

Genetic Distance and its Association with Heterosis in Cacao

Luiz Antônio dos Santos Dias^{1*}, Jane Marita², Cosme Damião Cruz³, Everaldo Gonçalves de Barros¹ and Tânia Maria Fernandes Salomão¹

¹ Departamento de Biologia Geral; BIOAGRO; Universidade Federal de Viçosa; 36571-000; Viçosa - MG - Brazil.

² Department of Forestry; 1925 Linden Drive; USDFRC; WI, 53706; Madison - USA. ³ Departamento de Biologia Geral; Universidade Federal de Viçosa; 36571-000; Viçosa - MG - Brazil

ABSTRACT

The efficiency of cacao breeding program can be increased by choosing superior crosses to be made between divergent clones. We assessed the genetic distance among five clones with RAPD data (genetic distance - GD) and with yield component data (Mahalanobis distance - MD). The clones were evaluated in a diallel, during five years, for five yield components. A total of 130 RAPD bands were scored. GD and MD were used to determine the correlation between genetic distances among clones and the performance of their hybrids. The correlation between GD and MD was 0.67 (P=0.03). Both distances were related to heterotic performance of hybrids for wet seed weight/plant and wet seed weight/fruit. The average hybrid performance for the same two yield components was correlated with only MD. Hence, genetic distances measured by RAPD and yield components can be used as a guide to the choice of the superior crosses.

Key words: Genetic distances, heterosis, performance data, RAPD data, *Theobroma cacao* L.

INTRODUCTION

Cacao (*Theobroma cacao* L.) is one of the most important Neotropical crops. It is a typical commodity produced by developing countries and consumed by industrialized countries. Only six countries account for about 80% of the world market supply, while seven industrialized countries account for about 60% of world's consumption. In 1997/98, world production of cocoa beans reached 2.68 million tons.

The development of superior hybrids has made a significant contribution to cacao productivity, particularly in Brazil. There has been a consensus among experts that hybridization is the main breeding method applied to cacao and would be so for next decade (Gotsch, 1997). Dias et al. (1998)

have demonstrated the need for using cacao tree hybrids in the plantations; in a large field trial, cacao tree hybrids showed wide adaptability, low interaction with years and an outstanding performance for yield and its components, when compared to unimproved traditional local cultivars. Unfortunately, no data regarding productivity is available from the 250,000 ha of hybrid cacao plantations in Brazil. Hybrid cacao trees, which are non-conventional hybrids synthesized from non-inbred clonal progenitors, allowed for the renewal of plantations with high and early yielding material that also showed some resistance to diseases (Dias and Kageyama, 1995). The major worldwide drawback in cacao hybridization is its empiricism. Historically, the first criterion to be applied to cacao hybridization

* Author for correspondence

programs was the geographic divergence, as an indicator for genetic divergence. This is still a common practice, in many instances, however, geographic and genetic divergence do not correlate positively (Dias et al., 1997). Another frequently adopted criterion is based on the complementation of traits of interest, aiming at the correction of each parent's deficiencies. Crosses are conducted randomly, primarily when information about the parent clones is not available. Because the yield ability of the clones *per se* is not associated with their performance as progenitors, combining ability tests have been introduced to overcome part of the randomness of the hybridization process. Nevertheless, according to Bos and Sparnaaij (1993), the use of a combining ability test in a diallel crossing scheme to study a complex trait like yield has lost its appeal because of: i) the restricted resources in terms of land and labour which limit the number of parents that can be tested, and ii) the almost invariable conclusion that the variance of general combining ability effects exceeds the variance for specific combining ability effects. Apart from being costly and time consuming, these tests do not prevent the cacao breeder making crosses and field evaluations, and for this reason they are not predictive. Since the long juvenile period of cacao lasts from three to five years and a single selection cycle might last one decade, new predictive tools provided by quantitative and molecular genetics must be used more intensively.

The efficiency of a cacao breeding program could be increased by choosing the superior crosses to be made between divergent clones, and so reducing the number of crosses required for field trial. Dias and Kageyama (1997a) were able to associate the average and heterotic performance of cacao hybrids with the genetic divergence among their parents, estimated by the Mahalanobis D^2 statistic applied to performance data. The comparison of clusters of parents formed from the D^2 distance matrix based on pooled (5-year data) analysis and taken as standard, with clusters obtained from each year, showed a stable clustering pattern in the most favourable years. In other words, the divergence was stable in time, at least on a single climatically favourable crop year (Dias and Kageyama, 1997b). Inferences about adaptation and coadaptation patterns can be assessed by computing genetic distance estimates based on morphological data, since natural selection affects morphological traits linked to adaptive

characteristics. In addition, this category of traits can be particularly useful for classification below the species level (Camussi et al., 1985) and this strategy has a predictive nature because it avoids making and evaluating hundreds of undesirable crosses (Dias and Kageyama, 1997a).

