Genetic variability in food-type soybean accessions assessed by morphoagronomic traits

D.M. Zeffa[1], L.J. Perini[2], G.H. Freiria[2], G.M. da Silva[2], G.R. Gomes[2], L.V. Constantino[2], S.M. Alves[2], R. Koyama[2] and S.S. de Oliveira Neto[3]

[1] Departamento de Agronomia, Universidade Estadual de Maringá, Maringá, PR, Brasil
[2] Departamento de Agronomia, Universidade Estadual de Londrina, Londrina, PR, Brasil
[3] Departamento de Agricultura, Universidade Estadual Paulista, Botucatu, SP, Brasil

Corresponding author: D.M. Zeffa
E-mail: douglas.mz@hotmail.com

ABSTRACT. The estimation of genetic variability in germplasm banks is important not only for the conservation of genetic resources, but also for their use in plant breeding. However, a large number of traits from different categories (qualitative and quantitative) can make the analysis and interpretation of the results difficult, often resulting in an incomplete distinction among accessions. We characterized 45 accessions of food-type soybean (*Glycine max*) from the Soybean Germplasm Bank of Londrina State University based on morphoagronomic traits. The experiment was carried out in the university farm using a randomized block design with three replicates. Twelve traits were evaluated: flower color, pubescence color, tegument color, hilum color, days until flowering, days until maturation, plant height, height of insertion of the first pod, mass of 100 grains, grain yield per plot, agronomic value, and lodging index. The quantitative trait data were submitted to deviance analysis. The genetic divergence among the accessions was examined by Gower's distance and clustering of the accessions was made based on Ward's hierarchical method. Morphoagronomic descriptors proved efficient in detecting the levels of genetic variability among the accessions maintained in the germplasm collection. Gower's distance was efficient in the discrimination of accessions, showing that simultaneous analysis of quantitative and

qualitative traits is useful to evaluate the variability in a germoplasm bank. These results can be used as an additional source of information to be exploited in food-type soybean breeding programs.

INTRODUCTION

Soybean (*Glycine max*) is a highly nutritious food, which can play an important functional role in the human body, with beneficial effects for health and prevention of diseases (Hea and Chen, 2017). According to FAOstat (2018), world soybean production is 337 million tons per year. Although less than 5% of world soybean production is for human consumption as food (fresh and processed foods), niche markets of soybeans for food use are growing worldwide, mainly due to their innumerable products, such as sprouts, tofu (soybean curd), natto (fermented soybean), edamame (unripe soybean), kuromame (sweet soybean), and salads (Yokomizo et al., 2000; Hirakuri and Lazarotto, 2014; Carrão-Panizzi et al., 2016).

Notably in Paraná state, southern Brazil, the production of soybean for food use is a very common activity by Asian ancestry family farmers, who usually grow this species using seeds acquired from relatives or neighbors and maintaining them for generations. The seeds used by these farmers are not transgenic and, in most cases, these are not commercial varieties. The metropolitan region of Londrina has a large number of food-type soybean farmers who frequently market their grains and/or soy products in free markets (Júnior Perini et al., 2018). In this context, we collected, characterized and evaluated these local varieties to promote the conservation and maintenance of genetic diversity, as well as their availability and use in soybean breeding programs for human consumption (Li et al., 2008; Oliveira et al., 2016).

The characterization of germplasm banks produces a massive amount of data from different categories (quantitative and qualitative), which may hinder or limit the analysis and interpretation of the results and, consequently, result in an incomplete distinction among accessions. Joint analysis of quantitative and qualitative traits may provide a more accurate indication of the genetic variability. Gower's (1971) distance is a cluster mixed data analysis of low complexity that has been used in various studies of characterization of germplasm banks (Quintal et al., 2012; Kyriakopoulou et al., 2014; Cabodevila et al., 2017).

We characterized accessions of food-type soybeans in a germplasm collection by means of morphoagronomic traits. The results of this study will increase the knowledge about the variability found in food-type soybean accessions, contributing to the conservation of the species, as well as the identification of potential accessions to be used in breeding programs.

