

Distinctness of cacao cultivars using yield component data and RAPD markers

Luiz Antônio dos Santos Dias^{1*}, Rodrigo Barros Rocha¹, and Edgard Augusto de Toledo Picoli¹

Received 2 November 2004

Accepted 2 March 2005

ABSTRACT - *The identification of cacao cultivars is crucial now that new cacao clones are available to producers. We compared the efficiency of the RAPD markers in relation to yield components for the discrimination of five cultivars ('Maranhão', 'Pará', 'Parazinho', the open-pollinated 'ICS 1', and the commercial hybrid mixture). Three yield components evaluated over 10 years and 182 polymorphic bands were scored for genetic distances and multidimensional scaling analyses. We found a close relationship between the genetic distance (GD) of RAPD data and Mahalanobis distance (MD) of yield component data, and between the biplots of GD and MD, although only RAPD markers were able to distinguish the five cultivars from each other. Biplots were able to discriminate cultivars with different improvement levels. Thirty-one primers were needed for an exclusive identification of each cultivar. In future identifications of cacao clones these primers will allow the construction of clone specific-SCAR primers of high reproducibility.*

Key words: *Theobroma cacao* L., cultivar differentiation, RAPD markers, yield components.

INTRODUCTION

Cacao (*Theobroma cacao* L.), despite its economical importance, lacks a formal committee in charge of the registration and documentation of new cultivars as well as the information on and supervision of supply of genetic materials to producers. For the constitution of such a committee, the procedures to be used for cultivar identification must be established, defined, and standardized. This study focuses on the distinctness of cultivars through information of both agronomic

descriptors and molecular markers.

The determination of varietal distinctness is an essential procedure to warrant the uniqueness of a variety. An unequivocally identified variety is appropriate for protection and registration, which gives the authorities conditions to control and monitor its trade and breeders to assure their intellectual property rights. In many countries the controlling legislation follows the Convention of the International Union for the Protection of New Varieties of Plants - UPOV (UPOV 1991). Candidate varieties for protection and registration must be tested for

¹BIOAGRO, Universidade Federal de Viçosa, 36.571-000, Viçosa, MG, Brasil. *E-mail: lasdias@ufv.br

distinctness, uniformity and stability (DUS). Brazil is signatory of the UPOV's Convention 1978. The Brazilian law (number 9.456; April 25, 1997) guarantees the intellectual property of the breeder who develops a variety and characterizes the descriptors that differentiate it from others. As descriptor, the Brazilian law defines any morphologic, physiologic, biochemical or molecular characteristic that is genetically inherited. However, in the practice, only morphologic descriptors are accepted in the description of a variety (Rocha et al. 2002). In spite of their easiness of identification, their limited number on the background of the increasing number of already registered varieties impairs the identification of essentially derived varieties (Becher et al. 2000). Since many of these descriptors can only be evaluated at adult age, lawsuits over fraudulent appropriations of protected variety may drag on over years (Wunsch and Hormaza 2002).

The use of molecular genetic markers, together with morphologic descriptors, may help overcome many of these limitations. Molecular markers appear in practically limitless number and have the potential to provide a clear identification of varieties (Becher et al. 2000, Wunsch and Hormaza 2002) at any stage of the plant development. Molecular differentiation is assumed to reflect differences of the DNA sequences among varieties (Staub et al. 1996). However, the identification and quantification of genetic divergence among varieties is a very complex problem. There is no technique that could provide an absolute measure for genetic divergence, partly because this would require comparisons among all genome sequences. The obtained measures are result of the analyzed genetic material, the sampling strategy and the sensitivity degree of the applied technique (Roldán-Ruiz et al. 2001).

Random Amplified Polymorphic DNA (RAPD) markers (Williams et al. 1990) have come to be commonly used in studies of varietal/accession differentiation of cacao (Wilde et al. 1992, Figueira 1998, Dias 2001, Dias et al. 2001, Faleiro et al. 2001). RAPD markers have the advantages of relatively low costs, simplicity, speed, genome-wide coverage and operational automation for a large number of single or bulk DNA samples. Further details on the impact of DNA markers on cacao breeding are available in Dias (1995). In a practical way, genetic polymorphism among populations can be evaluated using bulks of DNA. The bulk technique consists in the use of equal DNA amounts of individuals from the same population. Genetic polymorphism among bulks is directly related to the genes that are present in one population and absent in another.

