Research Article

Genetic polymorphism among 14 elite *Coffea arabica* L. cultivars using RAPD markers associated with restriction digestion

Tumoru Sera[1], Paulo Maurício Ruas[2]*, Claudete de Fátima Ruas[2], Leandro Eugênio Cardamone Diniz[2], Valdemar de Paula Carvalho[2], Leandro Rampim[2], Eduardo Augusto Ruas[2] and Sheila Recepute da Silveira[2]

[1]*Instituto Agronômico do Paraná (IAPAR), Londrina, Paraná, Brazil.*
[2]*Universidade Estadual de Londrina, Londrina, Paraná, Brazil.*

Abstract

Knowledge of the genetic variability among genotypes is important for the transfer of useful genes and to maximize the use of available germplasm resources. This study was carried out to assess the genetic variability of 14 elite *Coffea arabica* cultivars using random amplified polymorphic DNA (RAPD) associated with a prior digestion of genomic DNA with restriction endonucleases. The accessions were obtained from the *Coffea* collection maintained at the Instituto Agronômico do Paraná (IAPAR), located in Londrina, Paraná, Brazil. Twenty-four informative RAPD primers, used in association with restriction enzymes, yielded 330 reproducible and scorable DNA bands, of which 224 (68%) were polymorphic. The amplified products were used to estimate the genetic variability using Dice's similarity coefficient. The data matrix was converted to a dendrogram and a three-dimensional plot using principal coordinate analysis. The accessions studied were separated into clusters in a manner that was consistent with the known pedigree. The associations obtained in the dendrogram and in the principal coordinate analysis plot suggest the probable origin of the Kattimor cultivar. The RAPD technique associated with restriction digestion was proved to be a useful tool for genetic characterization of *C. arabica* genotypes making an important contribution to the application of molecular markers to coffee breeding.

Key words: *Coffea arabica*, RAPD, restriction endonucleases.

Received: October 21, 2002; accepted: December 10, 2002.

Introduction

The wild coffee plant (*Coffea arabica* L.) is indigenous to Ethiopia. It was discovered in about 850 AD and was spread to other parts of the Islamic world by pilgrims. During the 17th century, coffee beverage consumption spread all over Europe, with great profit for the only producing country, the Yemen. In 1616, a Dutch trader succeeded in stealing seeds and in cultivating them in the colony of Java. Progenies from this first introduction were planted in the Surinam colony in South America (Berthaud and Charrier, 1988) and brought from there to French Guiana. The introduction of some coffee plants to Northern Brazil in 1727 was an important landmark in the economic activity of Brazil. In few years with the establishment of plantations in many states such as Rio de Janeiro, São Paulo, and Minas Gerais, Brazil became the main coffee producer and exporter in the world. Coffee was very important for the foundation and development of several new cities at the end of the nineteenth and during the last century. Nowadays, coffee is the second most important agricultural product of Brazil.

Until recently, genetic diversity among species or cultivars was determined with morphological or isozyme markers. However, these markers are often unsuitable to measure genetic variation. The analysis of six isozyme patterns in different *C. arabica* accessions revealed absence of polymorphism contrasting with the high level of morphological variation and suggesting that isozymes are not appropriate for the study of genetic diversity and for *C. arabica* accession identification (Berthou and Trouslot, 1977). Located in southern Brazil, Paraná state is an important coffee producer. In this state, the Research Center named Instituto Agronômico do Paraná (IAPAR) maintains a germplasm bank with eight *Coffea* species and several varieties, cultivars, and more than a thousand progenies of *C. arabica* and *C. canephora*. Despite its great importance, this *Coffea* collection lacks information about its genetic variability. Thus, much of the applied research is conducted without the genetic knowledge mainly when it concerns the DNA level.

Send correspondence to Paulo Maurício Ruas. Departamento de Biologia Geral, Universidade Estadual de Londrina, 86.051-990 Londrina, Paraná, Brazil. E-mail: ruas@sercomtel.com.br.

In the last ten years, techniques based on DNA markers have been used to detect variation at DNA level and have proven to be very effective for distinguishing between closely related genotypes. The development of the random amplified polymorphic DNA (RAPD) markers (Williams *et al.*, 1990; Welsh *et al.*, 1991) allowed genome characterization in several plant groups. The RAPD technique has been useful in studying polymorphism, identifying genes of interest, and characterizing genetic resources. In *Coffea,* RAPD markers were used to identify polymorphism in different coffee accessions (Orozco-Castillo *et al.*, 1994; Lashermes *et al.*, 1996) and to analyse spontaneous and subspontaneous *C. arabica* trees from Ethiopia (Anthony *et al.*, 2001). In this study, the RAPD technique were associated with restriction digestion of genomic DNA and used to identify the genetic polymorphism among 14 *C. arabica* cultivars acquired from the IAPAR *Coffea* Germplasm Collection.

