Selection of progenitors for increase in oil content in soybean[1]

Josiane Isabela da Silva Rodrigues[2], Klever Márcio Antunes Arruda[3], Newton Deniz Piovesan[2], Everaldo Gonçalves de Barros[4], Maurilio Alves Moreira[5]*

ABSTRACT

The low genetic diversity brings limitation to breeding, because genetically similar genotypes share alleles in common, causing little complementarity and low vigor due to the low levels of heterozygosity in crosses. The objective of this work was to analyze the oil content and genetic diversity of soybean genotypes (*Glycine max* (L.) Merrill) based on QTL regions of this trait for choice of progenitors for increase in oil content. Twenty-two genotypes with wide variation in oil content, including cultivars with high oil contents, were cultivated in different Brazilian conditions and the oil content of the grains was quantified by infrared spectrometry. Microsatellite markers selected based on QTL regions for oil content in soybean were analyzed to estimate the genetic diversity. In these studies, a wide variation in oil content (17.28-23.01%) and a reasonable diversity among the genotypes were observed, being PI181544 the most divergent genotype, followed by Suprema. The genotypes PI371610/Suprema and Suprema/CD01RR8384 showed genetic distance and higher oil contents in the grains, while the cultivars Suprema and CD01RR8384 had the highest oil contents and proved to be little genetically related. These genotypes are promising progenitors for selection of high oil content in soybean.

Key words: *Glycine max*; genotypes; genetic variability.

RESUMO

Seleção de progenitores para aumento no teor de óleo em soja

A baixa diversidade genética traz limitação ao melhoramento de plantas, pois genótipos semelhantes compartilham alelos em comum causando pouca complementaridade e baixo vigor, devido aos baixos níveis de heterozigosidade obtidos nos cruzamentos .Objetivou-se, com este trabalho, analisar o teor de óleo e a diversidade genética de genótipos de soja com base em regiões de QTL desta característica para a escolha de progenitores visando ao aumento no teor de óleo em soja. Vinte e dois genótipos com ampla variação no teor de óleo incluindo cultivares com altos teores foram cultivados em diferentes condições brasileiras e o teor de óleo nos grãos foi quantificado por espectrometria do infravermelho. Para estimar a diversidade genética foram analisados marcadores microssatélites selecionados com base em regiões de QTL para o teor de óleo em soja. Nestes estudos observou-se ampla variação quanto ao teor de óleo nos grãos e razoável diversidade genética entre os genótipos, sendo a PI181544 a mais divergente, seguida por Suprema. Os genótipos PI371610/Suprema e Suprema/CD01RR8384 mostraram distância genética e maiores teores de óleo nos grãos. Já as cultivares Suprema e CD01RR8384 tiveram os mais altos teores de óleo e se mostraram pouco relacionadas geneticamente. Estes genótipos são progenitores promissores para a seleção de alto teor de óleo em soja.

Palavras chave: *Glycine max*; genótipos; variabilidade genética.

[2] Federal University of Viçosa, Viçosa, Minas Gerais, Brazil. josianeisabela@gmail.com; kleverantunes@yahoo.com.br; piovesan@ufv.br

[3] Agronomic Institute of Paraná, Londrina, Brazil. kleverantunes@yahoo.com.br

[4] Catholic University of Brasília, Post Graduation program in Genomic Sciences and Biotechnology , Brasília, Distrito Federal, Brasil. everaldodebarros@gmail.com

[5] *In memoriam.*

*Corresponding author: josianeisabela@gmail.com

INTRODUCTION

The use of elite germplasm in crosses maintains a narrow genetic base of soybean [*Glycine max* (L.) Merrill] varieties cultivated in Brazil. Historically, the genetic variability has been little explored by breeding programs. The low genetic diversity brings limitation to breeding, because genetically similar genotypes share alleles in common, causing little complementarity and low vigor due to the low levels of heterozygosity in crosses. On the other hand, superior genotypes are most likely in producing populations with high genetic variability, which are resulted from crosses between genetically distant progenitors and with high phenotypic values.

