Evaluation of genetic diversity among soybean (*Glycine max*) genotypes using univariate and multivariate analysis

M.M. Oliveira[1], L.B. Sousa[2], M.C. Reis[2], E.G. Silva Junior[2], D.B.O. Cardoso[1], O.T. Hamawaki[2] and A.P.O. Nogueira[3]

[1]Programa de Pós-Graduação em Agronomia, Instituto de Ciências Agrárias, Universidade Federal de Uberlândia, Uberlândia, MG, Brasil
[2]Instituto de Ciências Agrárias, Universidade Federal de Uberlândia, Uberlândia, MG, Brasil
[3]Instituto de Genética e Bioquímica, Universidade Federal de Uberlândia, Uberlândia, MG, Brasil

Corresponding author: M.M. Oliveira
E-mail: morony91@gmail.com

ABSTRACT. The genetic diversity study has paramount importance in breeding programs; hence, it allows selection and choice of the parental genetic divergence, which have the agronomic traits desired by the breeder. This study aimed to characterize the genetic divergence between 24 soybean genotypes through their agronomic traits, using multivariate clustering methods to select the potential genitors for the promising hybrid combinations. Six agronomic traits evaluated were number of days to flowering and maturity, plant height at flowering and maturity, insertion height of the first pod, and yield. The genetic divergence evaluated by multivariate analysis that esteemed first the Mahalanobis' generalized distance (D^2), then the clustering using Tocher's optimization methods, and then the unweighted pair group method with arithmetic average (UPGMA). Tocher's optimization

method and the UPGMA agreed with the groups' constitution between each other, the formation of eight distinct groups according Tocher's method and seven distinct groups using UPGMA. The trait number of days for flowering (45.66%) was the most efficient to explain dissimilarity between genotypes, and must be one of the main traits considered by the breeder in the moment of genitors choice in soybean-breeding programs. The genetic variability allowed the identification of dissimilar genotypes and with superior performances. The hybridizations UFU 18 x UFUS CARAJÁS, UFU 15 x UFU 13, and UFU 13 x UFUS CARAJÁS are promising to obtain superior segregating populations, which enable the development of more productive genotypes.

Key words: *Glycine max*; Soybean; Genetic dissimilarity; Clustering analysis

INTRODUCTION

The soybean, *Glycine max* (L.) Merrill, is the most important oilseed crop grown worldwide. It has been domesticated for five thousand years by the Chinese people, is an important grain for human consumption and animal feeding with good percentage of proteins (40%), lipids (20%), and essential in animal nutrition (Bellaver et al., 2002).

Soybean is the most cultivated culture in Brazil with a planted area at about 33 million hectares in 2015/2016, producing great volumes of grains that in this same season exceeded 95 million tons, what happens in all regions of the country (CONAB, 2016).

Soybean productivity increased significantly since its introduction in Brazil, around 1882, in Bahia (Câmera, 2012), due to the culture expansion, more modern plantation and management techniques, which has been generating an increasing need to introduce cultivars adapted to the different edaphoclimatic conditions of the several soybean producing regions in the country (Anselmo et al., 2011).

The soybean breeding is crucial in the process of new cultivar development, and is responsible for studying and developing genotypes resistant to different undesirable factors and with desirable agronomic traits (EMBRAPA, 2007). Knowing the genetic variability degree between different genotypes is essential for the efficiency of breeding programs (Rabbani et al., 2010).

The genetic diversity study allows a better understanding of the evolutionary relationships between the accessions and develops strategies to incorporate genes of interest in the progenitors, which will be the segregating population (Pervaiz et al., 2010). For getting these populations, which enclosing a huge genetic variability, the choice of genitors with high degree of dissimilarity to be the parental of the hybridization process is vital for its success (Cruz et al., 2011).

The clustering multivariate technique of analysis selects the individuals indirectly by characteristics of interest evaluations (Hair et al., 2009). The clustering methods of analysis let individual segregation in homogeneous groups, maximizing the heterogeneity between them (Cruz et al., 2014).