Material and Methods

Fourteen *C. arabica* genotypes representing elite cultivars widely used in breeding programs were used in the current study. The accessions were obtained from the *Coffea* Germplasm Collection of the Instituto Agronômico do Paraná (IAPAR), Londrina, Brazil (Table 1).

Genomic DNAs were isolated from fresh leaf tissue following the CTAB method (Doyle and Doyle, 1987), except that CTAB was replaced by MATAB (Mixed Alyltrimethylammonium Bromide, Sigma) in the extraction buffer. DNA concentration was estimated using a fluorometer (DyNA Quant 200, Höefer-Pharmacia), according to the manufacturer's instructions. DNA samples, obtained from at least five different plants of each accession, were adjusted to 10 ng/μL. The individual DNA samples were bulked (Michelmore *et al.*, 1991) by accessions and used in the amplification reactions in a 15 μL volume containing 1x PCR buffer (75 mM Tris-/HCl, 50 mM KCl, 2.0 mM MgCl$_2$, and 20 mM (NH$_4$)$_2$SO$_4$); 0.2 mM each of dATP, dTTP, dCTTP, and dGTP; 0.4 μM primer (Operon Technologies); 0.9U *Taq* DNA polymerase (Biotools); and 20 ng template DNA. For restriction digestion, genomic DNA was incubated for 1 h with one of the following enzymes, *Bam* HI, *Eco* RI, or *Hae* III, which was added directly in the amplification reaction. DNA amplification was carried out using a PTC 100 (MJ Research) thermal cycler, programmed with 3 min at 94 °C for initial DNA denaturation, followed by 48 cycles of 1 min at 94 °C, 1 min 45 s at 38 °C, and 2 min at 72 °C. The final cycle was followed by a 7 min extension at 72 °C. The samples were stored at 4 °C until electrophoresis. Amplified products were resolved in 1.2 % agarose gel in TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0) at 120 V for 3 hs and stained with ethidium bromide. The RAPD profiles were visualized under UV light and photographed with a video camera.

Table 1 - Genotypes, accession numbers and origin types of 14 arabica coffee studied.

Genotypes	Accession number[b]	Type
1 - Mundo Novo IAC 376-4	93187	Improved line resulted from Bourbon Vermelho x Sumatra
2 - Catuaí Vermelho IAC-81[a]	92223	Caturra x Mundo Novo
3 - Caturra Vermelho	85135	Mutant of a Bourbon type coffee
4 - Caturra Amarelo	85134	Mutant of a Bourbon type coffee
5 - Villasarchi	75121	Mutant of Bourbon Vermelho
6 - Colombia Amarelo	93223	Catimor germplasm
7 - IAPAR 77028	93224	Sarchimor germplasm
8 - IAPAR-59	94013	Sarchimor germplasm
9 - Tupi LC 1669-33	94001	Sarchimor germplasm
10 - IAPAR 75163-21-10	95059	Sarchimor germplasm
11 - IAPAR 75163-12	94187	Sarchimor germplasm
12 - Kattimor	94190	Unknown origin
13 - F$_1$	83050	(IAPAR 59 x Mundo Novo IAC 376-4)
14 - F$_2$	92201	(IAPAR 59 x Mundo Novo IAC 376-4)

[a]IAC - Instituto Agronômico de Campinas.
[b] Number in the active germplasm collection at Instituto Agronômico do Paraná (IAPAR), Londrina, Paraná, Brazil.

Data analysis

DNA markers were scored for the presence (1) and absence (0) of homologous DNA bands. The genetic similarity among accessions was estimated using the Dice coefficient of the NTSYS package (Numerical Taxonomy and Multivariate Analysis for personal computer), version 2.1 (Rohlf, 2000). A dendrogram was constructed using the UPGMA (unweighted pair-group method using arithmetic averages) method. The matrix of genetic similarity was also used to obtain a principal coordinate analysis (PCOORDA) plot to resolve the patterns of variation among genotypes. The bootstrap method was employed to evaluate the reliability of tree topology. The calculations were performed with the BOOD software, version 3.0 (Coelho, 2001). The cophenetic coefficient between the matrix of genetic similarity and the dendrogram were computed using an appropriate routine of the NTSYS-pc software. The significance of the cophenetic correlation was tested by using the Mantel correspondence test (Mantel, 1967).

