Soybean Parent Selection Based on Genetic Diversity

Valéria Carpentieri-Pípolo*[1], Antônio Eduardo Pípolo[2], Flávio André Martins da Silva[3] and Marcos Rafael Petek[3]

[1] Departamento de Agronomia da UEL, Caixa Postal 6001, CEP 86051-990, Londrina, PR. [2] Embrapa Soja, Caixa Postal 231, CEP 86001-970, Londrina, PR. [3] Graduando em Engenharia Agronômica da UEL, Caixa Postal 6001, CEP 86051-990, Londrina – PR, Brazil

ABSTRACT

Thirty-four soybean lines were assessed for twelve traits. The genetic distances were estimates using multivariate techniques, to identify parents to be included in breeding programs for hybridization. Grouping by the Tocher method, from generalized Mahalanobis distances, divided the 34 lines into four groups. The most important agronomic traits, weight of seeds per plot, plant height, height of first pod and days to maturity were considered when recommending for crossing. The following crosses were recommended based on the genetic divergence and the key agronomic traits: lines 23, 10, 2, 27 and 25 (group I) with genotype 6 (group II) and genotype 16 (group III). Thus only ten crosses would be made, representing only 2% of the total crosses which could be made in the partial diallel among the 34 lines assessed, which would allow up to 561 combinations.

Key words: soybean, genetic divergence, Mahalanobis, multivariate analysis

INTRODUCTION

The most important step in a breeding program is the choice of parents with good performance and wide genetic base. Analysis of the parents' isolated behavior and the result of diallel crosses are options in the choice of parents. But the need to carry out a large number of artificial crosses and experiments involving a large number of hybrids limit the use of diallel crosses.

Thus measures of the genetic divergence, ahead of making any cross is made, may help breeders to concentrate their efforts only on the most promising combinations. Heterosis, manifested in the crosses, is directly related to genetic divergence among the parents (Falconer, 1981).

Destro (1991) observed that high D^2 (Mahalanobis generalized distance) in soybean was not always associated with high yield, nevertheless, low D^2 was generally associated with low yield. The best results were obtained in the crosses with intermediate D^2. Therefore, in the choice of parents to be used in hybridization, emphasis should be placed both on the cultivars performance and on its genetic distances (Miranda *et al*., 1988; Cruz, 1990; Rangel *et al*, 1991; Destro, 1991).

Carvalho *et al.* (1995) used the genetic divergence method to access a diallel cross among seven cotton cultivars for resistance to ramulose (*Colettotrichum gossypii* var. *cephalosporioides*). Forty nine treatments were divided into seven groups and the traits such as plant height, fiber percentage and short cycle most contributed to divergence among the genotypes.

Pípolo *et al.* (1995) assessed the performance of 25 genotypes for eight traits to identify suitable parents for a guandu breeding program. The genotypes were divided into five groups and 24 crosses were recommended when the traits plant height, protein percentage and oil percentage were considered, 8% of the total possible crosses involving the 25 suitable genotypes.

Amaral Júnior *et al.* (1996) estimated the genetic divergence among eight strawberry genotypes and indicated the most promising combinations and described the most important morphoagronomic traits in the characterization of variability. Four

---

[*] Author for correspondence

groups were established by the Tocher method and four crossing combinations were recommended to widen the genetic base of the species.

Souza (1996) used enzymatic and agronomic data to estimate the genetic diversity among Caribbean cherry accessions. Seven enzymatic systems and six agronomic traits were observed to characterize the 16 Caribbean cherry accessions. The divergence estimates by enzymatic standards did not correlate significantly with the agronomic characteristics because the accessions grouped in four groups for the isozymatic systems and in only three for the agronomic traits. However, for breeding programs, two crossing combinations were recommended among three accessions because of the high genetic diversity found both in the isozymatic and the agronomic data.

This study was carried out to estimate the genetic divergence of 34 soybean lines to identify suitable parents for genetic breeding programs.

MATERIAL AND METHODS

Thirty-four soybean lines were accessed (Table 1) from the National Center for Soybean Research (Embrapa-Soja), The Agronomic Institute (IAC), CODETEC and FT – Pesquisa e sementes) at the Faculty of Agrarian and Veterinary Sciences, UNESP Jaboticabal, SP, during 1990.

A randomized complete block design was used with 34 treatments and three replications. The plots were four 5m rows, spaced 0.50m. The soil in the experimental area was classified as distrophic Dark Red Latossol with medium texture and fertility.