More recently, data from biochemical and DNA markers have been used in estimating genetic distances and predicting heterosis in plant breeding. Biochemical markers, such as isozymes, allow the screening of only a limited number of loci. DNA-based markers, such as RFLP (Restriction Fragment Length Polymorphism) and VNTR (Variable Number of Tandem Repeats), are expensive, laborious and usually need radioactivity for labelling. Randomly Amplified Polymorphic DNA (RAPD) markers (Williams et al., 1990) overcome these disadvantages. With RAPD molecular markers a large number of loci can be screened. This technique has also shown to be adequate for studying the cacao genome (Ronning et al., 1995). RAPD uses arbitrary decamer primers to amplify specific portions of the genome. In addition, due to the random nature of the primers, RAPD markers, in contrast with RFLP, VNTR and isoenzymes, tend to represent a random sample from the entire genome. Further details on the impact of DNA markers on cacao breeding are available in Dias (1995).

In this work, our objectives were *i*) to investigate genetic diversity among five non-commercial cacao clones using yield components and RAPD data, *ii*) to compare genetic distance estimates based on highly heritable yield components data with those estimated from RAPD data and, *iii*) to determine the degree of association between genetic distance among parental entries and the performance of their hybrids. It is expected that genetic distance estimates based on performance and RAPD data may be utilized as a guide to identify cacao hybrids heterotic for yield and others important traits.

MATERIALS AND METHODS

Materials: This study was based on a 5 x 5 complete diallel cross involving five non-commercial clonal cacao cultivars tested jointly with their 20 possible hybrids. Seeds from crosses were obtained by controlled pollination of clonal cultivars and seeds from parental clones were produced by selfing. The five self-compatible

clonal cultivars, their racial groups and their origins were: 1) 'CC 41' - Trinitario, originating from Costa Rica, Centro de Cacao; 2) 'SIAL 169' - Lower Amazon, originating from southern Bahia and selected by the former Instituto Agronômico do Leste; 3) 'CEPEC 1' - Lower Amazon, from southern Bahia, corresponding to the clone Catongo, a cultivar selected as a spontaneous mutant with white seeds and named by the Centro de Pesquisas do Cacao; 4) 'ICS 1' - Trinitario, selected in Trinidad by the Imperial College of Tropical Agriculture; 5) 'SIC 19' - Lower Amazon, selected in southern Bahia by the former Instituto de Cacao da Bahia.

Field evaluation: The 25 entries (10 F_1 s, 10 reciprocal F_1 s and five parents) were grown in a randomized complete block design with four replications and square 16-plant plots. On assessing the divergence between the parents, the data representing five yield components (Table 2) collected over a period of five years was used. Combining ability analysis, agronomic practices, as well as details of performance data records can be found in a previous study (Dias and Kageyama, 1995).

Mahalanobis' genetic distance: Details of the parental divergence, estimated by the D^2 statistic of Mahalanobis' distance (Rao, 1952), and its association with mid-parent heterosis in their hybrids are described elsewhere (Dias and Kageyama, 1997a). When associating the 10 D^2 values of parental distances with their 20 hybrids, the 20 F_1 (10 F_1 s and 10 reciprocal F_1 s) data were pairwise combined to give 10 F_1 cross means. Cluster analysis was performed on the Mahalanobis distance (MD) with the unweighted pair-group method based on arithmetic averages (UPGMA).

RAPD analysis: DNA extractions were done from leaf tissue using five cacao clonal cultivars. Details of the DNA extraction and isolation procedures, RAPD primer selection and RAPD reactions are described elsewhere (Marita, 1998).

RAPD genetic distance: Data from the five cacao clonal cultivars came from the study performed by Marita (1998), who selected 38 primers to evaluate 270 clones from CEPEC's collection in Itabuna, Bahia, Brazil. As for the five cultivars, a total of 130 RAPD bands were scored as 1 for presence or

0 for absence of the band. Genetic distance (GD) was calculated using Nei and Li coefficient (Nei and Li, 1979). It was computed as follows:

$$GD_{ij} = 1 - 2N_{ij}/(N_i + N_j)$$

where N_{ij} was the number of bands shared by cultivars i and j ("positive" matches, type 11), and N_i was the number of bands in cultivar i , and N_j was the number of bands in cultivar j .