The clustering analysis application is necessary to choose a coefficient that quantifies how much two individuals are dissimilar between each other (Ferraudo, 2014). The Mahalanobis' distance is the most used measure to esteem this statistical parameter,

thereforeallows the experimental analysis with repetitions (Silva, 2012). The generated matrix should use non-hierarchical clustering methods (Tocher) or hierarchical clustering methods (unweighted pair group method with arithmetic mean, UPGMA) in accordance with the similarities and the breeder's choice (Ferraudo, 2014).

The hierarchical method UPGMA esteems the distance between two groups, defined as the average of distances between all pairs of value in a group related to the other group. The results of these procedures are in a diagram of two dimensions known as dendrogram, in which the objects clustered in accordance with the similarities (Ferraudo, 2014; Leite, 2015). The Tocher's non-hierarchical method differs from the hierarchical for not supplying a graphical representation of the segregation of groups, therefore the number of groupings formed is determined before the analysis, thereby minimizing the variability inside the group and maximizing it between them (Hair et al., 2009).

This study characterized the genetic divergence between 24 soybean genotypes through their agronomic traits, using clustering multivariate methods to select the potential genitors for the promising hybrid combinations.

MATERIAL AND METHODS

The experiment was carried out at Fazenda Capim Branco, a research station of Universidade Federal de Uberlândia, between the geographical coordinates 18°52'S; 48°20'W and 805 m of altitude, in Uberlândia, Minas Gerais, during 2011/2012. The soil characterization was Latossolo Vermelho-Amarelo (Brazilian classification system). The preparation of soil was based on a conventional tillage, with one plowing and two harrowings.

Twenty-four soybean genotypes were evaluated, and of these, 17 were lineages proceeding from the Soybean Breeding Program of Universidade Federal de Uberlândia, and 7 were commercial cultivar witnesses: EMGOPA 316, UFUS GUARANI, UFUS RIQUEZA, UFUS TIKUNA, M-SOY6101, M-SOY8001, and UFUS CARAJÁS. The experimental design was the randomized complete block with three replications. The plots consisted of four rows of soybean plants 5 m long with row spacing of 0.5 m. The plot floor areas were the two central rows, discarding the soybean plants to 0.5 m from the borders of each row.

The seed rate was 20 seeds per linear meter planted 2.5 cm in depth. The thinning occurred 7 days after seedling emergence, with the goal to establish a plant density of 10 plants per linear meter, adopting the population density of 200,000 plants per hectare. Carrying out the soil sampling of the area allowed the chemical fertilization. The control of weed, pest, and disease was made in accordance with the monitoring area.

The agronomic performance of soybean genotypes evaluated was according to visual observation and characteristic mensuration of five randomly plants at the plot floor area as the stages of development for the crop. The following traits evaluated were:

a) Number of days to flowering (NDF): corresponds to the number of days from seedling emergence to flowering. According to the phenological staging scale of Fehr and Caviness (1977), NDF is the phenological stage symbolized by R1 and defined as 50% of the plants with an open flower on the main stem.

b) Number of days to maturity (NDM): considered the number of days between emergences to the grain physiological maturity in the field, defined by 50% of plants with 95% of mature pods, symbolized as the phenological stage R8, according to phenological staging scale of Fehr and Caviness (1977).

c) Plant height at flowering (PHF): at the beginning of flowering (R1 stage; Fehr and Caviness, 1977), five plants were evaluated at random in each plot, measuring the plant height with a metric ruler help, which consists in the height between the soil surface and the main stem apex of each plant.

d) Plant height at maturity (PHM): in physiological maturity stage of the grains (R8 stage; Fehr and Caviness, 1977), five plants at random were measured with a metric ruler help, considering the distance between the soil surface and the main stem apex of each plant.

e) Insertion height of the first pod (IHFP): stands for the distance (cm) from the soil surface to the insertion point of first pod on the plant main stem. This trait was evaluated in five plants at random in each plot with a metric ruler help.

f) Yield: the plants of plot floor area were harvested and threshed manually. The mass of the harvested and threshed grains was determined with an electronic precision balance help. The grain mass value was converted, according to the previous results, to esteem the productivity in kg/ha of each genotype, and adjusting the seed mass to 13% moisture.

After the tabulation of data, the averages were submitted to analysis of variance (F-test and significant at 0.05) to investigate the genetic variability. Then, the Scott-Knott test clustered the averages, also at the 0.05 level of significance.