MATERIAL AND METHODS

The experiment was carried out at Londrina State University (UEL), located in the municipality of Londrina, Paraná, Brazil (23°19'42" S, 51°12'11" W), in the agricultural harvest of 2012/2013. A total of 41 soybean accessions from the UEL Soybean Germplasm Bank and four commercial cultivars (BRS 213, BRS 216, BRS 257, and BRS 258) were

evaluated. The accessions were obtained from visits to small farmers and free markets in the metropolitan region of Londrina between 2010 and 2011. The 45 accessions were sown in four crop rows of 4 m spaced 0.5 m; we evaluated the two central rows and neglected 0.5 m from each end. The experiment was performed using a completely randomized block design with four repetitions

Twelve morphoagronomic descriptors were evaluated, encompassing four qualitative and eight quantitative traits. The qualitative traits evaluated were: i) flower color; ii) pubescence color; iii) tegument color, and iv) hilum color. Regarding the quantitative traits, we evaluated: i) days until flowering (DF), ii) days until maturation (DM), iii) plant height (PH, in cm), iv) height of insertion of the first pod (IP, in cm), v) mass of 100 grains (MG, in g), vi) grain yield per plot (GY, in g), vii) agronomic value (AV), through a visual scale of grades ranging from 1 (worst agronomic traits) to 5 (best agronomic traits), and viii) lodging index (LI), through a visual scale ranging from 1 (upright plants) and 5 (lodging of the crop).

The quantitative traits data was analyzed by restricted maximum likelihood (REML) and best linear unbiased prediction (BLUP) methods using the software Selegen-REML/BLUP (Resende, 2016). The predicted genotypic values were calculated after verification of normality and homogeneity of data by Shapiro and Wilk (1965) and Hartley (1950) tests, respectively. The data that did not follow the assumptions of normality and homogeneity were subjected to Box-Cox transformation (1964). The deviance analysis was performed considering the following model:

$$y = Xb + Za + e \qquad \text{(Eq. 1)}$$

where: y is the vector of data; a is the vector of the effects of blocks (assumed to be fixed) added to the general mean, b is the vector of the genotypic effects (assumed to be random), e is the vector of error (random), X and Z represent the incidence matrices for b and a, respectively.

Pearson linear correlation and Singh's relative importance (Singh, 1981) analysis were performed using the predicted genotypic values. A genetic distance matrix among the accessions was generated through a joint analysis of qualitative and quantitative descriptors based on Gower's distance (Gower, 1971), according to the following formula:

$$Sij = \sum_{k=1}^{p} Wijk . Sijk \Big/ \sum_{k=1}^{p} Wijk \qquad \text{(Eq. 2)}$$

where: k is the number of variables (k = 1, 2, ...); p is the total number of descriptors evaluated; i and j, any two individuals; W_{ijk} is a weight given the comparison ijk, assigning value 1 for valid comparisons and 0 for invalid comparisons (when the value of the variable is absent in one or both individuals); S_{ijk} is the contribution of variable k in the similarity between individuals i and j, with values between 0 and 1.

The Ward method (1963) was used in the hierarchical cluster of the accessions. The clusters were validated based on the cophenetic correlation coefficient (Sokal and Rohlf, 1962), using a thousand permutations at 5% probability level. The optimal number of groups, formed in the dendrograms, was determined using the methods described by Milligan and Cooper (1985). Statistical analyses were performed using the Genes (Cruz, 2016) and R software (www.r-project.org) by the NbClust (Charrad et al., 2014), dendextend (Galili, 2015), corrplot (Wei et al., 2017), and factoMineR (Lê et al., 2008) packages.

RESULTS AND DISCUSSION

Polymorphism was observed for all the qualitative traits evaluated in food-type soybean accessions (Table 1). Regarding flower color, 29 of the 45 accessions had purple flowers, while 16 had white flowers. For pubescence color, 35 of the 45 accessions had a gray color and 10 were brown. The yellowish tegument was the most predominant, being observed in 37 of the 45 accessions, but also green (n = 4), black (3), and brown (1) teguments were found. The color of the hilum descriptor had the largest diversity, with brown (n = 16), light brown (14), black (9), dark black (3), yellow (1), green (1), and gray (1) colorations.

Table 1. Identification and qualitative descriptors of 45 food-type soybean accessions.

