Single-locus inheritance and partial linkage map of *Coffea arabica* L.

Terezinha Aparecida Teixeira-Cabral[1], Ney Sussumu Sakiyama[2*], Laércio Zambolim[2], Antonio Alves Pereira[3], and Ivan Schuster[4]

ABSTRACT - *In a backcross population of the allotetraploid* Coffea arabica *L. the loci with diploid-like segregation were predominant, although a few loci with tetrassomic inheritance or distortion of the expected segregation were also observed. A partial genetic map of* Coffea arabica *L. was constructed with 82 RAPD loci scored in this backcross population of 104 individuals. It covered the estimated length of 540.6 cM in eight linkage groups. The linkage group size was highly correlated with the number of markers, indicating random distribution of the markers in the groups. The average distance between two markers was 7.3 cM.*

Key words: coffee, molecular markers, allotetraploid, map

INTRODUCTION

Arabica coffee (*Coffea arabica* L.) is an allotetraploid (2n=44) and self-fertile species (Charrier and Berthaud 1985). It is probably originated from the diploid species *C. eugenioides* and *C. congensis* (Raina et al. 1998). *C. arabica* displays diploid-like meiotic behavior that can be observed in the segregation of many morphological traits (Krug and Carvalho 1951, Carvalho et al. 1991). However, tetrassomic inheritance was also observed in a segregant populations derived from Arabusta (*C. arabica* x *C. canephora*) (Lashermes et al. 2000, Herrera et al. 2002).

The development of molecular markers has opened new possibilities for genetic studies on Arabica coffee. An advantage of these markers is that they can be obtained in a greater number than the morphological markers, allowing the construction of genetic maps. A partial map for *C. arabica* was reported with 16 major and 15 small linkage groups, covering 1802.8 cM of the genome, with an average distance of 10.2 cM between adjacent markers (Pearl et al. 2004). Partial linkage maps of *C. canephora*, a diploid and self-incompatible species, were also constructed (Paillard et al. 1996, Lashermes et al. 2001). Another partial linkage map of the genus *Coffea* was obtained by using a backcross population derived from two diploid and self-incompatible species *C. pseudozanguebariae* and *C. liberica* var. *dewevrei* (Ky et al. 2000). The construction of linkage maps in *C. arabica* (allotetraploid and self-fertile species) was expected

to be difficult due to the low level of polymorphism (Paillard et al. 1993, Orozco-Castillo et al. 1994, Lashermes et al. 1996, 1999) and polyploidy (Paillard et al. 1996). This article described the single-locus inheritance analysis and the construction of a partial molecular linkage map for *C. arabica* L.

MATERIAL AND METHODS

Segregant population

A first backcross generation (BC$_1$) population of 104 coffee plants was obtained as follows:

Mundo Novo IAC 464-18 x Híbrido de Timor CIFC 2570

↓

H 464-2 x Híbrido de Timor CIFC 2570

↓

H 789 (104 individuals of BC$_1$)

Mundo Novo IAC 464-18 is a *Coffe arabica* cultivar and Híbrido de Timor CIFC 2570 was derived from a natural hybridization between *C. arabica* and *C. canephora*. Híbrido de Timor was used as recurrent parent, so that the dominant alleles present in the Mundo Novo cultivar could be analyzed in the backcross population by a dominant DNA marker.

DNA extraction

Young leaves were crushed in liquid nitrogen and transferred to Eppendorf tubes containing extraction buffer (sorbitol 0.35 M, Tris HCl 0.1 M pH 8.0, EDTA Na$_2$ 0.005 M), nuclear lysis buffer (Tris HCl 0.2 M pH 8.0, EDTA Na$_2$ 0.05 M, NaCl 2M, CTAB 2%), sodium bisulphate 0.02 M, Sarcosil 0.8 % (w/v) and activated carbon 1% (w/v). After incubation for one hour at 65 ^{0}C and a chloroform extraction procedure, the liquid phase was transferred to a new tube and the DNA was precipitated by an equal volume of cold isopropanol. The DNA was washed with cold 70% ethanol and dissolved in TE buffer containing RNAse (50 μg mL^{-1}).