In the 80's, the soybean varieties cultivated in Brazil were introductions of the Southern U.S. and the remaining came from hybridization of the North American introductions. Seventy-nine cultivars at that time descended from 26 ancestors and 11 formed 89% of the gene pool (Hiromoto & Vello, 1986). Other 69 cultivars showed an average value of relatedness of 0.16 and effective population size estimated between 11 and 15 (Vello *et al.*, 1988). One hundred cultivars released between 1984 and 1998 had an average value of relatedness of 0.21 (Bonato *et al.*, 2006a). In addition, the same degree of relatedness and an effective population size equal to 11 were estimated among 90 elite cultivars (Miranda *et al.*, 2007).

Genetic relationships among Brazilian cultivars based on molecular markers are also reported in the literature. Three hundred and seventeen cultivars released between 1962 and 1998 had similarity indices between 0.17 and 0.97 and average similarity of 0.6 (Bonato *et al.*, 2006b). Vieira *et al.* (2009) observed similarities between 0.27 and 0.98 and average similarity of 0.53 in another group of 53 genotypes. Other 168 cultivars showed similarity indices between 0.01 and 0.9 and average similarity of 0.42 (Priolli *et al.*, 2010). Therefore, the low values of effective size, the high estimates of relatedness, and the elevated similarity indices indicate a high similarity in the improved cultivars. Moreover, by the estimates, a relatively low level of diversity is maintained in soybean germplasm from different breeding programs in Brazil.

Molecular markers, in general, can be useful in predicting genetic variability for the development of populations, and greater predictive ability for molecular markers is expected when the diversity estimates consider the positions of QTLs (quantitative trait loci). These estimates, when predicting genetic variability, are probably most effective when based on molecular markers located in QTL regions of the target trait in relation to random regions in the genome (Rodrigues *et al.*, 2015). The objective of this work was to identify progenitors

genetically divergent and with high oil contents in the grains for increase in oil content and in genetic basis of soybean breeding programs.

MATERIAL AND METHODS

Genetic material and determination of oil content

Twenty-two soybean genotypes from the germplasm bank of the Soybean Quality Breeding Program of the Federal University of Viçosa, with wide variation in oil content including cultivars with high oil contents, were cultivated in Viçosa, MG (Dec 2009, 20°452 S, 42°522 W), Visconde do Rio Branco, MG (Feb 2010, 21°00' S, 42°50' W) and São Gotardo, MG (Feb 2010, Oct 2011, 19°18' S, 46°02' W) using the randomized block design with three replications. In the trials, fifteen seeds were sown per row of 1 m with 0.5 m spacing between rows . The grains were ground in an industrial mill (model MA020, Marconi) and the soybean flour was analyzed for oil content by infrared spectrometry using a FT-NIR spectrometer (model Antaris II, Thermo Scientific). A combined analysis of variance was performed, considering the effect of genotypes (G_i) as fixed and the effect of environments (E_j) as random. The components of variance were estimated according to Cruz *et al.* (2014). The percentages of the simple and complex parts of the mean square of the genotype x environment interaction (MSGxEjj') were calculated according to Cruz & Castoldi (1991).

DNA extraction and amplification of microsatellite markers

The DNA was purified using the Miniprep DNA Purification System kit (Promega) and thirty-three microsatellite primer pairs were amplified (Satt006; Satt468; Satt510; Satt020; Satt196; Satt562; Satt239; Satt308; Satt250; Satt156; Satt496; Satt313; Satt398; Satt184; Satt458; Satt166; Satt229; Satt540; Satt274; Satt317; Satt420; Satt479; Satt277; Satt212; Satt571; Satt144; Satt257; Satt551; Satt141; Satt355; Satt384; Satt200; and Satt259). These microsatellite loci were selected based on QTL regions of the oil content trait in soybean. The respective QTLs for oil content in the regions of the molecular markers are described by Mansur *et al.* (1996), Specht *et al.* (2001), Csanádi *et al.* (2001), Tajuddin *et al.* (2003), Chung *et al.* (2003), Fasoula *et al.* (2004), Hyten *et al.* (2004), Panthee *et al.* (2005), Reinprecht *et al.* (2006), Chen *et al.* (2007), Li *et al.* (2007), Shibata *et al.* (2008), Rodrigues *et al.* (2010), and Eskandari *et al.* (2013), on the linkage groups A1, B2, C2, D1, D1a, D1b, D2, E , F, H, I, K, L, M, N, and O.

