

Varietal Identification of Coffee Seeds by RAPD Technique

Maria Lúcia Crochemore*, Liliane Moreira Nunes, Giselly Aparecida Andrade, Hugo Bruno Correa Molinari and Maria Elizabeth Vasconcellos

Instituto Agronômico do Paraná - IAPAR; C. P. 48; 86001-970; Londrina - PR - Brazil

ABSTRACT

*This study aimed the identification of cultivars and/or lines of *Coffea arabica* of commercial interest, using PCR-RAPD markers. The DNA of ground seeds lots of 12 cultivars and/or lines were evaluated with five primers (Operon OPA 01, OPA 04, OPG 11, OPY 16, and OPX 09) were obtained from a selection of 56 primers. The electrophoretic profiles allowed distinction among eight cultivars and/or lines as well as heterogeneity between and within lots of IAPAR59.*

Key words: *Coffea*, cultivar identification, molecular marker

INTRODUCTION

The institution of the Cultivar Protection Law (# 9456, 1997) is imposing profound changes regarding varietal identification and distinction. Such protection system is based on DUS test. Cultivars must be distinct (D), uniform (U) and stable (S), considering the characteristics used to distinguish them. DUS test involves the comparison of new cultivars to the already existing ones solely by the observation of phenotypic characteristics. If the genetic identity controls in the field are not efficient, and the morphological characteristics of the seeds are similar, varietal mixtures will not be perceived in the laboratory control. In fact, physical purity of a seed lot is easily detected by the seed analyst in the laboratory. However, as far as genetic purity is concerned, there will always be doubts when two or more cultivars present identical or very close characteristics, thus making the identification practically impossible.

Nowadays, techniques that utilize DNA structure as a tool for genotypes identification (lines, varieties, populations) are available. The use of these descriptors, which are neutral in relation to the environment, precise under the genetic point of view, and rapidly manipulated, reduces the time and labor of evaluation in characterization and identification. Because of that they eliminate any possible subjectivity that are common in the classic evaluation. Genetic fingerprint becomes an essential tool for all species of commercial interest as well as for the cultivars, as far as their DUS characterization is concerned.

In *C. arabica*, the genetic uniformity of populations is higher due to predominant autogamy. Levels up to 10% of cross-pollinating have been reported by Carvalho et al. (1991) and Anthony et al. (2001). Intraspecific variation was easily detected by RAPD markers, in *C. canephora* and *C. liberica*, whereas the primers assayed failed to reveal polymorphism between *C. arabica* accessions (Lashermes et al., 1993). RAPD markers did not detect any within-

* Author for correspondence

collection polymorphism in accessions derived from spontaneous and subspontaneous trees in Etiopia (Anthony et al. 2001). On the other hand, material originating from Ethiopia and the arabica sub-groups - *C. arabica* var. *typica* and *C. arabica* var. *bourbon* were clearly distinguished by RAPD markers (Orozco-Castillo et al., 1994). RAPD method appeared to be effective in resolving genetic variations between cultivated, wild accessions and in distinct groups of *C. arabica* germplasm, suggesting an East-West differentiation in Ethiopia, the primary centre of diversification of this species (Lashermes et al., 1996). The aim of this research was to identify and to distinguish lines and cultivars of *C. arabica* of commercial interest using RAPD markers.

MATERIALS AND METHODS

Seed lots of 12 different *C. arabica* commercial cultivars and/or lines were evaluated (Table 1). DNA was extracted from a single seed of each cultivar and/or line. Thirty milligrams of each seed were ground to a fine powder in liquid nitrogen. The powder was re-suspended in 600 µL of extraction buffer [100 mM Tris- HCl (pH 8.0), 50 mM EDTA (pH 8.0), 500 mM NaCl, 140 mM β-mercaptoethanol, 40 µL SDS 20%]. DNA was re-suspended in 30 µL of TE+RNAse, stored at -20°C temperature (Dellaporta et al., 1983), and quantified by fluorimetry (DyNa Quant 200 fluorimeter Hoefer).

