SCREENING SOYBEAN GERMPLASM FOR ALUMINIUM TOLERANCE USING CLUSTER ANALYSIS[1]

CARLOS ROBERTO SPEHAR[2]

ABSTRACT - Principal component and cluster analyses were carried out to assess diversity on soybean germplasm selected to cultivation in the Brazilian Savannas (Cerrados). The results indicated that aluminium tolerance was attained in the breeding material which showed variability when tested in the high-Al environment. The ample tolerance may have been a consequence of breeding and selecting soybeans in improved soils that were still Al-rich in deeper layers, although the varieties were homogeneous for agronomic characters, as shown by the analysis of data from the low-Al environment. This method can be successfully employed in genetic studies and in breeding programmes for crop improvement when parents of different clusters are chosen for crossing.

Index terms: genetics, breeding, variety, adaptation, cerrados, mineral element, interaction.

EMPREGO DA ANÁLISE DE AGRUPAMENTO QUANTO À TOLERÂNCIA AO ALUMÍNIO EM GERMOPLASMA DE SOJA

RESUMO - Efetuou-se a análise dos componentes principais e de agrupamento dos genótipos em classes, para avaliar a diversidade do germoplasma de soja adaptada ao cultivo nos cerrados. Os resultados do teste em solo parcialmente corrigido indicaram que tolerância ao alumínio foi preservada no processo de seleção. Essa variabilidade foi conseqüência do cruzamento e seleção de variedades em solos corrigidos onde ainda é abundante o alumínio nas camadas inferiores do solo, embora o material genético seja homogêneo quanto às características agronômicas, conforme mostram os resultados do teste em solo corrigido. Esta metodologia pode ser empregada com sucesso em estudos genéticos e em programas de melhoramento, para ampliar a adaptação da soja, quando variedades parentais de diferentes classes são utilizadas em hibridação.

Termos para indexação: genética, melhoramento, variedade, adaptação, cerrados, elementos químicos, interação.

INTRODUCTION

There are vast areas in the Tropics which have natural limiting factors to develop sustainable agriculture, such as toxic acidity and lack of nutrients (calcium) in the soil. These are the main features of highly weathered soils of the Brazilian Savannas or Cerrados and prevent deep root growth into the soil of intolerant species and varieties of plants (Spehar, 1989). To compound these soil problems, there are frequent dry spells during the rainy season that might affect crop yields.

The Cerrados occupy one-fourth of the country's surface and, for farming systems to be effective and environmentally balanced, genetic tolerance to these soil hindrances is necessary in the adaptation of crops.

The first step in the adaptation of crops to a new environment is germplasm evaluation. The more diverse the germplasm, the higher the chances of identifying genotypes which carry the desirable traits. When acidity and low fertility of the soil are major components of the environment, the task of identifying and selecting desirable phenotypes is not simple due to mineral element interactions (Foy et al.,

[1] Accepted for publication on August 4, 1993.

[2] Eng.-Agr., PhD., EMBRAPA/Centro de Pesquisa Agropecuária dos Cerrados (CPAC), Caixa Postal 08223, CEP 73301-970 Planaltina, DF, Brazil.

1969; Camargo, 1985; Saneoka et al., 1986; Wilkinson & Duncan, 1993).

To achieve progress in breeding for these problem soils, the adapted genotypes should yield economically. However, grain yield, per se, as the end product, is the result of many physiological paths involved in the synthesis and translocation of compounds. If each of these paths is controlled by at least one gene, it is not difficult to admit that grain yield is quantitatively inherited, as has already been shown (Allard, 1960). However, if a selected variety for high yield does not have tolerance to aluminium, for instance, its whole genome, although being superior, is inhibited from expressing itself in the Al-stress environment. Thus, it is desirable to identify genotypes carrying relevant characters in relation to other members of a germplasm collection.

One method which seems to be efficient in the identification of desirable genotypes is the use of multivariate and cluster analyses (Pielou, 1984). These analyses, first developed for ecological research, allow the separation of genotypes into distinct classes, and the identification of the variables that characterize these groupings. They are appropriately used as a first step in evaluating te genetic diversity of germplasm collections when error due to soil variability is eliminated either by replications of the experiment or by the use of augmented block designs (Spehar, 1993).