In the amplification reactions, we used 10 mM Tris-HCl, pH 8.3; 50 mM KCl; 2 mM $MgCl_2$; 0.1% Triton X100; 100 mM of each deoxynucleotide; 0.3 mM of each primer;

one unit of *Taq* DNA polymerase and 30 ng of DNA. The PCR reactions had an initial step of 94°C for 4 min, 30 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 2 min, with a final step of 72°C for 7 min. The amplification products were separated by electrophoresis on 10% polyacrylamide gels, using 1X TAE buffer (40 mM Tris-acetate and 1 mM EDTA) and running period of three hours at 140 volts. The polyacrylamide gels were stained with 2% silver nitrate and photodocumented with the equipment Loccus Biotechnology (model L-PIX EX).

Analysis of genetic diversity

The polymorphism information content of the molecular markers was calculated according to Cruz *et al.* (2011). Three matrices of dissimilarity were calculated through the complements of the unweighted and weighted similarity indices and Smouse and Peakall d² index. Using the estimates of the complement of the weighted similarity index, the genotypes were grouped by the clustering methods UPGMA, Tocher and modified Tocher (Vasconcelos *et al.*, 2007), and the two and three-dimensional projections were obtained. The estimates of the complement of the weighted similarity index based on the thirty-three microsatellite markers were compared to the estimates obtained by the same index with only microsatellites of linkage group I, the most related with oil content in the literature (Rodrigues *et al.*, 2010; Li *et al.*, 2011, Qi *et al.*, 2011, Rodrigues *et al.*, 2013; Leite *et al.*, 2016), where the distances in both cases were compared. All analyses were performed with the program Genes (Cruz, 2013).

RESULTS AND DISCUSSION

The analysis of variance showed difference in the oil contents of the soybean genotypes in the different environments, indicating genetic variability, variation in the environments and differential response of the genotypes in the environments (Table 1). The genotype x environment interaction was predominantly complex, except in one of the pairs of environments (Visconde do Rio Branco/Feb 2010 and São Gotardo/Feb 2010) (Table 2). The coefficients of variation showed precision in controlling the causes of experimental variation (2.84-6.69%) and the ratio among the largest and the smallest residual mean square indicated homogeneity in residual variances (Falconer & Mackay, 1996). The ratio CVg/CVe was greater than 1, indicating favorable condition to selection (Araújo *et al.*, 2014).

The variation in the oil contents along the four field trials was from 16.0 to 24.2%. Suprema showed the greatest oil content based on the average of the environments (23.01%), followed by CD01RR8384 (22.91%), while BR8014887 presented the lowest oil content (17.28%) (Table 3).

The microsatellite loci of QTL regions for oil content showed to be effective in distinguishing the twenty-two genotypes. A total of 108 alleles was observed in this study. The number of alleles per locus ranged from two to six with an average of 3.3 alleles per locus. The size of alleles was between 100 and 600 base pairs. Altogether, we observed four heterozygote genotypes, which are compatible with the level of inbreeding expected in the autogamous species. The polymorphism information content, which estimates the informativeness of each locus, ranged from 0.08 to 0.77 with an average of 0.44. These estimates are close to the values reported by Mian *et al.* (2009) and Mulato *et al.* (2010), who also evaluated microsatellite markers in elite cultivars and accessions of soybean germplasm.

The dissimilarity estimates based on the complements of the unweighted and weighted similarity indices and based on Smouse and Peakall d² index showed high values of Person's correlation (> 0.94), indicating concordance in the estimates of the indices and small difference in their use. For all indices, the highest dissimilarity was estimated between PI181544 and the cultivar CD224. After this pair of genotypes, higher estimates for the last two indices were observed between each one of the introductions PI371611 and PI371610 and the cultivar Suprema.