Cacao cultivars planted in Brazil for the last three

centuries belong to the Amazon Forastero racial group, known as 'Comum', 'Pará', and 'Maranhão' (Vello et al. 1969). There are other types in-between 'Pará' and 'Maranhão' such as 'Parazinho', probably derived from 'Pará', as well as 'Maranhão Liso' and 'Maranhão Rugoso', both derived from 'Maranhão' (Vello and Garcia 1971). According to Soria (1963) and Vello and Garcia (1971), these types arose from mutation, segregation and from known or unknown introductions. The main yield components from these cacao cultivars were shown by Vello and Garcia (1971). The expansion of cacao cultivation in recent decades demanded the creation of high yield hybrid cultivars involving crosses between local and introduced clones. The commercial use of heterosis in Brazil and Trinidad began in the sixties (Dias et al. 1998). Nowadays, clonal plantings in Malaysia and Indonesia are being used with great success. In Brazil, where approximately 60000 hectares have already been replanted with clones, the expectations of cacao producers and breeders, who use this strategy, are high (Dias et al. 2001). Cloning at a commercial scale allows the fixation of desirable hybrid mixtures, in particular the fixation of genes for resistance to witches' broom - the most destructive cacao disease of the Americas. Clonal planting requires the development of procedures and methods for the differentiation and registration of clonal cultivars - objectives to which this study is dedicated and may contribute.

Cacao cultivars and types have been differentiated, basically, according to their yield, fruit (form, weight, and size), and seed traits (weight and number), besides others (Vello et al. 1969). The characterization of materials was based on univariate analyses of mean data obtained during three years, from one-tree plots in 10 replications (Vello and Garcia 1971). The application of univariate tests for a differentiation of the referred materials, however, has presented limitations. Vello and Garcia (1971) argued that it is not possible to identify each cultivar or type by the exam of a simple trait. In such cases, the use of multivariate distances, which consider different measure scales of characteristics and the correlation among them, may be advantageous. Nevertheless, if the simultaneous application of molecular markers and multivariate distances provides similar results, the possibility of using the former to identify cacao varieties is open.

Goals of this paper are: i) characterization of five cacao varieties by yield component data and RAPD markers, and ii) evaluation of efficiency of the RAPD markers for varietal identification as compared with yield component data.

Yield component data

Cacao cultivars evaluated for distinctness in a comparative yield trial and the procedures of data collection have been reported earlier (Dias et al. 1998). Briefly, the five assessed cultivars were 'Maranhão', 'Pará', 'Parazinho', the open-pollinated 'ICS 1', and the commercial hybrid mixture, which is known as 'Hybrid' and whose seeds are distributed free of charge by CEPLAC (Comissão Executiva do Plano da Lavoura Cacaueira). The hybrid comprised at least five clonal parent combinations, consisting of clonal parents from Bahia (series SIC and SIAL) and Espírito Santo (serie EEG), in crosses with introduced clones, such as SCA, UF, IMC, DR, among others (Dias et al. 1998). The experiment was set up in February 1982, in a 5 x 5 Latin square design, with 196-plant plots. The harvest was monitored monthly for over 10 years (1984 to 1993). Three yield components were evaluated: number of healthy fruits per plant, wet seed weight per fruit (g fruit⁻¹), and wet seed weight per hectare (kg ha⁻¹); these data were shown by Dias et al. (1998).