The main objective of this work is to classify the genotypes from soybean germplasm collection in groups to facilitate their use in genetic studies and breeding for aluminium tolerance.

MATERIALS AND METHODS

One hundred and eighty four soybean breeding lines and varieties, originating from public and private research organizations, were grown in an augmented randomized complete block design and the data were corrected for block effect (Spehar, 1993), on the basis of standard variety performance. There were four standard varieties of wide range of maturity in Central Brazil per block. They were chosen because they were extensively used in the farmers, fields and were included in uniform trials in the Cerrado region. Previous information on the performance of some of them for grain yield in acid soils was also available (Relatório... 1980). The blocks were extended and the gaps were filled by the entries to be evaluated.

The experiment was repeated for two years at two levels of fertilization on a cerrado soil, a variety of oxisol classified as Dark-Red Latosol (Typic Haplustox fine, kaolinitic, isohyperthermic in the U.S. soil taxonomy), at the Centro de Pesquisa Agropecuária dos Cerrados (EMBRAPA-CPAC), Planaltina, Brazil, which is located at 15°36' S and 47°12' W at an elevation of 1,000 m.

The two areas in which the experiment was repeated were fertilized in the following manner: one received 1,000 kg/ha dolomitic limestone (100% CaCO3 equivalent), 75 kg/ha P, 75 kg/ha K and 40 kg/ha of slow release micronutrientes source, namely, FTEBR-12; the other 4,000 kg/ha dolomitic limestone (100% $CaCO_3$ equivalent), 150 kg/ha P, 75 kg/ha K and the same amount and source of micronutrients. They were named high-Al and low-Al areas, for experimental purpose, respectively. All amendments were applied broadcast on the two areas and disc incorporated to approximately 20 cm. The soil in the high-Al area been broken prior to the experiment by ploughing a grazing area previously covered with cerrado vegetation. The low-Al area had already been cultivated with maize for one crop season and left fallow for two years. The physical characteristics of the soil are: sand= 34%, silt= 19% and clay= 46%; the results of soil chemical analysis are: pH (H_20 1:1) 4.7, 5.1 and 5.6, Al (meq/100g) 1.9, 0.8, 0,16, Ca+Mg 0.40, 1.98 and 4,47, P (ppm) 0.9, 3.4 and 9.1, K 16, 30 and 30, for the virgin, high-Al and low-Al areas, respectively.

The seeds of standard varieties and those of the breeding lines were all produced in the same cropping season in an experimental field of CPAC. Before the experiment started, germination tests were carried out with the seeds, to ensure a density of 25 plants/m in the row after emergence. The seeds were inoculated with *Bradyrhizobium japonicum*-peat inoculant at sowing time, in the first year of the experiment only.

Foliar samples were collected at flowering time for each of the experimental plots. Twenty fully expanded leaves with petioles comprised each sample. A fully expanded leaf corresponded to the fourth leaf down from the top of the plant. This method was adapted from the one described by Bataglia (1991). The samples were oven-dried at 100°C. For 72 hours and milled for mineral analysis. A constant weight of the milled material was wet-digested using a mixture of nitric and perchloric acids. Nitrogen was determined by a colorimetric technique that measures

ammonia produced in a modified Kjeldahl digestion. Potassium was determined by flame photometry. The elements P, Ca, Mg, Mn and Zn were determined by atomic absorption.

At physiological maturity, that is, when 95 % of the plants were ripe, determinations were made on plant height, first pod height and lodging. The harvested area consisted of the central row, the ends of which were trimmed to 3 m long. After harvesting the plants were threshed and the seeds collected and weighed for yield and seed weight evaluations. The seeds moisture content was assessed and all yield values were expressed in kg/ha at 13 % moisture.