Table 1: Combined analysis of variance of oil content of the twenty-two soybean genotypes cultivated in Viçosa, MG (Dec/ 2009), Visconde do Rio Branco, MG (Feb/2010) and São Gotardo, MG (Feb/ 2010 and Oct/2011)

Source of variation	Df	Mean Square
Blocks/Environments	8	3.79
Genotypes	21	22.97**
Environments	3	42.62**
Genotypes x Environments	63	2.77**
Residue	168	0.92
Mean	-	20.26
Coefficient of variation (%)	-	4.73
Relation higher/lower MRS	-	3.80
CVg/CVe	-	1.35

**significant at 1% probability by the F-test; Df: degree of freedom.

Table 2: Percentages of the simple and complex parts of the mean square of genotype x environment interaction (MSGxEjj')

% of the simple part of MSGxEjj'	% of the complex part of MSGxEjj'
1 x 2 = 32.76	1 x 2 = 67.24
1 x 3 = 46.43	1 x 3 = 53.57
1 x 4 = 37.99	1 x 4 = 62.01
2 x 3 = 52.41	2 x 3 = 47.59
2 x 4 = 38.89	2 x 4 = 61.11
3 x 4 = 44.12	3 x 4 = 55.88

1, 2, 3, 4 refers to Viçosa, MG (Dec 2009), Visconde do Rio Branco, MG (Feb 2010), São Gotardo, MG (Feb 2010), and São Gotardo, MG (Oct 2011), respectively.

The distances based on the complement of the weighted similarity index ranged between 0.06 and 0.81 and had an average of 0.61, higher variation and average than those reported by Vieira *et al.* (2009) with microsatellite markers in soybean. The greatest distance value (0.81) was estimated for the pairs of genotypes PI181544/CD224, Suprema/PI371611, PI371610/Suprema, PI371610/CD224, and Garantia/CD225RR and the lowest distance value was estimated between PI371611/PI371610 (0.06). The pairs of genotypes Suprema/PI371610 (0.81), CD01RR8384/Suprema (0.63) and Suprema/CD219RR (0.63) showed higher distances and oil content greater than or equal to 21.5%.

In relation to the mean dissimilarity of each genotype in relation to the twenty-one remaining, the highest value was estimated for PI181544 (0.68), followed by Suprema (0.67) and CD01RR8384 (0.65) and the lowest value was estimated for CD219RR (0.56), followed by CD983321RR, CD226RR and Luziânia (0.57) (Table 2). Despite being the most divergent, the genotype PI181544 showed low oil content, while the cultivars Suprema and CD01RR8384 showed oil contents considered elevated. When the overall mean dissimilarity was calculated for the plant introductions (PIs) (0.65), for the cultivars with the initials CD (0.61) and for the remaining genotypes (0.60), the group of PIs was the most dissimilar group, which was expected.

In the genotype clustering by the UPGMA method, a cut at 95.77% of the dissimilarity (value indicated by the method of Mojema using k = 1.25) establishes four groups (Mojema, 1977). One group gathers only PIs, another group gathers CD01RR8376 and CD01RR8384 and the remaining group gathers the other cultivars, except Suprema that is not grouped. The values of cophenetic correlation, distortion and stress were 0.76, 1.2% and 10.9%, respectively, which evidence good adjustment of the original and graphic values and small distortion of the distances in the dendrogram (Figure 1).

The Tocher method establishes six groups with more than one genotype and shows relationships also observed in the dendrogram obtained by the previous method. The cultivars CD01RR8376 and CD01RR8384 are kept together and Suprema follows ungrouped again – relationships that are also observed in the modified method (Table 4).

In the bi and tridimensional projections, there was a relative distance of PIs and the cultivars CD224 and Suprema in relation to most genotypes, despite the low adjustment between the original distances and those obtained in the projections. The values of cophenetic correlation, distortion and stress were 0.52 and 0.61, 29.07 and 13.22%, and 42.62 and 31.26%, respectively, and indicate a low effectiveness in representing the distance matrix. Based on these measurements, the clustering by the UPGMA method was the most effective.