The GD matrix was submitted to cluster analysis following the UPGMA algorithm. To evaluate the robustness of the clusters formed, the binary dataset was submitted to bootstrapping using Winboot software (Yap and Nelson, 1996). The matrix data was reconstructed 2,000 times (replications) by resampling of bands with replacement. The accuracy of a bootstrap estimate is a function of both the value of the bootstrap and the number of bootstrap replications. To ensure the accuracy of the bootstrap estimate is $\pm 5\%$, for example, it is necessary to do 2000 replications if the bootstrap estimate is $\pm 95\%$. The frequency with which a particular group was clustered, was considered to reflect the robustness of the cluster. The association between GD and MD as well as between each one and the average and heterotic performance realized in the hybrids, were quantified using Spearman's rank correlation coefficient (r).

RESULTS AND DISCUSSION

Divergence among parental clones: A partial agreement between the genetic distance matrix based on performance data and the one based on RAPD data was found (Table 1), in terms of the two most divergent and the two closest pairs. Based on the performance data, CC 41 and ICS 1, and SIAL 169 and ICS 1 were the most divergent pairs, while CC 41 and SIC 19, and CEPEC 1 and SIC 19 were the most similar ones. As for RAPD data, SIAL 169 and ICS 1, and CEPEC 1 and ICS 1 were the most divergent, while SIAL 169 and CEPEC 1, and CC 41 and SIC 19 were the closest ones. Both kinds of data showed that ICS 1 was the most divergent clone in relation to the other four clones. It is important to bear in mind that not only the divergence between parental cultivars should be considered as a criterion for making crossing decisions, but also the performance *per se*

of such cultivars for traits of high economic importance, as illustrated by Dias and Kageyama (1997a). These authors showed that ICS 1 had the highest average for wet seed weight/fruit, even though it ranked fourth best in wet seed weight/plant. SIAL 169 had the second highest wet seed weight/fruit along with the highest wet seed weight/plant. However, the best performing hybrids with regard to wet seed weight/plant and wet seed weight/fruit were those involving ICS 1, where the SIAL 169 x ICS 1 hybrid and its

reciprocal were outstanding (Dias and Kageyama, 1995; 1997a). In addition, ICS 1 had a high general combining ability for the same components (Dias and Kageyama, 1995), suggesting a relationship between combining ability effects and parental genetic divergence. Nevertheless, this not expected apparent relationship, as will be shown in the next section, reflects the high additive genetic value of the ICS 1 parent.

Table 1 - Genetic distance (GD - above diagonal) between five cacao cultivars based on RAPD data and Mahalanobis distance (MD - below diagonal) based on performance data.

Cultivar	CC 41	SIAL 169	CEPEC 1	ICS 1	SIC 19
CC 41	0.00	0.23	0.25	0.28	0.18
SIAL 169	11.77	0.00	0.07	0.36	0.20
CEPEC 1	7.96	11.95	0.00	0.34	0.22
ICS 1	32.16	23.51	13.10	0.00	0.31
SIC 19	1.94	8.50	5.10	20.86	0.00

The corresponding dendrograms (Fig.1) showed the same divergence scenery. Cacao clones were classified into two main clusters: one cluster formed by ICS 1, while SIAL 169, CEPEC 1, SIC 19 and CC 41 together formed cluster two. CC 41 remains in the Lower Amazon cluster by both measures of distance, but the performance data (Fig. 1B) placed it much closer to SIC 19 than the RAPD data (Fig. 1A). Printed at each node of the dendrogram (Fig. 1A) was the percentage of times

the cluster to the left of that node was found. The percentage could be considered a nonparametric statistical test (confidence limits) on the validity of the various groups. The higher the percentage, the greater the confidence that a particular cluster was true rather than an artifact of the clustering algorithm. In a stringent way, Felsenstein (1985) suggested that only clusters with bootstrap values of 95% or greater could be considered significant.