Principal component and cluster analyses were carried out using the above mentioned data. Cluster analysis describes the relationship between the entries by combining them into clusters, and clusters into larger clusters to produce a hierarchical branching pattern or dendrogram. At each stage the two most similar clusters are combined. In the present analysis the centroid method was used. This method is described as falling between the nearest-neighbour and farthest-neighbour methods, of clusterings and overcomes the limitation of these two methods, in which the distance between two points decides the outcome of each step in the clustering. In the centroid method two clusters are separated by their respective centroids which are the hypothetical quadrats containing the average individual over all the cluster members (Pielou, 1984). The centroids are their centres of gravity in the multi-dimensional space of the variables measured. Entries are allocated to clusters so as to maximize this distance.

In order to assess the relative contribution of each measured variable to the clustering, principal component analysis was performed. The analysis was conducted on the results from the high-Al environment, the low-Al environment and also on the ratios between them in both years.

RESULTS AND DISCUSSION

The previously adjusted data according to augmented design procedure (Spehar, 1993), were analysed for a principal component analysis. The results for the analysis of the high-Al experiment, the low-Al experiment and the ratios high-/low-Al are presented in Table 1 to Table 3.

The analysis of the data from the high-Al environment allowed to say that at the first level of ralationship, which accounts for 32.5 % of the variance, days to first flower, days to maturity, plant and first pod heights, seed yield, and, with negative sign, calcium and magnesium were the characters that most contributed to the first principal component. This component measured the effect maturity groupings in the dimension of variation. For the second principal component, which accounted for 16.8 % of the variance, phosphorus and nitrogen with negative sign, followed by zinc with positive sign, were the characters that explained the association encountered in the germplasm at this level. At a third level, which accounted for 10.5 % of the variance, negative values of calcium, manganese and seed weight, and, a positive value of potassium were the characters that weighed highest in explaining the relationship found in the germplasm. The last two components were independent of the maturity grouping.

Considering the same three components in the low-Al experiment, it can be said that negative values of days to first flower, days to maturity, plant height, and, positive values of magnesium, calcium and potassium played an important role in the association found in the germplasm at the first component. It is possible to observe an inverse relationship at this level when the two aluminium experiments are compared. At the second level of relationship, similarly to the high-Al experiment, negative phosphorus and nitrogen explained better the association in the germplasm, followed by negative seed yield, which contrasted with its corresponding value in the high-Al experiment. At a third level negative zinc and manganese, and, positive seed weight and seed yield were the characters that better explained the relationship in the soybean germplasm. There is a clear indication that the maturity groupings affected the first principal component at both levels of aluminium, while the other two components were more independent, even though the accumulated percentage variance for the three classes in both environments did not seem to be different.

Significant negative correlation coefficients for calcium, magnesium and phosphorus in the leaves were obtained for maturity in the high-Al experiment, although the r^2 was not high (Spehar, 1993). This relationship may have been a result of leaf sampling method, i. e., the fourth fully expanded leaf at

flowering is not collected from the same node in varieties of different maturity groups. This implies that the maximum level of nutrients is different in the individual leaves.

To avoid the effect of leaf position in the plant it is suggested that, in further work, a modification to the method be introduced. This consists of harvesting the seventh and the eighth fully expanded leaves for

TABLE 1. Principal component analysis for soybean germplasm grown in high-Al area.

Principal Component	1	2	3	4	5
Latent Root	4.2314	2.1801	1.3651	1.1578	0.8306
Percentage Variance	32.5496	16.7700	10.5012	8.9065	6.3895
			Variate Loading		
Yield	0.2828	-0.0789	-0.2338	0.3821	-0.1403
Seed Weight	-0.0201	0.0286	-0.4453	0.6551	-0.1038
Plant Height	0.3896	-0.2622	-0.0567	0.0323	0.1040
First Pod Height	0.3534	-0.2245	-0.0634	-0.0143	0.0886
First Flower	0.4146	-0.2165	0.0025	-0.1983	0.0988
Maturity	0.3919	-0.1143	-0.1117	-0.1095	0.2011
Phosphorus	-0.2285	-0.4924	0.2928	0.1167	-0.0760
Potassium	-0.1124	-0.2123	0.4198	0.4071	0.6779
Calcium	-0.2811	-0.0497	-0.4517	0.0129	0.3018
Magnesium	-0.3397	-0.2731	-0.2677	-0.0486	0.1111
Nitrogen	0.0084	-0.5226	-0.0797	0.0636	-0.3001
Manganese	-0.1460	-0.2068	-0.4160	-0.4181	0.3037
Zinc	-0.1919	-0.3653	-0.1012	-0.1076	-0.3869

TABLE 2. Principal component analysis for soybean germplasm grown in low-Al area.