The microsatellites of linkage group I indicated genetic variability in the main region of the linkage group. When the distances obtained by the complement of the weighted similarity index by analysis of the 33 microsatellite markers were compared to those obtained by the same index based on the analysis of microsatellites of linkage group I only, the correlation between the estimates was r = 0.61. In both cases, the greatest distance value was observed in the pairs of genotypes PI371610/CD224, Garantia/CD225RR, Suprema/PI371611 and PI371610/Suprema and the lowest value in the pair of genotypes PI371611/PI371610. Therefore, there is genetic variation in the main linkage group that controls oil content and the concordance in distance relationships in these two cases suggests the presence of QTLs of greater effect in the linkage group.

The linkage group I has strong association with oil content. QTLs mapped on the main region of the linkage group explained from 6 to 24% of the variation of oil content in the studied populations, the most reported and involved region with oil content in soybean (Csanádi *et al.*, 2001; Chung *et al.*, 2003; Nichols *et al.*, 2006).

Based on the distance estimates and clusters obtained, CD224/PI181544, PI371611/Suprema and PI371610/Supre-

Table 3: Average percentages of oil in the four environments on a dry basis and average dissimilarity estimates of the twenty-two soybean genotypes

Genotype	Oil (%)	Diss.[1]
1-BRSMG Garantia	19.74	0.62
2-FMT Tucunaré	20.82	0.58
3-BRSGO Luziânia	19.42	0.57
4-MG/BR 46 (Conquista)	19.47	0.60
5-Suprema	23.01	0.67
6-BARC-8	18.27	0.60
7-CS303TNKCA	20.86	0.60
8-BRSMG 68 (Vencedora)	19.79	0.59
9-PI181544	18.26	0.68
10-PI371611	20.66	0.63
11-PI371610	21.50	0.64
12-CD225RR	20.22	0.64
13-CD224	20.71	0.62
14-CD219RR	21.61	0.56
15-PI235347	19.48	0.65
16-CD226RR	20.13	0.57
17-UFV20 (Florestal)	19.67	0.58
18-CD222	20.63	0.61
19-CD01RR8376	21.42	0.63
20-CD983321RR	19.84	0.57
21-BR8014887	17.28	0.62
22-CD01RR8384	22.91	0.65

Diss.[1], average dissimilarity based on the complement of the weighted similarity index of each genotype in relation to the twenty-one remaining.

ma were the most divergent pairs of genotypes, with PI371610/Suprema showing greater oil contents than the previous genotypes. The group of PIs, with greater genetic distances in relation to most groups, showed percentages of oil considered low or normal, except PI371610 that exhibited a median content. Although they are very similar, PI371610 presented oil content 1% higher than PI371611. The genotypes Suprema and CD01RR8384, in turn, showed certain genetic distance and the highest oil contents. These genotypes are promising progenitors, because genetically divergent accessions and with high oil contents can be simultaneously donor sources of additive genes for increase in oil content in soybean.

According to Cruz *et al*. (2011), greater gain with selection is expected in populations with higher averages and genetic variances. The population average seems related to averages of parents, while genetic variance seems related to genetic diversity (Rodrigues *et al*., 2013). Hughes *et al*. (2008) consider the diversity and/or genetic distance as a measure of genetic variability. In this way, genetic diversity estimates can be useful to breeding, once crosses between genetically divergent genotypes are most likely in producing greater genetic variability and heterotic effect in progenies (Filho *et al*., 2010; Riaz *et al*., 2008). Thus, genetic diversity can be considered in predicting the potential of populations in the phase of selection of parents, this way avoiding populations with low genetic variability. In this case, greater predictive capacity of genetic variability for a particular trait is expected when the estimate of genetic

distance considers QTL regions of the trait, instead of random regions of the genome (Melchinger *et al*., 1992; Charcosset *et al*., 1991).

Molecular markers have been the preferably used methodologies to assess genetic relationships between cultivars because information of access genealogy is incomplete, or not available or detailed enough and because of the absence of environmental influence of the molecular markers alternatively to most agronomic traits. (Mulato *et al*., 2010). Moreover, between molecular markers, the microsatellite markers are the most used tools in genetic diversity studies, due to the abundance, high level of polymorphism, multiallelism, and codominant inheritance (Rodrigues *et al*., 2015).