DNA markers

Random Amplified Polymorphic DNA – RAPD (Williams et al. 1990), a dominant DNA marker, was obtained by the Polymerase Chain Reaction – PCR amplification and electrophoresis separation of DNA fragments. DNA fragment amplification was performed in a Perkin-Elmer 9600 thermocycler with 25 μL reaction volume, containing 25 ng genomic DNA, 1 unit AmpliTaq DNA polymerase, 0.1 mM dNTP, 0.2 M primer, 50 mM KCl, 10 mM Tris HCl pH 8.3, 2 mM MgCl$_2$, and ultrapure water. The PCR was performed with one cycle of DNA denaturation (95 ^{0}C for 1 minute), 39 cycles of DNA fragment amplification (a denaturation step of 15 sec at 94 ^{0}C, an annealing step of 30 seconds at 35 ^{0}C, and an extension

step of 60 seconds at 72 ^{0}C, and one final extension step of 7 minutes at 72 ^{0}C. The amplified DNA fragments were separated by 1.4% agarose gel electrophoresis, stained with ethidium bromide, and exposed to UV light for image capturing. The total of 680 primers (Operon Technologies kits OPA-OPZ, OPBA, and OPBH) were tested with both parents and F$_1$ plants. Only sharp polymorphic bands with bands derived from Mundo Novo IAC 464-18 were selected to produce RAPD markers that segregate in the backcross population. The reproducibility of these selected RAPD markers was tested with a sub-sample (12 plants) of the backcross population, parents, and F$_1$ plants. This test consisted in duplicate data of the RAPD markers. Reproducible markers only were scored for the entire population.

Data analysis

RAPD markers were recorded as dominant markers: the presence of the band corresponded to the dominant allele (from Mundo Novo IAC 464-18), and the absence of the band corresponded to the recessive allele (from Híbrido de Timor CIFC 2570). RAPD loci of the backcross population were tested for 1:1.5:1, and 2:1 segregation ratio using the chi-square (χ^2) test. Genetic aspects of Arabica coffee, an allotetraploid species, were discussed. A partial linkage map was constructed using the Mapmaker v. 3 software (Lander et al. 1987). A two point analysis identified linkage groups with maximum recombination values (θ) of 0.4 and a minimum 3.0 LOD score. The Kosambi function (Kosambi 1944) was used to convert the recombination rates into map distances (CentriMorgan - cM).

RESULTS AND DISCUSSION

Ninety-three polymorphic RAPD loci were obtained with 80 selected primers (1.2 markers per primer) and presented the following segregation ratios: 1:1 for 87 loci, 2:1 for four loci, and 5:1 for two loci, according to the chi-square (χ^2) test (P > 0.01). In a double-haploid population of *C. canephora*, 81% of RAPD and 85% of RFLP markers showed 1:1 segregation ratio (Paillard et al. 1996). In a backcross population of a tetraploid species, 1:1 ratio is expected for both disomic and tetrasomic inheritance if the F$_1$ plant carries only one dominant allele (i.e., Aa and Aaaa, respectively); 5:1 ratio is expected for tetrasomic inheritance, which is a result of either tetravalents formed by preferential pairing of homologues or bivalents formed by random pairing among the four chromosomes; and 2:1 ratio is a distortion of the 1:1 segregation ratio. The Aaaa F$_1$ genotype is possibly a result of a heterozygous locus of Mundo Novo IAC 464-18

with at least one recessive allele (i.e., AAAa, Aaaa, or Aaaa) since the recurrent parent Hibrido de Timor CIFC 2570 has only recessive alleles for these RAPD loci.

As expected for this allotetraploid species the majority of Arabica coffee RAPD loci (87 loci - 93.5% of the total) showed 1:1 ratio, confirming the morphological and cytogenetic evidences of its diploid-like meiotic behavior. Lashermes et al. (2000) observed that the segregation of 3 RFLP heterozygous loci in 14 F_2 plants was also in agreement with the expectations for disomic inheritance. The predominance of the disomic inheritance (87 out of 93 loci) strongly indicates that *C. arabica* may present genetically controlled preferential pairing of homologous chromosomes as in wheat (Riley et al. 1960, Feldman 1993), controlled by the gene *Ph*1 (homologous pairing).

