

Genetic divergence in soybean parents for backcrossing programs

Rita Maria Alves de Moraes^{1*}, Cosme Damião Cruz², Everaldo Gonçalves de Barros², and Maurilio Alves Moreira³

Received 3 August 2004

Accepted 18 June 2005

ABSTRACT - Parents are generally selected for their performance and genetic divergence. The selection of the genetically closer plants to the recurrent parent reduces the number of backcrossings required to recover the genome. Our study aimed at the selection of plants genetically closer to each other for backcrossings. Nine soybean genotypes with a normal protein content and 18 high protein donor lines of the Breeding program of the BIOAGRO/UFV were evaluated with 57 pairs of primer microsatellites. The genetic distances (GD) varied from 0.08 to 0.74 and allowed the identification of the closest genotypes. These were clustered according to their GD by the UPGMA (five groups), Single Linkage (six groups) and Tocher methods (nine groups). The donor genotypes closer to the recurrent parent were chosen to reduce the number of backcrossings needed to achieve high protein, productive soybean lines.

Key words: fingerprinting, microsatellites, genetic distances, molecular markers.

INTRODUCTION

To evaluate superior hybrid combinations, biometric techniques based on the quantification of heterosis as in the diallel analyses and predictive procedures are commonly being used. Diallel analyses have the disadvantage of requiring the evaluation of each parent in all possible combinations, which may be impossible in the case of studies involving many parents. The predictive methods which dispense with the previous establishment of the hybrid combinations have attracted interest. These methods base upon the morphological, physiological and

molecular differences, among others, which the parents present in the evaluation of divergence, generally quantified by a measure of dissimilarity. The choice of the most adequate method is determined in function of the desired precision, of the easiness of analysis and by the way the data were obtained (Cruz 2001).

When molecular genetic maps first came up (Bernatzky et al. 1986, Shoemaker et al. 1992, Morgante et al. 1994, Akkaya et al. 1995, Shoemaker and Specht 1995), map-based molecular markers became a powerful tool in genetic improvement. The term molecular improvement is used to describe improvement programs that are assisted

¹Embrapa Trigo, BR 285, km 174, C. P. 451, 99.001-970, Passo Fundo, RS, Brasil. *E-mail: rita@cnpt.embrapa.br

²Departamento de Biologia Geral, BIOAGRO, Universidade Federal de Viçosa (UFV), 36.571-000, Viçosa, MG, Brasil

³Departamento de Bioquímica e Biologia Molecular, BIOAGRO, UFV

by DNA-based diagnostics. One of the most commonly used classes of markers in improvement is the class of microsatellites or SSRs (Simple Sequence Repeats) which are sequences of 1 to 6 pb repeated across the entire genome. SSRs are PCR-amplified with complementary primers to the conserved regions that flank the SSR region. The polymorphism of the fragments (alleles) results in variations in the length of the SSR replications which are separated in agarose gel or polyacrylamide (Narvel et al. 2000). Owing to their vast distribution across the chromosomes, the SSR markers allow a more in-depth evaluation of the entire genome, resulting in an enhanced monitoring of the gene introgression during the backcrossing process. Besides, these markers provide support for the selection of base populations for improvement programs, generating information for genotype clustering and crossing schedules (Brown-Guedira et al. 2000). Depending on the number of evaluated plants, the estimates of genetic distances are somewhat unmanageable while the evaluation of relationships among plants hampers the work with selection. To solve this problem, one can use statistical procedures such as clustering analysis, construction of dendrograms, estimate of principal components or multidimensional scaling (Santos 1994). Based on the underlying genetic dissimilarities which were determined by the microsatellite analysis, estimated by the coefficient of coincidence simples and grouped by the UPGMA, Single linkage and Tocher methods, our study had the objective to select the soybean parents genetically closest to each other, thus aiming at the introgression of high-protein alleles into elite soybean varieties by backcrossings.