At harvest, stage R8 (Fehr & Cavines 1977), 15 plants from two central rows of each plot were collected to determine the following agronomic traits: YP: yield per plot (g); PH: plant height (cm); HFP: height of first pod (cm); NPP: number of pods per plant; NP1G: number of pods with one grain; NP2G: number of pods with two grains; NP3G: number of pods with 3 grains; SD: stem diameter (cm); SWB: stem weight plus bark (g); WS: weight of a hundred seeds (g); DM: days to maturity (number of days from emergence to 90% ripe pods); DF: days to flowering (number of days from emergence to 90% open flowers).

In addition to the analyses of variance, multivariate analysis and comparison of measurements using the criteria proposed by Duncan, the following were also obtained: a) estimate of Mahalanobis generalized distance (1936); b) establishment of similar groups using the optimization method proposed by Tocher, quoted by Rao (1952) with a reduced example of this method presented by Singh & Chaudhary (1977) and c) graphic analysis of divergence, using canonic variables. These procedures will supply information on genetic distance, by the scattering of the scores of each genotype in the cartesian system, whose axes are established by the first canonic pair (Rao, 1952).

Table 1 - Means for 12 agronomic traits evaluated in 34 soybean lines.

Lines	YP	PH	HFP	NPP	NP1G	NP2G	NP3G	SD	SWB	WS	DF	DM
23. BR-88 40.242	3562.00 a	91.33 abc	16.66 abc	313.00 abc	103.66 abcde	181.00 abc	28.33 abc	0.603 abcde	3679.66 abc	24.33 abcdef	46	123
10. BR-88 40.152	3423.00 a	87.66 abcd	15.33 bc	291.66 abcd	130.00 abcde	116.33 cd	45.00 ab	0.527 cde	3146.66 abcd	24.66 abcde	50	131
2. IAC-Foscarim	3407.33 ab	87.66 abcd	20.00 abc	388.66 a	130.66 abcde	211.66 ab	46.33 a	0.600 abcde	4888.66 a	23.66 bcdefg	51	126
27. BR-88 40.354	3332.00 ab	85.00 abcd	14.00 bc	295.00 abcd	126.00 abcde	133.33 abcd	35.66 abc	0.583 bcde	3404.33 abcd	21.66 ghi	49	124
9. BR-88 40.150	3164.33 ab	87.33 abcd	14.33 bc	266.66 abcd	92.66 bcde	146.33 abcd	27.66 abc	0.517 cde	4461.66 abc	24.66 abcde	52	129
25. BR-88 40.255	3129.66 abc	89.00 abc	20.00 abc	285.33 abcd	139.00 abcde	126.66 bcd	19.66 abc	0.450 de	2693.00 bcd	20.33 i	51	128
1. IAC-16	3113.00 abc	83.33 abcde	12.66 bc	319.66 abc	136.66 abcde	162.33 abcd	20.66 abc	0.523 cde	3670.66 abc	26.00 ab	48	126
14. BR-88 40.192	3081.00 abc	79.66 abcde	20.00 abc	279.66 abcd	118.33 abcde	135.33 abcd	26.00 abc	0.443 de	3555.33 abcd	25.00 abcd	49	132
22. BR-88 40.219	3030.66 abc	93.66 ab	15.33 bc	386.66 ab	114.33 abcde	224.33 a	48.00 a	0.750 ab	4515.00 abc	22.66 defghi	47	129
12. BR-88 40.187	2969.66 abc	73.66 de	13.33 bc	336.00 abc	128.00 abcde	174.00 abc	34.00 abc	0.710 abc	3928.66 abc	23.00 cdefgh	50	124
34. BR-88 40.103	2936.33 abc	88.66 abc	21.33 ab	320.66 abc	125.00 abcde	164.33 abcd	31.33 abc	0.583 bcde	3724.00 abc	22.00 fghi	48	130
8. BR-88 40.062	2919.66 abc	92.33 abc	19.00 abc	345.00 abc	157.00 ab	160.00 abcd	28.00 abc	0.587 bcde	4764.00 ab	21.66 ghi	50	130
18. BR-88 40.204	2877.66 abc	87.33 abcd	17.66 abc	324.33 abc	141.66 abcd	153.66 abcd	29.00 abc	0.577 bcde	3751.00 abc	23.33 cdefgh	48	126