Mahalanobis' genetic distance

Details of the varietal differentiation using the multivariate criterion of Weatherup (1994) are described elsewhere (Dias 2001). Briefly, the D² statistic of Mahalanobis' distance (Mahalanobis 1936) was performed for all cultivar pairs, using the three mentioned yield components averaged over years. Weatherup's (1994) criterion consists in applying the T² statistics of Hotelling (Hotelling 1931) to the Mahalanobis matrix to generate the critical D². The Mahalanobis distance (MD matrix) was submitted to multidimensional scaling (MDS) algorithm. To evaluate how well a particular spatial configuration (biplot) reproduces the observed distance matrix we used the stress statistics, a goodness-of-fit measure. The smaller the stress value, the better is the fit of the reproduced distance matrix. The MDS algorithm and stress statistics were described by Dias (1998).

RAPD data

For the RAPD analysis, fresh and healthy ripe leaves of all five cultivars were collected from the same comparative yield trial where the yield components had been evaluated. These leaves, properly identified, were obtained at the Instituto de Biotecnologia aplicada à Agropecuária - BIOAGRO - at the Universidade Federal de Viçosa (UFV), Brazil. A DNA bulk composed of 6 to 10 trees, selected randomly on each plot, was formed for each cultivar and the DNA extracted in agreement with Doyle and Doyle's (1990) method, with minor modifications

(adding 1% insoluble PVP and 0.4% β-mercaptoethanol to the extraction buffer). Amplification reactions were performed according to Williams et al. (1990). They were carried out in a thermal cycler (Programmable Thermal Controller-100, MJ Research Inc.) programmed for an initial step of two minutes at 94 °C and 40 cycles of 30 seconds at 94 °C, of 30 seconds at 35 °C and of 1 minute at 72 °C, followed by a final extension step of 7 minutes at 72 °C.

The amplified products were analyzed by electrophoresis in 1.5% agarose gel and stained with ethidium bromide. The bands were visualized on an ultraviolet transilluminator, using the Eagle Eye Video System (Stratagene). A total of 217 random primers (Operon Technologies Inc.) were used in this analysis, which generated a mean of 7 bands per primer ranging from 4 to 13.

RAPD genetic distance

A total of 182 reliable polymorphic bands were used for the differentiation of cultivars, coded as 1 for the presence and 0 for the absence of bands. Only the bands within an interval of defined size (from 2027 to 493 bp) (see Figure 1) were taken into account to determine each individual's characteristic band pattern. The genetic similarities (GS_{ij}) were estimated for all pairs of cultivars i and j by the coefficient of Jaccard (Dias 1998). The genetic distance (GD_{ij}) was obtained as the arithmetic complement of the genetic similarity, in other words,

$$GD_{ij} = 1 - GS_{ij}$$

Next, the GD matrix was submitted to the MDS algorithm (Dias 1998).

All statistical analyses described for yield component and RAPD data were performed using SAS (SAS Institute Inc. 1989) and Statistica (StatSoft 1997) software, respectively. To evaluate the efficiency of a given primer for the purpose of cultivar identification the D statistics of Tessier (Tessier et al. 1999) was performed. This statistic reflects the probability that two randomly chosen individuals have different band patterns, besides selecting the optimal combination of primers needed to identify a set of cultivars.

RESULTS AND DISCUSSION

The GD among the cultivars varied from 0.392 for 'Pará' and 'Parazinho' to 0.835 between 'Maranhão' and 'ICS 1' (Table 1). The MD among the cultivars varied from 0.621 for 'Pará' and 'Parazinho' to 63.620 between 'Pará' and 'ICS 1' (Table 1). These results show a great amount of variability in the set of cultivars and are in agreement in relation to the nearest pair of cultivars.