Principal Component	1	2	3	4	5
Latent Root	4.7137	1.6936	1.2924	1.0743	0.8454
Percentage Variance	36.2591	13.0280	9.9413	8.2639	6.5033
			Variate Loading		
Yield	-0.0079	-0.3821	0.2978	0.4271	-0.6055
Seed Weight	0.1101	-0.3054	0.3145	0.3858	0.6945
Plant Height	-0.3552	0.0256	0.1490	-0.0844	0.1765
First Pod Height	-0.2702	0.1359	-0.0115	-0.3029	-0.0789
First Flower	-0.4189	0.1129	-0.0494	0.1044	0.0577
Maturity	-0.3541	0.0424	0.0908	0.2994	0.1881
Phosphorus	-0.1435	-0.6248	-0.1623	-0.2161	-0.0339
Potassium	0.3378	-0.2392	0.1214	-0.2081	0.0024
Calcium	0.3221	0.1344	-0.1760	0.2381	-0.0534
Magnesium	0.3716	-0.0079	-0.1576	0.1518	0.0329
Nitrogen	-0.2971	-0.4196	-0.1632	0.0475	-0.0775
Manganese	-0.1402	0.1472	-0.4892	0.5361	-0.0984
Zinc	0.0575	-0.2524	-0.6453	-0.1001	0.2330

TABLE 3. Principal component analysis for the ratio high-Al/low-Al of soybean germplasm.

Principal Component	1	2	3	4	5
Latent Root	2.2519	2.0091	1.6205	1.3741	1.0832
Percentage Variance	17.3220	15.4548	12.4651	10.5704	8.3324
			Variate Loading		
Yield	-0.4354	-0.2396	0.0721	-0.2033	0.2583
Seed Weight	-0.3146	-0.2011	0.1876	0.1826	0.5178
Plant Height	-0.4820	-0.0017	-0.0026	-0.1436	0.0589
First Pod Height	-0.2134	0.3417	-0.0299	-0.1858	-0.0586
First Flower	-0.0324	-0.1114	-0.0533	0.6494	-0.0719
Maturity	-0.2692	-0.2035	-0.2306	0.5531	-0.1640
Phosphorus	-0.2111	0.4728	0.3610	0.0999	-0.1839
Potassium	-0.4046	-0.0468	-0.0428	-0.1342	-0.5297
Calcium	-0.0579	0.2167	-0.5471	-0.1157	0.3397
Magnesium	-0.2028	0.1779	-0.5887	0.0136	0.0522
Nitrogen	-0.1732	0.3770	0.3323	0.2059	0.3274
Manganese	0.2610	0.3151	-0.0892	0.1718	0.2379
Zinc	-0.0958	0.4311	-0.0794	0.1705	-0.1727

all the plants to standardize the samples for comparisons, irrespective of maturity groupings.

When the principal component analysis was carried out on the ratios high-/low-Al, it is possible to observe a more spread out variance,in contrast with the analysis on the data for each environment (Table 3). This is interpreted that some of the correlations due to maturity groupings, which may have affected the level of nutrient in the leaves, were eliminated in the calculated relative values and no component predominated. The first three classes of principal components for the high- and for the low-Al environments explained 59.8 and 59.1 % in contrast with the same classes for the ratio high/low-Al, which together explained 45.2 %. Negative values of plant height, yield and potassium explained most of the relationship in the first component. In the second level, phosphorus, zinc, nitrogen and first pod height explained the association amongst the varieties. At a third level, negative calcium and magnesium were followed by phosphorus and nitrogen as the characters that better explained the associations amongst the varieties. It is interesting to note that calcium and magnesiun are still associated with each other though no longer with yield. This is interpreted as seed yield being also a function of other elements in the plant.