Accession	Flower color	Pubescence color	Tegument color	Hilum color
SF1	Purple	Brown	Black	Black
SF2	Purple	Brown	Yellow	Black
SF3	Purple	Brown	Black	Black
SF4	Purple	Brown	Green	Black
SF5	Purple	Brown	Green	Black
SF6	Purple	Brown	Green	Green
SF7	White	Brown	Yellow	Black
SF8	White	Gray	Yellow	Brown
SF9	Purple	Brown	Yellow	Black
SF10	Purple	Brown	Green	Black
SF11	Purple	Gray	Yellow	Brown
SF12	Purple	Gray	Yellow	Light brown
SF13	Purple	Gray	Yellow	Brown
SF14	Purple	Gray	Yellow	Light brown
SF15	Purple	Gray	Yellow	Light brown
SF16	White	Brown	Black	Black
SF17	Purple	Gray	Yellow	Brown
SF18	White	Gray	Yellow	Brown
SF19	Purple	Gray	Yellow	Dark black
SF20	Purple	Gray	Yellow	Brown
SF21	Purple	Gray	Yellow	Light brown
SF22	Purple	Gray	Yellow	Brown
SF23	Purple	Gray	Yellow	Brown
SF24	Purple	Gray	Yellow	Brown
SF25	Purple	Gray	Yellow	Light brown
SF26	Purple	Gray	Yellow	Light brown
SF27	White	Gray	Yellow	Brown
SF28	Purple	Gray	Yellow	Light brown
SF29	White	Gray	Yellow	Light brown
SF30	Purple	Gray	Yellow	Brown
SF31	White	Gray	Yellow	Brown
SF32	Purple	Gray	Yellow	Light brown
SF33	Purple	Gray	Yellow	Gray
SF34	Purple	Gray	Yellow	Light brown
SF35	White	Gray	Yellow	Brown
SF36	White	Gray	Brown	Brown
SF37	White	Gray	Yellow	Brown
SF38	Purple	Gray	Yellow	Dark black
SF39	Purple	Gray	Yellow	Dark black
SF40	White	Gray	Yellow	Brown
SF41	White	Gray	Yellow	Light brown
BRS 213	White	Gray	Yellow	Yellow
BRS 216	White	Gray	Yellow	Light brown
BRS 257	White	Gray	Yellow	Light brown
BRS 258	White	Gray	Yellow	Light brown

These qualitative traits have a great importance in soybean breeding programs. The occurrence of natural cross-pollination, varietal mixture or mutations can be confirmed by the presence of plants that differ from the expected traits in segregating generations, since they are easily distinguishable traits in the field or laboratory (Vernetti and Vernetti Junior, 2017). Additionally, qualitative traits are considered to be less influenced by environmental conditions (Dashek and Harrison, 2006).

The analysis of deviance was significant (P < 0.01) for all descriptors, which denotes a wide genetic variation among the food-type soybean accessions (Table 2). The coefficients of variation (CV) ranged from 2.11 (DM) to 9.17% (AV) and are considered low and similar to those reported by Silva et al. (2016) and Silveira et al. (2016), indicating high experimental quality and, consequently, greater reliability of the results.

Table 2. Analysis of deviance and coefficients of variation (CV) of the following traits: days until flowering (DF), days until maturation (DM), plant height (PH), height of the first pod insertion (IP), mass of 100 grains (MG), grain yield (GY), agronomic value (AV) and lodging index (LI) of 45 food-type soybean accessions.

Traits	Deviance	Mean	CV (%)
DF	243.67*	60.12	2.56
DM	235.66*	136.38	2.11
PH	135.61*	80.33	9.08
IP	145.56*	9.80	9.17
MG	233.83*	22.47	8.85
GY	211.34*	449.62	9.12
AV	137.44*	3.43	5.56
LI	122.46*	2.17	6.92

* = significant at a 1% probability level by the Chi-squared test.

The mixed models methodology (REML/BLUP) is an accurate procedure for prediction of genetic values and estimation of components of variance in genotypic diversity studies, since most of studies show an imbalance of the data due to various factors (e.g. reduced number of seeds, non-germination or death of the plants not adapted to the evaluation sites) (Resende et al., 2017).

Through the predicted genotypic values based on the quantitative descriptors, high amplitude can be observed within the evaluated traits, which denotes a genetic variability among the accessions (Figure 1). DF values ranged from 51.9 (SF6) to 82.4 days (SF3), while mean DM ranged from 127.6 (SF7) to 156.5 days (SF3). Despite the variation, in general, the accessions presented a crop production cycle similar to most of the soybean cultivars adapted to Brazilian conditions, between 90 and 150 days (Bezerra et al., 2017).