Results and Discussion

The RAPD technique associated with prior digestion of genomic DNA with restriction enzymes was used for detection of polymorphism in 14 elite *C. arabica* cultivars (Table 1). A total of 330 highly reproducible markers was analyzed of which 224 (68.0%) were polymorphic. The data matrix with the values of genetic similarities, the resulting dendrogram, and the plot of the principal coordinate analysis with the graphic distribution of the genotypes, are shown in Table 2 and Figures 1 and 2, respectively. The high value of cophenetic correlation (r = 0,91) between the matrix of genetic similarity and the dendrogram indicates the extent to which the clustering of genotypes accurately represented the estimates of genetic similarities among the coffee accessions studied.

The results revealed that the combination of restriction digestion and RAPD amplification of genomic DNA represented an excellent approach for the identification of polymorphism in elite *C. arabica* cultivars. According to Williams *et al.* (1990), restriction digestion of genomic DNA prior to PCR amplification simplifies the pattern of amplified products and changes the relative intensities of the bands. Moreover, digestion of genomic DNA may reveal a restriction site polymorphism located between the primer-binding sites in an otherwise monomorphic band. Koebner (1995) showed that in wheat, restriction digestion of genomic DNA with different endonucleases before PCR amplification improved the level of polymorphism. These results were probably due to a greater efficiency in primer annealing along shorter DNA fragments, where a simplified secondary DNA structure is less likely to interface with the process. In this way, different priming sites may become accessible depending on the restriction enzyme em-
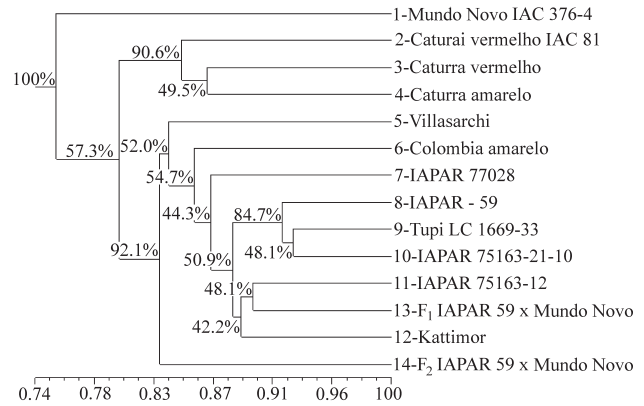
Figure 1 - UPGMA dendrogram of 14 *C. arabica* accessions based on Dice's genetic similarity. Numbers on branches are bootstrap values (%) generated after 1000 permutations.
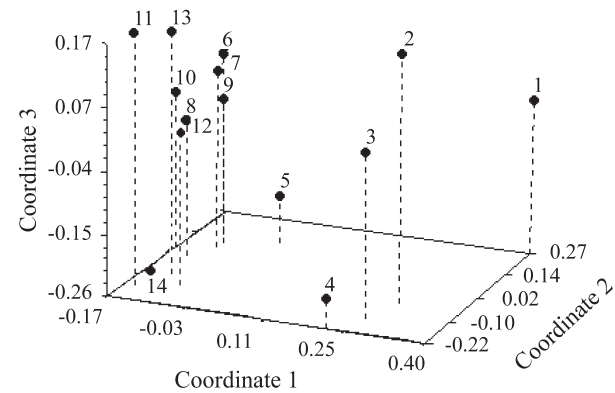
Figure 2 - Principal coordinate analysis of genetic distance among 14 elite *C. arabica* cultivars. The first, the second, and the third principal coordinates explain 22.6%, 11.6%, and 10.6% of the total variation, respectively. The numbers correspond to those listed on Table 2.

Table 2 - Genetic similarity based on the Dice coefficient applied to 14 *C. arabica* cultivars.