The results of the cluster analysis are presented in Fig. 1 to Fig. 3 and correspond to the clustering of the germplasm for the high-Al environment, the low-Al environment and the ratio high-/low-Al, respectively. The range of the levels of clustering was higher at the high-Al environment than the range at the low-Al and the ratio. This supports the idea that observations of individual characters in high-Al is more informative and strongly suggests that high variability is present in the germplasm but its expression is a direct function of environmental stress. Relationship amongst the varieties in any cluster for the high-Al environment did not correspond to that obtained for the same varieties analysed in low-Al. When the ratios were clustered by the same method, an independent set of clusters was obtained and only a few entries would be classified as related throughout the analyses. The fact that in the high Al environment the diversity was more expressed, suggests that variability for aluminium tolerance in the germplasm was attained by indirect selection. In selecting soybean varieties with better agronomic characters for adaptability to the Cerrados of Brazil, where acid soils predominate, a build-up of more favourable genes either for aluminium tolerance or for efficient use of nutrients seemed to have occurred. The lower variability as detected by the clustering of data on the

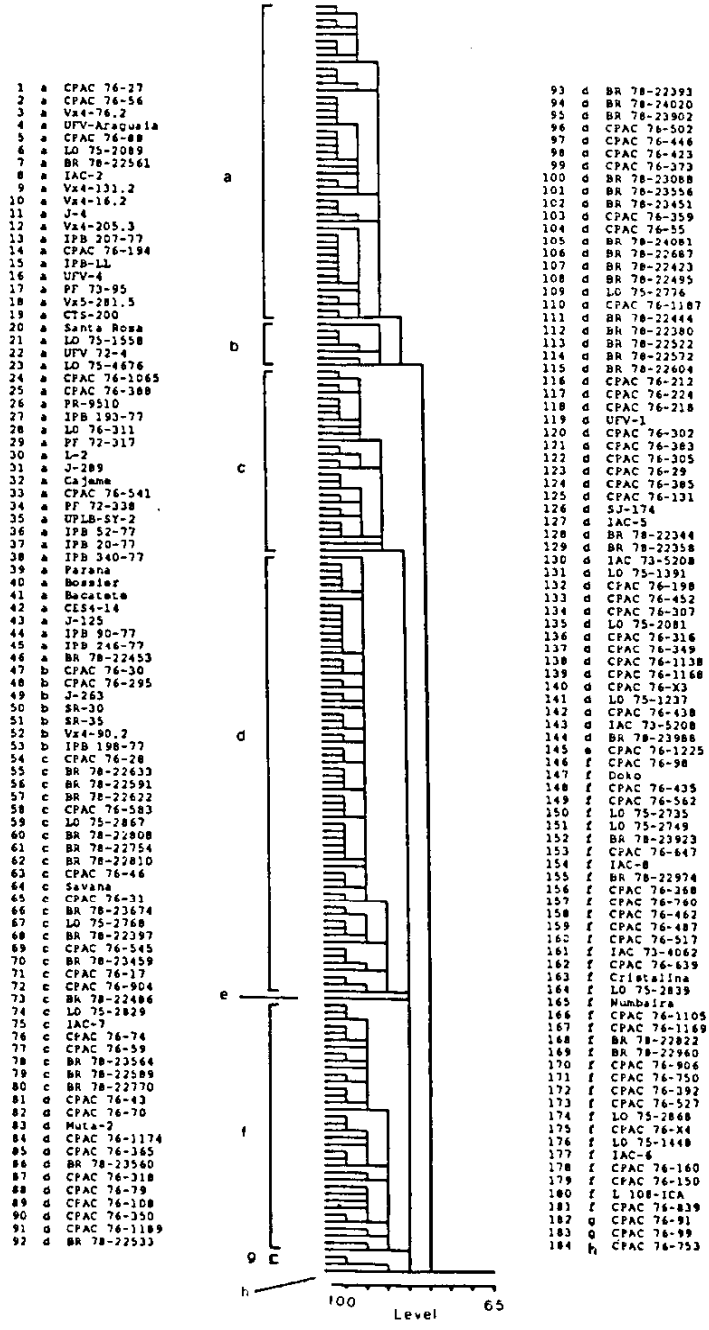

FIG. 1. Dendrogram from the cluster analysis of 184 soybean varieties grown in the high Al area.