The 5:1 ratio can be ascribed to tetrasomic segregation. A few RAPD loci showed 5:1 segregation ratio (four loci, 2.2%) in the backcross population. It is admitted that *C. arabica*, an allotetraploid species (Grassias and Kammacher 1975, Medina-Filho et al. 1984, Pinto-Maglio and Cruz 1998), may show tetrasomic segregation in addition to disomic segregation. Bivalents were predominant while tetravalents were rare in the meiosis of *C. arabica* (Pinto-Maglio and Cruz 1998). Nine out of 11 RFLP loci analyzed by Lashermes et al. (2000) followed tetrasomic inheritance in the arabusta interspecific hybrid (*Coffea arabica* x *Coffea canephora*). It was suggested that the preferential chromosome pairing in this hybrid might not occur, but random association instead. Further studies indicated that recombination in the tetraploid arabusta hybrid is not significantly affected by the genetic differentiation between chromosomes belonging to the different genomes and that the four sets of chromosomes have no difficulty in recombining (Herrera et al. 2002). In the present study, it might be expected that Híbrido de Timor CIFC 2570 derived from an interspecific cross between *C. arabica* and *C. canephora* could affect the chromosome pairing in the same way. However, a low number of tetrasomic inheritance loci was observed in the backcross population on focus. This fact could be a result of: *i*) Híbrido de Timor CIFC 2570 was derived from one or more natural backcross generations to *C. arabica* and have partially recovered the preferential pairing characteristic of this species; *ii*) a possible additional preferential pairing recovery in F_1 after crossing with the Arabica cultivar Mundo Novo IAC 464-18; *iii*) heterozygous loci of Mundo Novo IAC 464-18 parent, which would engender the same segregation ratio for disomic and tetrasomic inheritance in the backcross population; and *iv*) origin of the RAPD dominant alleles from the Mundo Novo IAC 464-18 cultivar. All 93 loci considered in this study are dominant (RAPD) and all dominant alleles were derived from the Arabica coffee parent Mundo Novo IAC 464-18. Since the Hibrido de Timor CIFC 2570 was derived from an interspecific cross between *C. arabica*
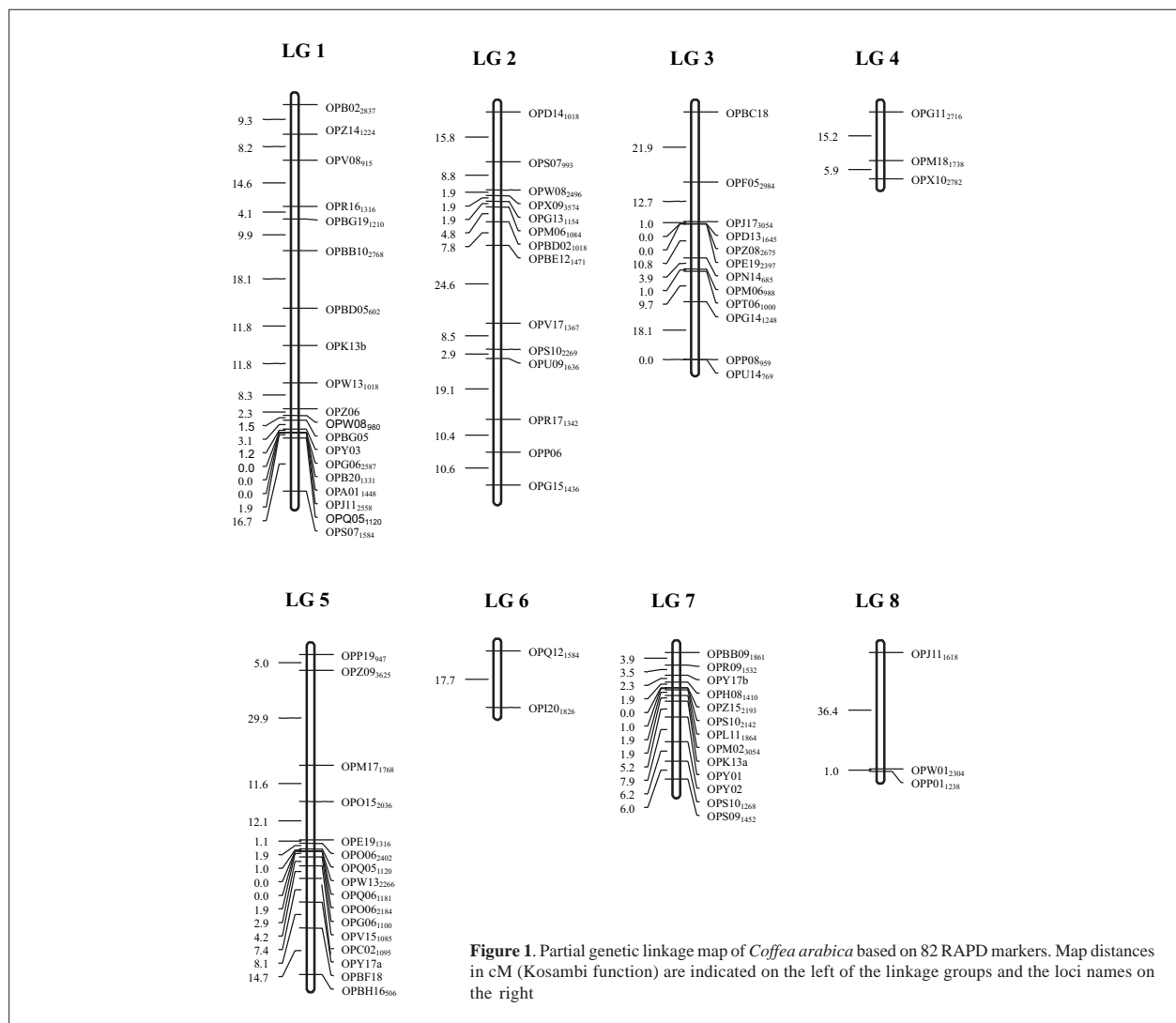
and *C. canephora* it is possible that the frequency of tetraploid inheritance is higher for the heterozygous loci that have dominant alleles from the Hibrido de Timor CIFC 2570. This hypothesis should further be tested by using a backcross population where Mundo Novo is used as recurrent parent, so that the dominant alleles found in the Híbrido de Timor parent can be analyzed with a dominant DNA marker.