The average of the traits evaluated resulted in the matrix of Mahalanobis' generalized distance (D^2). Thereafter, clustering method of optimization (Rao, 1952) and UPGMA (Cruz et al., 2014) were applied.

The criterion of Singh (1981) quantified the relative contribution of traits in the dissimilarity among genotypes. Statistical analyzes were conducted in the GENES software (Cruz, 2013).

RESULTS

The results demonstrated significant differences between averages of the genotypes for all the traits, at the probability level of 5% on the F-test, which indicates the genetic variability (Table 1).

Table 1. Analysis of variance results according to six agronomic traits evaluated in 24 soybean genotypes in Uberlândia-MG, during 2011/2012.

GV	d.f.	Mean squares					
		NDF	NDM	PHF (cm)	PHM (cm)	IHFP (cm)	Y (kg/ha)
Blocks	2	13.62	21.29	39.42	29.12	3.09	611,212.66
Genotypes	23	64.88*	160.70*	164.24*	234.12*	16.49*	1,336,926.98*
Error	46	4.90	30.61	55.01	76.26	7.71	662,394.34
Average	-	65.62	140.04	50.02	60.52	10.67	3,256.5
CV (%)	-	3.37	3.95	14.82	14.42	26.03	24.99

*Significant 5% of probability for the F-test. GV: genetic variation; d.f.: degrees of freedom; NDF: number of days to flowering; NDM: number of days to maturity; PHF: plant height at flowering; PHM: plant height at maturity; IHFP: insertion height of the first pod; Y: yield; CV: coefficient of variation.

The coefficient of variation oscillated from 3.37 (NDF) to 26.03 (IHFP), which demonstrates an acceptable experimental precision, since the traits evaluated have quantitative nature and have been influenced by environmental conditions.

In Table 2, the Scott-Knott test results shows that for the trait NDF, the lineage UFU 7 and the witnesses M-SOY6101, EMGOPA 316, and M-SOY8001 presented earliness, with

averages below 57.67 days. Of these cultivars that presented lower NDF, UFU 7 and EMGOPA 316, also showed lower values than the number of days, and therefore, were early, with 119 days for both genotypes. However, M-SOY 6101 and M-SOY8001 did not demonstrate this relation, had 148 and 137.67 days, respectively, and were characterized as average and logging more long-lived results among genotypes. The lineages UFU 3, UFU 5, UFU 18, and witnesses UFUS GUARANI and UFUS RIQUEZA had late flowering, with averages above of 69 days and had also kept the biggest values in NDM, in other words, growing during for more time and, is characterized, therefore, as late.

Table 2. Averages of the evaluated agronomic traits in 24 soybean genotypes, in Uberlândia, Minas Gerais, during 2011/2012.