Accessions	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1-Mundo Novo IAC 376-4	1.00													
2-Catuaí Vermelho IAC-81	0.81	1.00												
3-Caturra Vermelho	0.81	0.87	1.00											
4-Caturra Amarelo	0.78	0.83	0.87	1.00										
5-Villasarchi	0.75	0.79	0.79	0.81	1.00									
6-Colombia Amarelo	0.74	0.79	0.80	0.77	0.82	1.00								
7-IAPAR 77028	0.73	0.80	0.82	0.79	0.86	0.85	1.00							
8-IAPAR-59	0.76	0.81	0.82	0.82	0.85	0.87	0.87	1.00						
9-Tupi LC 1669-33	0.78	0.81	0.80	0.80	0.86	0.86	0.89	0.92	1.00					
10-IAPAR 75163-21-10	0.76	0.80	0.82	0.80	0.85	0.88	0.88	0.93	0.93	1.00				
11-IAPAR 75163-12	0.71	0.80	0.81	0.79	0.80	0.83	0.85	0.88	0.86	0.90	1.00			
12-Kattimor	0.75	0.80	0.83	0.85	0.84	0.85	0.86	0.89	0.87	0.89	0.87	1.00		
13-F1[a]	0.74	0.81	0.81	0.80	0.82	0.85	0.86	0.89	0.88	0.89	0.90	0.89	1.00	
14-F2[b]	0.70	0.74	0.77	0.79	0.79	0.82	0.81	0.87	0.85	0.87	0.82	0.83	0.82	1.00

[a,b]IAPAR-59 x Mundo Novo IAC 376-4.

ployed (Koebner, 1995). As observed in wheat, the RAPD profiles obtained in *C. arabica* with certain primers were not affected by restriction digestion of the template DNAs. Other primers, however, produced distinct profiles for each restriction enzyme assayed (data not shown). The high level of polymorphism detected in the present study contrasts with previous results regarding to the use of molecular markers to detect genetic variation in *C. arabica*. Isozyme and RAPD analyses revealed extremely low polymorphism among *C. arabica* accessions (Berthou and Trouslot, 1977; Anthony, *et al.*, 2001). Nevertheless, Orozco-Castillo *et al.* (1994) successfully used RAPD markers to discriminate *Coffea* genotypes belonging to different subgroups. AFLP also exposed low genetic variability in arabica accessions, while a considerable level of polymorphic markers was revealed among Timor Hybrid-derived genotypes, which were associated to introgression of genes from the *C. canephora* gene pool (Lashermes *et al.*, 2000).

Genetic variability

The molecular profile obtained by combining restriction DNA digestion and RAPD was found efficient enough to reveal usable levels of DNA polymorphism among 14 elite *C. arabica* cultivars. The dendrogram (Figure 1) generated from the matrix of Dice's genetic similarity revealed two genetic groups comprising the Bourbon-derived genotypes and the IAPAR cultivars derived from the Sarchimor germplasm.

A relatively close relationship was revealed among the IAPAR cultivars, which clustered together with the values of genetic similarities ranging from 0.93, between IAPAR 75163-21-10 and the cultivars IAPAR 59 and Tupi LC 1669-33, to 0.85, between IAPAR 77028 and IAPAR 75163-12. These values are in agreement with the common origin of these cultivars since all of them were derived from selected progenies of the Sarchimor (IAC LC 1669) germplasm. Lower values of genetic similarity were found between the Sarchimor-derived genotypes and other accessions (Tables 1 and 2, Figures 1 and 2). The genetic similarity of the IAPAR cultivars ranged from 0.71 to 0.78 with the cultivar Mundo Novo, 0.80 to 0.81 with Catuaí Vermelho, 0.79 to 0.82 with the Caturra cultivars. As expected, there was a higher genetic similarity among cultivars of the same gene pool, such as the Sarchimor-derived genotypes, even though they still display enough usable genetic variation.

A relatively high variation was detected among the Bourbon-type coffee (Tables 1 and 2). The values of genetic similarities ranged from 0.79 between Villasarchi and both Catuaí Vermelho IAC 81 and Caturra Vermelho to 0.87 between the Caturra genotypes and between Catuaí Vermelho IAC 81 and Caturra Vermelho. The genetic similarity (0.87) between Caturra Vermelho and Caturra Amarelo points to the existence of genetic variation that is not reflected in plant morphologies. The Caturra coffee displays the compact architecture (CtCt) that was originated by spontaneous mutation in plants of Bourbon Vermelho (ctct) (Carvalho *et al.*, 1991). While the Caturra germplasms still demonstrate a heritability of 0.03 for productivity (Sera, 1980), Caturra Vermelho and Caturra Amarelo are very similar and can only be separated by the fruit color (red or yellow).