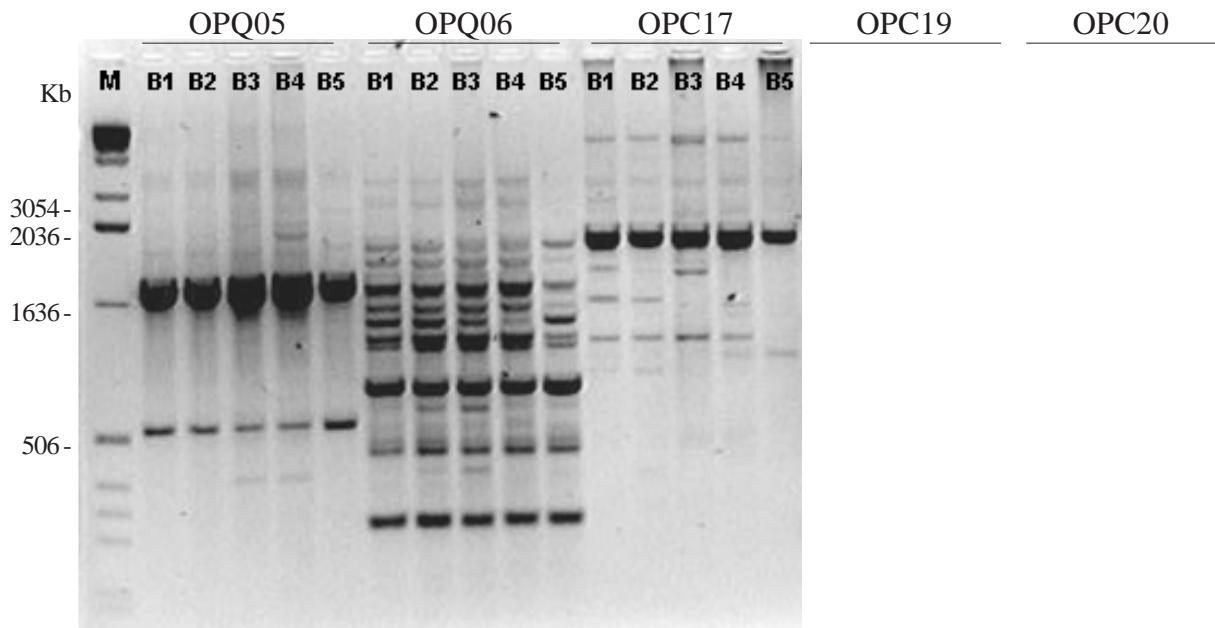


Figure 1. Agarose gel (1.5%) electrophoresis of five DNA bulks of RAPD-amplified cacao cultivars (see code in Table 1) using the primers OPQ05, OPQ06, OPC17, OPC19, and OPC20. The first column M correspond to the DNA fragment size pattern

Cultivar ICS 1, a Trinitarian, was the most distant from the others (Table 1 and Figure 2). Dias and Kageyama (1997), when studying the morphological divergence among local cacao selections from Bahia (CEPEC 1, SIAL 169 and SIC 19) and introduced selections from Costa Rica (CC 10) and Trinidad (ICS 1) also showed that the latter cultivar was the most distant among them. For the previous cultivars, a similar pattern of molecular divergence by RAPD was also achieved (Dias et al. 2003). In turn, the 'Hybrid' was distant from 'ICS 1' and at an equal distance from the cultivars Pará and Parazinho (Table 1 for MD). This result was coherent, since D^2 indicated that 'Maranhão' is at the same distance from 'Pará' and 'Parazinho' (Table 1), which does not distinguish cultivar Pará from Parazinho (Table 1 for MD). On the other hand, the Hybrid was close to cultivar Maranhão (critical D^2 not significant at the 1% probability level, as shown in Table 1 for MD), suggesting that improved material retains genes from the local cultivar at a higher proportion than from 'Pará' and 'Parazinho'.

The distinctness of the cultivars was visualized in MDS biplots (Figure 2). In both biplots, the cultivars were distinct from each other, except for 'Pará' and 'Parazinho' for MD (Figure 2b). Although these two cultivars were allocated close to each other in the same quadrant in both biplots (Figures 2a and 2b) built with the GD and MD matrices, only RAPD markers (Figures 2a) were able to distinguish 'Pará' from 'Parazinho'. In general,

the biplots show the formation of two clusters, one composed of traditional, unimproved cultivars (Maranhão, Pará, and Parazinho) on the left quadrant, and the other formed by two improved cultivars (open-pollinated ICS 1 and Hybrid) on the right quadrant (Figure 2). The 'Pará'- 'Parazinho' cluster (Figure 2b) was confirmed based on the critical D^2 (Table 1). Traditional cultivars belong to the Amazon Forastero racial group, whereas ICS 1 is a Trinitarian - a hybrid group between Forastero and Criollo. The Hybrid contains genes from local and introduced Trinitarian clones. 'ICS 1' and the Hybrid expressed better performance and temporal stability of yield than the other three cultivars throughout the 10 study years (1984-1993), as shown by Dias et al. (1998). These results suggest that the RAPD markers and yield components were able to distinguish cacao cultivars of different improvement levels. Besides this exception only RAPD markers were able to distinguish the five cultivars from each other.