MATERIAL AND METHODS

Plant material and crossings

Three varieties and six lines of soybean from the Cooperativa Central Agropecuária de Desenvolvimento Tecnológico e Econômico Ltda - COODETEC (Cooperative

Center for Technological and Economical Development in Agriculture and Animal Husbandry) were used as recurrent parents (Table 1). Eighteen high-protein lines (47 to 52%) of the soybean breeding program of the BIOAGRO/UFV (Table 1) were used as donors of the trait high protein content/concentration with the objective of incorporating alleles for the trait in question into the COODETEC varieties. After determining the proteins in the seeds, the donors were selected according to the Kjeldahl method to quantify the total nitrogen described by the Association of Official Analytical Chemists (AOAC 1975) with modifications. The genotypes were fingerprinted to define the closest ones in order to accelerate the backcrossing program.

DNA extraction and purification

DNA samples were extracted by the method described by McDonald et al. (1994) with modifications. To 50 mg seeds 1000 μL of extraction buffer were added containing Tris-HCl 0.2 M, NaCl 0.28 M, 25 mM EDTA and 10% SDS, extracted in Polytron and centrifuged immediately for 10 min at 14.000 rpm. The supernatants were transferred to new tubes, adding 10 μL proteinase K 10 mg mL^{-1} and 10 μL CaCl_2 1 mM, and immersed in a 55 °C water bath for 30 min. The samples were then supplemented with 900 μL isopropanol and left to settle for 2 min. After this time the samples were centrifuged for 10 min at 14.000 rpm. The supernatants were discarded, the precipitate was dried for 15 min at room temperature and resuspended in 10 mM Tris-EDTA solution and 1 mM with 60 $\mu\text{g mL}^{-1}$ RNase and immersed in a water bath for 90 min. The samples were precipitated again by adding 900 μL isopropanol and left to settle during 2 min. Thereafter, the samples were centrifuged once more for 10 min at 14.000 rpm and the supernatants discarded. The precipitates were finally resuspended in TE (10 and 1 mM). The DNA concentration was estimated spectrophotometrically by a reading of the absorbance at 260 nm (each absorbance

Table 1. Parents used in the backcrossing program aiming at the incorporation of high-protein alleles into COODETEC soybean varieties. Numbers in brackets represent the number used in the analyses and tables

Parent	Genotype
Donor lines	OC 128PT258(1); OC 14PT260(2); CD 203PT30-1;(3) CD 203PT30-3(4)CD 203PT30-2(5); CD 202PT31-3(6); CD 202PT31-1(7); CD 202PT31-2(8); CD 202PT111-1(9); OC 671PT100-1(10); OC 671PT46-1(11); OC 671PT46-2(12)OC 13PT158-2(13); OC 13PT158-3(14) CD 201PT82-3(17); B1PTA25(25)B1PTA97(26); B1PTA217(27)
Recurrent lines	CD 983321(15); CD 985015(16); OC 953194(18); CD 211(19); CD 210(20); OC 953006(21); OC 953312(22); CD 983343(23) CD 204(24)

unit corresponded to 50 $\mu\text{g mL}^{-1}$ double-stranded DNA (Sambrook et al. 1989).

Amplification reactions

The microsatellites were amplified in a total volume of 15 μL , containing 12.5 μM Tris-HCl (pH 8.3), 62.5 μM of KCl, 2.5 mM of MgCl_2 , 125 μM of each one of the deoxynucleotides (dATP, dTTP, dGTP and dCTP), 0.6 μM of each primer, one unit of the enzyme Taq-polymerase and 18 ng DNA. The amplifications were performed in a Perkin Elmer 9600 thermal cycler programmed for an initial phase of 7 min at 72 °C, followed by 30 cycles of 1 min at 94 °C, 1 min at 50 °C and 2 min at 72 °C plus a final step of 7 min at 72 °C. The amplified fragments were separated by 10% nondenaturant polyacrylamide gel electrophoresis during four hours at 120 V, and stained with ethidium bromide solution (0.2 $\mu\text{g mL}^{-1}$). After the electrophoresis, the gels were photographed by the digital Eagle Eye II (Stratagene) image system.