19. BR-88 40.205	2849.66 abc	80.33 abcde	20.66 ab	317.66 abc	110.00 abcde	168.66 abc	39.00 abc	0.580 bcde	3990.66 abc	21.66 ghi	49	128
13. BR-88 40.189	2845.33 abc	89.33 abc	18.33 abc	360.00 ab	172.00 a	164.66 abcd	23.33 abc	0.597 abcde	3635.33 abc	25.00 abcd	46	127
26. BR-88 40.257	2817.33 abc	83.33 abcde	16.33 abc	296.33 abcd	125.00 abcde	145.33 abcd	26.00 abc	0.517 cde	3768.33 abc	21.33 ghi	49	127
5. BR-88 40.055	2807.33 abc	70.00 e	18.66 abc	328.66 abc	113.33 abcde	178.66 abc	36.66 abc	0.560 bcde	3964.00 abc	23.00 cdefgh	46	126
29. BR-88 40.502	2807.33 abc	81.33 abcde	25.66 a	269.66 abcd	109.00 abcde	129.33 bcd	31.33 abc	0.543 cde	4070.66 abc	22.33 efghi	47	130
24. BR-88 40.249	2733.00 abc	87.33 abcd	17.00 abc	295.00 abcd	128.00 abcde	137.66 abcd	29.33 abc	0.550 cde	1550.33 d	21.33 ghi	51	130
4. FT-Cometa	2699.66 abc	78.00 cde	21.33 ab	297.66 abcd	136.66 abcde	135.66 abcd	25.33 abc	0.560 bcde	3475.33 abcd	23.66 bcdefg	47	125
6. BR-88 40.057	2628.66 abc	90.33 abc	12.33 bc	201.66 cd	81.33 cde	98.33 cd	22.00 abc	0.423 e	2479.66 cd	26.00 ab	49	130
20. BR-88 40.210	2596.33 abc	81.33 abcde	15.00 bc	294.66 abcd	153.00 abc	110.00 cd	31.66 abc	0.560 bcde	3884.00 abc	26.33 a	48	130
7. BR-88 40.058	2568.66 abc	81.00 abcde	18.33 abc	284.33 abcd	123.00 abcde	122.66 bcd	38.66 abc	0.593 abcde	3359.66 abcd	22.66 defghi	49	131
33. BR-88 40.052	2553.00 abc	94.00 a	21.33 ab	256.33 abcd	109.66 abcde	114.00 cd	32.66 abc	0.550 cde	2897.33 abcd	22.66 defghi	46	130
32. BR-88 40.013	2426.33 abc	87.33 abcd	15.33 bc	238.33 bcd	80.00 de	116.33 cd	42.00 abc	0.620 abcde	3857.33 abc	23.66 bcdefg	49	130
21. BR-88 40.217	2400.66 abc	81.66 abcde	17.66 abc	303.66 abcd	105.00 abcde	155.66 abcd	44.66 ab	0.523 cde	4746.33 ab	24.33 abcdef	49	130
28. BR-88 40.407	2348.66 abc	88.00 abcd	15.33 bc	267.33 abcd	135.66 abcde	111.33 cd	20.33 abc	0.553 cde	3608.66 abc	21.33 ghi	49	120
17. BR-88 40.202	2324.00 abc	79.33 bcde	17.33 abc	242.00 abcd	106.33 abcde	103.33 cd	32.33 abc	0.567 bcde	3857.66 abc	22.33 efghi	44	132
15. BR-88 40.197	2276.33 abc	88.66 abc	17.66 abc	280.33 abcd	133.00 abcde	117.00 cd	30.33 abc	0.450 de	4319.66 abc	22.33 efghi	48	130
16. BR-88 40.199	2264.33 abc	88.33 abc	17.66 abc	279.33 abcd	123.33 abcde	155.33 abcd	37.33 abc	0.513 cde	4186.33 abc	21.00 hi	48	130
30. BR-88 40.002	2109.66 bcd	81.33 abcde	14.00 bc	304.00 abcd	132.00 abcde	141.66 abcd	30.33 abc	0.623 abcd	3093.00 abcd	26.00 ab	47	131
31. BR-88 40.007	1833.00 cd	80.33 abcde	16.33 abc	154.66 d	67.66 e	73.00 d	14.00 bc	0.513 cde	2542.00 cd	23.00 cdefgh	51	126
3. Ocepar-3	1827.33 cd	78.00 cde	10.33 c	291.66 abcd	171.00 a	104.00 cd	16.66 abc	0.780 a	3288.66 abcd	23.00 cdefgh	47	131
11. BR-88 40.174	1032.33 d	93.00 ab	18.00 abc	263.66 abcd	116.66 abcde	135.33 abcd	11.66 c	0.550 cde	3244.00 abcd	25.33 abc	48	130

YP: yield per plot (g), PH: plant height (cm), HFP: height of first pod (cm), NPP: number of pods per plant, NP1G: number of pods with one grain, NP2G: number of pods with two grains, NP3G number of pods with 3 grains, SD stem diameter (cm), SWB stem weight plus bark (g), WS: weight of a hundred seeds (g), DF: flowering (days), DM: days to maturity (days).