The Amplification protocol was based on Williams et al. (1990). The PCR procedure was performed with 25 ng of genomic DNA, 2.5 mM buffer 10X [200 mM Tris (pH 8.4), 500 mM KCL], gelatin 0.01%, 1.5 mM MgCl₂, 1.25 mM dNTP, 5 picomols of primer (decamer oligonucleotide purchased from Operon Technologies Inc. - OPA 01, OPA 04, OPG 11, OPY 16 and OPX 09), 1 U of Taq polymerase and ultra pure water up to 25 µL. The amplifications were performed in a thermocycler PTC 100™ (MJ Research), using the temperature program described by Lashermes et al. (1996): 1 cycle of 4 min. at 94°C followed by 43 cycles of 1 min. at 94°C, 1 min. at 37°C and 2 min. at 72°C, with a final extension of 6 min. at 72°C. The amplifications were analyzed by electrophoresis at 80 V (5V/cm) in agarose gel (1.8%) in the presence of a 100 pb molecular weight marker.

Gel was stained in a 0.5 µg/mL ethidium bromide solution, visualized and photographed by the KODAK EDAS 120 system (Electrophoresis Documentation and Analysis System 120). The experiment was repeated once.

In order to verify the correlation between cultivars and markers, as well as the tendency of grouping, a Factorial Correspondence Analysis (FCA) was performed. Sokal and Michener coefficient similarity (Sokal and Michener, 1958) were calculated with data provided by the observation of presence (1) and absence (0) of bands between pairs of accessions using the formula $D_{(ij)} = a + d / (a+b+c+d)$, where: **a** is the number of fragments shared by accessions **i** and **j**; **b** is the number of fragments present in **i** and absent in **j**; **c** is the number of fragments present in **j** and absent in **i** and **d** is the number of fragments absents in **i** and **j**. A hierarchical classification (dendrogram) was obtained from the similarity matrix using the UPGMA (Unweighted Pair Group Mathematical Average) method. FCA e Sokal and Michener index were performed using SAS Statistical Program (SAS Institute, 1998).

Table 1 - *Coffea arabica* cultivars/lines evaluated (groups and genetic origin).

Cultivar/Line	Group	Origin
Icatu Precoce IAC 3282 - L101	Icatu	<i>C. canephora</i> x <i>C. arabica</i> x
Icatu Precoce IAC 3282 - L102	Icatu	<i>C. arabica</i> ²
Icatu Precoce IAC 3282 - L103	Icatu	
Catuaí Amarelo IAC 86	Catuaí	Caturra Amarelo x Mundo Novo
Mundo Novo IAC 388-17-1	Mundo Novo	Sumatra x
Mundo Novo IAC 388-6	Mundo Novo	Bourbon
Mundo Novo IAC 464 - 12	Mundo Novo	Vermelho
IAPAR 59 - L 104	Sarchimor	Villa Sarchi x Hít
IAPAR 59 - L 105	Sarchimor	Timor
IPR 99	Sarchimor	
Tupi IAC 1669-33	Sarchimor	
Obatã IAC 1669-20	Sarchimor	

RESULTS AND DISCUSSION

From 56 primers pre-selected in preliminary studies, four primers yielded differentiable amplification products that were able to generate polymorphic reproducible bands with sizes varying from 600 to 1400 pb, allowing the differentiation of cultivars/lines of *C. arabica*. Primers OPA 01, OPA 04, OPG 11 and OPY 16 generated 36 bands and among them six were polymorphic (16.7%). An example of banding patterns given by DNA amplification products after agarose gel electrophoresis is presented in Fig. 1. The number of bands produced per primer varied from 1 to 3 (Table 2).

The distribution of bands and cultivars in the first plane of FCA represented 65% of the total variation. The first and second dimensions of FCA absorbed respectively 41 and 24 %, of the total inertia (Fig. 2). Tupi was the only cultivar that presented the band OPA01-1250 bp and did not present the band OPY16-1400 bp.