Table 4: Clustering by the Tocher method based on the complement of the weighted similarity index by analysis of microsatellite markers on QTL regions for oil content

G[1]	Genotypes
1	PI371611, PI371610, CD983321RR
2	Garantia, Conquista, Luziânia, BARC-8, Tucunaré, UFV20, CD226RR, CD219RR, Vencedora
3	CD01RR8376, CD01RR8384
4	CD222, BR8014887
5	CS303TNKCA, CD225RR
6	PI181544, PI235347
7	Suprema
8	CD224

G[1], groups.
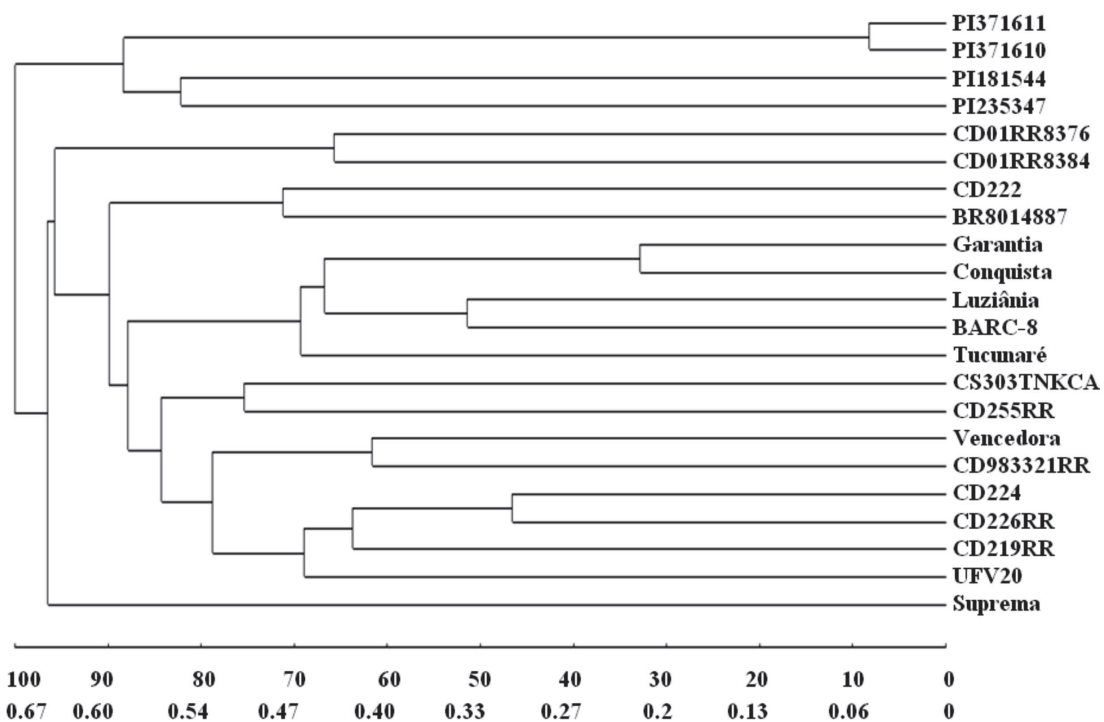
Figure 1: Clustering by the UPGMA method according to the distance estimates based on the complement of the weighted similarity index by analysis of microsatellite markers on QTL regions for oil content.

CONCLUSIONS

The genotypes PI371610/Suprema and Suprema/CD01RR8384 showed genetic distance and higher oil contents in the grains. Suprema and CD01RR8384 have the highest oil contents and are little genetically related. These pairs of genotypes are most likely in producing genetic variability and heterotic effect for new gains in oil content in future populations.

ACKNOWLEDGEMENTS

REFERENCES

Araújo BL, Arnhold E, Oliveira Júnior EA & Lima CF (2014) Parâmetros genéticos em cultivares de sorgo granífero avaliados em safrinha. Revista Trópica, 8:51-59.

