, UFV-17 and UFV-19
7	Doko and Bragg
8	FT-12 Nissei
9	UFVS-2301
10	Improved Pelican and Conquista
Parentage Coefficient	
Group	Cultivars
1	Doko and UFVS-2301
2	Viçoja, UFV-10, UFV-1, Bragg, FT-Cristalina RCH, FT-12 Nissei, Conquista, Improved Pelican, BR-IAC-21, FT-Cristalina and Santa Rosa
3	IAC-8 and UFVS-2007
4	Davis. ValiosaRR and UFV-19
5	UFV-17
6	UFV-18
7	BRS-Tuiuiú
Phenotypic Characters	
Groups	Cultivars
1	Doko, UFV-19, Bragg and UFVS-2301
2	FT-Cristalina, FT-Cristalina RCH, BRS-Tuiuiú, UFV-10
3	UFV-1, UFV-17, UFV-18 and UFVS-2007
4	FT-12 Nissei, ValiosaRR and Santa Rosa
5	Viçoja. IAC-8, Conquista and BR-IAC-21
6	Improved Pelican
7	Davis

selected for their informativeness. Furthermore, there is still enough genetic variability in the Brazilian soybean germplasm to be exploited by breeding programs (Vieira et al. 2009).

Fu, Peterson and Morrison (2007) when analyzing 45 soybean cultivars from Canada, released from 1934 to 2001 and 37 accessions of exotic germplasm through 37 SSR markers, concluded that they maintained large genetic diversity. This is revealed, according to the authors, by the fact that the cultivars released after 1990 presented a little more diversity when compared to those taken before 1970.

From the dissimilarity measures using information from microsatellite markers, genealogy and phenotypic traits, dendrograms were obtained using the UPGMA method. To facilitate the analysis of dendrograms, they were compiled in a table form (Table 6). Based on the Tocher optimization method, eight mutually exclusive groups from molecular data, 15 groups using information of parentage coefficient and three groups using phenotypic characters were formed (Table 7).

Cultivars with the same genealogy have not been grouped using information of parentage coefficient, which is the case of UFV-17 and UFV-19 (Table 8). On the other hand, by using molecular information, they are placed in the same group, while according to their phenotypic characteristics they should not be in the same group since they present flower color, hilum color

Table 7. Groups obtained for the 21 soybean cultivars by Tocher grouping based on the dissimilarity measures using information from microsatellites, parentage coefficient and phenotypic characters.

---------- *Microsatellite Markers* ----------	
Groups	Cultivars
1	Viçoja, UFV-1, BRS-Tuiuiú, UFVS-2007, FT-Cristalina, FT- Cristalina RCH, UFV-18 and Davis
2	IAC-8, Santa Rosa, UFV-19 and UFV-17
3	Improved Pelican and Conquista
4	Doko and Bragg
5	FT-12 Nissei and BR-IAC-21
6	UFVS-2301
7	UFV-10
8	Valiosa RR
---------- *Parentage Coefficient* ----------	
Groups	Cultivars
1	Improved Pelican, Viçoja, UFV-1, Davis, FT-12 Nissei, FT-Cristalina and Doko
2	FT-Cristalina RCH
3	IAC-8
4	UFV-17
5	UFV-19
6	UFV-18
7	BRS-Tuiuiú
8	Santa Rosa
9	UFV-10
10	UFVS-2301
11	UFVS-2007
12	Conquista
13	Valiosa RR
14	Bragg
15	BR-IAC-21
---------- *Phenotypic Characters* ----------	
Groups	Cultivars
1	Viçoja, Improved Pelican, UFV-1, IAC-8, UFV-17, UFVS-2007, BRS-Tuiuiú, UFV-10, FT-Cristalina, FT-Cristalina RCH, UFV-18, Conquista, ValiosaRR and BR-IAC-21
2	FT-12 Nissei, Santa Rosa, UFV-19, Doko, UFVS-2301 and Bragg
3	Davis

Table 8. Comparison of the groups, obtained by the UPGMA method and Tocher grouping, between the assessments by microsatellites, parentage coefficient and phenotypic characters for 21 soybean cultivars[1].