Genotypes	NDF	NDM	PHF (cm)	PHM (cm)	IHFP (cm)	Y (kg/ha)
UFUS TIKUNA	68.67[a]	144.67[a]	45.13[b]	47.07[b]	8.53[b]	2770.67[b]
UFU 2	68.67[a]	142.67[a]	45.87[b]	61.13[a]	9.87[b]	3173.33[a]
UFU 3	69.33[a]	141.00[a]	56.73[a]	68.40[a]	11.73[b]	2862.67[b]
UFU 4	68.67[a]	142.00[a]	49.80[b]	59.73[b]	11.87[b]	3197.33[a]
UFU 5	69.33[a]	143.67[a]	46.80[b]	50.87[b]	8.40[b]	2518.67[b]
UFUS CARAJÁS	68.67[a]	142.00[a]	50.40[b]	52.73[b]	8.80[b]	4296.00[a]
UFU 7	56.00[d]	119.00[c]	42.80[b]	57.13[b]	9.20[b]	2120.00[b]
UFU 8	64.67[b]	142.33[a]	58.87[a]	69.00[a]	8.33[b]	3792.00[a]
UFU 9	68.67[a]	141.00[a]	63.40[a]	67.07[a]	11.33[b]	3349.33[a]
UFU 10	68.67[a]	141.00[a]	55.53[a]	74.80[a]	10.33[b]	4021.33[a]
UFU 11	62.00[c]	139.00[a]	45.00[b]	50.23[b]	10.20[b]	3176.00[a]
UFU 12	68.67[a]	146.33[a]	44.13[b]	57.93[b]	9.53[b]	2550.67[b]
UFU 13	64.67[b]	138.33[a]	47.53[b]	65.00[a]	17.00[a]	4408.00[a]
UFU 14	64.00[b]	145.00[a]	60.53[a]	73.80[a]	11.73[b]	3456.00[a]
UFU 15	61.67[c]	131.00[b]	45.40[b]	52.00[b]	9.73[b]	3700.00[a]
UFU 16	68.67[a]	140.33[a]	59.20[a]	65.40[a]	14.60[a]	3510.67[a]
UFU 17	66.67[b]	144.00[a]	56.67[a]	64.33[a]	13.93[a]	3754.67[a]
UFU 18	69.33[a]	140.33[a]	57.27[a]	73.13[a]	14.73[a]	4108.00[a]
UFU 19	67.33[a]	144.67[a]	58.00[a]	71.53[a]	9.93[b]	3476.00[a]
M-SOY6101	57.67[d]	148.00[a]	42.23[b]	47.93[b]	9.60[b]	3252.00[a]
EMGOPA 316	56.00[d]	119.00[c]	40.47[b]	54.40[b]	8.33[b]	1925.33[b]
UFUS GUARANI	69.33[a]	145.00[a]	49.27[b]	64.07[a]	11.00[b]	2822.67[b]
UFUS RIQUEZA	70.00[a]	143.00[a]	43.00[b]	56.13[b]	9.27[b]	2341.33[b]
M-SOY8001	57.67[d]	137.67[a]	36.67[b]	48.73[b]	8.13[b]	3573.33[a]
General average	65.62	140.04	50.02	60.52	10.67	3256.50

Averages followed by equal letters belong to the same group according the Scott-Knott test at 0.05 of significance. Number of days to flowering (NDF) and maturity (NDM), plant height at flowering (PHF) and maturity (PHM), insertion height of the first pod (IHFP), and Yield (Y).

For PHF, the averages varied from 36.67 cm (M-SOY8001) to 63.40 cm (UFU 9). There was a standard in the PHF with PHM groups according to the observations, where the genotypes that demonstrated higher averages in PHF, also had the same result in PHM. The opposite was shown true for lower averages, the exceptions were UFU 2, UFU 13, and UFUS GUARANI, which for PHF stayed in less height group and in maturity they grew more vigorously, so it continued in the group with higher height average.

For PHM, the lineages UFU 10, UFU 14, and UFU 18 demonstrated higher averages with 74.80, 73.80, and 73.13 cm, respectively. The less averages of PHM were of the witnesses UFUS TIKUNA, M-SOY6101, and M-SOY8001 with 47.07, 47.93, and 48.73 cm, respectively.

When evaluating the IHFP, the genotypes presented averages between 10 and 15 cm, among the acceptable minimum interval, according to Silva et al. (2010). The exceptions were

the genotypes UFUS TIKUNA (8.53 cm), UFU 5 (8.40 cm), UFUS CARAJÁS (8.80 cm), UFU 8 (8.33 cm), EMGOPA 316 (8.33 cm), and M-SOY8001 (8.13 cm), whose averages are below the minimum demanded. The IHFP group, UFU 13, UFU 16, UFU 17, and UFU 18 got the higher averages with heights of 17.0, 14.6, 13.93, and 14.73 cm, respectively, demonstrating that this trait influenced their production, since these genotypes presented the best productivities. The genotype UFUS 13 had the highest insertion height of the first pod (17 cm) and the highest productivity with 4408.00 kg/ha.

The genotypes UFUS CARAJÁS, UFU 10, UFU 13, and UFU 18 registered the best productive incomes, above of 4.000 kg/ha. The lower productivity averages were shown by genotypes EMGOPA 316 with 1925.33 kg/ha and UFU 7 with 2120.00 kg/ha. As a result, these genotypes were also the ones which have obtained the lowest average for both NDF and NDM traits. The productivity average variability observed presented values between 1925.33 and 4408.00 kg/ha.