There was an excess of heterozygous for four out of 93 markers (4.3%) that segregated 2:1, probably due to gametophytic selection or lethal zygotes. Plants derived from the backcross of arabusta F_1 hybrid to *C. arabica* also exhibited segregation distortion for three out of 11 RFLP loci (Lashermes et al. 2000) and one out of 13 microsatelite loci (Herrera et al. 2002).

There is therefore evidence of the predominance of the diploid-like meiotic behavior of *C. arabica*. The predominance of the disomic inheritance strongly indicates that *C. arabica* may present genetically controlled preferential pairing of homologous chromosomes. However, tetrasomic inheritance was also observed, which may either be due to tetravalents formed in the meiosis of *C. arabica* or of bivalents formed by random pairing among four homeologous chromosomes. A distortion of the expected segregation ratio occurred, probably due to the gametophytic selection or lethal zygotes. In spite of the worldwide importance of *C. arabica*, a lot of relevant genetic information about this species is still undiscovered. Efforts should be made to establish the genetic analysis of Arabica coffee for a better understanding of the inheritance pattern of the genes and its application in breeding programs, especially those that involve interspecific hybrids. In this context, DNA markers play an important role.

A partial molecular map with eight linkage groups covering 540.6 cM was constructed with 82 RAPD markers with a 1:1 segregation ratio (Figure 1). Five markers with the same segregation ratio showed no linkage. The linkage group size was highly correlated with the number of markers (r = 0.887), indicating random distribution of the markers in the groups, as observed also in *Coffea canephora* (r = 0.959) (Paillard et al. 1996). The minimum average distance between two markers was 3.47 cM in group 7 and the maximum 18.7 cM in 8. The highest interval between two markers was 36.4 cM in group 8, and 94.6% of the intervals were lower than 20 cM. The average distance between two markers was 7.3 cM, which was comparable to 10.2 cM reported for *C. arabica* (Pearl et al. 2004) and 10 cM for *C. canephora* (Paillard et al. 1996).

Arabica coffee is an allotetraploid and the mapping studies should be directed to identify 22 linkage groups (n=22) and the aspects related to the homeologous chromosomes. Linkage maps of related diploid species might be necessary for comparisons in order to distinguish homeologous chromosomes. A partial map for *C. arabica* was reported with 16 major linkage groups and

LG 1

9.3	OPB02$_{2837}$
8.2	OPZ14$_{1224}$
14.6	OPV08$_{915}$
4.1	OPR16$_{1316}$
9.9	OPBG19$_{1210}$
18.1	OPBB10$_{2768}$
11.8	OPBD05$_{602}$
11.8	OPK13b
8.3	OPW13$_{1018}$
2.3	OPZ06
1.5	OPW08$_{980}$
3.1	OPBG05
1.2	OPY03
0.0	OPG06$_{2587}$
0.0	OPB20$_{1331}$
0.0	OPA01$_{1448}$
1.9	OPJ11$_{2558}$
16.7	OPQ05$_{1120}$
	OPS07$_{1584}$