---------- *UPGMA Method* ----------			
Groups	Microsatellites	Parentage coeficiente	Phenotypic characters
1	7, 12, 13, 11, 14, 2 and 3	8 and 21	8, 10, 18 and 21
2	4	2, 20, 3, 18, 14, 5, 15, 1, 19, 7 and 17	7, 14, 13 and 20
3	16	6 and 12	3, 9, 10 and 12
4	19	4, 16 and 10	5, 16 and 17
5	20	9	2, 6, 15 and 19
6	6, 17, 9 and 10	11	1
7	8 and 18	13	4
8	5		
9	21		
10	1 and 15		
---------- *Tocher Grouping Method* ----------			
Groups	Microsatellites	Parentage coeficiente	Phenotypic characters
1	2, 3, 13, 12, 7, 14, 11 and 4	1, 2, 3, 4, 5, 7 and 8	2, 6, 1, 3, 9, 12, 13, 20, 7, 14, 11, 15, 16 and 19
2	6, 17, 9 and 10	14	5, 17, 10, 8, 21 and 18
3	1 and 15	6	4
4	8 and 18	9	
5	5 and 19	10	
6	21	11	
7	20	13	
8	16	17	
9		20	
10		21	
11		12	
12		15	
13		16	
14		18	
15		19	

[1] Cultivars: 1 = Improved Pelican, 2 = Viçoja, 3 = UFV-1, 4 = Davis, 5 = FT-12 (Nissei), 6 = IAC-8, 7 = FT-Cristalina, 8 = Doko, 9 = UFV-17, 10 = UFV-19, 11 = UFV-18, 12 = UFVS-2007, 13 = BRS – Tuiuiu, 14 = FT - Cristalina RCH, 15 = Conquista, 16 = Valiosa RR, 17 = Santa Rosa, 18 = Bragg, 19 = BR-IAC-21, 20 = UFV-10 (Uberaba) and 21 = UFVS-2301.

and different cycles. It was found that the FT-Cristalina and FT-Cristalina RCH cultivars are always in the same group regardless of the evaluation method. The same goes for the Viçoja and UFV-1 cultivars, which used only molecular and coefficient of parentage information. Despite the fact that the Conquista and Viçosa cultivars did not show any parentage with other cultivars and that the Valiosa RR is originated from Conquista, they were not grouped in any of the evaluation methods used.

In the clustering obtained by the Tocher optimization method from information on parentage coefficient, the formation of 15 mutually exclusive groups was detected, despite the presence of some cultivars with common parental. As for the phenotypic characters, the formation of only three groups was observed. UFV-17 and UFV-19 cultivars are present in different groups using parentage coefficient and in the same group using molecular characteristics. The justification for cultivars with parents in common to be allocated to different groups by the coefficient of parentage is that the estimated coefficient of parentage assumes that the parents are not related, which leads to the fact that the similarity and dissimilarity estimates obtained by the parentage coefficient are quite biased.

By using microsatellite markers information, a group of cultivars with common parents and without parents in common was detected since microsatellite markers are inherited in a co-dominant manner and the estimates generated by them are more informative and less biased than estimates generated by coefficient of parentage and phenotypic characters. Moreover, the estimates generated by the microsatellites take into account the effects of selection practiced over generations of breeding.

CONCLUSIONS

It was possible to verify the contribution of cultivars considered as old cultivars for the cultivars considered as intermediate and newer, stating that within the group of cultivars using the genetic variability remained the same over nearly 40 years of improvement.

As for the newer/recent cultivars, there is still genetic variability useful to soybean genetic improvement.

With the combination of the 6 microsatellite primers Satt192, Satt263, Satt070, Satt100, Satt108 and Satt215 it was possible to distinguish the 21 cultivars that were used in this study.

For genetic diversity studies, the microsatellite markers were those that showed less biased estimates when compared with estimates obtained by the coefficient of parentage and phenotypic characters.

ACKNOWLEDGMENTS

The authors wish to thank the National Council for Scientific and Technological Development (CNPq) and the Coordination for Improvement of Higher Education Personnel (CAPES) for granting fellowships on Master and productivity research.

REFERENCES

Alcântara Neto F (2005) Caracterização genético-molecular de um banco ativo de germoplasma de soja. Thesis, Universidade Federal de Viçosa.