Selection of microsatellite primer pairs

To obtain the molecular data of the genotypes, 57 pairs of polymorphic SSR primers based on the integrated map of soybean (Cregan et al. 1999) were selected as listed below: (1)Sat009, (2)Satt005, (3)Satt066, (4)Satt070, (5)Satt079, (6)Satt082, (7)Satt115, (8)Satt131, (9)Satt135, (10)Satt136, (11)Satt146, (12)Satt157, (13)Satt163, (14)Satt173, (15)Satt179, (16)Satt181, (17)Satt183, (18)Satt186, (19)Satt187, (20)Satt197, (21)Satt200, (22)Satt212, (23)Satt215, (24)Satt235, (25)Satt238, (26)Satt242, (27)Satt250, (28)Satt260, (29)Satt262, (30)Satt263, (31)Satt269, (32)Satt270, (33)Satt275, (34)Satt290, (35)Satt304, (36)Satt302, (37)Satt309, (38)Satt311, (39)Satt314, (40)Satt320, (41)Satt322, (42)Satt323, (43)Satt353, (44)Satt358, (45)Satt354, (46)Satt362, (47)Satt373, (48)Satt380, (49)Satt381, (50)Satt385, (51)Satt415, (52)Satt512, (53)Satt527, (54)Satt536, (55)Satt545, (56)Satt551 and (57)Satt574.

Estimators of genetic distances

The genetic distances between the parents were estimated based on the data of the analyses with the microsatellites, using the complement of the simple coincidence (SC) coefficient as measure of dissimilarity. The SC between two plants was obtained by dividing the total of microsatellite loci containing common alleles by the total number of evaluated loci. The data matrix was established attributing value 2 to the microsatellite locus

with two copies of the homozygous dominant allele, value 1 to each allele when the locus was heterozygous, value 0 when the locus had two copies of the recessive homozygous allele and value 9 when the reaction failed.

Clustering methods

Three clustering methods were used for an enhanced visualization of the groups for crossings, since, depending on the number of evaluated genotypes, the matrix of genetic distance does not allow a general vision of the relationship among the genotypes. These were the agglomerative, UPGMA (unweighted pair group method with arithmetic mean) and Single linkage methods, and to obtain mutually exclusive groups the optimization method of Tocher.

RESULTS AND DISCUSSION

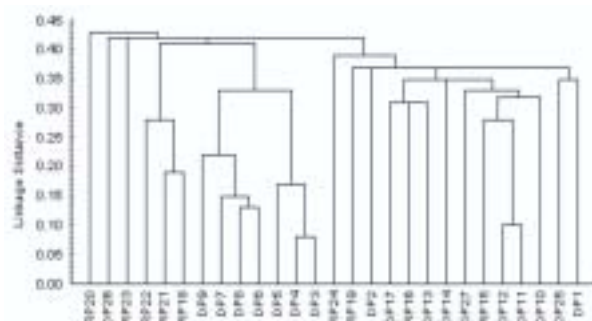
The studies realized into genetic divergence in plants are generally related to the objectives of characterizing varieties (Priolli et al. 2002), seeking information on the genetic constitution of germplasm (Bao-Rong L 2004) and having access to the variability in the crops (Kudryavtsev et al. 2004), and, from a practical point of view, used to verify the genealogies, in the planning of crossings with superior genotypes and the exploitation of heterosis.

The use of microsatellite markers (SSR) for the characterization of plant genotypes has become quite common in breeding programs (Narvel et al. 2000, Tanya et al. 2001). Figure 1 presents the amplification products of primer Satt181 in 10% polyacrylamide gel. Table 2 shows the distribution of the 57 SSR pairs selected in the present study across the linkage groups according to information from the integrated map of soybean (Cregan et al. 1999), four of which were not mapped. Eighteen of the 24 linkage groups of the integrated map for soybean were mapped indicating that the genome of the soybean was well sampled by the SSR markers. In this sense, Hospital et al. (1992) observed that molecular markers are useful in gene introgression programs provided that they are well distributed across the genome and their positions known.