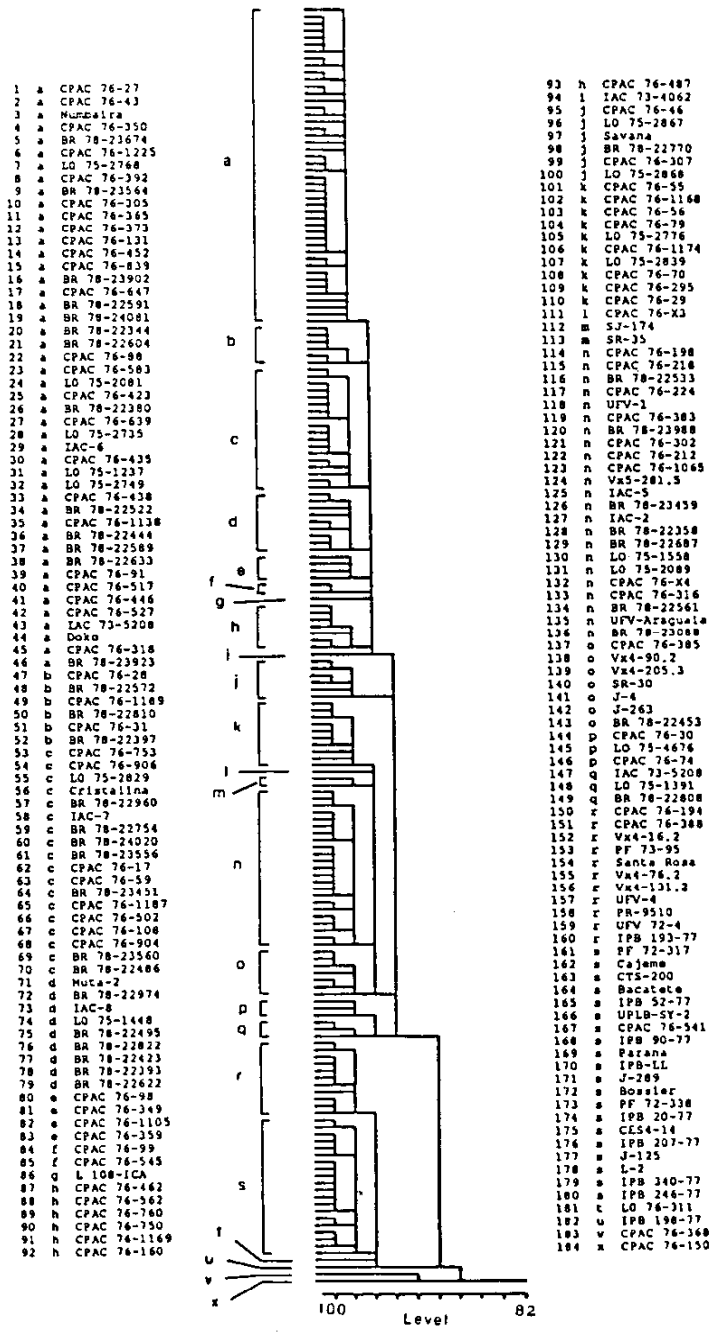

FIG. 2. Dendrogam from the cluster analysis of 184 soybean varieties grown in the low-Al area.