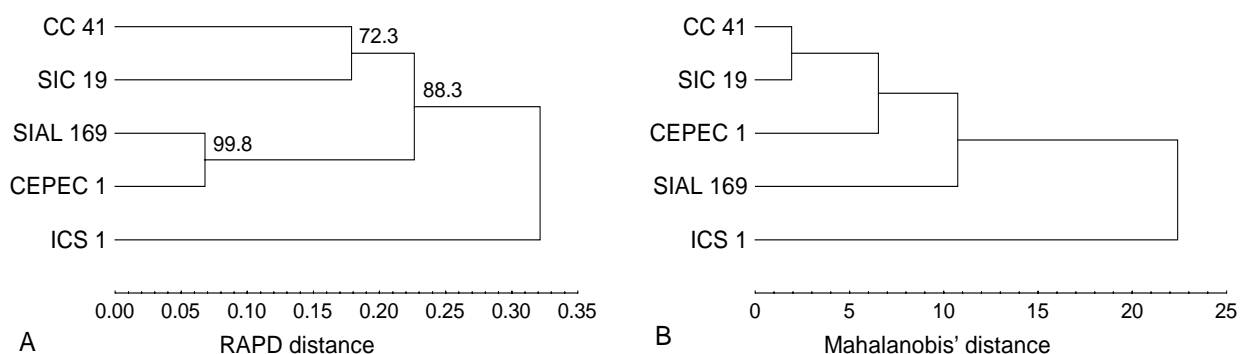


Figure 1 - Dendrograms of genetic distances of cacao cultivars based on RAPD data (A) and of Mahalanobis distance based on performance data (B). Bootstrap values (in %) are indicated at the corresponding node for each cluster in A.

Following this rule, and based on RAPD data (Fig. 1A), only the cluster (SIAL 169, CEPEC 1) was truly robust, with a bootstrap value of 99.8%. The cluster (CC 41, SIC 19) occurred in 72.3% of the bootstrap trees, and the cluster (CC 41, SIC 19, SIAL 169, CEPEC 1) occurred 88.3% of the time. Based on both cluster analyses, the divergence in germplasm was only partially in agreement with the conventional racial groups classification proposed by Cheesman (1944). By Cheesman's classification, cacao species are divided in two main racial groups: Criollo - originally cultivated by Maya people in Mesoamerica, some 3,000 years ago -, and 'Non-Criollo' or Amazon Forastero, which is found in a wild state in the Amazon basin, the putative center of origin of the species. The Trinitario complex is included into the Amazon Forastero group. It results from hybridization between Criollo and Amazon Forastero, and includes a mixture of hybrid types. Geographically, Forasteros are divided into the Lower and Upper Amazon subgroups. CC 41, a Trinitario clone, was shown to be closely related to the Lower Amazon Forastero SIC 19, and very distant from ICS 1, which is also a Trinitario clone. The CC 41 Trinitarian has originated from an open-pollinated progeny of the UF 676 clone. Raboin *et al.* (1993) based on morphological traits, and Laurent (1993) based on RFLP markers detected the similarity of UF 676, a Trinitario clone, to the Lower Amazon cultivars. Dias and Kageyama (1997a) observed that CC 41 could have received a higher contribution from Amazonian genes and, as a result, became similar to the clones from that group.

Association between distance measures: The correlation coefficient between GD and MD was moderate (Fig.2). Hence, both distance measures were sufficiently consistent to be used in the cacao hybridization programs for the identification of superior heterotic combinations. However, this correlation was greatly influenced by ICS 1, a parent with high general and specific combining ability (Dias and Kageyama, 1995). After removing ICS 1 from the diallel crossing the correlation disappeared ($r = 0.08$, $P = 0.87$). According to Sekhon and Gupta (1995), a perfect correspondence between the two types of distance measures was not expected since each data set represented a somewhat different part of the genome and was subject to different sources of error. While the number of markers and the

variation of the reference populations assayed were the strongest factors affecting GD, MD could be affected by environment, genotype x environment interaction, non-additive genetic effects, and the selection criteria used for the development of the cultivars. In agreement with previous reports a poor correlation between the two types of distances was found (Dillmann *et al.*, 1997; Sant *et al.*, 1999).