The genetic distances, when using the complement of the coefficient of coincidence simples (SC), varied from 0.08 to 0.74, demonstrating the great capacity of SSR markers to evaluate the genetic diversity in soybean. Using 12 SSR markers, Priolli et al. (2002) observed genetic distances which varied, in the mean, from 0.4 to 0.9 among 186 Brazilian soybean cultivars. The smallest genetic distance (0.08) was found between two sibling donor lines

Table 2. Distribution of the 57 SSR markers selected for the analysis of genetic divergence in the linkage groups of the integrated map of soybean of the USDA, University of Utah and University of Nebraska (from Cregan et al. 1999)

Linkage group	Microsatellite marker
A1/U07	Satt200; Satt385; Satt545
A2	Satt187
B1/U04	Satt415
P	Satt197
B2/U26	Satt066; Satt070; Satt304
C2/U09	Satt322; Satt079
D1a	Satt179; Satt320
D1b	Satt290; Satt005; Satt157
D2	Satt311; Satt082; Satt186; Satt135
E	Satt212; Satt263
F	Satt146; Satt269; Satt362
G/U05/CH11+CH14	Satt131; Satt163; Satt275; Satt235; Satt115; Satt309
H/U10	Satt181; Satt302; Satt353; Satt314
I	Satt270; Satt354
J/U01	Satt183; Satt215; Satt380
K	Satt242; Satt381; Satt260
L/U14	Satt238; Satt527; Satt373
M/U11	Satt250; Satt536; Satt551; Satt323
O	Satt358; Satt173; Satt262
C1; Q; W; N and Y	were not mapped
Not linked	Sat009; Satt136; Satt512 and Satt574

**Figure 1.** Amplification products of primer Satt181 in the 27 parents in 10% polyacrylamide gel. The last column on the right represents the pattern of molecular mass of pUC18 plasmid digested with enzyme Msp I**Figure 2.** Single linkage containing the 27 parents used in the backcrossing program. RP indicates a recurrent and DP a donor parent

for high-protein alleles CD 203PT30-1(3) and CD 203PT30-3 (4). The largest genetic distance (0.74) was observed between a donor line of high-protein alleles CD 203PT30-3 (4) and the variety with normal protein content CD 983321 (15). There is a large genetic variability among the donor parents of high-protein alleles (0.08 to 0.70). This result is important since it indicates that these parents, besides representing excellent sources of high-protein alleles, have a great genetic diversity, which is desirable in parents for allele introgression by means of backcrossings. A large

genetic variability (between 0.28 and 0.72) was also observed among the recurrent parents.

The matrix of dissimilarity was used for the cluster analyses by the agglomerative UPGMA and Single linkage methods. A point of significance for the dendrograms of the Single linkage of a linkage distance of 0.40 was defined (Figure 2), and of 0.50 for UPGMA (Figure 3). Next, the groups formed in each dendrogram were established according to the given limit. The dendrogram obtained by the Single linkage method (Figure 2) allowed the formation

of six groups with the following constitutions: Group 1 formed by 14 parents; Group 2 by 7; Group 3 by 3; Group 4 by 1 parent; Group 5 by 1 parent and Group 6 formed by 1 parent. The dendrogram obtained by the UPGMA method (Figure 3) united the parents in 5 distinct groups as discriminated in the following: Group 1 formed by 10 parents; Group 2 by 3; Group 3 by 3; Group 4 by 4; and Group 5 formed by 7 parents.

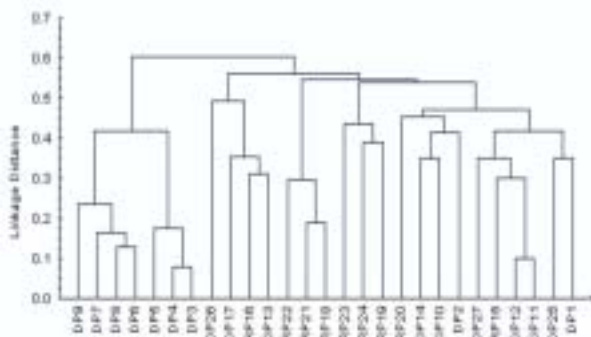


Figure 3. UPGMA containing the 27 parents used in the backcrossing program. RP indicates a recurrent and DP a donor parent

To obtain mutually exclusive groups, the optimization method of Tocher was used which identified nine groups (Table 3): group 1 formed by 7 parents, group 2 by 6; group 3 by 3; group 4 by 3; group 5 by 3; group 6 by 2; group 7 by 1 parent; group 8 by 1 parent, and group 9 formed by 1 parent. A comparison of the three clustering methods (Table 3) demonstrated that the Tocher method was the one that differentiated the largest number of groups, subdividing group 1. Nevertheless, the three methods agreed in the clustering of most of the analyzed genotypes; the Single linkage was closer to Tocher than UPGMA.