RESULTS AND DISCUSSION

Table 2 shows that there are significant differences among the means of the 34 soybean lines for all the traits. The genotypic determination coefficients (b) for the traits YP, PH, SD and WS have values greater than 30%, suggesting the possibility of fairly efficient discrimination of superior parental genotypes among the 34 assessed lines. Except for HFP, NP1G, NP2G, NP3G and SWB with a high C.V. of over 30%, the C.V. of the other traits varied from 6,27% to 28,35% in acceptable experimental accuracy levels for field experiments.

Table 2 - Analysis of variance for ten traits in 34 soybean genotypes, with the respective means, variation coefficient (C.V.) and genotypic determination (b), in Jaboticabal, SP, 1990.

		Mean Square*									
Sources of Variation	df	YP	PH	HFP	NPP	NP1G	NP2G	NP3G	SD	SWB	WS
Plots	2	1028578,41	74,954	11,186	31,294	76,713	1711,92	547,54	0,0323	2883637	6,303
Genotype	33	797857,37	98,978*	29,418	6346,11	1597,23	3071,91	247,51	0,01741	481320	7,908*
Residues	66	523140,66	63,314	30,762	6873,43	1613,90	2749,54	302,73	0,0120	1335686	2,132
Means		2696,92	84,97	17,19	293,51	122,78	141,39	30,45	0,56	3647,11	23,27
C.V.(%)		26,82	9,36	32,27	28,25	32,72	37,09	57,14	19,46	31,69	6,27
b(%)		34,43	36,03	0,00	0,00	0,00	10,49	0,00	30,81	9,83	73,04

*significant at 1% probability by the F test

YP: yield per plot (g); PH: plant height (cm); HFP: height of first pod (cm); NPP: number of pods per plant; NP1G: number of pods with one grain; NP2G: number of pods with two grains; NP3G number of pods with 3 grains; SD stem diameter (cm); SWB stem weight plus bark (g); WS: weight of a hundred seeds (g)

The Tocher grouping method (quoted by Rao, 1952) from the Malhalanobis generalized distances (D^2) enabled the division of the 34 lines into four groups (Fig.1). Group I contained 30 genotypes, corresponding to 88.24% of the studied genotypes; group II contained 2 genotypes, corresponding to 5.88% and groups III and IV were made up of one genotype each, representing 2,94% of the lines. The BR-88 40.058 (7) and BR-88 40.202 (17) were the most closely related genetically as they had the smallest distance between them (D^2 = 1,0901). The most genetically divergent genotypes were BR-88 40.057 (6) and BR-898 40.199 (16).

Many authors have recommended the use of parents with the greatest possible divergence to maximize the heterosis shown in the hybrids, increase the probability of superior segregants in advanced generations and widen the genetic base. Therefore, the information in Fig.1 should be used for cross recommendation. Using the multivariate statistic as a base, a high degree of similarity could be expected to be found among the genotypes belonging to the same group. Thus crosses within the same group should be avoided (Destro, 1991).The identification of genotypes based on genetic divergence, without considering their behavior, may not be a good strategy for a breeding program. Crosses among divergent genotypes that have a superior performance for the main characteristics of agronomic importance seem to be more recommended for breeding programs.

Three of the twelve traits studied in soybean breeding were especially important: WS, PH and DM. The WS trait was considered as representing the yield for each genotype in the plot, and genotypes with yield greater than 2500 kg\ha were selected. The PH, HFP and DM traits were considered by their importance to quantify the adaptation of the genotypes to the region; the earliest were selected with a cycle up to 130 days and with plant height above 80cm and first pod height above 15cm. Based on the established criteria (Table 1) the genotypes which stood out were 23, 10, 2, 9 and 25, although genotype 27 had HFP of 14,00 cm, it was also selected due to if its short cycle (120 day cycle). Fig. 1 shows these genotypes as belonging to group I, where the divergences can be seen graphically.