The clustering for Tocher's optimization method considered the formation of eight groups (Table 3). The constituents of each formed group are similar. In accordance with Tocher's method, Group I with the genotypes that presented lower values of productivity averages, below 2.200 kg/ha, showed earlier maturity, on average 119 days. Group II, formed by genotypes with huge variation on the traits evaluated, presented productivity above and below the average, for example, UFU 19 (3476.00 kg/ha) and UFUS RIQUEZA (2341.33 kg/ha). Groups III, IV, and VIII showed good productivity averages: 3601.11, 3624.00 and 4296.00 kg/ha, respectively. Group V, represented by the genotype UFU 15, had productivity above the general average, 3700.00 kg/ha, cycle of 131 days, and values below the general average. Group VII, represented by the genotype UFU 13, presented the higher productivity average with 4408.00 kg/ha and IHFP with 17.00 cm.

Table 3. Twenty-four soybean genotypes clustering by Tocher's method based on the Mahalanobis (D^2) dissimilarity measure, in Uberlândia, Minas Gerais, during 2011/2012.

Group	Genotypes
I	UFU 7, EMGOPA 316
II	UFU 2, UFU 4, UFUS GUARANI, UFUS RIQUEZA, UFU 12, UFU 5, UFUS TIKUNA, UFU 19
III	UFU 16, UFU 18, UFU 3, UFU 10, UFU 9, UFU 17
IV	UFU 8, UFU 14
V	UFU 11, UFU 15
VI	M-SOY6101, M-SOY8001
VII	UFU 13
VIII	UFUS CARAJÁS

Table 4 shows the dissimilarity matrix between the soybean genotypes based on the Mahalanobis (D^2) general distance. The minimum distance occurred between lineages UFU 7 and witness EMGOPA 316 ($D^2 = 0.25$), and the maximum distance occurred between the lineage UFU 9 and the witness M-SOY6101 ($D^2 = 69.14$), therefore, these genotypes presented the highest degrees of divergence. See values in bold in Table 4.

The group determination according the hierarchical method, UPGMA, with cut at 20% dissimilarity (Figure 1) formed seven distinct groups.

The relative contribution of the traits according to Singh method (1981) showed that the NDF was the highest in expressivity, with the genetic divergence between the genotypes of 45.66% (Table 5). The yield (in kg/ha) contributed little for the genetic divergence, with

Table 4. Genetic dissimilarity measurement between 24 soybean genotypes using Mahalanobis generalized distance (D^2), Uberlândia, Minas Gerais, during 2011/2012.

Genotypes	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24
G1																								
G2	5.00																							
G3	14.69	5.58																						
G4	4.89	1.13	4.04																					
G5	1.17	2.87	8.61	2.94																				
G6	6.25	10.88	19.05	8.89	8.43																			
G7	54.16	50.19	53.35	50.61	53.05	50.29																		
G8	12.29	9.59	11.54	9.01	10.98	9.43	28.47																	
G9	15.00	10.27	3.32	6.52	10.39	12.80	47.88	7.34																
G10	15.89	5.06	3.10	4.93	11.09	14.34	47.45	7.37	4.94															
G11	10.36	15.61	28.67	15.18	14.47	11.58	26.93	9.96	25.05	24.89														
G12	5.25	1.95	9.76	4.00	3.51	17.54	59.91	15.76	17.24	11.96	18.31													
G13	18.35	11.95	19.42	9.63	19.78	16.49	47.21	16.44	20.95	14.80	13.94	17.35												
G14	14.45	8.07	8.56	7.16	12.26	17.45	35.95	3.84	8.94	7.43	12.93	11.26	11.63											
G15	19.69	22.94	32.39	21.56	22.77	11.99	16.22	10.06	24.48	25.08	5.07	31.58	18.58	18.88										
G16	13.68	7.78	3.21	3.49	10.18	12.83	53.15	10.77	2.47	5.08	23.33	13.83	11.78	8.53	25.52									
G17	8.25	5.30	7.62	2.37	7.77	9.01	48.17	7.28	7.02	7.30	12.13	9.10	5.61	4.30	18.29	3.15								
G18	19.17	7.95	3.77	5.09	14.64	17.27	58.96	14.01	5.86	2.97	29.83	14.98	10.97	10.48	31.30	2.00	5.65							
G19	10.38	3.49	2.88	3.19	6.74	12.83	43.98	4.01	4.15	2.13	17.66	7.43	14.43	2.69	22.13	4.91	4.23	5.74						
G20	34.66	46.18	71.70	48.00	44.44	42.15	54.41	39.13	**69.14**	65.95	13.52	42.72	37.92	38.29	29.15	64.38	40.48	72.93	51.17					
G21	53.43	50.66	55.42	51.73	52.70	50.83	**0.25**	29.71	50.34	49.68	26.47	59.60	49.46	37.77	16.53	55.91	50.11	62.17	45.59	52.33				
G22	8.33	1.41	3.78	2.08	4.64	17.74	59.91	13.90	10.62	6.24	22.90	1.75	15.51	8.78	33.12	7.54	6.60	7.48	3.98	54.97	60.77			
G23	6.10	1.72	6.63	3.41	2.81	17.34	59.94	16.63	14.06	9.72	22.65	1.20	20.25	13.72	33.00	11.71	10.55	12.63	7.19	54.70	59.80	1.33		
G24	32.61	40.21	63.63	42.78	40.69	32.85	30.36	29.20	59.41	53.75	8.27	42.41	31.93	34.40	13.49	57.21	37.50	63.49	44.49	6.37	28.57	52.20	50.77	