Almeida RD, Peluzio JM and Afférri FS (2011) Divergência genética entre cultivares de soja, sob condições de várzea irrigada, no sul do Estado Tocantins. Revista Ciência Agronômica 42: 108-115.

Anderson JA, Churchill GA, Autrique JE, Tanksley SD and Sorrels ME (1993) Optimizing parental selection for genetic linkage maps. Genome 36: 181-186.

Bertini CHM (2004) Caracterização e análise de diversidade genética em algodoeiro herbáceo por marcadores microssatélites e genealogia. Thesis, Universidade Federal de Viçosa.

Bharadwaj CH, Satyavathi CT, Tiwari SP and Karmakar PG (2002) Genetic base of soybean (*Glycine max*) varieties released in India as revealed by coefficient of parentage. Indian Journal of Agricultural Sciences 72: 467-469.

Bizari EH, Unêda-Trevisoli SH, Vianna VF, Meyer AS and Di Mauro AO (2014) Genetic diversity in early-maturing soybean genotypes based on biometric and molecular parameters. Journal of Food, Agriculture & Environment 12: 259-265.

Bonato ALV, Calvo ES, Geraldi IO and Arias CAA (2006) Genetic similarity among soybean (*Glycine max* (L.) Merrill) cultivars released in Brazil using AFLP markers. Genetics and Molecular Biology 29: 692-704.

Bowman DT, May OL and Calhoun DS (1997) Coefficient of parentage for 260 cotton cultivars released between 1970 and 1990, USDA Technical Bulletin 1852.

CONAB – Companhia Nacional de Abastecimento. 2014. Acompanhamento de safra brasileira: grãos, sexto levantamento, março 2014. Disponível em: <http://www.conab.gov.br> Acesso em: 25 de março de 2014.

Cox TS, Kiang YT, Gorman M and Rodgers DM (1985a) Relationship between coefficient of parentage and genetic similarity indices in the soybean. Crop Science 25: 529-532.

Cox TS, Lookhar GL, Walker DE, Harrell LG, Albers LD and Odgers DM (1985b) Genetic relationship among hard red winter wheat cultivars as evaluated by pedigree analysis and gliadin polyacrylamide gel electrophoretic patterns. Crop Science 25: 1058-1063.

Cruz CD (2013) GENES - a software package for analysis in experimental statistics and quantitative genetics. Acta Scientiarum. Agronomy 35: 271-276.

Cunha MCG, Hamawaki OT and Sousa LB (2013) Genetic variability among 79 soybean progenies from ufu breeding program. Bioscience Journal 29:340-349.

Doldi ML, Vollmann J and Lelley T (1997) Genetic diversity in soybean as determined by RAPD and microsatellite analysis. Plant Breeding 116: 331-335.

Dong D, Fu X, Yuan F, Chen P, Zhu S, Li B, Yang Q, Yu X and Zhu D (2014) Genetic diversity and population structure of vegetable soybean (*Glycine max* (L.) Merr.) in China as revealed by SSR markers. Genetic Resources and Crop Evolution 61: 173-183.

Fu YB, Peterson GW and Morrison MJ (2007) Genetic Diversity of Canadian Soybean Cultivars and Exotic Germplasm Revealed by Simple Sequence Repeat Markers. Crop Science 47: 1947-1954.

Gizlice Z, Carter JT and Burton JW (1993) Genetic diversity in North American soybean: I – Multivariate analysis of founding stock and relation to coefficient of parentage. Crop Science 33: 614-620.

Gizlice Z, Carter JT and Burton JW (1994) Genetic base for North American public soybean cultivars released between 1947 and 1988. Crop Science 34: 1143-1151.

Gizlice Z, Carter JT, Gerig T and Burton JW (1996) Genetic diversity patterns in North Americans public soybean cultivars based on coefficient of parentage. Crop Science 36: 753-765.

Graner A, Ludwing W and Melchinger A (1994) Relationships among European barley germplasm: Comparison RFLP and

pedigree data. Crop Science 34: 1199-1205.

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

Hiromoto DM and Vello NA (1986) The genetic base of Brazilian soybean (*Glycine max* (L.) Merrill) cultivars. Brazilian Journal of Genetics 9: 295-306.

Lanza MA, Schuster I and Guimarães CT (2000) Aplicação de marcadores moleculares no melhoramento genético. Informe Agropecuário 21: 97-108.