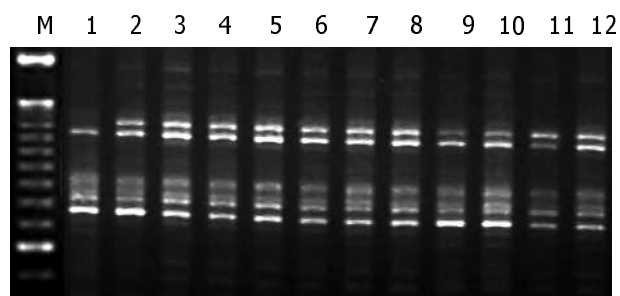


Figure 1 - Random amplified polymorphic DNA from 12 *Coffea arabica* cultivars using primer OPY-16. Lane M: 100bp molecular weight. Tupi IAC 1669-33 (1), Obatã IAC 1669-20 (2), Catuaí Amarelo IAC86 (3), IPR99 (4), Mundo Novo IAC 388-6 (5), Icatu Precoce IAC 3282-L103 (6), Icatu Precoce IAC 3282-L101 (7), Icatu Precoce IAC 3282-L102 (8), IAPAR-59-L105 (9), IAPAR-59-L104 (10), IAC Mundo Novo 464-12 (11), IAC Mundo Novo 388-17-1 (12).

Obatã was the only to present the band OPA01-1050 bp. The presence of the band OPA01-600 pb was only observed in both IAPAR59 lots and Icatu Precoce 3282 L101. It was important to notice that presence and absence of specific bands in some cultivars were considered only for the cultivars analysed in this study.

Sokal and Michener's coefficient similarity (1958) was used to determine genetic similarity, consider the double absence of bands. According to

Peltier et al. (1994), this mathematical method is recommended to analyse closer genotypes. In the dendrogram (Fig. 3), the genetic distance observed among cultivars and lines are shown.

Table 2 - Number of polymorphic bands and polymorphism rate obtained by four RAPD primers in the 12 *Coffea arabica* cultivars/lines analysed.

Primer sequence (5' - 3')	Number of bands	Bands Polymorphics	Level of polymorphism
OPA01 (CAGGCCCTTC)	10	3	30
OPA04 (AATCGGGCTG)	10	1	10
OPG11 (TGCCCGTCGT)	07	1	14
OPY16 (GGGCCAATGT)	09	1	12
Total	36	6	17

Tupi was not presented in Sarchimor group, as expected. It was genetically closer to Catuaí Amarelo IAC 86, Mundo Novo IAC 388-6 and Icatu Precoce 382-L103 and Icatu Precoce 3282-L102. Tupi did not present the same genetic origin of this four cultivars. These cultivars were not distinguished among one another however, such cultivars did not belong to the same genetic origin, therefore they were not genetically identicals, as showed. These results suggested that RAPD markers were not efficient enough to distinct *C. arabica* cultivars as detected by Lashermes et al. (1993,1996) and Anthony et al. (2001), who reported that the low level of molecular polymorphism detected in *C. arabica* cultivars was most probably related to the predominant autogamy, the recent allotetraploid origin of the species that involve a limited number of plants from which the cultivars were selected. On the other hand, Obata (Sarchimor) was the cultivar that showed the highest genetic distance, closer to Sarchimor group [IPR99, IAPAR59 (L105), IAPAR 59 (L104)], as expected. This grouping confirmed the same genesis (Villa Sarchi x Híbrido de Timor) of this cultivars.

The homogeneity of IAPAR 59 was evaluated with primer OPX 09 using 15 individualized seeds. The presence of polymorphic bands found within this cultivar indicated a low heterogeneity in the lot (data not shown). Differences were also detected between two IAPAR59 lots. Lot 104 presented bands OPA01-600 bp and OPA01-700

bp whereas lot 105 presented only OPA01-600 bp. This variability could be associated to the sampling, probable mixtures occurred during propagation in the field or in harvesting and processing of seeds. Also, the possibility of cross-pollinating in *arabica*, which can occurs up to 10% (Carvalho et al., 1991), could be considered as an important factor involved in this results.

It could be necessary to emphasize that *arabica* presented a narrow genetic base (Lashermes et al., 1993), indicating the need of techniques that could detect higher levels of polymorphism to obtain better distinction, primary for DUS cultivars evaluations.