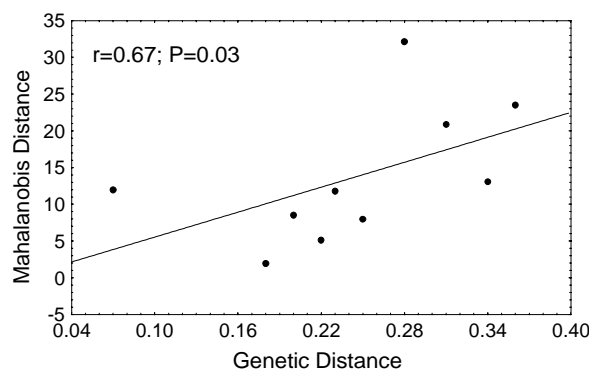


Figure 2 - Scatterplot of Genetic distance vs Mahalanobis distance for five cacao clones.

In our case, there are good reasons for the strong association between the GD and MD matrices. Firstly, the number of markers generated by the RAPD technique was not limiting, and we assayed cacao clones from different racial groups, which increased the diversity. Secondly, with regard to the phenotypic distance (MD), we used yield components data coming from 1,600 observations (5 clones x 5 years x 4 replications x 16-mature plants per plot) resulting in good field accuracy. Finally, for the same diallel crossing, lack of genotype x year interaction was found, and non-additive genetic effects were detected for the number of healthy fruits and the number of collected fruits/plant, and wet seed weight/plant (see Dias and Kageyama, 1995; 1997a), in terms of the average squared effects.

Association between heterosis and parental genetic distance: Both kinds of genetic distances (GD and MD) were related to heterotic performance of hybrids for wet seed weight/plant and wet seed weight/fruit (Table 2). Nevertheless, the average hybrid performance for the same two yield components was correlated with only MD. In general, the correlation values were high for traits which best reflect the commercial cacao yield such

as wet seed weight/plant and wet seed weight/fruit, and for which high mid-parent heterosis was found (55.4% and 8.8%, respectively, as reported by Dias and Kageyama, 1995). These correlations, while not perfect, seem sufficiently high to choose parents, which would produce crosses with superior performance, thus reducing the number of crosses required for field trial. The positive association between genetic distance and heterosis for a given trait depends on the presence of dominance effects and/or of differences in the frequency of the alleles controlling the trait considered (Falconer, 1960). In a previous analysis of the same diallel (Dias and Kageyama, 1995), dominance effects for the number of healthy fruits and the number of collected fruits/plant and, wet seed weight/plant were demonstrated, while additivity was held for wet seed weight/fruit. In addition, Arunachalam (1977) has also proven the role of additive x additive epistatic interactions in causing heterosis.

Findings on good correlations between the genetic distance of parents (from performance and molecular marker data) and the performance of

their hybrids were similar to those reported in cacao and several species (Ghaderi et al., 1984; Shamsuddin, 1985; Cox and Murphy, 1990; Cruz et al., 1994; Ali et al., 1995; Diers et al., 1996; Xiao et al., 1996; Dias and Kageyama, 1997a; Kisha et al., 1997; Lanza et al., 1997; Manjarrez-Sandoval et al., 1997; Saghai Maroof et al., 1997; Burkhamer et al., 1998). However, there are several reports describing poor correlations (Vaillancourt et al., 1995; Cerna et al., 1997; Helms et al., 1997; Chowdari et al., 1998; Fabrizius et al., 1998; Perenzin et al., 1998; Tenkouano et al., 1998; Bohn et al., 1999; Joyce et al., 1999; Kidwell et al., 1999; Sant et al., 1999). Such contrasting results showed that association between parental divergence and heterosis remained an unsolved question. Although genetic distance does not affect heterosis in a linear fashion, it is still important for obtaining heterosis in crosses. In many circumstances, the expression of heterosis is partly due to genetic diversity although it also depends on several other factors not completely elucidated, and for this reason, it has been unpredictable.

Table 2 - Spearman correlation coefficients between parental genetic distances (GD and MD) and average and heterotic hybrid performance for five traits.

Traits	Spearman correlation coefficients			
	Average hybrid performance		Heterotic hybrid performance	
	GD	MD	GD	MD
Number of healthy fruits/plant	-0.00 (0.03)	0.11 (-0.03)	0.56 (0.14)	0.43 (-0.26)
Number of collected fruits/plant	0.02 (0.14)	0.09 (-0.08)	0.61 (0.37)	0.42 (-0.20)
Wet seed weight/plant	0.61 (-0.26)	0.68* (0.08)	0.73* (0.14)	0.66* (-0.26)
Wet seed weight/fruit	0.62 (-0.77)	0.76** (0.03)	0.65* (-0.60)	0.66* (-0.37)
Percentage of diseased fruits/plant	-0.26 (-0.03)	-0.30 (-0.03)	-0.10 (0.14)	-0.20 (-0.20)

In parentheses are Spearman correlation coefficients, after removing ICS 1 clone from analysis;

*P < 0.05 and ** P < 0.01, respectively.