Similar results were observed by Tanya et al. (2001), who managed to unite 16 soybean varieties (genotypes) in 5 distinct groups using 20 SSR markers by the UPGMA method. In the same sense, Baranek et al. (2002), using RAPD markers, succeeded in uniting 19 soybean genotypes evaluated in five distinct groups, using the coefficient of Nei and the UPGMA clustering method. The donor lines with a common origin mostly grouped in the same cluster, as was the case of the donor lines CD 202PT (6, 7, 8 and 9) and CD 203 PT (3, 4, and 5), while the donor lines derived from OC 671 PT (10, 11 and 12) grouped into another cluster. The case of some sibling lines which grouped in different clusters, as for instance the lines derived from B1PTA (25, 26 and 27), can be explained by the reduced number of backcrossings, three, (without realizing fingerprinting of the backcrossed plants) which was insufficient to recover the genome of the recurrent

parent of around 93.75%. The discussions on the recovery of the genome of the recurrent parent during the backcrossing generations has focused on the proportional expected values of the recurrent parent without taking the existing variation around the expected mean value into consideration (Openshaw et al. 1994). In this sense, the SSR markers (provided that they are well-distributed) can define the genetic load of the genotypes with more accuracy an increase the effectivity of the backcrossing method.

The donor parents for high-protein alleles were selected based on their genetic dissimilarities, obtained from the matrix of genetic distance with the recurrent parents (Table 4). As the table shows, between 3-5 donor parents with genetic distances varying from 0.28 to 0.57 were selected for crossings with each recurrent parent. The maintenance of the genetic diversity with more than one donor for high-protein alleles is extremely important. In spite of being a trait controlled by few genes, it is a complex trait and extremely influenced by the environment, indicating the action of smaller genes on the trait. The numbers in brackets in Table 4 correspond to the codes assigned to the genotypes in Table 1. Thus, five donor parents (distances from 0.28 to 0.48) were selected for the crossing with CD 983321 (15), four donors (distances from 0.31 to 0.56) for CD 985015 (16), three donors (distances from 0.49 to 0.51) for OC 953194 (18), three donors (distances from 0.37 to 0.52) for CD 211 (19), four donors (distances from 0.43 to 0.47) for CD 210 (20), four donors (distances from 0.41 to 0.47) for OC 953006 (21), four donors (distances from 0.41 to 0.43) for OC 953312 (22), four donors (distances from 0.48 to 0.54) for CD 983343 (23), and four donors (distances from 0.42 to 0.57) for CD 204 (24). Faleiro et al. (2004) used RAPD markers to assist a backcrossing program that aimed at the introduction of rust and anthracnose-resistance genes into common bean. The DNA fingerprinting of the resistant plants was used to select the genetically closest plants to the recurrent parent. The genetic distances in relation to the recurrent parental of the resistant plants varied from 9 to 59% for BC_1 , from 33% for BC_2 and from 0 to 7% for BC_3 . Five rust and anthracnose-resistant lines were obtained after only three backcrossings with a genetic distance of approximately 0% in relation to the recurrent parental.

The next stage in the backcrossing program will be the evaluation and selection of BC_1F_2 plants for high protein content/concentration and fingerprinting of these plants for a later selection of the genetically closest to the recurrent parent. This way, our study aims to obtain high-protein soybean lines with a smaller number of backcrossings, oriented by SSR markers in fingerprinting.