Bonato ALV, Calvo ES, Arias CAA, Toledo JFF & Geraldi IO (2006b) Prediction of genetic variability through AFLP-based measure of genetic distance in soybean. Crop Breeding and Applied Biotechnology, 6:30-39.

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

Charcosset A & Essioux L (1994) The effect of population structure on the relationship between heterosis and heterozygosity at marker loci. Theoretical and Applied Genetics, 89:336-343.

Chen O, Zhang Z, Liu C, Xin D, Qiu H, Shan D, Shan C & Hu G (2007) QTL analysis of major agronomic traits in soybean. Agricultural Sciences in China, 6:399-405.

Chung J, Babka HL, Graef GL, Staswick PE, Lee DJ, Cregan PB, Shoemaker RC & Speech JE (2003) The seed protein, oil and yield QTL on soybean linkage group I. Crop Science, 43:1053-1067.

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

Cruz CD, Carneiro PCS & Regazzi AJ (2014) Modelos Biométricos Aplicados ao Melhoramento Genético. Viçosa, Editora UFV. 668p.

Cruz CD & Castoldi FL (1991) Decomposição da interação genótipos x ambientes em partes simples e complexa. Revista Ceres, 38:422-430.

Cruz CD, Ferreira FM & Pessoni LA (2011) Biometria aplicada ao estudo da diversidade genética. Viçosa, Editora Suprema. 620p.

Csanádi G, Vollmann J, Stift G & Lelley T (2001) Seed quality QTL identified in a molecular map of early maturing soybean. Theoretical and Applied Genetics, 103:912-919.

Eskandari M, Cober ER & Rajcan I (2013) Genetic control of soybean seed oil: I. QTL and genes associated with seed oil concentration in RIL populations derived from crossing moderately high-oil parents. Theoretical and Applied Genetics, 126:483-495.

Falconer DS & Mackay TFC (1996) Introduction to quantitative genetics. 4th ed. New York, Longman. 464p.

Fasoula VA, Harris DK & Boerma HR (2004) Validation and designation of quantitative trait loci for seed protein, seed oil, and seed weight from two soybean populations. Crop Science, 44:1218-1225.

Filho JMC, Geraldi IO & Barona MAA (2010) Heterose e distâncias genéticas moleculares para a produção de grãos em soja. Ciência Agrotécnica, 34:940-945.

Hiromoto DM & Vello NA (1986) The genetic base of Brazilian soybean *Glycine max* (L.) Merril) cultivars. Revista Brasileira de Genética, 9:295-306.

Hughes AR, Inouye BD, Johnson MTJ, Underwood N & Vellend M (2008) Ecological consequences of genetic diversity. Ecology Letters, 11:609-623.

Hyten DL, Pantalone VR, Sams CE, Saxton AM, Landau-Ellis D, Stefaniak TR & Schmidt ME (2004) Seed quality QTL in a prominent soybean population. Theoretical and Applied Genetics, 109:552-561.

Leite DC, Pinheiro JB, Campos JB, Di Mauro AO & Unêda-Trevisoli SH (2016) QTL mapping of soybean oil content for marker-assisted selection in plant breeding program. Genetics and Molecular Research, 15:01-11.

Li H, Zhao T, Wang Y, Yu D, Chen S, Zhou R & Gai J (2011) Genetic structure composed of additive QTL, epistatic QTL pairs and collective unmapped minor QTL conferring oil content and fatty acid components of soybeans. Euphytica, 182:117-132.

Li W, Sun D, Du Y, Chen Q, Zhang Z, Qiu L & Sun G (2007) Quantitative trait loci underlying the development of seed composition in soybean (*Glycine max* L. Merr.). Genome, 50:1067-1077.

Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB & Lark KG (1996) Genetic mapping of agronomic traits using recombinant inbred lines of soybean. Crop Science, 36:1327-1336.

Melchinger AE, Boppenmaier J, Dhillon BS, Pollmer WG & Herrmann RG (1992) Genetic diversity for RFLPs in European maize inbreds: Relation to performance of hybrids within versus heterotic groups for forage traits. Theoretical and Applied Genetics, 84:672-681.