The interpretation of the association between parental genetic distance and hybrid yield performance in genetic marker terms is not easy because markers should be located across the whole genome to accurately predict heterosis. However, convenient estimates of genetic distances from molecular marker data can be obtained by using a suitable DNA sampling technique. In turn, the RAPD markers represent

random samples from the entire genome (Williams et al., 1990). Hence, it is possible to reliably predict heterosis for yield and its components by using a randomly selected set of RAPD markers. Moreover, the technical simplicity and the less costly and time-consuming RAPD technique accounts for its use in developing countries. In view of these countries' limited financial resources, as discussed by Dias (1995), scientists

of such countries should concentrate efforts on the use of low cost techniques in their research centers.

When 'ICS 1' was removed from the present diallel, the observed correlations disappeared (Table 2, values in the parenthesis), showing the specificity of this clone. It follows that associations between parental genetic distance and heterosis are highly dependent on the racial groups that are considered. In fact, this association has been shown to be dependent on the type of germplasm employed (Kisha et al., 1997; Helms et al., 1997; Lanza et al., 1997; Saghai Maroof et al., 1997). Since the association procedure worked for this small diallel, used as a model, it could be expected that it might also work for the prediction of hybrid performance from crosses among a larger number of cultivars. The small diallel that could be analyzed limits the scope of the conclusions, but the results desired more studies.

ACKNOWLEDGMENTS

L.A.S. Dias was supported by a grant from the CNPq/Brasil. Special thanks are given to Dr. James Nienhuis and his laboratory at the University of Wisconsin, Madison where the RAPD reactions were completed. Thanks are also extended to Almirante Centro de Estudos de Cacao who sponsored the analysis.

RESUMO

A eficiência do programa de melhoramento de cacau incrementará pela seleção de cruzamentos superiores entre clones divergentes. Foram estimadas as distâncias genéticas (GD), com dados de RAPD, e de Mahalanobis (MD), com dados de componentes de produção, entre cinco clones. Os clones foram avaliados em um dialelo, por cinco componentes de produção, durante cinco anos. Um total de 130 bandas RAPD foram avaliadas. GD e MD foram utilizadas para cálculo da correlação entre distâncias entre clones e a performance de seus híbridos. A correlação entre GD e MD foi de 0.67 ($P=0.03$). Ambas as distâncias foram correlacionadas à performance heterótica dos híbridos para peso de sementes frescas/planta e peso de sementes frescas/fruto. A performance híbrida média para os mesmos componentes de

produção foi correlacionada somente com MD. Assim, distâncias genéticas medidas por RAPD e componentes de produção podem ser usadas como um guia para a escolha de cruzamentos superiores.

REFERENCES

- Ali, M.; Copeland, L. O.; Elias, S. G. and Kelly, J. D. (1995), Relationship between genetic distance and heterosis for yield and morphological traits in winter canola (*Brassica napus* L.). *Theor. Appl. Genet.*, **91**, 118-121.
- Arunachalam, V. (1977), Heterosis for characters governed by two genes. *J. Genet.*, **63**, 15-24.
- Bohn, M.; Friedrich Utz, H. and Melchinger, A. E. (1999), Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their use for predicting progeny variance. *Crop Sci.*, **39**, 228-237.
- Bos, I. and Sparnaaij, L. D. (1993), Component analysis of complex characters in plant breeding. II. The pursuit of heterosis. *Euphytica*, **70**, 237-245.
- Burkhamer, R. L.; Lanning, S. P.; Martens, R. J.; Martin, J. M. and Talbert, L. E. (1998), Predicting progeny variance from parental divergence in hard red spring wheat. *Crop Sci.*, **38**, 243-248.
- Camussi, A.; Ottaviano, E.; Calinski, T. and Kaczmarek, Z. (1985), Genetic distances based on quantitative traits. *Genetics*, **111**, 945-962.
- Cerna, F. J.; Cianzio, S. R.; Rafalski, A.; Tingey, S. and Dyer, D. (1997), Relationship between seed yield heterosis and molecular marker heterozygosity in soybean. *Theor. Appl. Genet.*, **95**, 460-467.
- Cheesman, E. E. (1944), Notes on the nomenclature, classification and possible relationships of cacao populations. *Trop. Agric.*, **21**, 144-159.
- Chowdari, K. V.; Venkatachalam, S. R.; Davierwala, A. P.; Gupta, V. S.; Ranjekar, P. K. and Govila, O. P. (1998), Hybrid performance and genetic distance as revealed by the (GATA)₄ microsatellite and RAPD markers in pearl millet. *Theor. Appl. Genet.*, **97**, 163-169.
- Cox, T. S. and Murphy, J. P. (1990), The effect of parental divergence on F₂ heterosis in winter wheat crosses. *Theor. Appl. Genet.*, **79**, 241-250.
- Cruz, C. D.; Carvalho, S. P. and Vencovsky, R. (1994), Estudos sobre divergência genética. II Eficiência da predição do comportamento de híbridos com base na divergência de progenitores. *Rev. Ceres*, **41**, 183-190.
- Dias, L. A. S. (1995), Biotecnologia e melhoramento genético do cacau (Theobroma cacao L.). *Agrotropica*, **7**, 1-14.
- Dias, L. A. S. and Kageyama, P. Y. (1995), Combining-ability for cacao (*Theobroma cacao* L.) yield components under southern Bahia conditions. *Theor. Appl. Genet.*, **90**, 534-541.