Table 3. Comparison between the clustering methods

Cluster	Single linkage	UPGMA	Tocher
1	1, 10, 11, 12, 15, 25, 27, 14, 13, 16, 17, 2, 19, 24	1, 11, 12, 15, 25, 27, 2, 10, 14, 20	3, 4, 5, 6, 7, 8, 9 1, 11, 12, 15, 25, 27
2	3, 4, 5, 6, 7, 8, 7, 9	19, 24, 23	18, 21, 22
3	18, 21, 22	18, 21, 22	13, 16, 17
4	23	13, 16, 17, 26	2, 10, 14
5	26	3, 4, 5, 6, 7, 8, 9	19, 24
6	20		23
7			20
8			26
9			

^aIndividuals that are common to the three clusters appear in bold

Table 4. Selection of the donor parents of high-protein alleles based on their genetic dissimilarities with the recurrent parents

Recurrent parent	Donor parent	Genetic distance
CD 983321 (15)	OC 128PT258 (1)	0.48
	OC 14PT260 (2)	0.48
	OC 671PT46-1 (11)	0.28
	OC 671PT46-2 (12)	0.32
	OC 13PT158-3 (14)	0.45
CD 985015 (16)	CD 202PT31-3 (6)	0.55
	OC 13PT158-3 (13)	0.31
	CD 203PT30-1 (3)	0.56
	CD 201PT82-3 (17)	0.31
OC 953194 (18)	OC 14PT260 (2)	0.49
	CD 202PT31-1 (7)	0.49
	CD 202PT111-1 (9)	0.51
CD 211 (19)	OC 671PT46-1 (11)	0.40
	OC 671PT46-2 (12)	0.37
	OC 671PT100-1 (10)	0.52
CD 210 (20)	OC 671PT100-1 (10)	0.43
	OC 13PT158-3 (14)	0.44
	OC 13PT158-2 (13)	0.47
	OC 671PT46-2 (12)	0.46
OC 953006 (21)	CD 202PT31-1 (7)	0.42
	CD 202PT111-1 (9)	0.41
	OC 671PT100-1 (10)	0.42
	OC PT13158-2 (14)	0.47
OC 953312 (22)	CD 202PT31-3 (6)	0.42
	CD 202PT31-1 (7)	0.41
	CD 202PT31-2 (8)	0.43
	CD 202PT111-1 (9)	0.41
CD 983343 (23)	OC 128PT258 (1)	0.48
	CD 202PT31-1 (7)	0.53
	CD 202PT31-1 (8)	0.54
	OC 13PT158-2 (13)	0.52
CD 204 (24)	OC 128PT258 (1)	0.49
	OC 671PT46-1 (11)	0.57
	OC 671PT46-2 (12)	0.54
	B1PTA25 (25)	0.42

The number assigned to each parent in Table 1 is indicated in brackets

ACKNOWLEDGEMENTS

This study was supported by grants from CNPq and FAPEMIG to MA Moreira. RMA Moraes was recipient to DS scholarship from CNPq.

Divergência genética entre genitores de soja para uso em programas de retrocruzamentos

RESUMO - Genitores geralmente são selecionados pelo seu desempenho e divergência genética. A seleção de indivíduos geneticamente mais próximos do genitor recorrente reduz o número de retrocruzamentos necessários à recuperação do genoma. Este trabalho objetivou selecionar indivíduos geneticamente mais próximos entre si para uso em retrocruzamentos. Nove genótipos de soja com teor normal de proteína e dezoito linhagens doadoras de alto teor de proteína foram avaliados utilizando 57 pares de primers microssatélites. As distâncias genéticas (DG) variaram de 0,08 a 0,74 permitindo identificar os genótipos mais próximos entre si. Os genótipos foram agrupados de acordo com suas DG utilizando-se os métodos UPGMA (cinco grupos), Vizinho Mais Próximo (seis grupos) e Tocher (nove grupos). Para cada genitor recorrente foram escolhidos os genótipos doadores com as menores distâncias, de modo a reduzir o número de retrocruzamentos necessários para obtenção de linhagens de soja com alto teor de proteína e produtivas.

Palavras-chave: fingerprinting, microssatélites, distâncias genéticas, marcadores moleculares.