The use of differentiation procedures based on molecular markers benefits future identifications of already described clones with speed and practice. Studies reveal that molecular marker techniques can be highly informative and effective in the differentiation of varieties/accessions in cacao (Wilde et al. 1992, Figueira 1998, Lanaud et al. 1999, Dias 2001, Dias et al. 2001, Faleiro et al. 2001, Motilal and Butler 2003, Swanson et al. 2003). However, an agreement among morphological and molecular divergence cannot always be expected (Lerceteau et al.

1997, Grant et al. 2001, Róldan-Ruiz et al. 2001). Róldan-Ruiz et al. (2001) suggest that such disparities should not just be understood as a result of different methodologies, since the morphological divergence is not necessarily a function of genetic differences; different gene pools can be manipulated to generate similar phenotypes. In our study we found a close relationship between the two types of divergence. Main factors causing this agreement were: the use of an experimental design, a robust measure of distance (see Dias and Kageyama 1998), yield component data averaged over ten years and the evaluation of five cultivars only. For the molecular divergence we use a marker such as RAPD with wide genome coverage, in a reasonable number of polymorphic bands, which exceeds the mean number (160) used in 139 studies of divergence (Dias et al. 2004). In this context the usefulness of molecular marker techniques for variety identification should be evaluated case by case, taking the aforementioned statements into consideration.

Although the Brazilian law on variety protection foresees the use of molecular markers as descriptors, the fingerprint of any cultivar protected in Brazil is characterized (<http://www.agricultura.gov.br/snpc/>). The lack of standardization of molecular characterization procedures has hampered the use of these techniques. A commonly used procedure to reduce the effort for genotyping is to

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

Cultivars (code)	Maranhão (B1)	Pará (B2)	Parazinho (B3)	ICS 1 (B4)	Hybrid (B5)
Maranhão	0	0.489	0.476	0.835	0.777
Pará	6.699**	0	0.392	0.744	0.700
Parazinho	6.185**	0.621	0	0.750	0.772
ICS 1	29.079**	63.620**	60.192**	0	0.649
Hybrid	2.126*	15.056**	15.262**	18.857**	0

$D_{critical}^2 = 2.808$; **($P < 0.01$)

$D_{critical}^2 = 1.840$; *($P < 0.05$)

Table 2. Number of primer combinations for discrimination in the set of five cacao cultivars

Cultivars	Number of trees for bulking cultivars	Number of primers for unique identification
Maranhão	10	4
Pará	10	2
Parazinho	06	4
ICS 1	07	18
Hybrid	10	3
Total		31

bulk individuals from each cultivar (bulks of DNA); a quite common strategy used in studies of genetic divergence, which results in a larger sampling potential of the population. Based on the assumption that different populations should present different gene frequencies, Fu (2000) argues that a DNA bulk composed of a minimum of 5 to 10 individuals randomly selected from either population is enough to represent and characterize the genetic divergence among populations, using dominant or co-dominant markers. We normally bulk 6 to 10 trees from each cultivar (Table 2). DNA bulks can be analyzed by means of several molecular marker techniques such as RFLP, RAPD, AFLP, and SSR.