Genotypes number 6 with 2.628,66 kg/ha, PH 90,33cm, HFP 12,33 cm and DM 128 days belonging to group II and the genotype number 16 with 2.264,33 kg/ha, PH 88,33, HFP 17,66 cm and DM 127 days (Table 1, Fig. 1) were selected for their agronomic qualities.

Both the conglomeration technique based on Mahalanobis generalized distance and the analysis of the canonic variables were used for the quantification of the genetic divergence among the parents. In the latter, the genetic divergence was assessed by the graphic scattering of the scores for the studied genotypes on cartesian axes (Fig. 1). Generally researchers have opted for graphic representation when the first two canonic variables are greater than 70% of total available variability (Cruz 1990).

In this study, the dispersion in Fig. 1 was made for the two first canonic variables, which explain 75% of the total variation, which was satisfactory for this study. Table 1 shows the means of the 34 soybean lines in the 12 traits studied.

The yield values of the genotypes in Table 1 showed that the greatest yields occurred in more plants with pods with two grains (NP2G) and not in plants which had a greater total number of pods (NPP), a greater number of pods with 3 grains (NP3G) and nor a greater weight per hundred seeds (WS). Table 1 further showed that the stem diameter values (SD) varied from 0,432 cm to 0,870 cm, which was more favorable because lower values caused less wear on machinery during harvest. The values of the stem plus bark varied from 1.550,33g to 488,66g the varieties with higher SWB values covered the soil more adequately and produced more green material.

The most important agronomic traits were recommended in the crosses, that is, yield per plot (YP), plant height (PH), height of the first pod (HFP) and days to maturity (DM). The following crosses were recommended based on the genetic divergence and on the key agronomic traits: genotypes 23, 10, 2, 27 and 25 (group I) with genotype 6 (group II) and genotype 16 (group III). Thus, only ten crosses would be made, representing only 2% of the possible total in a partial diallel among the 34 lines assess, which would result in up to 561 combinations.

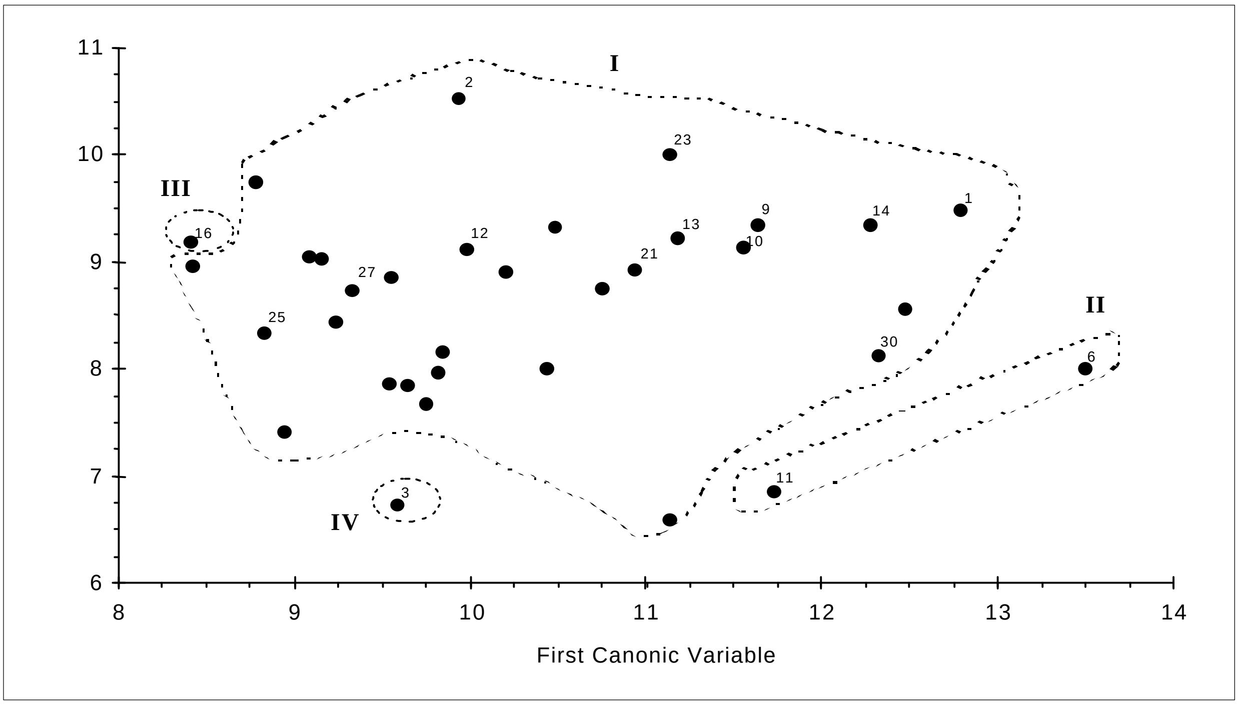

Figure 1 - Score scattering of the 34 soybean genotypes for the first two canonic variables.