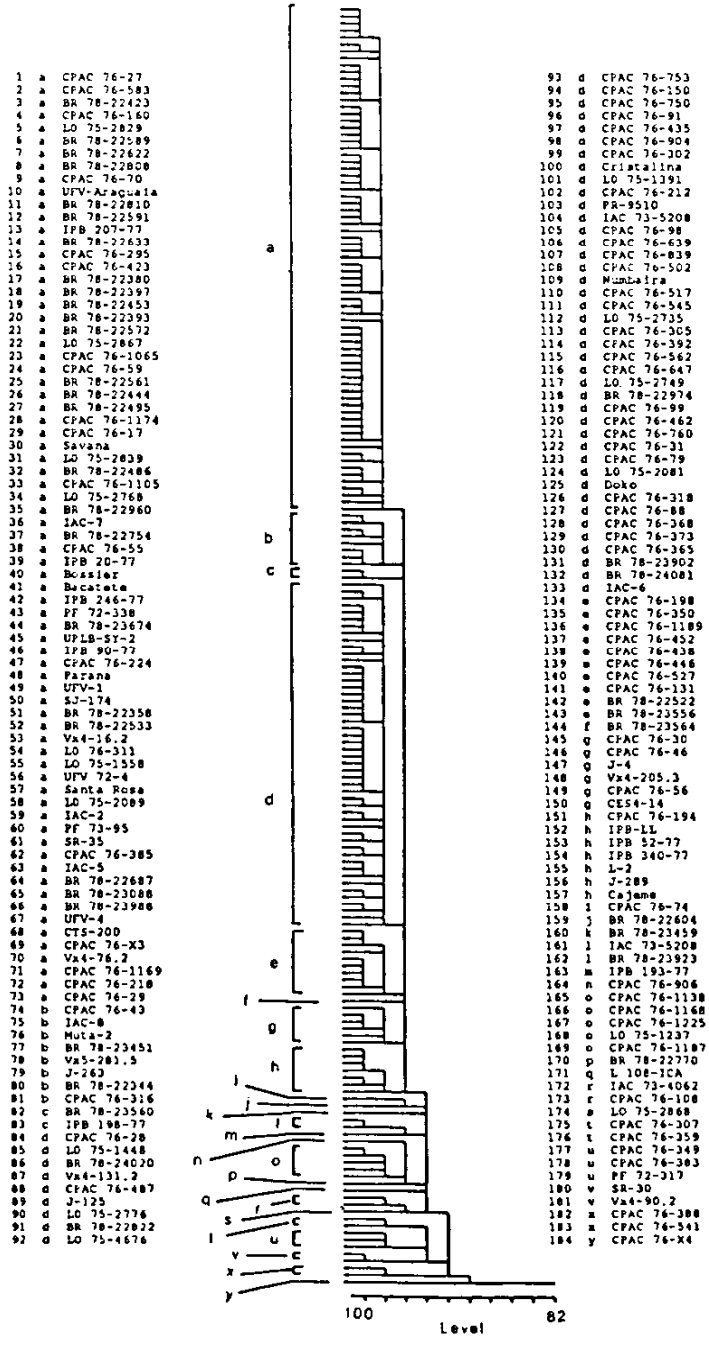

FIG. 3. Dendrogram from the cluster analysis on the ratio high/low-Al for 184 soybean varieties.

low-Al environment can be explained by the high adaptability of the germplasm to the environment at which selection is practiced. The agronomic uniformity and smaller magnitude of mineral differences at this environment reduced the expression of possible genetic differences among the soybean varieties.

The breeder deals in his empiricism with natural selection, as shown by the present cluster analysis. These results confirm the information reported elsewhere (Foy et al,. 1978; Bilski & Foy, 1987; Takagi et al., 1983), which emphasizes the influence of the environment in directing selection, i.e., aluminium tolerant varieties originate where the element is abundant. The fact the breeder conducts experiments in several locations to assess genotype x environment interaction before releasing superior lines, takes into account the forces of natural selection, even though they are not necessarily measured. The variability for acid soil cultivation, identified by the cluster analysis in the germplasm, was buffered in more favourable genetic combinations. The more stable genotypes from variety trials in improved acid soils were probably the ones which exhibited deeper rooting, due to Al-tolerance, and were less exposed to dry spells.