Mian MR, Sung-Taeg K & Margaret GR (2009) Microsatellite diversity of soybean genotypes differing in bean pod mottle virus leaf symptom. Canadian Journal of Plant Science, 89:359-367.

Miranda ZFS, Arias CAA, Prete CEC, Kiihl RAS, Almeida LA, Toledo JFF & Destro D (2007) Genetic characterization of ninety elite soybean cultivars using coefficient of parentage. Pesquisa Agropecuária Brasileira, 42:363-369.

Mojema R (1977) Hierarquial grouping methods and stopping rules: an evaluation. The Computer Journal, 20:359-363.

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

Nichols DM, Glover KD, Carlson SR, Specht JE & Diers BW (2006) Fine mapping of a seed protein QTL on soybean linkage group I and its correlated effects on agronomic traits. Crop Science, 46:834-839.

Panthee DR, Pantalone VR, West DR, Saxton AM & Sams CE (2005) Quantitative trait loci for seed protein and oil concentration and seed size in soybean. Crop Science, 45:2015-2022.

Priolli GR, Pinheiro JB, Zucchi MI, Bajay MM & Vello NA (2010) Genetic diversity among brazilian soybean cultivars based on SSR loci and pedigree data. Brazilian Archives of Biology and Technology, 53:519-531.

Qi Z, Wu Q, Han X, Sun Y, Du X, Liu C, Jiang H, Hu G & Chen Q (2011) Soybean oil content QTL mapping and integrating with meta-analysis method for mining genes Euphytica, 179:499-514.

Reinprecht Y, Poysa VW, Yu K, Rajcan I, Ablett GR & Pauls KP (2006) Seed and agronomic QTL in low linolenic acid, lipoxygenase free soybean germoplasm. Genome, 49:1510-1527.

Riaz A, Li G, Quresh Z, Swati MS & Quiros CF (2008) Genetic diversity of oilseed *Brassica napus* inbred lines based on sequence-related amplified polymorphism and its relation to hybrid performance. Plant Breeding, 120:411-415.

Rodrigues JIS, Arruda KMA, Cruz CD, Piovesan ND, Barros EG & Moreira MA (2013) Associação de marcadores microssatélites com teores de óleo e proteína em soja. Pesquisa Agropecuária Brasileira, 48:255-262.

Rodrigues JIS, Arruda KMA, Cruz CD, Piovesan ND, Barros EG & Moreira MA (2015) Divergência em QTLs e variância genética para teores de proteína e óleo em soja. Pesquisa Agropecuária Brasileira, 50:1042-1053.

Rodrigues JIS, Miranda FD, Borges LL, Silva MF, Good-God PIV, Piovesan ND, Barros EG, Cruz CD & Moreira MA (2010) QTL mapping for protein and oil content in soybeans in a Brazilian germplasm. Pesquisa Agropecuária Brasileira, 45:472-480.

Shibata M, Takayama K, Ujiie A, Yamada T, Abe J & Kitamura K (2008) Genetic relationship between lipid content and linolenic acid concentration in soybean seeds. Breeding Science, 58:361-366.

Specht JE, Chase K, Macrander M, Graef BL, Chung J, Markwell JP, Germann M, Orf JH & Lark KG (2001) Soybean response to water: a QTL analysis of drought tolerance. Crop Science, 41:493-509.

Tajuddin T, Watanabe S, Yamanaka N & Harada K (2003) Analysis of quantitative trait loci for protein and lipid contents in soybean seeds using recombinant inbred lines. Breeding Science, 55:133-140.

Vasconcelos ED, Cruz CD, Bhering LL & Resende Junior MFR (2007) Alternative method for clustering analysis. Pesquisa Agropecuária Brasileira, 42:1421-1428.

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

Vieira ESN, Schuster I, Silva RB & Oliveira MAR (2009) Genetic variability in soybean cultivars determined with microsatellite markers in agarose gel. Pesquisa Agropecuária Brasileira, 44:1460-1466.