Plant height ranged from 65.2 (SF7) to 101.3 cm (SF2), while mean values of IP varied between 7.1 (SF39) and 18.1 cm (SF3). Very tall or extremely short plants present great difficulty for mechanized harvesting. In relatively flat areas, plant heights varying between 50 and 60 cm are reported to be satisfactory in mechanical harvesting. However, in areas with uneven topography, the ideal height is 70 to 80 cm (Sediyama et al., 2009). As for IP, the cultivars most efficiently harvested are those whose insertion varies between 10 and 15 cm from the ground (Bezerra et al., 2017).
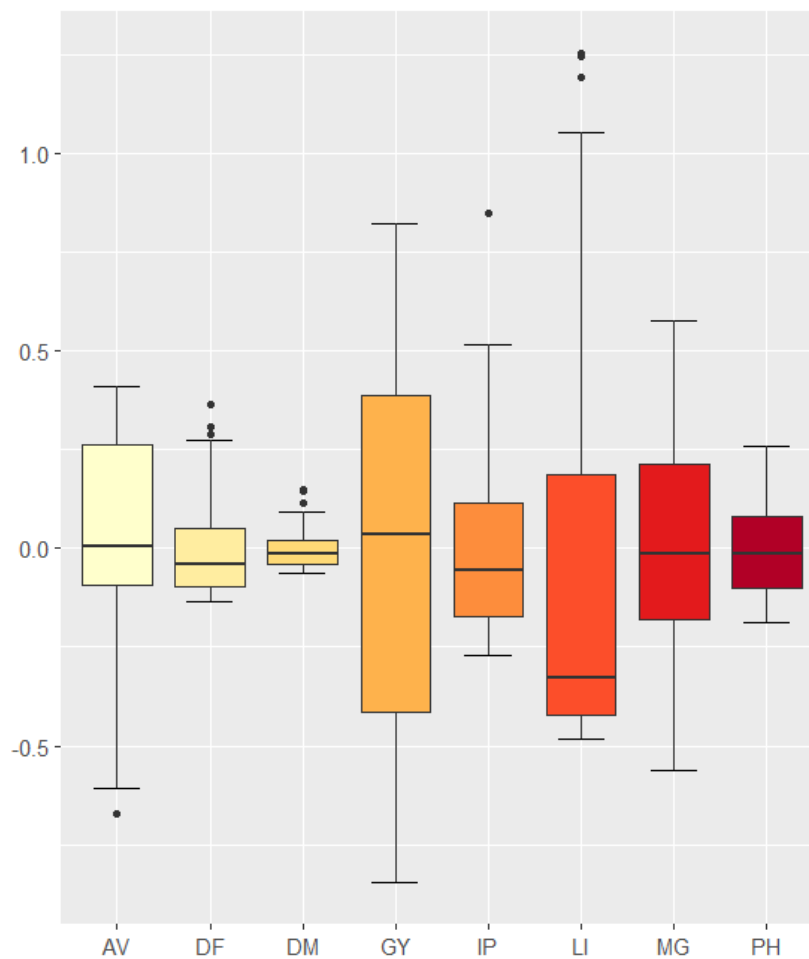
Figure 1. Boxplot of agronomic traits of 45 food-type soybean accessions. Data scaled to have mean zero and unit variance. Days until flowering: DF, days until maturation: DM, plant height: PH, height of the first pod insertion: IP, mass of 100 grains: MG, grain yield: GY, agronomic value: AV, and lodging index: LI.

GY ranged from 75.1 (SF1) to 812.3 g (BRS 213), while MG ranged from 9.8 (SF39) to 35.4 g (SF13). Food-type soybeans can be classified into two categories: (i) small seeds when the MG is less than 10 g; this category is destined for consumption as sprouts or natto (fermented), and (ii) large seeds (greater than 10 g), a category used in direct human consumption, such as unripe soybeans (edamame), sweet soybeans (kuromame) and in salads (Yokomizo et al., 2000; Cantelli et al., 2017).