McDonald MB, Elliot LJ, Burr B and Cantrell RG (1994) DNA extraction from dry seeds for RAPD analysis in varietal identification studies. Seed Science and Technology 22: 171-176.

Miranda ZFS, Arias CAA, Prete CEC, Kiihl RAS, Almeida LAA, Toledo JFF and Destro D (2007) Genetic characterization of ninety elite soybean cultivars using coefficient of parentage. Pesquisa Agropecuária Brasileira 42: 363-396.

Narvel JM, Fehr WR, Chu WS, Grant D and Shoemaker RC (2000) Simple sequence repeat diversity among soybean plant introductions and elite genotypes. Crop Science 40: 1452-1458.

Nelson RL, Amdor PJ, Orf JH and Cavins JF (1988) Evaluation of the USDA soybean germplasm collection: Maturity groups 000 to IV (PI427136 to PI445845). USDA Technical Bulletin 1726.

Nelson RL, Amdor PJ, Orf JH, Lambert JW, Cavins JF, Kleiman R, Laviolette FA and Athow KA (1987) Evaluation of the USDA soybean germplasm collection: Maturity groups 000 to IV (PI273483 to PI427107). USDA Technical Bulletin 1718.

Priolli RHG, Mendes-Junior CT, Arantes CE and Contel EPB (2002) Characterization of Brazilian soybean cultivars using microsatellite markers. Genetics and Molecular Biology 25:185-193.

Priolli RHG, Mendes-Junior CT, Sousa SMB, Sousa NEA and Contel EPB (2004) Diversidade genética da soja entre períodos e entre programas de melhoramento no Brasil. Pesquisa Agropecuária Brasileira 39: 967-975.

Rongwen J, Akkaya MS, Bhagwat AA, Lavi U and Cregan PB (1995) The use of microssatelite DNA markers for soybean genotype identification. Theoretical and Applied Genetics 20: 43-48.

Sambrook J, Fritsch EF and Maniatis T (1989) Molecular cloning: a laboratory manual. 2.ed. v.3, Cold Spring Harbor Laboratory Prees, New York.

Sediyama T, Teixeira RC and Barros HB (2009) Origem, evolução e importância econômica. In: Sediyama T (eds) Tecnologias de produção e usos da soja. Mecenas, Londrina. p. 1-5.

Setotaw TA, Caixeta ET, Pereira AA, Oliveira ACB, Cruz CD, Zambolim EM, Zambolim L and Sakiyama NS (2013) Coefficient of Parentage in *Coffea arabica* L. Cultivars Grown in Brazil. Crop Science 53: 1237-1247.

Sneller CH (1994) Pedigree analysis of elite soybean lines. Crop Science 34: 1515-1522.

Van der Wouw M, Van Hintum T, Kik C, van Treuren R and Visser B (2010) Genetic diversity trends in twentieth century crop cultivars: a meta analysis. Theoretical and Applied Genetics 120: 1241-1252

Vello NA, Hiromoto DM and Azevedo Filho AJBV (1988) Coefficient of parentage and breeding of Brazilian soybean germplasm. Revista Brasileira de Genética 11: 679-697.

Velusamy V, Toan VD and Park E (2013) Evaluation of genetic diversity in wild soybean (*Glycine soja* Sieb. and Zucc.) accessions based on SSR markers. International journal of Agronomy and Plant Production 4: 1241-1246.

Vieira ESN, Schuster I, Silva RB and Oliveira MAR (2009) Variabilidade genética em cultivares de soja determinada com marcadores microssatélites em gel de agarose. Pesquisa Agropecuária Brasileira 44: 1460-1466.

Wysmierski PT and Vello NA (2013) The genetic base of Brazilian soybean cultivars: evolution over time and breeding implications. Genetics and Molecular Biology 36: 547-555.

Yamanaka N, Sato H, Yang Z, Xu DH, Catelli LL, Binneck E, Arias CAA, Abdelnoor RV and Nepomuceno AL (2007) Genetic relationships between Chinese, Japanese, and Brazilian soybean gene pools revealed by simple sequence repeat (SSR) markers. Genetics and Molecular Biology 30: 85-88.

Received: January 16, 2015.

Accepted: February 18, 2015.

Published: April 02, 2015.