REFERENCES

- Akkaya MS, Shoemaker RC, Specht JE, Bhagwat AA and Cregan PB (1995) Integration of simple sequence repeat DNA markers into a soybean linkage map. **Crop Science** **35**: 1439-1445.
- Association of Official Analytical Chemists – AOAC (1975) **Official methods of analysis**. Washington, DC, 1094p.
- Baranek M, Kadlec M, Raddova J, Vachun M and Pidra M (2002) Evaluation of genetic diversity in 19 *Glycine max* (L.) Merr. accessions included in the Czech national collection of soybean genotypes. **Czech Journal Genetics Plant Breeding** **38**: 69-74.
- Bernatzky R and Tanksley SD (1986) Toward a saturated map of tomato based on isozymes and random cDNA sequences. **Genetics** **112**: 887-898.
- Bao-Rong L (2004) Conserving biodiversity of soybean gene pool in the biotechnology era. **Plant Species Biology** **19**: 115-125.
- Brown-Guedira GL, Thompson RL and Warburton ML (2000) Evaluation of genetic diversity of soybean introductions and north American ancestors using RAPD and SSR markers. **Crop Science** **40**: 815-823.
- Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, Van Toai TT, Lohnes DG, Chung J and Specht JE (1999) An integrated genetic linkage map of the soybean genome. **Crop Science** **39**: 1464-1490.
- Cruz CD (2001). **Programa Genes: versão Windows; aplicativo computacional em genética e estatística**. UFV, Viçosa, 648p.
- Faleiro FG, Ragagnin VA, Moreira MA and Barros EG (2004). Use of molecular markers to accelerate the breeding of common bean lines resistant to rust and anthracnose. **Euphytica** **138**: 213-218
- Hospital F, Chevalet C and Mulsant P (1992) Using markers in gene introgression breeding programs. **Genetics** **132**: 1199-1210.
- Kudryavtsev AM, Martynov SP, Broggio M and Buiatti M (2004) Evaluation of polymorphism at microsatellite loci of spring durum wheat (*Triticum durum* Desf.) varieties and the use of SSR-based analysis in phylogenetic studies. **Russian Journal of Genetics** **40**: 1102-1110.
- McDonald MB, Elliot EJ and Sweeney PM (1994) DNA extraction from seeds for RAPD analyses in varietal identification studies. **Seed Science & Technology** **22**: 171-176.
- Morgante M, Rafalski A, Biddle P, Tingey S and Olivieri AM (1994) Genetic mapping and variability of seven soybean sequence repeat loci. **Genome** **37**: 763-769.
- Narvel JM, Fehr WR, Chu WC, Grant D and Shoemaker RC (2000) Simple sequence repeat diversity among soybean plant introductions and elite genotypes. **Crop Science** **40**: 1452-1458.

- Openshaw SJ, Jarboe SG and Beavis WD (1994) Marker-assisted selection in backcross breeding. In: Lower R (ed.) **ASHS/CSSA Joint Plant Breeding Symposium on Analysis of Molecular Marker Data**. Oregon State University, Corvallis, p. 41-43.
- Priolli RHG, Mendes-Junior CT, Arantes NE and Contel EPB (2002) Characterization of Brazilian soybean cultivars using microsatellite markers. **Genetics and Molecular Biology** **25**: 185-193.
- Sambrook J, Fritsch Ef and Maniatis T (1989) **Molecular cloning - a laboratory manual**. 2nd ed., v.3, Cold Spring Harbor Laboratory, New York.
- Santos JB (1994) Emprego de marcadores moleculares no melhoramento de plantas. **Horticultura Brasileira** **12**: 282-286.
- Shoemaker RC, Guffy RD, Lorenzen LL and Specht JE (1992) Molecular genetic mapping of soybean: map utilization. **Crop Science** **32**: 1091-1098.
- Shoemaker RC and Specht JE (1995) Integration of the soybean molecular and classical genetic linkage groups. **Crop Science** **35**: 436-446.
- Tanya P, Srinives P, Toojinda T, Vanavichit A, Bo-Keun H, Jeong-Suk B, Jung-Kyung M and Suk-Ha L (2001) Evaluation of genetic diversity among soybean genotypes using SSR and SNP. **Korean Journal Crop Science** **46**: 334-340.