Mundo Novo IAC 376-4 was the most divergent cultivar, showing a genetic similarity ranging from 0.75 to 0.81 with cultivars belonging to the Bourbon gene pool and from to 0.71 to 0.78 with the cultivars of the Sarchimor germplasm. The closer relationship between the cultivars of the Bourbon-type coffee and Mundo Novo IAC 376-4 (Bourbon x Sumatra) is well supported by the origin of this cultivar. Moreover, these associations open the possibility for promising crosses between the Mundo Novo IAC 376-4 with selections of the Sarchimor germplasm and the Bourbon-derived genotypes that could result in hybrids with better performance and productivity. Fontes *et al.*, (2000) pointed out that high genetic divergence between *C. arabica* parents increased the possibility of combining different alleles, resulting in higher heterozigosity. Hybrid plants derived F_1 and F_2 generations of the Mundo Novo IAC 376-4 x IAPAR 59 cross were also included in this study. The F_1 showed similarity coefficients of 0.74 and 0.89 with Mundo Novo IAC 376 and IAPAR 59, respectively. Similar genetic similarities values were observed between the F_2 and Mundo Novo IAC 376 (0.70) and IAPAR 59 (0.87). The results of molecular data are in agreement with morphological characters. The F_1 hybrid displays a compact architecture and is resistant to all known leaf rust physiological races which are characters of the IAPAR 59 cultivar while Mundo Novo IAC-376-4 exhibits normal architecture and presents high susceptibility to leaf rust. The F_2 exhibits compact architecture for 75% of the plants and high productivity. The high performance of the F_1 and F_2 hybrids was also demonstrated with the genetic analysis that showed a heterosis of 25% for the F_1 and about a half for the F_2 plants (Sera, personal communication).

The considerable polymorphism detected in this study also illustrated that it is possible to find genetic divergence among arabica accessions of the same origin. Important characteristics have been incorporated into arabica germplasms. For instance, the leaf rust factors SH_6, SH_7, SH_8, SH_9, and SH (?) were selected in the F_3 and F_5 generations from crosses between *C. arabica* cv Villasarchi x Híbrido de Timor CIFC 832-2 (Bertrand *et al.*, 1999). The cultivar Villasarchi belongs to the Caturra germplasm and, as a Caturra variety, it is a mutant for small stature (CtCt). However, the RAPD analysis revealed a considerable low genetic similarity between Villasarchi and the Caturra-derived cultivars. On the other hand, the results are in conformity with the higher genetic similarities observed between Villasarchi and most selections of the Sarchimor

(Villasarchi x Timor Hybrid) germplasm (Table 1; Figures 1 and 2).

All of the arabica genotypes used in the present study were distinguished from each other. The examination of the dendrogram showed that the molecular data are consistent with pedigree information. For instance, the cultivar Colombia Amarelo is a multilinea that resulted from a mixture of several progenies of the Híbrido de Timor CIFC 1343 x Caturra CCC 135 which were selected for leaf rust resistance and other important agronomic traits, such as productivity, plant architecture, quality, and precocity for fruit maturation (Bertrand *et al.*, 1999). The estimative of genetic similarity revealed a close relationship between Colombia Amarelo and the Sarchimor-derived germplasm (IAPAR cultivars). On the other hand, the values of genetic similarity demonstrated that Colombia Amarelo is considerable different from the cultivars of the Bourbon germplasm (Table 2, Figures 1 and 2). These results suggest a higher incorporation of genes from the Híbrido de Timor gene pool.

Precise identification of arabica cultivars is required for variety registration, for preventing misappropriation, for protection of plant breeders, and for farmer's rights. Along with morphological and agronomical characters, the approaches that take DNA data into account are became valuable tools for defining distinctiveness, uniformity, and stability. The DNA data obtained in this study enabled the identification of a cultivar named Kattimor. In spite of its unknown origin, Kattimor is used as a commercial cultivar. It is supposed that Kattimor was selected from a mixture of best plants of the cultivars IAPAR 59, IAPAR-75163-21-10, and IAPAR 75163-12 (Sarchimor germplasms) and from plants that were originated by natural crossing with Catuaí Vermelho. Morphological traits, such as compact architecture together with RAPD data (Table 1, Figures 1, 2) indicate that Kattimor is very close to the IAPAR cultivars. Leaf rust resistance in 80% of the plants also characterizes it, which is a characteristic of Catuaí plants. Nevertheless, the molecular data (Table 1, Figures 1, 2) showed that Kattimor displays more genes from the Sarchimor germplasm than from the Catuaí cultivars.