LG 2

15.8	OPD14$_{1018}$
8.8	OPS07$_{993}$
1.9	OPW08$_{2496}$
1.9	OPX09$_{3574}$
1.9	OPG13$_{1154}$
4.8	OPM06$_{1084}$
7.8	OPBD02$_{1018}$
	OPBE12$_{1471}$
24.6	OPV17$_{1367}$
8.5	OPS10$_{2269}$
2.9	OPU09$_{1636}$
19.1	OPR17$_{1342}$
10.4	OPP06
10.6	OPG15$_{1436}$

LG 3

21.9	OPBC18
12.7	OPF05$_{2984}$
1.0	OPJ17$_{3054}$
0.0	OPD13$_{1645}$
0.0	OPZ08$_{2675}$
10.8	OPE19$_{2397}$
3.9	OPN14$_{685}$
1.0	OPM06$_{988}$
9.7	OPT06$_{1000}$
18.1	OPG14$_{1248}$
0.0	OPP08$_{959}$
	OPU14$_{769}$

LG 4

15.2	OPG11$_{2716}$
5.9	OPM18$_{1738}$
	OPX10$_{2782}$

LG 5

5.0	OPP19$_{947}$
29.9	OPZ09$_{3625}$
11.6	OPM17$_{1768}$
12.1	OPO15$_{2036}$
1.1	OPE19$_{1316}$
1.9	OPO06$_{2402}$
1.0	OPQ05$_{1120}$
0.0	OPW13$_{2266}$
0.0	OPO06$_{1181}$
1.9	OPO06$_{2184}$
2.9	OPG06$_{1100}$
4.2	OPV15$_{1085}$
7.4	OPC02$_{1095}$
8.1	OPY17a
14.7	OPBF18
	OPBH16$_{506}$

LG 6

17.7	OPQ12$_{1584}$
	OPI20$_{1826}$

LG 7

3.9	OPBB09$_{1861}$
3.5	OPR09$_{1532}$
2.3	OPY17b
1.9	OPH08$_{1410}$
0.0	OPZ15$_{2193}$
1.0	OPS10$_{2142}$
1.9	OPL11$_{1864}$
1.9	OPM02$_{3054}$
5.2	OPK13a
7.9	OPY01
6.2	OPY02
6.0	OPS10$_{1268}$
	OPS09$_{1452}$

LG 8

36.4	OPJ11$_{1618}$
1.0	OPW01$_{2304}$
	OPP01$_{1238}$

Figure 1. Partial genetic linkage map of *Coffea arabica* based on 82 RAPD markers. Map distances in cM (Kosambi function) are indicated on the left of the linkage groups and the loci names on the right

15 small linkage groups, covering 1802.8 cM of the genome (Pearl et al. 2004). In *C. canephora* a genetic linkage map comprising 160 loci identified 11 linkage groups that putatively correspond to the 11 gametic chromosomes with the total length of 1041 cM (Lashermes et al. 2001).

The low level of polymorphism among Arabica coffee cultivars (Paillard et al. 1993, Orozco-Castillo et al. 1994, Lashermes et al. 1996, 1999) has been a bottleneck for mapping the genome. Híbrido de Timor CIFC 2570 was used here to increase the number of polymorphic markers to map the Arabica coffee genome. Besides, Híbrido de Timor was used as recurrent parent in the backcross to study the dominant alleles derived from Mundo Novo IAC 464-18, a standard Arabica coffee cultivar.

The partial linkage map of *C. arabica* presented here is a significant step towards the mapping of the entire genome using the strategy of the integrated map construction. This strategy was chosen due to the complexity of the genome and the low level of polymorphism in this species. The use of different populations might increase the number of polymorphic loci, besides the use of additional types of DNA markers, such as microsatellites. The integration of molecular linkage maps of *C. arabica* will allow geneticists and plant breeders to study and manipulate complex traits and obtain valuable information about the polyploid evolution process in plants.

ACKNOWLEDGMENTS

Herança de loco simples e mapa parcial de ligação de *Coffea arabica* L.

RESUMO - *Os locos com segregações típicas de diplóides foram predominantes em uma população de retrocruzamento do alotetraplóide* Coffea arabica *L., embora alguns locos com herança tetrassômica ou distorção da segregação esperada foram também observados. Um mapa parcial de ligação gênica de* Coffea arabica *L. foi construído com 82 locos de RAPD nesta população de retrocruzamento de 104 indivíduos. O mapa cobriu o comprimento estimado de 540.6 cM em oito grupos de ligação. O tamanho do grupo de ligação foi altamente correlacionado com o número de marcas, indicando uma distribuição aleatória das marcas nos grupos. A distância média entre duas marcas foi de 7,3 cM.*

Palavras-chave: café, marcadores moleculares alotetraplóide, mapa

REFERENCES

Carvalho A, Medina-Filho HP, Fazuoli LC, Guerreiro-Filho O and Lima MMA (1991) Aspectos genéticos do cafeeiro. **Revista Brasileira de Genética 14**:135-183.