The present results suggest that the analysis by multivariate methods and by cluster analysis is an efficient assessment of desirable germplasm to utilize in breeding programmes for the adverse soil conditions prevalent in the tropics. The overall comment worth making is that field testing of germplasm in high-Al environment will allow to group the similar germplasm to be used in soybean breeding for better Al-tolerance. The progenies of these crosses are tested in laboratory for root elongation in hydroponics containing aluminium; the superior recombinants to both parents will be advanced and tested in the field for agronomic characters; the agronomically better breeding lines will be included in the uniform trials and the more stable released as new varieties. It is expected that more tolerant varieties to this and other limiting factors will develop a root system to exploit deeper layers of the soil where water and leached nutrients are present. This will result in more stable yields even when dry spells occur and shall contribute to develop sustainable cropping systems for the Cerrados.

CONCLUSIONS

1. Cluster analyses were useful to separate the germplasm in classes of similarity.

2. The range of clustering was wider in the high aluminium environment than in the low-Al environment, indicating that variability for the trait was preserved in agronomically adapted germplasm.

3. It is evident that the effect of soil fertility is great and the expression of genotypic variability is strongly a function of the characters used in the comparison.

4. It is concluded that a standardization of the sampling method should allow a more appropriate assessment of the differences among the genotypes, irrespective of their maturity groups.

REFERENCES

ALLARD, R. W. **Principles of plant breeding.** New York: Wiley, 1960.

BATAGLIA, O.C. Análise química de plantas. In: SIMPÓSIO SOBRE MICRONUTRIENTES NA AGRICULTURA, 1. 1988, Jaboticabal, SP. **Anais...** Piracicaba, SP: POTAFOS/CNPq, 1991. p. 289-308.

BILSKI, J. J.; FOY, C. D. Differential tolerance of oat cultivars to aluminium in nutrient solutions and in acid soils of Poland. **Journal of Plant Nutrition,** v.10, p. 129-141, 1987.

CAMARGO, C. E. O. Effect of phosphorus in nutrient solution on the tolerance to aluminum toxicity in wheat cultivares. Campinas, SP, Brazil, **Bragantia,** v.40, p.21-31, 1985.

FOY, C. D.; FLEMING, A,; ARMIGER, W. H. Aluminum tolerance in soybean varieties in relation to calcium nutrition. **Agronomy Journal,** v.61, p.505-511, 1969.

FOY, C. D.; CHANEY, R. L.; WHITE, M. C. The physiology of metal toxicity in plants. **Annual Review of Plant Physiology,** v.29, p.511-566, 1978.

PIELOU, E. C. **The interpretation of ecological data.** New York: Wiley, 1984.

RELATÓRIO TÉCNICO ANUAL. Brasília: EMBRAPA-CPAC, 1980.170p.

SANEOKA, H.; KANADA, K.; OGATA, S. Differential

tolerance among tropical forage crops to problem soil conditions. I. effect of low pH and aluminium in culture medium of growth and nutrient uptake of several tropical forage crops. **Journal of the Japanese Society of Grassland Science**, v.32, p.251-260, 1986.

SPEHAR, C. R. **The genetics of aluminium tolerance in soya beans [*Glycine max* (L) Merrill]**. Cambridge (England): University of Cambridge, 1989. 123p. Ph.D. Thesis.

SPEHAR, C. R. Composição mineral da soja cultivada em solo sob cerrado com dois níveis de calagem. **Pesquisa Agropecúaria Brasileira, Brasília** v.28, n.5, p.645-648,1993.

TAKAGI, H.; NAMAI, H.; MURAKAMI, K. I. Aluminium tolerance of registered wheat *Triticum aestivum* varieties in Japan. **Japanese Journal of Breeding**, v.33, p.69-75, 1983.

WILKINSON, R. E.; DUNCAN, R. R. Calcium ($Ca\text{-}45_{2+}$) absorption inhibition by aluminum (Al_{3+}) in sorghum roots. **Journal of Plant Nutrition**, v.16, p.235-240, 1993.