Genotypic correlations among the traits are presented in Figure 2. Positive and significant indexes of moderate levels were observed between traits DF $\times$ DM ($r = 0.61$), PH $\times$ LI ($r = 0.62$), and AV $\times$ GY ($r = 0.63$). On the other hand, negative correlations were found between AV $\times$ LI ($r = -0.51$). These results indicate the possibility of indirect selection of a particular trait related to another, when they are correlated (Falconer and Mackay, 1996). Similar results were also reported by Yokomizo et al. (2000), Lopes et al. (2002), Nogueira et al. (2012), and Dalchianon and Passos (2012).
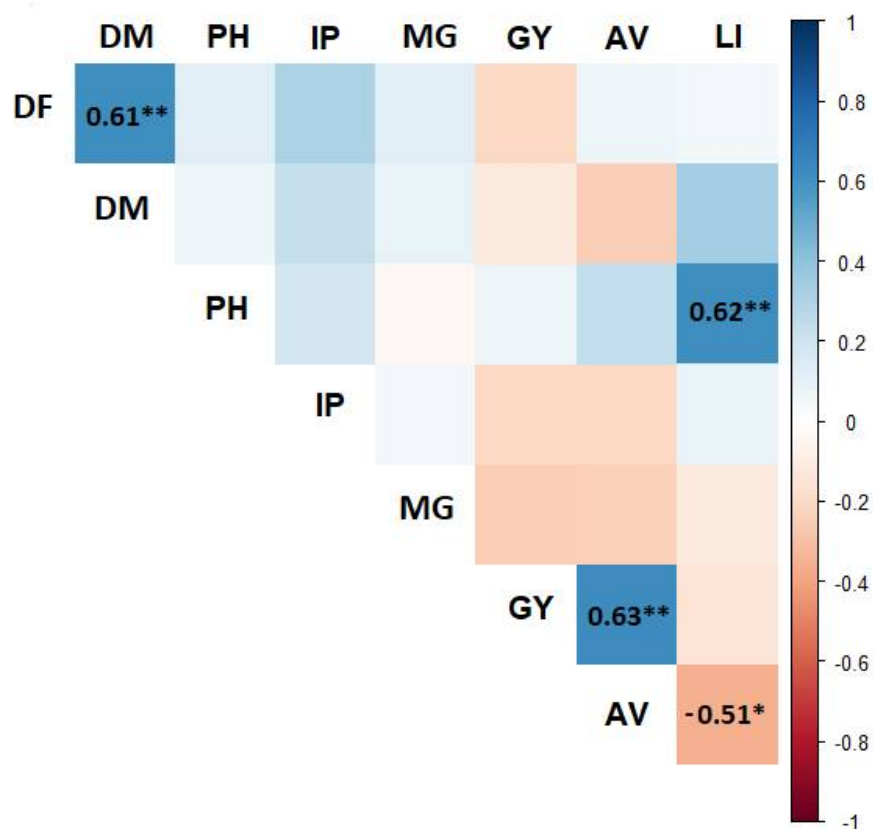
Figure 2. Genotypic correlations among the traits: days until flowering (DF), days until maturation (DM), plant height (PH), height of the first pod insertion (IP), mass of 100 grains (MG), grain yield (GY), agronomic value (AV) and lodging index (LI) of 45 food-type soybean accessions.

Using Singh's method (1981), we observed that all the descriptors contributed moderately to the discrimination of accessions, since the values of relative importance ranged from 10.11 (MG) to 16.72% (GY) (Figure 3). Almeida et al. (2011), studying the genetic diversity among soybean cultivars, verified that DM and MG were the traits presenting the greatest discriminatory power among the cultivars. Similarly, Rigon et al. (2012) reported that IP and DM were the variables that contributed the most to the dissimilarity of the soybean cultivars evaluated.

The genetic dissimilarity among soybean accessions presented an average Gower's distance of 0.37 ($\pm$0.08). The shortest distance was observed between SF23 and SF24 (0.12) accessions, while the BRS 213 and SF3 accessions were the most genetically distant (0.73). Ward's hierarchical cluster presented a high cophenetic correlation coefficient (CCC = 0.82), which indicates reliability between the dissimilarity matrix and the graphical representation obtained (Sokal and Rohlf, 1962). In addition, the CCC among the matrices was highly significant by the Mantel test ($P < 0.01$).