Charrier A and Berthaud J (1985) Botanical classification of coffee. In: Clifford MN and Willson KC (eds.) **Coffee: botany, biochemistry and production of beans and beverage. The AVI Publishing, Westport, p. 13-47.**

Feldman M (1993) Cytogenetic activity and mode of action of the pairing homoeologous (Ph1) gene of wheat. **Crop Science 33**: 894-897.

Grassias M and Kammacher P (1975) Observations sur la conjugaison chromosomique de *Coffea arabica* L. **Café Cacao Thé 19**: 177-190.

Herrera JC, Combes MC, Anthony F, Charrier A and Lashermes P (2002) Introgression into the allotetraploid coffee (*Coffea arabica* L.): segregation and recombination of the *C. canephora* genome in the tetraploid interspecific hybrid (*C. arabica* x *C. canephora*). **Theoretical and Applied Genetics 104**: 661-668.

Kosambi DD (1944) The estimation of map distances from recombination values. **Annals of Eugenics 12**: 172-175.

Krug CA and Carvalho A (1951) The genetics of *Coffea*. **Advances in Genetics 4**: 127-158.

Ky CL, Barre P, Lorieux M, Trouslot P, Akaffou S, Louarn J, Charrier A, Hamon S and Noirot M (2000) Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp). **Theoretical and Applied Genetics 101**: 669-676.

Lander ES, Green P, Abrahamson J, Bralow A, Daly MJ, Lincoln SE and Newburg L (1987) MAPMAKER: An interactive computer package for construction primary genetic linkage maps of experimental and natural populations. **Genomics 1**: 174-181.

Lashermes P, Combes MC, Prakash NS, Trouslot P, Lorieux M and Charrier A (2001) Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. **Genome 44**: 589-596.

Lashermes P, Combes MC, Robert J, Trouslot P, D'hont A, Anthony F and Charrier A (1999) Molecular characterisation and origin of the *Coffea arabica* L. genome. **Molecular & General Genetics 261**: 259-266.

Lashermes P, Paczek V, Trouslot P, Combes MC, Couturon E and Charrier A (2000) Single-locus inheritance in the allotetraploid *Coffea arabica* L. and interspecific hybrid *C. arabica* x *C. canephora*. **The Journal of Heredity 91**: 81-85.

Lashermes P, Trouslot P, Anthony F, Combes MC and Charrier A (1996) Genetic diversity for RAPD markers between cultivated and wild accessions of *Coffea arabica*. **Euphytica 87**: 59-64.

Medina-Filho HP, Carvalho A, Sondahl MR, Fazuoli LC and Costa WM (1984) Coffee breeding and related evolutionary aspects. In: Janick J (ed.) **Plant breeding reviews**. The AVI Publishing, Westport, vol 2, p. 157-193.

Orozco-Castillo C, Chalmers KJ, Waugh R and Powel W (1994) Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. **Theoretical and Applied Genetics 87**: 934-940.

Paillard M, Duchateau N and Pétiard V (1993) Diversité génétique de quelques groupes de caféiers: utilisation des outils moléculaires, RFLP et RAPD. In: Association Scientifique Internationale du Café (ed.) **Proceedings of the 15th International Scientific Colloquium on Coffee**. ASIC, Montpellier, p. 33-40.

Paillard M, Lashermes P and Pétiard V (1996) Construction of a molecular linkage map in coffee. **Theoretical and Applied Genetics 93**: 41-47.

Pearl HM, Nagai C, Moore PH, Steiger DL, Osgood RV and Ming R (2004) Construction of a genetic map for Arabica coffee. **Theoretical and Applied Genetics 108:** 829-835.

Pinto-Maglio CAF and Cruz ND (1998) Pachytene chromosome morphology in *Coffea* L. II. *C. arabica* L. complement. **Caryologia 51**: 19-35.

Raina SN, Mukai Y and Yamamoto M (1998) In situ hybridization identifies the diploid progenitor species of *Coffea arabica* (Rubiaceae). **Theoretical and Applied Genetics 97**: 1204-1209.

Riley R, Chapman V and Kimber G (1960) Position of the gene determining the diploid-like meiotic behaviour of wheat. **Nature 186**: 59-160.

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.