G1: UFUS TIKUNA; G2: UFU 2; G3: UFU 3; G4: UFU 4; G5: UFU 5; G6: UFUS CARAJÁS; G7: UFU 7; G8: UFU 8; G9: UFU 9; G10: UFU 10; G11: UFU 11; G12: UFU 12; G13: UFU 13; G14: UFU 14; G15: UFU 15; G16: UFU 16; G17: UFU 17; G18: UFU 18; G19: UFU 19; G20: M-SOY6101; G21: EMGOPA 316; G22: UFUS RIQUEZA, and G24: M-SOY8001.

approximately 5.74%, even so presented huge variation in the genotype averages (1905.33 and 4408.00 kg/ha).
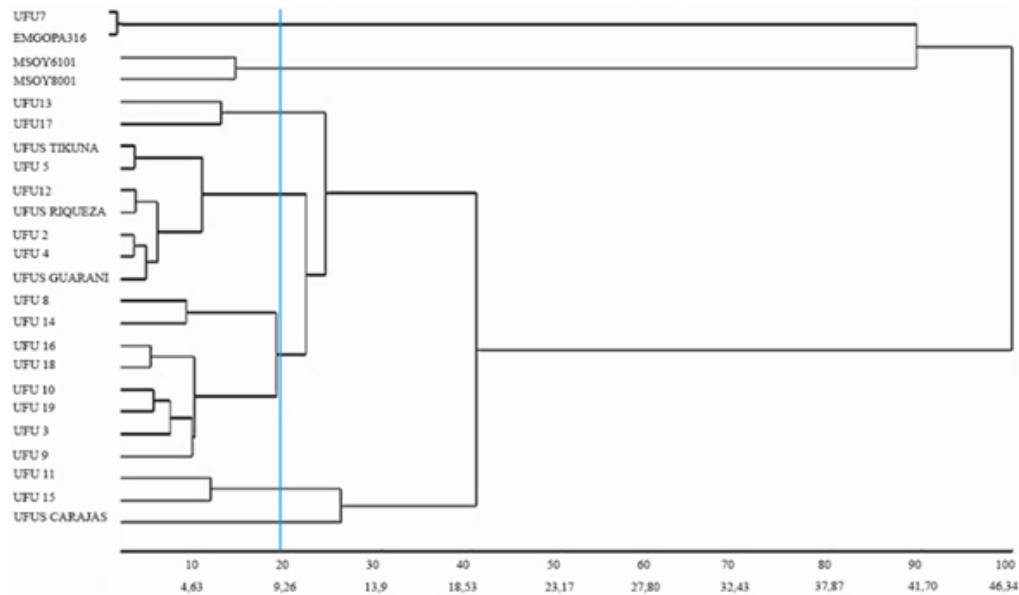
Figure 1. Dendrogram of the hierarchical clustering method of median distance (UPGMA) from the Mahalanobis's generalized distance (D²) for 24 soybean genotypes in Uberlândia, Minas Gerais, during 2011/212.