Gower's distance is a simple method and has shown robust results in several studies, even though it is not often used by researchers in genetic variability studies of germplasm banks (Oliveira et al., 2016). Some studies using such an approach have already been carried out on *Carica papaya* (Quintal et al., 2012*), Abelmoschus esculentus* (Kyriakopoulou et al., 2014), *Urochloa brizantha* (Torres et al., 2015), *Cucurbita moschata* (Oliveira et al., 2016), and *Solanum lycopersicum* (Cabodevila et al., 2017).
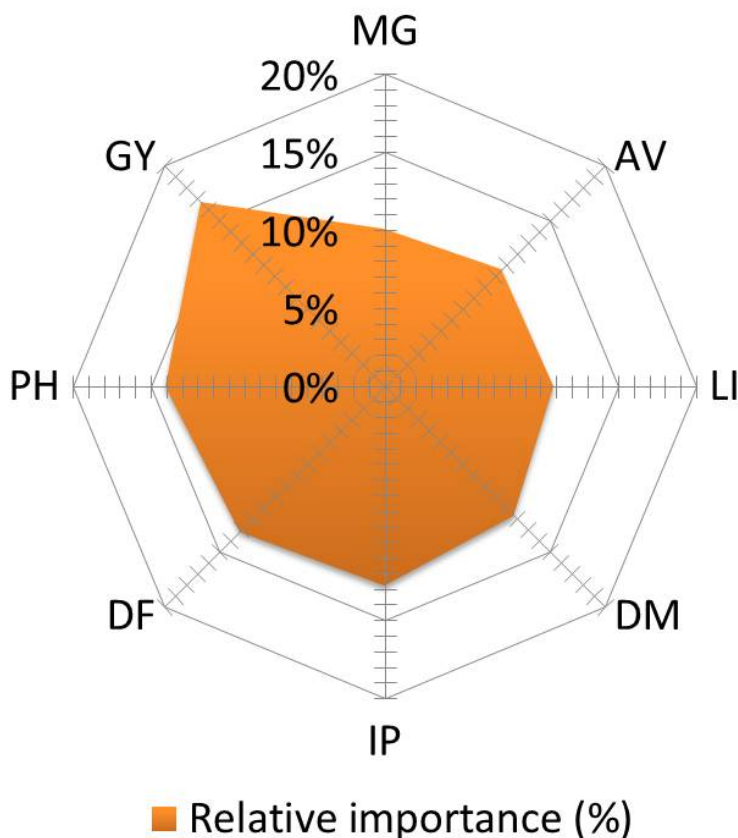

Figure 3. Relative importance of days until flowering (DF), days until maturation (DM), plant height (PH), height of the first pod insertion (IP), mass of 100 grains (MG), grain yield (GY), agronomic value (AV), and lodging index (LI) of 45 food-type soybean accessions.

In a dendrogram obtained using Ward's method, the formation of five distinct groups can be observed as suggested by the Milligan and Cooper (1985) method (Figure 4). Group I clustered accessions SF1 and SF3, which commonly presented purple flowers, black hilum and brown pubescence. Regarding the quantitative traits, these soybean accessions were characterized as the tallest and the longest maturity cycle, since they showed an average of 101.3 cm and up to 77.4 DM.

Figure 4. Dendrogram of 45 food-type soybean accessions by Gower's distance.

Group II clustered eight soybean accessions (SF2, SF4, SF5, SF6, SF7, SF9, SF10, and SF16), which showed predominantly black hilum and brownish pubescence, and presented green, black, or yellow tegument. DF values ranged from 51.9 to 76.6 days, while DM ranged from 127.6 to 138.4 days. There was large amplitude in GY, since the averages ranged from 120.9 to 624.4 g.

Group III clustered 21 soybean accessions, in which all presented gray pubescence and yellow tegument color. Regarding the quantitative traits, no pattern of a common trait was observed and it was not possible to make a clear relation among the clusters, since there was great variation between them [e.g. PG mean values ranged from 69.2 (SF21) to 819.8 g (SF14)].

The commercial cultivars BRS 213, BRS 216, BRS 257, and BRS 258, clustered with soybean accessions SF8, SF31, SF35, and SF41 in group IV, presenting the best required morphoagronomic traits. As qualitative traits in common, the accessions presented white flowers, gray pubescence, and yellow tegument. Regarding the quantitative traits, the accesses presented high GY values, ranging from 469.2 (SF8) to 812.3 g (BRS 213), low LI values, mostly close to 1 (upright plants), and high values for AV, ranging from 4.33 to 5.