Here, we have proposed, applied, and compared two procedures for the differentiation of cacao cultivars. The first was based on morphological data, in our case yield components appraised in field experiments. This procedure required 10 years of monthly data collection, besides the three years of juvenility before the trees entered the productive age. In spite of the considerable time consume and expenses with handwork, the achievement of morphological descriptors is necessary for the description of a cultivar. The use of differentiation procedures based on molecular markers will benefit future identifications of already described cultivars, indispensable for the protection of clones, with speed and practice. The use of

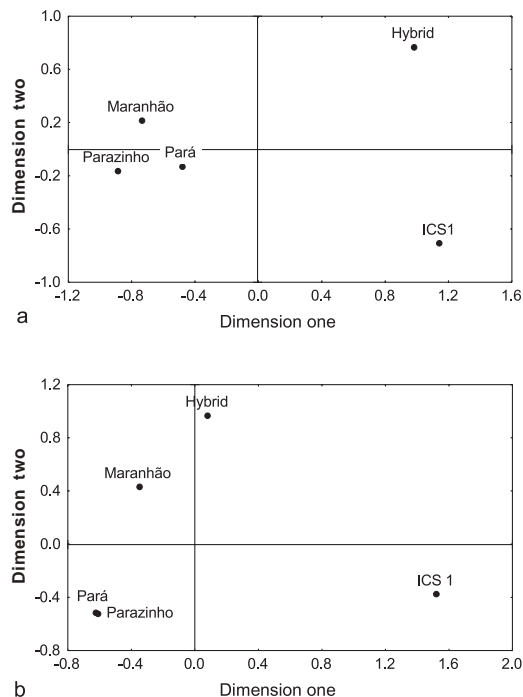


Figure 2. Biplots of multidimensional scaling built with the RAPD genetic distance (a; stress = 0.000000, after 0 iteration) and Mahalanobis distance (b; stress = 0.000345 after 100 iterations)

the RAPD together with the DNA bulk methodology seems to be a practical and quick strategy to evaluate the agreement between the morphological and genetic divergence. The agreement between these two divergence types facilitates the identification of cultivars. In this sense, the development of cultivar-specific SCAR markers is recommended for the standardization of the identification procedures due to its high reproducibility and specificity. The protection of an economically important cultivar justifies the use of the molecular marker technology together with morphological descriptors.

In theory, the low reproducibility of the amplification patterns of RAPD reactions does not indicate this methodology as the most suitable for future identifications of varieties. Microsatellites for varietal differentiation in cacao (Lanaud et al. 1999, Charters and Wilkinson 2000, Swanson et al. 2003) have been suggested as the most suitable. In the practice however, the wide genome coverage, low cost for cacao-producing countries and simultaneous evaluation of many loci make the RAPD an important tool, provided that previous studies on molecular characterization verify the coincidence between genetic and morphological divergence, as in our case. The development of SCAR markers, based on the RAPD, allows the achievement of highly reproducible specific markers. Cultivar-specific SCAR markers have already been developed successfully for several plant species (Zhang and Stommel 2001) and used for the identification of

varieties as in the case of grape (*Vitis vinifera* L.) (Vidal et al. 2000).

In the context of our search for an optimum primer combination for the identification of varieties and a determination of the minimum number of necessary primers to discriminate each one of the varieties, we have used the D statistics proposed by Tessier et al. (1999). Wilde et al. (1992) observed that a single RAPD primer was able to discriminate 10 clonal accessions of *T. cacao* L. unequivocally. We also found a primer (OPC17) that was able to discriminate the five cultivars simultaneously (Figure 1), and a total of 31 primers able to identify each one of the cultivars (Table 2).

The use of these primers of each one of the classes in any combination allows varietal identification and the construction of SCAR markers of high reproducibility, taking us one step closer to the standardization of procedures for a future identification of cacao clonal varieties.

ACKNOWLEDGEMENTS

This research was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG, Brazil).