Table 5. Relative contribution of the traits for genetic divergence among the 24 soybean genotypes, according Singh method (1981). Uberlândia, Minas Gerais, during 2011/2012.

Traits	Relative contribution (%)
Number of days to flowering	45.66
Number of days to maturity	19.05
Plant height at flowering	7.88
Plant height at maturity	13.97
Insertion height of the first pod	7.69
Yield	5.74

DISCUSSION

The soybean cultivars demand photoperiod minimum of 45-58 days to reach reproductive stage (Almeida et al., 2013). The NDF average was 65.62 days. Other researches, who evaluated soybean genotypes in similar edaphoclimatic conditions, obtained NDF average of 46.5 (Sousa, 2011), 49.62 (Polizel et al., 2011), and 47.43 days (Almeida et al., 2013).

The NDM general average was of 140.04 days. Sousa (2013) evaluated the same soybean genotypes and showed NDM average of 125.10 days, as well as Oliveira et al. (2014) in Ituverava-SP, who showed average of 121.66 days. Adopting the criteria preconized in this research (EMBRAPA, 2011), the classification of materials are according to their cycle. Thus, UFU 7 and EMGOPA 316 genotypes are classified as semi-early cycle (116-125 days),

UFU 15 as median cycle (126-137 days) and all the 21 remaining genotypes as median-late cycle (138-150 days), noting that this classification is valid for the edaphoclimatic condition evaluated.

The productivity average was of 3256.5 kg/ha, which is below the productivity average obtained by Sousa (2013) that was of 4251.30 kg/ha with variation of 2186.67 and 5884. 44 kg/ha, in Uberlândia-MG. Oliveira et al. (2014) also got a variation from 2083.00 to 4068.88 kg/ha in Ituverava, São Paulo, with some genotypes of the soybean breeding program of Universidade Federal de Uberlândia. It is worth mentioning that environment has extreme influence on yield expression, because it is a trait of quantitative nature.

The Tocher's optimization and the hierarchical UPGMA method were coherent among each other. The difference in the group number formation, eight groups according Tocher's method and seven according UPGMA, is also told by other researchers and it have occurred due to the clustering methodology of each method (Sousa, 2013).

The most relevant trait for the unfolding of dissimilarity between the evaluated genotypes was the NDF (45.66%), followed for the NDM (19.05%) and PHM (13.97%). Almeida et al. (2013) identified that among the evaluated traits in their study, the NDM (39.49%), the weight of 100 seeds (26.56%) and the NDF (13.59%) had higher representation to explain the dissimilarity between the cultivars. Oliveira et al. (2014) observed the grain yield was the trait of highest expressivity for genetic divergence (19.37%) and the NDF was the lowest (12.08%). The relative contribution depended on the genotypes studied; therefore, the characterization of dissimilarity based on the averages of their traits.

The Soybean Breeding Program of Universidade Federal de Uberlândia sought cultivars with high yield. This study allowed the identification of four potential genotypes: UFU 13, UFU 15, UFU 18, and UFUS CARAJÁS, and six possible combinations between them (UFU 13 x UFU 15; UFU 13 x UFU 18; UFU 13 x UFUS CARAJÁS; UFU 15 x UFU 18; UFU 15 x UFUS CARAJÁS; UFU 18 x UFUS CARAJÁS). Assuming that the crossings should be between divergent parents, there were only three parental crosses indicated, which are UFU 18 x UFUS CARAJÁS, UFU 15 x UFU 13, and UFU 13 x UFUS CARAJÁS.

The genotypes evaluated presented genetic diversity in the traits: NDF and NDM, PHF and PHM, YHFP, and yield. Coherence was evident through the clustering methods used in the formation of groups. The trait NDF was the one that most contributed for the detention of genetic divergence among the 24 genotypes evaluated. Aiming at high yield and genetic variability, the following crossings between the genotypes recommended were UFU 18 x UFUS CARAJÁS, UFU 15 x UFU 13, and UFU 13 x UFUS CARAJÁS, UFUS 10 x UFUS CARAJÁS, UFUS 10 x UFUS 15, and UFUS 10 x UFUS 13.