Six soybean accessions were clustered in group V (SF18, SF27, SF29, SF36, SF37, and SF40), which were characterized by yellowish and grayish pubescence. The GY averages ranged within the group, from 69.2 (SF1) to 559.4 g (SF39). The group presented high MG, between 28.56 (SF17) and 33.1 g (SF29). Accessions exhibited DM up to 135.5, while PH were lower than 87.6 cm.

The genetic diversity observed among food-type soybean accessions using the mixed model methodology (REML/BLUP), as well as the use of the Gower's distance associated with the Ward's clustering analysis, allowed a good discrimination by the evaluated descriptors and the clustering of the accessions based on their similarity, indicating that analysis of quantitative and qualitative traits allows accurate estimation of genetic variability.

CONCLUSIONS

Food-type soybean accessions showed genetic variability for all morphoagronomic descriptors evaluated, indicating a high potential of these accessions in food-type soybean breeding programs. The Gower's distance associated with Ward's method allowed the discrimination of accessions into five groups by clustering analysis based on quantitative and qualitative traits, which enables the selection of divergent parents for the breeding program.

REFERENCES

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

Bezerra ARG, Sediyama T, da Silva FL, Borém A, et al. (2017) Agronomical aspects of the development of cultivars. In: Soybean breeding (da Silva FL, Borém A, Sediyama T, Ludke WH, eds.). Springer, Switzerland.

Box GEP and Cox DR (1964). An analysis of transformations. *J. R. Stat. Soc. Series B Stat. Methodol*. 26: 211-252.

Cabodevila VG, Picardi LA and Pratta GR (2017). A multivariate approach to explore the genetic variability in the F$_2$ segregating population of a tomato second cycle hybrid. *BAG J. Basic. Appl. Genet*. 28: 7-18.

Cantelli KC, Schmitd JT, Oliveira MAD, Steffens J, et al. (2017). Sprouts of genetic soybean lines: evaluation of chemical-physical properties. *Braz. J. Food. Technol*. 40: 1-10.

Carrão-Panizzi MC, Bertagnolli PF, Moreira JUV, da Costa LC, et al. (2016). Melhoramento de soja para alimentação humana na Embrapa Trigo – safra agrícola 2015/2016. In: Soja: resultados de pesquisa 2015/2016 (Costamilan LM, Carrão-Panizzi MC, eds.). Embrapa Trigo, Passo Fundo.

Charrad M, Ghazzali N, Boiteau V and Niknafs A (2014). NbClust: an R package for determining the relevant number of clusters in a data set. *J. Stat. Softw.* 61: 1-36.

Cruz CD (2016). Genes Software - extended and integrated with the R, Matlab and Selegen. *Acta Sci Agron* 38: 547-552.

Dalchiavon FC and Passos M (2012). Correlação linear e espacial dos componentes de produção e produtividade da soja. *Semin Cienc. Agrar.* 33: 541-552.

Dashek WV and Harrison M (2006). Plant cell biology. 1st edn. Enfield NH: Science Publishers, United States.

Falconer DS and Mackay TFC (1996). Introduction to quantitative genetics. 4th edn. Addison Wesley Longman, Harlow.

FAOstat (2018). Food and Agriculture Organization of the United Nations, Rome, Italy.

Galili T (2015). Dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics.* 31: 3718-3720.

Gower JCA (1971). A general coefficient of similarity and some of its properties. *Biometrics.* 27: 857-871.

Hartley HO (1950). The maximum F-ratio as a short-cut test for heterogeneity of variance. *Biometrika.* 37: 308-312.

Hea FJ and Chen JQ (2017). Consumption of soya bean: the link to decreased breast cancer incidence. *Oilseeds Focus.* 3: 34-37.

Hirakuri MH and Lazzarotto J J (2014). O agronegócio da soja nos contextos mundial e brasileiro. 1st edn. Embrapa Soja, Londrina.

Kyriakopoulou OG, Arens P, Pelgrom KT, Karapanos I, et al. (2014). Genetic and morphological diversity of okra (*Abelmoschus esculentus* [L.] Moench.) genotypes and their possible relationships, with particular reference to Greek landraces. *Sci. Hortic.* 171: 58-70.