Coffea arabica is the most important species for good quality coffee, responding for about 70% of the coffee traded in the world. Whereas many varieties and mutants of *C. arabica* are present in several *Coffea* germplasm collections, established in many countries throughout the world, it has been reported that the genetic base of the cultivated varieties is rather narrow (Berthaud and Charrier, 1988). Under this circumstance, one would expect a reduced variability even at DNA level. Nevertheless, the molecular markers were able to distinguish among all 14 arabica cultivars used in this study. A high degree of molecular polymorphism was also revealed, suggesting that even the elite cultivars, which have been exposed to intensive selec-

tion, still show potential for genetic studies and for use in breeding programs.

References

Anthony F, Bertrand B, Quiros O, Wilches A, Lashermes P, Berthaud J and Charrier A (2001) Genetic diversity of wild coffee (*Coffea arabica* L.) using molecular markers. Euphytica 118:53-65.

Berthaud J and Charrier A (1988) Genetic Resources of *Coffea.* In: Clarke RJ and Macrae R (eds) *Coffea* Agronomy, Elsevier Applied Science Publishers Ltd., Amsterdam, pp 1-42.

Berthou F and Troslot P (1977) L'analyse du polymorphisme enzymatique dans de genre Coffea, adaptation d'une methode d'electrophorese en serie: Eighth Conference of ASIC, Abidjan, Ivory Coast.

Bertrand B, Etiene H, Santacreo R, Anzueto F and Anthony F (1999) El mejoramiento genético en América Central. In: IAPAR and IRD (eds) Proceedings of the III International Seminar on Biotechnology in the Coffee Agro industry, Londrina.

Carvalho A, Medina-Filho HP, Fazuoli LC, Guerreiro Filho O and Lima MMA (1991) Aspectos genéticos do cafeeiro (Genetic aspects of the coffee tree). Revista Brasileira de Genética 14:135-283.

Coelho ASG (2001) BOOD - Avaliação de dendrogramas baseados em estimativas de distâncias/similaridades genéticas através do procedimento de bootstrap, v. 3.0. Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Goiás. Goiânia, GO.

Doyle JJ and Doyle JL (1987) A Rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytoch Bull 19:11-15.

Fontes JRM, Cardoso AA, Cruz CD, Pereira AA, Zambolim L and Sakiama NS (2000) Study of combining ability and heterosis in coffee. In: Sera T, Soccol CR, Pandey A and Roussos S (eds) Coffee Biotechnology and Quality, Kluwer Academic Publishers, Dordrecht, pp 113-121.

Ko MK, Yang J, Jin YH, Lee CH and Oh BJ (1998) Genetic relationships of *Viola* species evaluated by random amplified polymorphic DNA analysis. The J of Horticultural Sci & Biotech 73:601-605.

Koebner RMD (1995) Predigestion of DNA template improves the level of polymorphism of random amplified polymorphic DNAs in wheat. Genetic Analysis: Biomolecular Engineering 12:63-67.

Lashermes P, Combes MC, Topard P, Graziozi G, Bertrand B and Anthony F (2000) Molecular breeding in *Coffea* (*Coffea arabica* L.) In: Sera T, C.R. Soccol A, Pandey A and Roussos S (eds) Coffee Biotechnology and Quality, Kluwer Academic Publishers, Dordrecht, pp 101-112.

Lashermes P, Trouslot P, Anthony F, Combes MC and Charrier A (1996) Genetic diversity for RAPD markers between cultivated and wild accessions of *Coffea*. Euphytica 87:59-66.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209-220.

Michelmore RW, Paran I and Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci 88:9828-9832.

Orozco-Castillo C, Chalmers KJ and Powell RW (1994) Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. Theor Appl Genet 87:934-940.

Rohlf FJ (2000) NTSYS-pc Numerical Taxonomy and Multivariate Analysis System version 2.1. *Owner manual*.

Sera T (1980) Estimação dos componentes da variância e do coeficiente de determinação genotípica de produção de grãos de café (*Coffea arabica* L.). Master Thesis, Escola Superior de Agricultura Luis de Queiroz, Universidade de São Paulo, São Paulo, pp 62.

Welsh J, Peterson C and McClelland M (1991) Polymorphisms generated by arbitrarily primed PCR in the mouse: Application to strain identification and genetic mapping. Nucleic Acids Research 20:303-306.

Williams JGK, Kubelik AR, Livak KJ and Rafalski JA (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18:6531-6535.

Associate Editor: Marcio de Castro Silva Filho