RESUMO

Classicamente, a identificação de cultivares do cafeeiro utiliza descritores morfológicos da semente e da planta em crescimento. O uso de marcadores moleculares, além de permitir uma

identificação genética precisa, possibilita a caracterização dos genótipos com menor custo de mão-de-obra e de tempo, pois a descrição é realizada com pequenas quantidades de DNA genômico, extraído em qualquer etapa de crescimento da planta. Este estudo visou a identificação de cultivares/linhagens de *Coffea arabica*, de interesse comercial para o IAPAR, utilizando-se marcadores do tipo PCR-RAPD. O DNA de sementes moídas de 12 lotes de cultivares e/ou linhagens foi avaliado com cinco *primers* (Operon OPA 01, OPA 04, OPG 11, OPY 16 e OPX 09) obtidos de uma pré-seleção de 56 primers. Os perfis eletroforéticos mostraram diferenças entre oito cultivares/linhagens e uma baixa heterogeneidade foi detectada dentro de um lote de IAPAR59 e entre dois lotes deste mesmo cultivar.

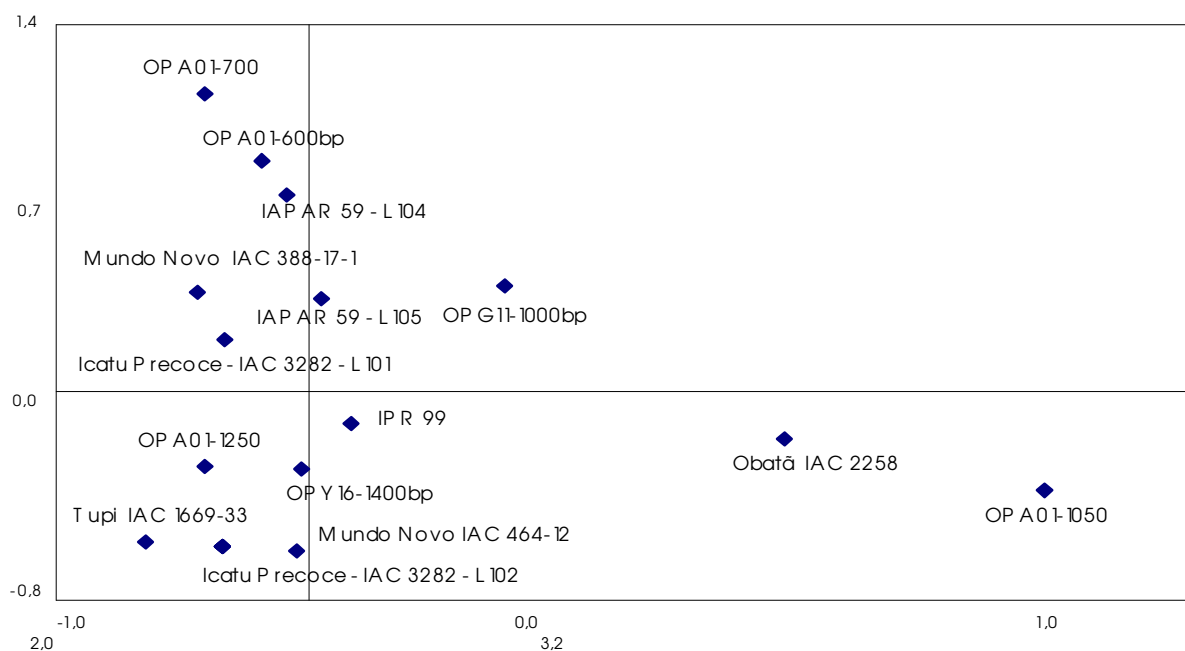


Figure 2 - First plane of a Factorial Correspondence Analysis based on 12 commercial lots of *C. arabica* cultivars and/or lines and 6 DNA/RAPD fragments