Silva FCS, Sediyama T, da Silva AF, Bezerra AERG, et al. (2016). Identification of new descriptors for differentiation of soybean genotypes by Gower algorithm. *Afri. J. Agric. Res.* 11: 961-966.

Silveira, DA, Pricinotto LF, Nardino M, Bahry CA, et al. (2016). Determination of the adaptability and stability of soybean cultivars in different locations and at different sowing times in Paraná state using the AMMI and Eberhart and Russel methods. *Semin. Cienc. Agrar.* 37: 3973-3982.

Lê S, Josse J and Husson F (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* 25: 1-18.

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

Lopes ACA, Vello NA, Pandini F, de Moura M, et al. (2002). Variabilidade e correlações entre caracteres em cruzamentos de soja. *Sci. Agri.* 59: 341-348.

Mantel N (1967). The detection of disease clustering and a generalized regression approach. *Canc. Res.* 27: 209-220.

Milligan G and Cooper M (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika.* 50:159–179.

Nogueira APO, Sediyama T, de Sousa LB, Hamawaki OT, et al. (2012). Análise de trilha e correlações entre caracteres em soja cultivada em duas épocas de semeadura. *Biosci. J.* 28: 877-888.

Oliveira RL, Gonçalves LSA, Rodrigues R, Baba VY, et al. (2016). Genetic divergence among pumpkin landraces. *Semin. Cienc. Agrar.* 37: 547-556.

Perini Junior L, Zeffa DM, Freiria GH, Novais PS, et al. (2018). Diversidade genética entre acessos de soja tipo alimento com base no algoritmo de Gower. *Colloquium Agrariae.* 14(4): 47-57.

Quintal SSR, Viana AP, Gonçalves LSA, Pereira MG, et al. (2012). Genetic divergence among papaya accessions by morphoagronomic traits. *Semin. Cienc. Agrar*. 33: 131-142.

Resende MDV (2016). Software Selegen-REML/BLUP: a useful tool for plant breeding. *CBAB*. 16: 330-339.

Resende MDV, Ramalho MAP, Nunes JAR, da Silva FL, et al. (2017). BLUP in the genetic evaluation of parents, generations, populations, and progenies. In: Soybean breeding (da Silva FL, Borém A, Sediyama T, Ludke WH, eds.). Springer, Switzerland.

Rigon JPG, Capuani S, de Brito Neto JF, da Rosa GM, et al. (2012). Dissimilaridade genética e análise de trilha de cultivares de soja avaliada por meio de descritores quantitativos. *Rev. Ceres.* 59: 233-240.

Sediyama T, Teixeira RC and Barros HB (2009). Cultivares. In: Tecnologias de produção e usos da soja (Sediyama T. (ed.). Editora Mecenas, Londrina.

Shapiro SS and Wilk MB (1965). An analysis of variance test for normality (complete samples). *Biometrika.* 52: 591-611.

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

Sokal RR and Rohlf FJ (1962). The comparison of dendrograms by objective methods. *Taxon.* 11: 33–40.

Sudré CP, Gonçalves LSA, Rodrigues R, Amaral Júnior AD, et al. (2010). Genetic variability in domesticated Capsicum spp. as assessed by morphological and agronomic data in mixed statistical analysis. *Genet. Mol. Res.* 9: 283-294.

Torres FE, Valle CBD, Lempp B, Teodoro PE, et al. (2015). Estimation of genetic divergence between braquiaria ecotypes based on quantitative and qualitative descriptors. *Cien.c Rural.* 45: 485-491.

Vernetti FJ and Vernetti Junior FJ (2017). Qualitative traits in breeding. In: Soybean breeding (da Silva FL, Borém, A, Sediyama T, Ludke WH, eds.). Springer, Switzerland.
Ward JH (1963). Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58: 236-244.
Wei T, Simko V, Levy M, Xie Y, et al. (2017). Package 'corrplot'. *Am. Stat.* 56: 316-324.
Yokomizo GKI, Duarte JB and Vello NA (2000). Correlações fenotípicas entre tamanho de grãos e outros caracteres em topo-cruzamentos de soja tipo alimento com tipo grão. *Pesqui. Agropecu. Bras.* 35: 2235-2241.