- Dias, L. A. S. and Kageyama, P. Y. (1997a), Multivariate genetic divergence and hybrid performance of cacao (*Theobroma cacao* L.). *Brazil. J. Genet.*, **20**, 63-70.
- Dias, L. A. S. and Kageyama, P. Y. (1997b), Temporal stability of multivariate genetic divergence in cacao (*Theobroma cacao* L.) in Southern Bahia conditions. *Euphytica*, **93**, 181-187.
- Dias, L. A. S.; Kageyama, P. Y. and Castro, G. C. T. (1997), Divergência fenética multivariada na preservação de germoplasma de cacau (*Theobroma cacao* L.). *Agrotropica*, **9**, 29-40.
- Dias, L. A. S.; Souza, C. A. S.; Augusto, S. G.; Siqueira, P. R. and Müller, M. W. (1998), Performance and temporal stability analyses of cacao cultivars in Linhares, Brasil. *Plant Rec. Develop.*, **5**, 343-355.
- Diers, B. W.; McVetty, P. B. E. and Osborn, T. C. (1996), Relationship between heterosis and genetic distance based on restriction fragment length polymorphism markers in oilseed rape (*Brassica napus* L.). *Crop Sci.*, **36**, 79-83.
- Dillmann, C.; Bar-Hen, A.; Guérin, D.; Charcosset, A. and Murigneux, A. (1997), Comparison of RFLP and morphological distances between maize *Zea mays* L. inbred lines. Consequences for germplasm protection purposes. *Theor. Appl. Genet.*, **95**, 92-102.
- Fabrizius, M. A.; Busch, R. H.; Khan, K. and Huckle, L. (1998), Genetic diversity and heterosis of spring wheat crosses. *Crop Sci.*, **38**, 1108-1112.
- Falconer, D. S. (1960), *Introduction to quantitative genetics*. The Ronald Press, New York.
- Felsenstein, J. (1985), Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783-791.
- Ghaderi, A.; Adams, M. W. and Nassib, A. M. (1984), Relationship between genetic distance and heterosis for yield and morphological traits in dry edible bean and faba bean. *Crop Sci.*, **24**, 37-42.
- Gotsch, N. (1997), Cocoa biotechnology: status, constraints and future prospects. *Biotechnology Adv.*, **15**, 333-352.
- Helms, T.; Orf, J.; Vallard, G. and McClean, P. (1997), Genetic variance, coefficient of parentage, and genetic distance of six soybean populations. *Theor. Appl. Genet.*, **94**, 20-26.
- Joyce, T. A.; Abberton, M. T.; Michaelson-Yeates, T. P. T. and Forster, J. W. (1999), Relationships between genetic distance measured by RAPD-PCR and heterosis in inbred lines of white clover (*Trifolium repens* L.). *Euphytica*, **107**, 159-165.
- Kidwell, K. K.; Hartweck, L. M.; Yandell, B. S.; Crump, P. M.; Brummer, J. E.; Moutray, J. and Osborn, T. C. (1999), Forage yields of alfalfa populations derived from parents selected on the basis of molecular marker diversity. *Crop Sci.*, **39**, 223-227.
- Kisha, T. J.; Sneller, C. H. and Diers, B. W. (1997), Relationship between genetic distance among parents and genetic variance in populations of soybean. *Crop Sci.*, **37**, 1317-1325.
- Lanza, L. L. B.; Souza Jr., C. L.; Ottoboni, L. M. M.; Vieira, M. L. C. and Souza, A. P. (1997), Genetic distance of inbred lines and prediction of maize single-cross performance using RAPD markers. *Theor. Appl. Genet.*, **94**, 1023-1030.
- Laurent, V. (1993), Etude de la diversité génétique du cacaoyer (*Theobroma cacao* L.) basée sur le polymorphisme de la longueur des fragments de restriction (RFLP). PhD Thesis, Université de Paris, Paris, France.
- Manjarrez-Sandoval, P.; Carter Jr., T. E.; Webb, D. M. and Burton, J. W. (1997), RFLP genetic similarity estimates and coefficient of parentage as genetic variance predictors for soybean yield. *Crop Sci.*, **37**, 698-703.
- Marita, J. (1998), Characterization of *Theobroma cacao* using RAPD marker based estimates of genetic distance and development of a core collection to maximize genetic diversity. MSc Thesis, University of Wisconsin, Madison, USA.
- Nei, M. and Li, W. H. (1979), Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci.*, **76**, 5296-5273.
- Perenzin, M.; Corbellini, M.; Accerbi, M.; Vaccino, P. and Borghi, B. (1998), Bread wheat: F₁ hybrid performance and parental diversity estimates using molecular markers. *Euphytica*, **100**, 273-279.
- Raboin, L. M.; Paulin, D.; Cilas, C. and Eskes, A. B. (1993), Analyse génétique de quelques caractères quantitatifs des fleurs du cacaoyer (*Theobroma cacao* L.). Leur intérêt pour l'évaluation de la diversité de l'espèce. *Café Cacao Thé*, **37**, 273-282.
- Rao, C. R. (1952), *Advanced statistical methods in biometric research*. John Wiley, New York.
- Ronning, C. M.; Schnell, R. J. and Kuhn, D. N. (1995), Inheritance of random amplified polymorphic DNA (RAPD) markers in *Theobroma cacao* L. *J. Am. Soc. Hort. Sci.*, **120**, 681-686.
- Saghai Maroof, M. A.; Yang, G. P.; Zhang, Q. and Gravois, K. A. (1997), Correlation between molecular marker distance and hybrid performance in U.S. Southern long grain rice. *Crop Sci.*, **37**, 145-150.
- Sant, V. J.; Patankar, A. G.; Sarode, N. D.; Mhase, L. B.; Sainani, M. N.; Deshmukh, R. B.; Ranjekar, P. K. and Gupta, V. S. (1999), Potential of DNA markers in detecting divergence and in analysing heterosis in Indian elite chickpea cultivars. *Theor. Appl. Genet.*, **98**, 1217-1225.
- Sekhon, M. S. and Gupta, V. P. (1995), Genetic distance and heterosis in Indian mustard: developmental isozymes as indicators of genetic relationships. *Theor. Appl. Genet.*, **91**, 1148-1152.