RESUMO

Trinta e quatro linhagens de soja foram avaliadas para doze características. As distâncias genéticas foram estimadas utilizando técnicas multivariadas com objetivo de identificar parentais a serem incluidos em um programa de melhoramento envolvendo hibridação. O agrupamento pelo método de Tocher, a partir das distâncias generalizadas de Mahalanobis, dividiu as 34 linhagens em 4 grupos. As caracterísiticas agronômicas mais importantes, peso de sementes por parcela, altura de planta, altura da primeira vagem e dias para maturação foram consideradas para a recomendação dos cruzamentos. Os seguintes cruzamentos foram recomendados baseado na divergência genética e nas características agronômicas chave: linhagens 23, 10, 2, 27 e 25 (grupo I) com genótipo 6 (grupoII) e com o genótipo 16 (grupo III). Portanto somente 10 cruzamentos poderiam ser realizados representando somente 2% do total de cruzamentos qu poderiam ser realizados em um dialelo parcial entre as 34 linhagens avaliadas as quais admitiriam até 561 combinações.

REFERENCES

Amaral Júnior, A.T.; Casali, V.W.D.; Cruz, C.D.; Finger, F.L. (1996), Utilização de variáveis canônicas e de análise de agrupamentos na avaliação da divergência genética entre acesso de moranga. *Horticultura Brasileira*, Brasília, **14**, 2, 182-184.

Carvalho, L.P.; Cruz, C.D.; Moraes, C.F. (1995) Genetic divergence in Brazilian cotton, *Gossypium hirsutum* var. *latifolium* Hutch. *Revista Brasileira de Genética*, Ribeirão Preto, **18**, 3, 439-443.

Cruz, C.D. (1990) Aplicações de algumas técnicas multivariadas no melhoramento de plantas. Piracicaba: ESALQ/USP, 188, Tese de Doutorado.

Destro, D. (1991) Capacidade de combinação de genótipos de soja *(Glycine max (L.) Merrill)* apropriados para o consumo humano. Piracicaba: ESALQ/USP, 158. Tese de Doutorado.

Falconer, D.S. (1981) *Introduction to quantitative genetics*. 2.ed. London: Longman, 340.

Fehr, W.R.; Cavines, C.E. (1977) *Stages of soybean development*. Ames, Iowa: State University of Science and Technology.

Mahalanobis, P.C. (1936) On the generalized distance in statistics. *Proceedings* National Institute of Science, **2**, 49-55.

Miranda, J. E. C.; Cruz, C.D.; Costa, C.P. (1998) Predição do comportamento de híbridos de pimentão (Capsicum annuum, L.) pela divergência genética dos progenitores. *Revista Brasileira de Genética***,** **11**, 929-937.

Pípolo, V.C.; Pípolo, A.E.; Destro, D.; Guerra, E.P. (1995) Seleção de genótipos de guandu baseada na divergência genética multivariada. *Pesquisa Agropecuária Brasileira*, Brasília, **30**, 7, 977-982.

Rangel, P.H.N.; Cruz, C.D.; Vencovski, R.; Ferreira, R.P. (1991) Selection of local lowland rice cultivars based on multivariante genetic divergence. *Revista Brasileira de Genética*, **14**, 2, 437-453.

Rao, R.C. (1952) *Advanced statistical methods in biometric research*. New York: John Wiley and Sons, 390.

Singh, R.K.; Chaudhary, B.D. (1977) *Biometrical methods in quantitative genetic analysis*. New Delhi: Kalyani Publ., 303.

Souza, J.C. (1996) Diversidade Genética entre acessos de acerola (*Malpighia* sp.) com base em dados isoenzimáticos e agronômicos, Viçosa:UFV, 67, Tese de Mestrado.

Received: April 05, 1999;
Revised: July 01, 1999;
Accepted: October 06,1999.