Diferenciação de cultivares de cacau usando dados de componentes de produção e marcadores RAPD

RESUMO - A identificação de cultivares de cacau é crucial agora que novos clones estão disponíveis aos produtores. Comparou-se então a eficiência dos marcadores RAPD em relação à componentes de produção para a diferenciação de cinco cultivares ('Maranhão', 'Pará', 'Parazinho', 'ICS 1' polinização livre e a mistura híbrida comercial). Três componentes de produção avaliados durante 10 anos e 182 bandas polimórficas foram utilizados para análises de distância genética e escala multidimensional. Verificou-se estreita relação entre distância genética (DG) de marcadores RAPD e distância de Mahalanobis (DM) de dados de componentes de produção, e entre os biplots de DG e DM, embora somente marcadores RAPD discriminaram todos os cultivares. Os cultivares com diferentes níveis de melhoramento foram diferenciados pelos biplots. Para a exclusiva identificação de cada cultivar foram necessários 31 "primers". Para a futura identificação de clones de cacau esses primers de RAPD permitirão a construção de primers de SCAR clones-específicos de alta reprodutibilidade.

Palavras-chave: *Theobroma cacao* L., diferenciação de cultivares, marcadores RAPD, componentes de produção.

REFERENCES

- Becher SA, Steinmetz K, Weising K, Boury S, Peltier D, Renou JP, Kahl G and Wolff K (2000) Microsatellites for varietal identification in *Pelargonium*. **Theoretical and Applied Genetics** **101**: 643-651.
- Charters YM and Wilkinson MJ (2000) The use of self-pollinated progenies as 'in-groups' for the genetic characterization of cocoa germplasm. **Theoretical and Applied Genetics** **100**: 160-166.
- Dias LAS (1995) Biotecnologia e melhoramento genético do cacauero (*Theobroma cacao* L.). **Agrotrópica** **7**: 1-14.
- Dias LAS (1998) Análises multidimensionais. In: Alfenas AC (ed.) **Eletroforese de isoenzimas e proteínas afins: fundamentos e aplicações em plantas e microrganismos**. Editora UFV, Viçosa, p. 405-475.
- Dias LAS (2001) Contribuições do melhoramento. In: LAS Dias (ed.) **Melhoramento genético do cacauero**. FUNAPE, Viçosa, p. 493-529.
- Dias LAS and Kageyama PY (1997) Multivariate genetic divergence and hybrid performance of cacao (*Theobroma cacao* L.). **Brazilian Journal of Genetics** **20**: 63-70.
- Dias LAS and Kageyama PY (1998) Comparison between multivariate methods applied for the evaluation of genetic divergence in cacao (*Theobroma cacao* L.). Brazil. **Archives of Biology and Technology** **41**: 199-206.
- Dias LAS, Souza CAS, Augusto SG, Siqueira PR and Muller MW (1998) Performance and temporal stability analyses of cacao cultivars in Linhares, Brazil. **Plantations, Recherche, Développement** **5**: 343-355.
- Dias LAS, Rocha RB, Muro-Abad JI, Salomão TMF and Alfenas AC (2001) Cocoa cultivar distinctiveness and hybrid prediction using RAPD markers. **INGENIC Newsletter** **6**: 19-21.
- Dias LAS, Marita J, Cruz CD, Barros EG and Salomão TMF (2003) Genetic distance and its association with heterosis in cacao. **Brazilian Archives of Biology and Technology** **46**: 339-347.
- Dias LAS, Picoli EAT, Rocha RB and Alfenas AC (2004) A *priori* choice of hybrid parents in plants. **Genetics and Molecular Research** **3**: 356-368
- Doyle JJ and Doyle JL (1990) Isolation of plant DNA from fresh tissue. **Focus** **12**: 13-15.
- Faleiro FG, Lopes UV, Yamada MM, Pires JL, Bahia RCS, Santos RS, Gomes LMC, Araújo IS, Faleiro ASG, Gramacho KP, Melo GRP, Monteiro WR and Valle RR (2001) Caracterização de variedades clonais de *Theobroma cacao* L. com base em marcadores RAPD, AFLP e microssatélites. **Agrotrópica** **13**: 79-86.
- Figueira A (1998) Homonymous genotypes and misidentification in germplasm collections of Brasil and Malaysia. **INGENIC Newsletter** **4**: 4-8.
- Fu YB (2000) Effectiveness of bulking procedures in measuring population-pairwise similarity with dominant and co-dominant genetic markers. **Theoretical and Applied Genetics** **100**: 1284-1289.
- Grant ML, Miller DM and Culham A (2001) An investigation into the identity of the perennial cultivars of *Lavatera* L. using morphology and RAPD markers. **Acta Horticulture** **546**: 407-408.
- Hotelling HO (1931) The generalisation of Student's ratio. **Annals of Mathematical Statistics** **2**: 360-378.
- Lanaud C, Risterucci AM, Pieretti I, Falque M, Bouet A and Lagoda PJJ (1999) Isolation and characterization of microsatellites in *Theobroma cacao* L. **Molecular Ecology** **8**: 2141-2152.
- Lerceteau E, Robert T, Pétiard V and Crouzillat D (1997) Evaluation of the extent of genetic variability among *Theobroma cacao* accessions using RAPD and RFLP markers. **Theoretical and Applied Genetics** **95**: 10-19.
- Mahalanobis PC (1936) On the generalized distance in statistics. **Proceedings of National Institute of Science of the India** **12**: 49-55.
- Motilal L and Butler D (2003) Verification of identities in global cacao germplasm collections. **Genetics Resources and Crop Evolution** **50**: 799-807.
- Rocha RB, Muro Abad JI, Pires IE and Araújo EF (2002) Fingerprint and genetic diversity analysis of *Eucalyptus* spp. genotypes using RAPD and SSR markers. **Scientia Forestalis** **62**: 24-31.
- Roldán-Ruiz I, van Eeuwijk FA, Gilliland TJ, Dubreuil P, Dillmann C, Lallemand J, De Loose M and Baril CP (2001) A comparative study of molecular and morphological methods of describing relationships between perennial ryegrass (*Lolium perenne* L.) varieties. **Theoretical and Applied Genetics** **103**: 1138-1150.
- SAS Institute Inc. (1989) **SAS/STAT user's guide**. Version 6, 4th ed., vol. 2. SAS Institute, Cary, 846p.
- Soria V (1963) Observaciones sobre las variedades y cultivares de cacao en Bahia, Brasil. **Cacao** **8**: 1-6.
- StatSoft (1997) **STATISTICA for Windows**. <http://www.statsoft.com>
- Staub JE, Gabert A and Wehner TC (1996) Plant variety protection: a consideration of genetic relationships. **HortScience** **31**: 1086-1091.