- Shamsuddin, A. K. M. (1985), Genetic diversity in relation to heterosis and combining ability in spring wheat. *Theor. Appl. Genet.*, **70**, 306-308.
- Tenkouano, A.; Crouch, J. H.; Crouch, H. K. and Ortiz, R. (1998), Genetic diversity, hybrid performance, and combining ability for yield in *Musa* germplasm. *Euphytica*, **102**, 281-288.
- Vaillancourt, R. E.; Potts, B. M.; Watson, W.; Volker, P. W.; Hodge, G. R.; Reid, J. B. and West, A. K. (1995), Detection and prediction of heterosis in *Eucalytus globulus*. *For. Genet.*, **2**, 11-19.
- Williams, J. G. K.; Kubeliak, A. R.; Livak, K. J.; Rafalski, J. A. and Tingey, S. V. (1990), DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.*, **18**, 6531-6535.
- Xiao, J.; Li, J.; Yuan, L.; Mc Couch, S. R. and Tanksley, S. D. (1996), Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. *Theor. Appl. Genet.*, **92**, 637-643.
- Yap, I. V. and Nelson, R. J. (1996), *Winboot: a program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms*. IRRI Discussion Paper Series 14, Manila.

Received: November 19, 2001;

Revised: June 04, 2002;

Accepted: October 24, 2002.