ACKNOWLEDGMENTS

REFERENCES

Almeida FA, Bruscke EL, Polizel AC, Petter FA, et al. (2013). Desempenho agronômico de linhagens e cultivares de soja frente a doenças foliares. *Rev. Cienc. Agrar. (Belem.)* 56: 88-94. http://dx.doi.org/10.4322/rca.2013.014

Anselmo JL, Andrade JAC, Lazarini E, Costa DS, et al. (2011). Estabilidade e adaptabilidade de cultivares transgênicas e convencionais de soja, na região dos Chapadões. *Científica* 39: 69-78.

Bellaver C, et al. (2002). Soja integral: processamento e uso. *Alimentação Animal* 7: 28-30.

Camara GMS (2012). Introdução ao agronegócio soja. USP/ESALQ - Departamento de Produção Vegetal - novembro/2011, São Paulo.

CONAB (Companhia Nacional de Abastecimento) (2016). Acompanhamento da safra brasileira de grãos, safra 2015/2016, oitavo levantamento. CONAB, Brasília.

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

Cruz CD, Ferreira FM and Pessoni LA (2011). Biometria aplicada ao estudo da diversidade genética. Suprema, Viçosa.

Cruz CD, Carneiro PC and Regazzi AJ (2014). Modelos biométricos aplicados ao melhoramento genético. 3 th edn. UFV, Viçosa.

EMBRAPA (2007). Tecnologias, de produção de soja. Região Central do Brasil. Embrapa Soja, Londrina.

EMBRAPA (Empresa Brasileira de Pesquisa Agropecuária) (2011). Tecnologias de Produção de Soja - Região Central do Brasil 2012 e 2013. EMBRAPA, Londrina.

Fehr WR and Caviness CE (1977). Stages of soybean development. Iowa State University, Ames (Special Report, 80).

Ferraudo AS Técnicas de Análise Multivariada - uma introdução. Treinamento. Jaboticabal, SP, 2014.

Hair JF, William B, Babin B and Anderson RE (2009) Análise multivariada de dados. 6th edn. Bookman, Porto Alegre.

Leite DC (2015). Mapeamento de QTL e análises de espectroscopia para teor de óleo visando aplicação em programas de melhoramento de soja. Doctoral thesis. Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Ciências Agrárias e Veterinárias.

Oliveira SM, Sousa LB, Nogueira AP, Hamawaki OT et al. (2014). Caracteres agronômicos e divergência genética entre genótipos de soja. *Enciclopédia biosfera*, 10: 759-768.

Pervaiz ZH, Rabbani MA, Khaliq I, Pearce SR, et al. (2010). Genetic diversity associated with agronomic traits using microsatellite markers in Pakistani rice landraces. *Electron. J. Biotechnol.* 13: 1-12.

Polizel AC, Silva MAP, Hamawaki OT, Silva EMB, Menezes PC (2011). Desenvolvimento de linhagens de soja de ciclo semiprecoce/médio para resistência à ferrugem asiática em Rondonópolis/MT. *Enciclopédia Biosfera* 7: 1-9.

Rabbani MA, Masood MS, Shinwari ZK and Shinozaki KY (2010). Genetic analysis of basmati and nonbasmati Pakistani rice (*Oryza sativa* L.) cultivars using microsatellite markers. *Pak. J. Bot.* 42: 2551-2564.

Rao CR (1952). Advanced statistical methodos in biometric research. John Willey, New York.

Silva AR (2012). Métodos de agrupamento: avaliação e aplicação ao estudo de divergência genética em acessos de alho. Master's thesis. Universidade Federal de Viçosa, Viçosa.

Silva JB, Lazarini E, Silva AM and Reco PC (2010). Ensaio comparativo de cultivares de soja em época convencional em Selvíria, ms: características agronômicas e produtividade. *Bioscience Journal, Uberlândia* 26: 747-754.

Singh D (1981). The relative importance of characters affecting genetic divergence. *Indian J. Genet. Plant Breed.* 41: 237-245.

Sousa LB (2011). Parâmetros genéticos de variabilidade em genótipos de soja. Master's thesis. Universidade Federal de Uberlândia, Uberlândia.

Sousa LB (2013). Diversidade fenotípica e molecular, correlações entre caracteres, adaptabilidade e estabilidade de genótipos de soja. Doctoral thesis. Universidade Federal de Uberlândia, Uberlândia.