- Swanson JD, Lee AC and Gultinan MJ (2003) USDA cacao DNA fingerprinting ring test: results from Penn State University. **INGENIC Newsletter 8**: 22-24.
- Tessier C, David J, This P, Boursiquot JM and Charrier A (1999) Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. **Theoretical and Applied Genetics 98**: 171-177.
- UPOV (1991) **International convention for the protection of new varieties of plants**. UPOV, Geneva. (Publication nr 221 E March 19).
- Vello F and Garcia JR (1971) Características das principais variedades de cacau cultivadas na Bahia. **Revista Theobroma 1**: 3-10.
- Vello F, Mariano AH, Garcia JR, Nascimento IF and Magalhães WS (1969). O programa de melhoramento genético do cacau na Bahia. In: **II Conferência Internacional de Pesquisas em Cacau**. CEPLAC, Ilhéus, p. 43-56.
- Vidal JR, Delavault P, Coarer M and Defontaine A (2000) Design of grapevine (*Vitis vinifera* L.) cultivar-specific SCAR primers for PCR fingerprinting. **Theoretical and Applied Genetics 101**: 1194-1201.
- Weatherup STC (1994) Use of Mahalanobis distance to measure varietal distinctness. **Plant Varieties and Seeds 7**: 107-119.
- Wilde J, Waugh R and Powell W (1992) Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. **Theoretical and Applied Genetics 83**: 871-877.
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA and Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. **Nucleic Acids Research 18**: 6531-6535.
- Wunsch A and Hormaza JI (2002) Cultivar identification and genetic fingerprinting of temperate fruit tree species using DNA markers. **Euphytica 125**: 59-67.
- Zhang Y and Stommel RJ (2001) Development of SCAR and CAPS markers linked to the *Beta* gene in tomato. **Crop Science 41**: 1602-1608.