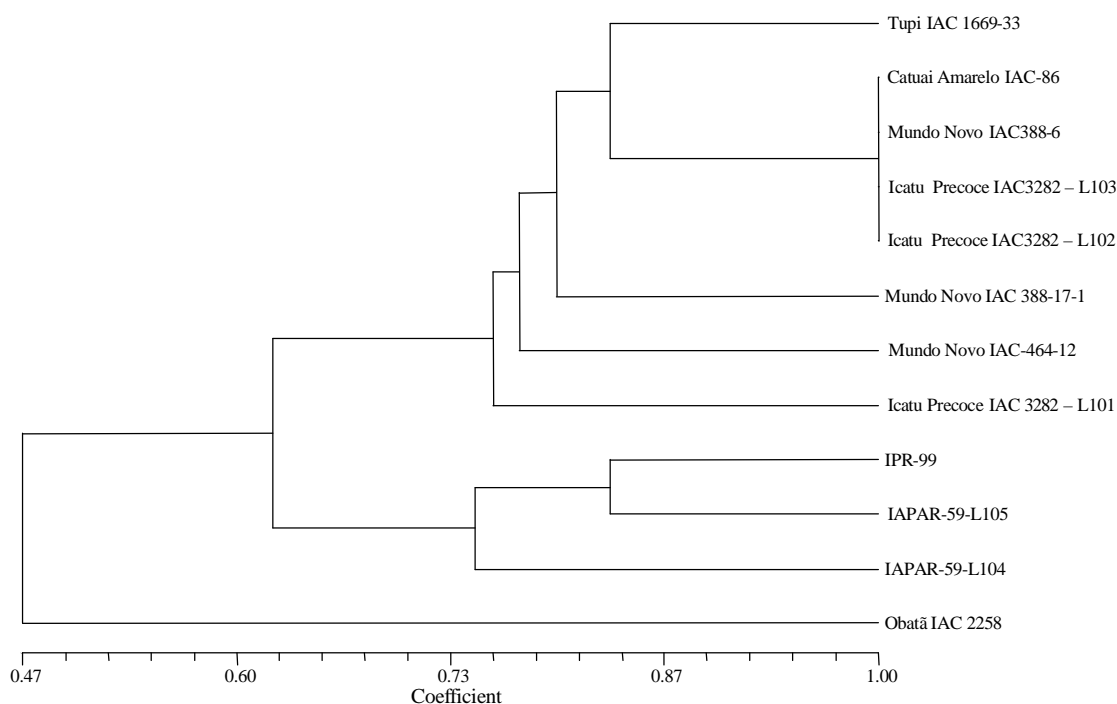


Figure 3 - Dendrogram of the 12 commercial lots of *Coffea arabica* cultivars and/or lines classified according to Sokal and Michener Similarity constructed by the UPGMA method

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.
- Carvalho, A.; Medina Filho, H. P.; Fazuoli, L. C.; Guerreiro Filho, O. and Lima, M. M. A. (1991). Aspectos genéticos do cafeeiro. *Rev. Brasil. Genet.*, **14**: (1), 135-183.
- Lashermes, P.; Trouslot, P.; Anthony, F.; Combes, M. C. and Charrier, A. (1996). Genetic diversity for RAPD markers between cultivated and wild accessions of *Coffea arabica*. *Euphytica*, **87**, 59-64
- Lashermes, P.; Berthaud, J. and Charrier, A. (2001). Genetic diversity of wild coffee (*Coffea arabica* L.) using molecular markers. *Euphytica*, **118**, 53-65.
- Orozco-Castillo, C.; Chalmers, K. J.; Waugh, R. and Powell, W. (1994). Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. *Theor. Appl. Genet.*, **87**, 934-940.
- Peltier, D.; Chacon, H.; Tersac, M.; Caraux, G.; Dulieu, H. and Bervill, A. (1995). Utilisation des RAPD pour la construction de phénogrammes et de phylogrammes chez *Petunia*. Techniques et utilisations des marqueurs moléculaires. Paris : INRA. In: *Les Colloques*, **72**, 188-202.
- SAS Institute. (1998), SAS/STAT User's Guide. Release 6. 12. ed. Cary, N.C., USA. SAS Institute Inc.
- Sokal, R. R. and Michener, M. A. (1958), Statistical method for evaluating systematic relationship. Univ. Kansas. *Sci. Bull.*, **28**, 1409-1438.
- Williams, J. G. K.; Kubelick, 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.

Received: December 04, 2001;

Revised: August 01, 2002;

Accepted: April 30, 2003.